Network: Diseases Due to Environmental Factors

Project: Sarcoidosis: Genome-wide Linkage Analysis and Association Mapping of Chromosomes 3p and 9q

Manfred Schürmann - University of Lübeck - schuerma@medinf.mu-luebeck.de

Introduction

Sarcoidosis is a multi-organ immune disorder of unknown aetiology. It is characterized by an exaggerated cellular immune response to one or more not yet identified agent(s), with increased inflammatory activity of macrophages and CD4 helper T cells, and accumulation to typical epithelioid cell granulomas. These structures resemble to some degree granulomas found in tuberculosis, but in sarcoidosis all attempts failed to detect the trigger within the immune cell clusters. Since the lung and adjacent lymph nodes are the primary site of affection, an air-borne and inhaled agent has been suspected to be the origin of sarcoidosis.

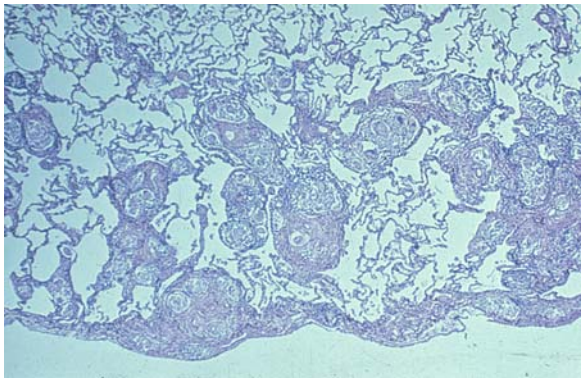

Fig 1: *Pulmonary sarcoidosis. Lung tissue with accumulation of typical granulomas. (Picture by courtesy of J. Müller-Quernheim, University of Freiburg)*

The annual incidence rate of sarcoidosis is approximately 10 – 15 / 100.000 in Germany, presumably with a considerable number of additional undiagnosed cases. In many patients is sarcoidosis a relatively benign self-limited disorder without need of medication. However, the clinical presentation of pulmonary sarcoidosis varies widely from acute inflammation (often in combination with arthritis, skin rush and good prognosis = Löfgren syndrome) to nonspecific illness with slowly progressive lung damage. Virtually every other organ can also be inflicted, especially the eyes, liver, and the skin. Life-threatening complications can result from cardiac sarcoidosis and affection of the central nervous system.

Sarcoidosis is mainly a disease of young and middle-aged adults, with a slight preponderance of females. Early onset sarcoidosis in children is a rare condition. Sarcoidosis is influenced by an inherited susceptibility of the patients. There is a significantly, up to twenty fold increased risk of sarcoidosis in close relatives of patients. The effect of the genetic background on the risk of sarcoidosis is also demonstrated by striking differences of disease prevalence and pattern of presentation between ethnic groups. For instance, severity of the disease and recurrence risk in families is more pronounced in US Americans of African origin than in US Whites. In Europe, sarcoidosis is more prevalent in the North, with a high proportion of Löfgren syndrome patients.

Sarcoidosis is less common than the other disorders of the disease oriented networks of the NGFN. With its hallmark of a unique immune reaction in the context of different clinical presentations and with a marked genetic susceptibility seems it to be a very promising condition to study the impact of predisposing genotype variability on the phenotype of the disease. Nevertheless, there is only limited genome-wide linkage information in sarcoidosis. The first whole genome scan, based on 138 German siblings with sarcoidosis, has been performed by our group [1], and the only second, in African Americans, has been published very recently. Numerous case-control association studies of possible functional candidate genes have been published, but many of them are conflicting. Most consistent evidence of a genetic contribution to the aetiology of sarcoidosis is linked to the major histocompatibility complex (MHC) gene clusters on the short arm of chromosome 6.

Project Status

We have performed in NGFN1 a high density single nucleotide polymorphism (SNP) mapping of the MHC gene region. We were able to identify a truncating splice site mutation of the BTNL2 gene as the primary cause of the major peak of the whole genome scan and as a new risk factor in sarcoidosis. An association of the neighbouring HLA-DR genes with sarcoidosis was confirmed as well by our results. Consecutive work clarified that BTNL2 and HLA-DR contribute independently to the risk of sarcoidosis [2]. These results – association with BTNL2 and independence from the HLA-DR effect – have recently been replicated and confirmed in a study of US Whites and African Americans as well. It is intriguing that it is not the BTNL2 splice site mutation but rather another variant of the gene that confers the BTNL2 associated risk in African Americans [3]. We have now analysed the BTNL2 and HLA-DR genotype data one by one and together with respect of differential association with various clinical presentations of sarcoidosis (phenotypes). There were clear differences that allowed us to propose a model of how the two risk factors interact (to be published).

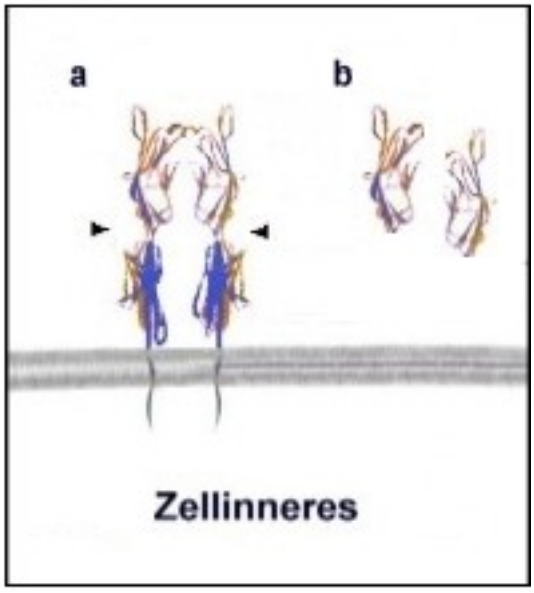

Fig 2: *Model of the C-terminal part of BTNL2 protein, (a) with and (b) without the transmembrane domain as a consequence of a truncating splice site mutation (see ref. 2).*

However, besides the obvious meaning of BTNL2 and HLA-DR in the aetiology of sarcoidosis, more susceptibility genes can be expected. For instance, a subgroup of nine more complex sarcoidosis families, each with three to six patients, contributed little to the whole genome linkage peak at the MHC gene region. Minor peaks of our scan, including those on chromosomes 3p and 9q were found in the study of African American siblings as well.

Chromosome 3p

The linkage peak region on chromosome 3p (cytogenetic band 3p21) contains a number of interesting candidate genes with a function in cellular immune response and cytokine networking. Among these, polymorphisms of the chemokine receptors CCR2 [4] and CCR5 have been reported as risk factors of sarcoidosis or markers of Löfgren's syndrome. Our attempt to replicate the CCR2 association failed, and our preliminary microsatellite linkage data argue for a more centromeric localisation of the linkage peak. In the context of discussion about the meaning of CCR2 as a marker of Löfgren's syndrome we have investigated the concordance of affected siblings with respect to sarcoidosis phenotypes. Though we found a significant degree of concordance, about one third of sibling pairs were discordant for acute versus chronic sarcoidosis [5]. This seems to show that either the effect of predisposing and modifying gene variants or the scope of sarcoidosis phenotypes is more complex.

Chromosome 9q

Another intriguing chromosomal candidate region is the whole genome scan linkage peak on the long arm of chromosome 9 (cytognetic band 9q33). One of the first genetic studies performed more than forty years ago in 518 sarcoidosis patients showed an increased risk associated with blood group A [6] that is encoded on chromosome 9 band q34. More recently, the toll like receptor (TLR) family has attracted attention as key molecules in innate and acquired immunity. TLR4 appears to be most interesting in the context of sarcoidosis, since it is involved in the recognition of mycobacteria and initiation of a granulomatous immune reaction. We have checked the loci of all members of the TLR gene family and found significant linkage with TLR4 on chromosome 9q33. However, a common missense mutation (asp299gly) that affects the extracellular domain of the TLR4 receptor and previously has been found in association with blunted response to inhaled lipopolysaccharide did not segregate with sarcoidosis in our families.

Whole genome analysis

In order to study the loci on chromosomes 3p and 9q in more detail and to identify further candidate loci, we performed a whole genome high density SNP association study by use of the chip genotyping technology (Affymetrix GeneChips 100 K set). The experiment involved 400 sarcoidosis patients and 400 healthy control individuals matched for sex and age. Care was taken to include approximately 200 unrelated patients with a family history of sarcoidosis. The rate of yield was close to 100 % and the total of approximately 80 million genotypes was analysed with respect to accordance with Hardy-Weinberg equilibrium (HWE), case-control differences on the level of allele, genotype and extended haplotype frequencies, etc. Applying cut-off conditions that incorporated significance of several parameters in the absence of violation of HWE, approximately 120 loci were identified besides the known MHC linkage peaks. These loci will now be investigated in detail by genotyping for additional SNP in the entire set of more than 2000 sarcoidosis patients. Due to the composition of the study population, affected sib pair analysis and transmission disequilibrium tests (TDT) will be possible as well.

Outlook

A great amount of genotype data is generated and analysed at the present state of the study. To our knowledge, there are no other comparably extensive studies of sarcoidosis on the way. With a total of more than 2000 patients involved, queries beyond the identification of mere associations can hopefully be answered. A major question to be addressed is to what extend is sarcoidosis a heterogeneous condition, with different causes leading to similar clinical consequences. Are different predisposing genes involved in the development of different sarcoidosis phenotypes? Can the sum of genotype information help to understand the crucial switches that at first lead to chronic versus acute disease and than - in some patients - to progressive lung fibrosis? To what extend can genotypes be used to predict the course of the disorder and to help making decisions about medication? Answers to these questions will have to be verified in an independent set of patients with very detailed clinical phenotype description and as possible long term follow up. Efforts to establish an international co-operation to this end, with inclusion of patients and controls with different ethic background are moving forward in a promising way.

*Lit.: **1.** Schürmann M et al. Results from a genome-wide search for predisposing genes in sarcoidosis. Am J Respir Crit Care Med. 2001 Sep 1;164(5):840-6. **2.** Valentonyte R et al. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. Nat Genet. 2005 Apr;37(4):357-64. **3.** Rybicki BA et al. The BTNL2 Gene and Sarcoidosis Susceptibility in African Americans and Whites. Am J Hum Genet. 2005 Sep;77(3):491-9. **4.** Spagnolo P et al. C-C chemokine receptor 2 and sarcoidosis: association with Lofgren's syndrome. Am J Respir Crit Care Med. 2003 Nov 15;168(10):1162-6. **5.** Valentonyte R 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 May 15;171(10):1136-41. **6.** Jörgensen G, Wurm K. ABO blood groups in sarcoidosis. Nature. 1964 Sep 5;203:1095.*