A Novel Sarcoidosis Risk Locus for Europeans on Chromosome 11q13.1

Annegret Fischer[1]*, Benjamin Schmid[1]*, David Ellinghaus[1], Michael Nothnagel[2], Karoline I. Gaede[3], Manfred Schürmann[4], Simone Lipinski[1], Philip Rosenstiel[1], Gernot Zissel[5], Kerstin Höhne[5], Martin Petrek[6], Vitezslav Kolek[7], Stefan Pabst[8], Christian Grohé[9], Johan Grunewald[10], Marcus Ronninger[11], Anders Eklund[10], Leonid Padyukov[11], Christian Gieger[12], H.-Erich Wichmann[13,14,15], Almut Nebel[1], Andre Franke[1], Joachim Müller-Quernheim[5], Sylvia Hofmann[1], and Stefan Schreiber[1,16,17]

[1]Institute of Clinical Molecular Biology and [2]Institute of Medical Informatics and Statistics, Christian-Albrechts University, Kiel, Germany; [3]Department of Pneumology, Research Center Borstel, Borstel, Germany; [4]Institute of Human Genetics, University of Lübeck, Lübeck, Germany; [5]Department of Pneumology, University of Freiburg, Freiburg, Germany; [6]Laboratory of Immunogenomics and Immunoproteomics and [7]Department of Respiratory Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic; [8]Medical Clinic II, Department of Pneumology, University of Bonn, Bonn, Germany; [9]Department of Respiratory Medicine, Evangelische Lungenklinik, Berlin-Buch, Germany; [10]Respiratory Medicine Unit and [11]Rheumatology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden; [12]Institute of Genetic Epidemiology and [13]Institute of Epidemiology I, Helmholtz Centre Munich, German Research Center for Environmental Health, Neuherberg, Germany; [14]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany; [15]Klinikum Grosshadern, Munich, Germany; and [16]Department of General Internal Medicine and [17]Popgen Biobank, University Hospital Schleswig-Holstein, Kiel, Germany

***Rationale*: Sarcoidosis is a complex inflammatory disease with a heterogeneous clinical picture. Among others, an acute and chronic clinical course can be distinguished, for which specific genetic risk factors are known.**

***Objectives*: To identify additional risk loci for sarcoidosis and its acute and chronic subforms, we analyzed imputed data from a genome-wide association scan for these phenotypes.**

***Methods*: After quality control, the genome-wide association scan comprised nearly 1.3 million imputed single-nucleotide polymorphisms based on an Affymetrix 6.0 Gene Chip dataset of 564 German sarcoidosis cases, including 176 acute and 354 chronic cases and 1,575 control subjects.**

***Measurements and Main Results*: We identified chromosome 11q13.1 (rs479777) as a novel locus influencing susceptibility to sarcoidosis with genome-wide significance. The marker was significantly associated in three distinct German case-control populations and in an additional German family sample with odds ratios ranging from 0.67 to 0.77. This finding was further replicated in two independent European case-control populations from the Czech Republic (odds ratio, 0.75) and from Sweden (odds ratio, 0.79). In a meta-analysis of the included European case-control samples the marker yielded a *P* value of 2.68×10^{-18}. The locus was previously reported to be associated with Crohn disease, psoriasis, alopecia areata, and leprosy. For sarcoidosis, fine-mapping and expression analysis suggest *KCNK4*, *PRDX5*, *PCLB3*, and most promising *CCDC88B* as candidates for the underlying risk gene in the associated region.**

***Conclusions*: This study provides striking evidence for association of chromosome 11q13.1 with sarcoidosis in Europeans, and thus identified a further genetic risk locus shared by sarcoidosis, Crohn disease and psoriasis.**

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

Sarcoidosis is a complex, granulomatous disease of the lung, for which several genetic risk loci are known. However, the fact that many genetic findings in sarcoidosis await replication or display conflicting results in independent populations complicates their interpretation and reduces their potential for clinical application.

What This Study Adds to the Field

Analysis of a genome-wide case-control genotype dataset identified a novel sarcoidosis risk locus on chromosome 11q13.1 with genome-wide significance, being an additional risk region shared with Crohn disease. The finding was replicated in four independent European sarcoidosis populations, which provides strong evidence for this new susceptibility locus. Initial functional studies on the transcripts encoded in this region revealed an allele-dependent expression of *CCDC88B* in bronchoalveolar lavage samples of patients with sarcoidosis, making the encoded protein (Gipie) an interesting functional candidate.

Keywords: sarcoidosis; genome-wide association study; imputation; Gipie; PRDX5

Sarcoidosis is a systemic inflammatory disease of unknown cause characterized by noncaseating epitheloid cell granulomas. It is a rare disease with a prevalence rate ranging from 1 to 60 per 100,000 inhabitants that decreases from northern to southern Europe. It mainly affects young adults (20–40 yr) with a slight

(*Received in original form April 18, 2012; accepted in final form July 5, 2012*)

* These authors contributed equally to this work.

Supported by grants of the Federal Ministry for Education and Research in Germany (BMBF) through the National Genome Research Network (NGFN), the Deutsche Forschungsgemeinschaft (DFG) through the Cluster of Excellence "Inflammation at Interfaces," grant no. FR 2821/2-1 and GenPhenReSa MU692/8-1, by the German Network for Diffuse Parenchymal Lung Disease (GOLD.net), in part by Palacky University IGA PU LF 2012_007, by the Swedish Heart-Lung Foundation, the Swedish Medical Research Council, and through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and the Karolinska Institutet.

Author Contributions: Study concept and design, A. Fischer and S.S. Acquisition of data and material support, all authors. Analysis and interpretation, A. Fischer, B.S., D.E., M.N., K.H., G.Z., and S.L. Drafting the manuscript, A. Fischer and B.S. Statistical analysis, A. Fischer, M.N., D.E., B.S., and G.Z. Critical revision of the manuscript for important intellectual content, all authors. Fundraising and supervision, S.S., S.H., A. Franke, J.M.-Q., and A.N.

Correspondence and requests for reprints should be addressed to Stefan Schreiber, M.D., Ph.D., Christian-Albrechts University, Institute of Clinical Molecular Biology, Schittenhelmstrasse 12, 24105 Kiel, Germany. E-mail: s.schreiber@mucosa.de.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 186, Iss. 9, pp 877–885, Nov 1, 2012

Originally Published in Press as DOI: 10.1164/rccm.201204-0708OC on July 26, 2012
Internet address: www.atsjournals.org

female preponderance (1). It is thought to be triggered by a complex combination of environmental and genetic factors with an estimated heritability of 66% (2). Genome-wide association studies (GWAS) and candidate studies identified a number of genetic risk loci, such as *BTNL2* (3–5), *ANXA11* (6–8), *Rab23* (9), the cytokine *TNF-α*, and most consistently several *HLA*-loci (for review see [10]).

Sarcoidosis is further characterized by great phenotypic heterogeneity with respect to the affected organ system and to the type of onset and the course of the disease. According to the latter, patients can be classified as being affected by acute or chronic sarcoidosis (11). The phenotypic heterogeneity of the disease is strongly reflected by the genetic susceptibility to certain sarcoidosis subphenotypes (12–14), especially to Löfgren syndrome (15–18).

With the present study, we aimed at the identification of further susceptibility loci for sarcoidosis, including subphenotype-specific risk loci, which may have been overlooked in a previous analysis of the same dataset, by now using imputed genotype data. With this approach we identified a novel locus influencing susceptibility to sarcoidosis on chromosome 11q13.1.

METHODS

Patients and Control Subjects

Patients with sarcoidosis were diagnosed according to international standards (11). When possible, patients with sarcoidosis were further classified as having chronic or acute sarcoidosis as previously described (6, 15). Subphenotype information was not available for the entire study population. All patients suffered from pulmonary sarcoidosis and all study participants were of German origin unless reported differently (*see* below panel C-III and C-IV). Analyses were performed in six independent panels (A, B, C-I, C-II, C-III, and C-IV), whereas panel D is derived from panels A + B. Figure E1 in the online supplement provides an overview. The following numbers refer to the population size before quality control (QC). The screening panel A comprised 646 sarcoidosis cases (including 194 acute and 414 chronically affected cases) and 1,770 control subjects. In a distinct validation panel B, genotypes of 1,530 patients with sarcoidosis, including 502 individuals with acute and 891 individuals with chronic sarcoidosis, were compared with those of 2,204 healthy individuals. Replication panel C-I comprised 307 patients with sarcoidosis and 285 control subjects from a sample described elsewhere (19) and had no overlap with any other panel. The family sample (nonoverlapping replication panel C-II) comprised 342 trios, comprising both parents with one offspring affected by sarcoidosis (12). Replication panel C-III consisted of 267 cases and 330 controls from the Czech Republic and substantially overlapped with the sample described elsewhere (8). The Swedish cohort (panel C-IV) comprised 1,066 cases and was complemented with 940 controls from the Swedish Epidemiological Investigation of Rheumatoid Arthritis study (20). Fine mapping was performed using panel D, with a total of 1,815 patients with sarcoidosis comprising all patients from panel B and parts of panel A, including 597 acute and 1,055 chronically affected patients, and the 2,204 control subjects from panel B. The online supplement provides a complete description of the study sample, including subphenotype information for the replication panels.

Genotyping, QC, Imputation, and Analysis of the GWAS Dataset

Genotyping for the screening stage (panel A) was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, CA). Individuals with more than 5% missing data were excluded (n = 195). We also excluded one individual from each pair of unexpected duplicates or relatives (n = 63) and individuals with outlier heterozygotic state of ± 5 SD away from the mean (five samples with heterozygotic state of <31% or >34%; mean heterozygosity for all individuals was 32.33%). The remaining individuals were tested for population stratification using principal components analysis (21) and 14 population outliers were excluded subsequently, leaving 2,139 samples (564 patients with sarcoidosis, including 176 acute cases and 354 chronic cases; 1,575 control subject) for analysis. Single-nucleotide polymorphism (SNP) imputation was performed using BEAGLE (22) and the HapMap3 reference samples from the CEU, TSI, MEX, and GJT cohorts. This approach allows for the inference of missing autosomal genotypes *in silico* based on linkage disequilibrium (LD), resulting in an extended dataset of 1,294,967 autosomal markers with at least moderate confidence (INFO score $r^2 > 0.3$) and a minor allele frequency greater than 1% in cases or in control subjects. Association analysis was performed using PLINK logistic regression framework for dosage data (23), including adjustment for potential confounding influences of population stratification (21). The online supplement provides complete information on imputation and analysis.

Genotyping, QC, SNP Selection, and Analysis in Validation, Fine Mapping, and Replication

Genotype data for validation, fine-mapping, and replication steps were generated using the Sequenom Mass-ARRAY iPlex (Sequenom, Inc., San Diego, CA) (24) and TaqMan (Applied Biosystems, Foster City, CA) technologies. In every analysis step, SNPs that had greater than 5% missing data (call rate [CR] ⩾ 95%), a minor allele frequency less than 1%, and exact Hardy-Weinberg equilibrium P less than 10^{-4} were excluded. Data filtering and statistical analysis of genotype data was performed using PLINK v.1.07 and single-marker allele-based association analyses were performed using χ^2 tests ($df = 1$) unless stated otherwise (23). The transmission disequilibrium test was used for family-based association analysis (25) and the Breslow-Day test was used to assess homogeneity of odds ratios (OR) in the different populations. SNPs to enter the validation stage were selected on ranking top with their P value in the association analysis and the presence of at least one "clumping" SNP and a positive visual inspection of regional plots (*see* Figure E2). HapMap tagging SNPs were selected for fine mapping using Haploview v.4.1 (26) with r^2 greater than 0.8. The online supplement provides additional information.

In Silico and Expression Analyses

For prediction of the functional consequences of certain SNPs, the regulatory potential was determined using mammalian genome sequences (27) as implemented in the UCSC genome browser (28). Information on amino acid changes was retrieved from dbSNP by the UCSC genome database (28). Further potential functional effects, including impact on splicing, on transcription factor and miRNA binding, and so forth, were investigated using the National Institute of Environmental Health Sciences (NIEHS) SNPinfo web server (29). The Genevar database and SNPexp were used to estimate potential allele-specific expression patterns from three cell lines and three tissue types (30) and from an independent set of in lymphoblastoid cell lines from the HapMap project (31), respectively.

Tissue-specific gene expression was investigated with an endpoint polymerase chain reaction (PCR) using a human cDNA tissue panel (Clontech, Palo Alto, CA), sequence-specific primers (*see* Table E1), and standard PCR conditions. Expression of selected candidate genes was studied in bronchoalveolar lavage (BAL) panel I comprising samples from healthy individuals and patients with sarcoidosis (each n = 12) that were selected for a high percentage of alveolar macrophages (>90%) and tested for differential expression using a nonparametric Mann-Whitney U test as implemented in the Graphpad statistical software (Graphpad, Inc., La Jolla, CA). A second series of BAL samples was obtained from 27 unselected patients with active sarcoidosis (BAL panel II). The percentage of alveolar macrophages varied from 7 to 96% in this panel. Real-time PCR was applied to measure relative transcript levels. Statistical analysis of the expression levels in the three different genotype groups was performed using StatView (SAS Institute, Cary, NC) using a Kruskal-Wallis rank test. The online supplement provides complete information.

RESULTS

GWAS Analyses

After imputation and the application of conservative and established quality filters to the data set (panel A; *see* METHODS section), 2,139 samples (564 sarcoidosis cases, 176 acute cases,

354 chronic cases, and 1,575 controls) and 1,294,967 SNPs were included in the GWAS. The measured population heterogeneity was moderate, with a genomic inflation factor of $\lambda_{GC} = 1.031$ based on a median χ^2 distribution (32), where $\lambda_{GC} = 1.0$ corresponds to no inflation (*see* Figure E3 for the corresponding quantile-quantile plot). Three distinct case-control analyses were conducted based on this imputed GWAS dataset: (*1*) sarcoidosis cases versus control subjects (denoted analysis on "overall" sarcoidosis from now on); (*2*) acute cases versus control subjects; and (*3*) chronic cases versus control subjects. Association signals for known sarcoidosis risk loci are given in the online supplement.

Validation of Lead Variants

Following the aforementioned selection criteria (*see* METHODS section), a total of 30 SNPs were selected for validation from the three analysis types (overall sarcoidosis: 13 SNPs; acute subphenotype: 6 SNPs; and chronic subphenotype: 11 SNPs) and were genotyped in the independent validation sample (panel B) using Sequenom technology (Table 1). After QC, 2,137 German control subjects and 1,486 patients with sarcoidosis, including 488 acute and 865 chronically affected patients, were included in the analysis. Three markers showed a nominally significant association in the case-control analysis, of which rs479777 from the "overall" analysis remained significantly associated after Bonferroni correction for multiple testing (corrected *P* value $[P_{corr}] = 9.79 \times 10^{-6}$; OR, 0.77; 95% confidence interval [CI], 0.70–0.85) (*see* Table 1 for results and Table E2 for complete results including genotype counts). No difference was observed for the acute and chronic subphenotypes in panel B (both OR, 0.78). In the meta-analysis of panels A and B, including all samples from the screening and validation stage, SNP rs479777 was associated with sarcoidosis with genome-wide significance ($P_{meta} = 2.05 \times 10^{-12}$; OR, 0.75). Based on the GWAS data (panel A), no significant SNP–SNP interaction of this SNP with known susceptibility variants for sarcoidosis in the *ANXA11* (rs2573351), *BTNL2* (rs2076530), *Rab23* (rs9396260), or the *IL23R* (rs11209026) gene loci was observed. The association signals of the lead SNPs identified in previous GWAS (6, 9, 33) from our group in the present dataset and the association signal of the novel loci in those datasets are given in the online supplement.

Replication in Four Independent European Populations

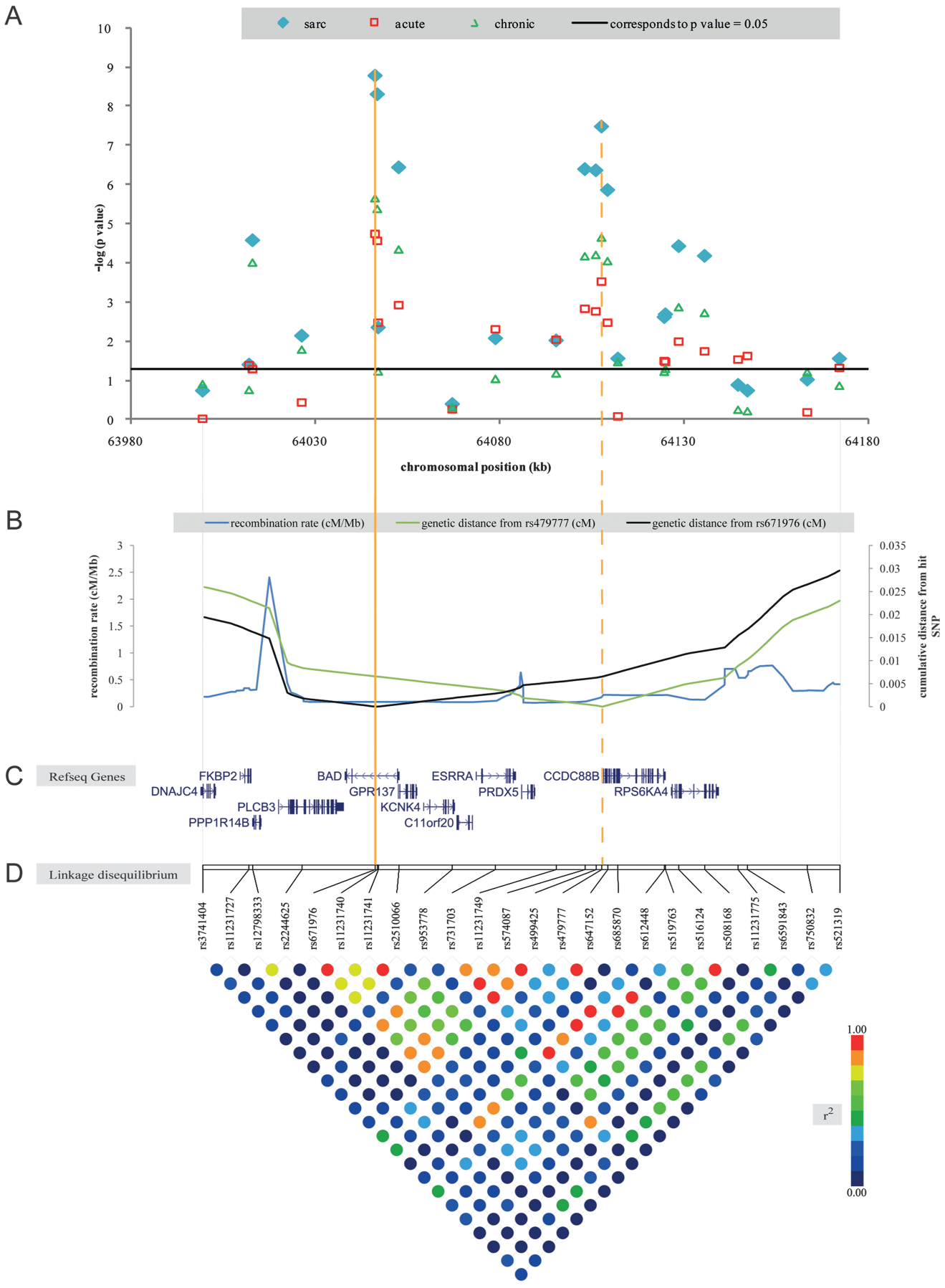
To substantiate this novel association, we analyzed SNP rs479777 in the European replication panels C-I to C-IV using TaqMan technology. In the German panel C-I genotyping was successful for 303 cases and 281 controls samples (CR = 98.6%). The marker was significantly associated with sarcoidosis in this panel ($P = 1.33 \times 10^{-3}$; OR, 0.67; 95% CI, 0.53–0.86) with minor allele frequencies of 0.29 in cases and 0.38 in control subjects. Of the 342 trios in the German family panel C-II, 11 displayed a mendelian error and were therefore excluded. Of the remaining 993 individuals, 31 failed genotyping (CR = 96.9%), leaving 301 affected offspring trios for analysis using transmission

TABLE 1. RESULTS FOR SINGLE-NUCLEOTIDE POLYMORPHISMS WITH NOMINALLY SIGNIFICANT *P* VALUES IN THE VALIDATION STAGE

Chr	Position (*bp*)	dbSNP ID	Analysis Type	Gene	A1	AF Ca GWAS	AF Co GWAS	P_{CCA} GWAS	OR (95% CI) GWAS	AF Ca val	AF Co val	P_{CORR} val	OR (95% CI) val
1	9,776,339	rs3934934	Overall	*PIK3CD, i*	G	0.14	0.11	5.06×10^{-6}	1.82 (1.41–2.35)	0.10	0.10	1	1.03 (0.89–1.20)
1	53,574,076	rs1679938	Overall	*SLC1A7, i*	G	0.28	0.23	2.74×10^{-5}	1.50 (1.24–1.81)	0.25	0.25	1	1.04 (0.93–1.16)
1	151,827,813	rs4845624	Overall	*THEM5, us*	G	0.37	0.32	4.71×10^{-5}	1.43 (1.20–1.70)	0.33	0.30	4.82×10^{-1}	1.13 (1.02–1.25)
2	27,083,489	rs6756245	Overall	*DPYSL5, i*	C	0.39	0.44	9.85×10^{-6}	0.69 (0.59–0.82)	0.42	0.42	1	1.00 (0.91–1.10)
2	203,266,610	rs16839127	Overall	*BMPR2, i*	A	0.09	0.06	2.92×10^{-5}	1.91 (1.41–2.59)	0.07	0.07	1	1.13 (0.94–1.35)
11	5,146,869	rs12577357	Overall	Intergenic	T	0.22	0.28	5.54×10^{-5}	0.67 (0.56–0.82)	0.26	0.26	1	1.01 (0.91–1.13)
11	**64,107,477**	**rs479777**	**Overall**	***CCDC88B, us***	**C**	**0.29**	**0.36**	$\mathbf{4.92 \times 10^{-6}}$	**0.67 (0.56–0.79)**	**0.30**	**0.36**	$\mathbf{9.79 \times 10^{-6}}$	**0.77 (0.70–0.85)**
11	82,065,238	rs7930387	Overall	Intergenic	G	0.23	0.19	4.20×10^{-5}	1.49 (1.23–1.80)	0.19	0.19	1	0.99 (0.88–1.11)
11	118,172,520	rs894678	Overall	Intergenic	A	0.31	0.38	3.11×10^{-5}	0.70 (0.59–0.83)	0.35	0.37	1	0.93 (0.85–1.03)
14	21,230,576	rs12895983	Overall	Intergenic	T	0.06	0.10	3.53×10^{-5}	0.53 (0.39–0.72)	0.10	0.08	2.49×10^{-1}	1.24 (1.06–1.46)
15	79,296,648	rs3825929	Overall	*RASGRF, i*	A	0.49	0.43	3.37×10^{-6}	1.46 (1.25–1.72)	0.45	0.43	1	1.07 (0.98–1.18)
17	27,871,025	rs17766675	Overall	*TAOK1, e*	G	0.24	0.19	5.35×10^{-5}	1.48 (1.22–1.72)	0.21	0.20	1	1.04 (0.92–1.16)
22	39,529,133	rs4820369	Overall	*CBX7, e*	A	0.18	0.13	4.09×10^{-5}	1.57 (1.26–1.94)	0.15	0.16	1	0.92 (0.81–1.05)
2	109,742,307	rs6542797	Acute	*SH3RF3, us*	C	0.33	0.41	1.86×10^{-4}	0.60 (0.46–0.78)	0.39	0.39	1	0.98 (0.85–1.13)
4	79,019,831	rs13106755	Acute	*FRAS1, i*	C	0.34	0.25	2.90×10^{-5}	1.81 (1.37–2.39)	0.27	0.28	1	0.95 (0.82–1.12)
5	6,710,771	rs274667	Acute	*POLS, us*	A	0.28	0.36	1.91×10^{-4}	0.59 (0.45–0.78)	0.36	0.37	1	0.92 (0.80–1.07)
7	84,294,717	rs10954758	Acute	Intergenic	C	0.49	0.40	1.63×10^{-4}	1.65 (1.27–2.14)	0.38	0.39	1	0.95 (0.83–1.10)
7	146,344,283	rs10262146	Acute	*CNTNAP2, i*	T	0.33	0.22	1.23×10^{-5}	1.84 (1.40–2.41)	0.20	0.21	1	0.93 (0.78–1.10)
16	54,963,926	rs7203657	Acute	*PNAS-108, ds*	T	0.30	0.21	4.78×10^{-6}	1.94 (1.46–2.57)	0.24	0.21	9.38×10^{-1}	1.20 (1.02–1.41)
1	162,139,692	rs7546353	Chronic	*NOS1AP, i*	C	0.23	0.18	2.03×10^{-4}	1.53 (1.22–1.91)	0.19	0.18	1	1.06 (0.92–1.23)
2	36,260,998	rs1387287	Chronic	Intergenic	T	0.11	0.07	4.36×10^{-5}	1.96 (1.42–2.72)	0.07	0.08	1	0.84 (0.68–1.04)
2	224,183,485	rs1517634	Chronic	Intergenic	G	0.30	0.23	2.06×10^{-5}	1.57 (1.27–1.93)	0.26	0.25	1	1.02 (0.90–1.16)
4	175,503,012	rs2250175	Chronic	Intergenic	A	0.52	0.46	7.61×10^{-5}	1.46 (1.21–1.75)	0.45	0.45	1	1.00 (0.90–1.12)
5	60,993,594	rs2030888	Chronic	*FLJ37543, i*	T	0.26	0.20	3.89×10^{-5}	1.59 (1.27–1.98)	0.20	0.23	1	0.88 (0.77–1.01)
8	33,853,868	rs7465297	Chronic	Intergenic	G	0.28	0.21	7.02×10^{-5}	1.54 (1.24–1.90)	0.22	0.21	1	1.03 (0.90–1.18)
10	46,146,746	rs17157941	Chronic	*ANUBL1, i*	C	0.06	0.03	3.50×10^{-5}	2.46 (1.61–3.78)	0.03	0.04	1.58×10^{-1}	0.63 (0.45–0.87)
11	5,168,952	rs6578556	Chronic	*OR52A1, us*	C	0.41	0.49	3.29×10^{-5}	0.67 (0.56–0.81)	0.48	0.47	1	1.02 (0.91–1.14)
11	107,420,192	rs638234	Chronic	*ABH8, i*	G	0.19	0.15	1.91×10^{-4}	1.58 (1.24–2.02)	0.15	0.16	1	0.98 (0.84–1.14)
15	58,877,599	rs1869133	Chronic	*ADAM10, us*	C	0.24	0.19	8.34×10^{-5}	1.55 (1.25–1.93)	0.22	0.21	1	1.08 (0.94–1.24)
19	17,182,017	rs7259348	Chronic	*HAUS8, i*	T	0.21	0.29	1.70×10^{-4}	0.66 (0.53–0.82)	0.27	0.29	1	0.92 (0.81–1.04)

Definition of abbreviations: Ca = cases; Chr = chromosome; CI = confidence interval; Co = controls; ds = downstream; e = exonic; GWAS = genome-wide association study; i = intronic; OR = odds ratio; us = upstream.

Thirty SNPs were successfully genotyped in the validation panel and tested for association with "overall" sarcoidosis (13 SNPs), and with the acute (six SNPs) and chronic (11 SNPs) subphenotypes in the validation panel B (val). *P* values obtained in the case-control analysis using an allele-based χ^2 test with one degree of freedom after Bonferroni corrected for multiple testing (n = 30) are listed (P_{CORR}). The significant result is highlighted in bold ($P < 0.05$). Nucleotide positions refer to human genome build 19. The analysis type refers to the phenotype, for which the SNP has been tested for association against controls. A1 denotes the less frequent allele in controls with the respective allele frequencies (AF). Allelic odds ratios and 95% confidence intervals for allele A1 are shown.


A
sarc
acute
chronic
corresponds to p value = 0.05
-log (p value)
chromosomal position (kb)
63980
64030
64080
64130
64180
B
recombination rate (cM/Mb)
genetic distance from rs479777 (cM)
genetic distance from rs671976 (cM)
recombination rate (cM/Mb)
cumulative distance from hit SNP
C
Refseq Genes
FKBP2
DNAJC4
PLCB3
PPP1R14B
BAD
GPR137
KCNK4
C11orf20
ESRRA
PRDX5
CCDC88B
RPS6KA4
D
Linkage disequilibrium
rs3741404
rs11231727
rs12798333
rs2244625
rs671976
rs11231740
rs11231741
rs2510066
rs953778
rs731703
rs11231749
rs574087
rs499425
rs479777
rs647152
rs685870
rs612448
rs519763
rs516124
rs508168
rs11231775
rs6591843
rs750832
rs521319
r2
1.00
0.00

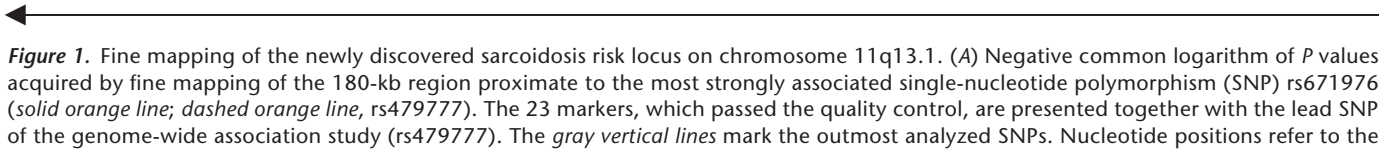


Figure 1. Fine mapping of the newly discovered sarcoidosis risk locus on chromosome 11q13.1. (*A*) Negative common logarithm of *P* values acquired by fine mapping of the 180-kb region proximate to the most strongly associated single-nucleotide polymorphism (SNP) rs671976 (*solid orange line*; *dashed orange line*, rs479777). The 23 markers, which passed the quality control, are presented together with the lead SNP of the genome-wide association study (rs479777). The *gray vertical lines* mark the outmost analyzed SNPs. Nucleotide positions refer to the human genome build 19. (*B*) The *blue graph* shows the recombination rate (cM per Mb) for the investigated region. The *green graph* depicts the genetic distance of the SNP rs479777 in centimorgan. The *black graph* shows the genetic distance for the hit SNP rs671976 (cM), all according to HapMap. (*C*) Intron and exon structure of the genes in this region according to Refseq. (*D*) Plot of pair-wise linkage disequilibrium between all markers analyzed in control individuals, measured by r^2.

disequilibrium test. The SNP was found to be significantly associated with sarcoidosis ($P = 1.95 \times 10^{-3}$; OR, 0.68; 95% CI, 0.55–87). We further replicated our novel finding in independent samples from the Czech Republic (panel C-III; CR = 98.7%; $P = 0.021$; OR, 0.75; 95% CI, 0.58–0.96) and Sweden (panel C-IV; CR = 96.7%; $P = 5.83 \times 10^{-4}$; OR, 0.79; 95% CI, 0.69–0.90). A subphenotype-specific analysis was inconclusive in replication panels C-I, C-II, and C-IV, whereas no subphenotype information was available for panel C-III. Strikingly, rs479777 was significantly associated in each of the tested populations with ORs ranging from 0.67 to 0.79 and yielded a *P* value of 2.68×10^{-18} in the meta-analysis of all available case-control panels (A, B, and C-I, -III, and -IV). The ORs obtained for the German (panel A, B, and C-I), the Czech (panel C-III), and the Swedish study populations (panel C-IV) did not differ significantly as examined with a Breslow-Day test for heterogeneity of the ORs. Tables E3 and E4 provide complete results including subphenotype analysis.

Fine Mapping of 11q13.1 (rs479777)

Based on the GWAS results, we defined an approximately 180-kb region around the lead SNP rs479777 as the region that will most likely harbor the underlying genetic risk factor (*see* Figure E2). To characterize this region in greater detail, we selected 25 tagging SNPs from HapMap in addition to the lead SNP rs479777 for fine mapping with panel D (*see* METHODS section). Twenty-four SNPs and 1,810 cases, including 596 acute and 1,050 chronically affected patients, and 2,182 control individuals passed the defined quality criteria. Nineteen markers yielded a nominal *P* value less than 0.05 in the association analysis with the strongest association signals for the lead SNP rs479777 ($P = 3.18 \times 10^{-8}$; OR, 0.76; 95% CI, 0.69–0.84) and rs671976 ($P = 1.60 \times 10^{-9}$; OR, 1.31; 95% CI, 1.20–1.43). The corresponding population attributable risk for the rs671976 risk allele was 0.27.

Figure 1 gives an overview of the association signals, the genes, recombination rates, and LD structure at the 11q13.1 locus, showing that the association signal is restricted to a 120-kb region, determined by a recombination hotspot at chromosomal position 64,200 kb and absence of association beyond 64,140 kb. Association results are given in Table 2; genotype counts and results for the acute and chronic subphenotypes are given in Table E5.

The locus shows a complex LD structure, with high LD between rs479777 and rs2510066, rs574087, rs499425, and rs647152 ($r^2 > 0.82$) in the control subjects of panel D. In contrast, only one marker (rs11231740) was in high LD with the most strongly associated SNP rs671976 ($r^2 = 0.98$). The two SNPs rs671976 and rs479777 showed only moderate LD with each other ($r^2 = 0.44$) in control subjects, but do not represent independent association signals, as shown by a likelihood-ratio test within a logistic regression model (data not shown). Haplotype analysis of the 10 most significantly associated SNPs revealed a significant difference in the haplotype frequencies between patients with sarcoidosis and control subjects, but none of the haplotypes had a smaller *P* value in the χ^2 test than rs671976 alone (data not shown).

TABLE 2. RESULTS OF THE FINE-MAPPING STAGE (PANEL D)

dbSNP ID	Position (*bp*)	Gene	A1	AF Ca	AF Co	P_{CCA}	OR (95% CI)
rs3741404	63,999,240	*DNAJC4, i*	G	0.33	0.32	1.76×10^{-1}	1.07 (0.97–1.17)
rs11231727	64,011,854	*PPP1R14B, us*	T	0.47	0.45	3.82×10^{-2}	1.10 (1.01–1.20)
rs12798333	64,012,784	*PPP1R14B, i*	A	0.18	0.15	2.55×10^{-5}	1.29 (1.15–1.45)
rs2244625	64,026,144	*PLCB3, i*	C	0.32	0.29	6.98×10^{-3}	1.14 (1.04–1.26)
rs671976	**64,046,029**	***BAD, i***	**G**	**0.53**	**0.46**	$\mathbf{1.60 \times 10^{-9}}$	**1.31 (1.20–1.43)**
rs11231740	64,046,641	*BAD, i*	T	0.52	0.45	4.78×10^{-9}	1.30 (1.19–1.42)
rs11231741	64,046,885	*BAD, i*	T	0.35	0.32	4.29×10^{-3}	1.15 (1.04–1.26)
rs2510066	64,052,447	*GPR137, 5′-UTR*	A	0.34	0.39	3.51×10^{-7}	0.79 (0.72–0.86)
rs953778	64,066,999	*KCNK4, nonsyn*	T	0.15	0.16	3.86×10^{-1}	0.95 (0.84–1.07)
rs731703	64,078,706	*ESRRA, i*	A	0.35	0.32	8.11×10^{-3}	1.13 (1.03–1.25)
rs11231749	64,095,178	Intergenic	C	0.36	0.33	9.21×10^{-3}	1.13 (1.03–1.24)
rs574087	64,102,948	Intergenic	C	0.34	0.40	3.89×10^{-7}	0.79 (0.72–0.86)
rs499425	64,105,929	*CCDC88B, us*	T	0.34	0.39	4.18×10^{-7}	0.79 (0.72–0.87)
rs479777	**64,107,477**	***CCDC88B, us***	**C**	**0.30**	**0.36**	$\mathbf{3.18 \times 10^{-8}}$	**0.76 (0.69–0.84)**
rs647152	64,109,118	*CCDC88B, nonsyn*	G	0.34	0.39	1.33×10^{-6}	0.80 (0.73–0.87)
rs685870	64,111,928	*CCDC88B, nonsyn*	T	0.31	0.28	2.67×10^{-2}	1.12 (1.01–1.23)
rs612448	64,124,515	*CCDC88B, syn*	G	0.44	0.47	2.34×10^{-3}	0.87 (0.80–0.95)
rs519763	64,124,823	*CCDC88B, 5′-UTR*	T	0.44	0.47	1.98×10^{-3}	0.87 (0.80–0.95)
rs516124	64,128,423	*RPS6KA4, i*	T	0.36	0.40	3.61×10^{-5}	0.83 (0.75–0.90)
rs508168	64,135,435	*RPS6KA4, i*	A	0.36	0.40	6.38×10^{-5}	0.83 (0.76–0.91)
rs11231775	64,144,525	Intergenic	C	0.41	0.42	1.26×10^{-1}	0.93 (0.85–1.02)
rs6591843	64,147,083	Intergenic	A	0.47	0.49	1.76×10^{-1}	0.94 (0.86–1.03)
rs750832	64,163,302	Intergenic	C	0.25	0.27	9.22×10^{-2}	0.92 (0.83–1.01)
rs521319	64,172,031	Intergenic	A	0.33	0.31	2.69×10^{-2}	1.11 (1.01–1.22)

Results for the 24 SNPs included in the fine-mapping stage. For description and abbreviations, see Table 1. In addition, the location or type of exonic SNPs are given (*5′-/3′-UTR*, synonymous [*syn*], nonsynonymous [*nonsyn*]).

The locus has been reported before to be associated with Crohn disease (CD), alopecia areata (AA), primary biliary cirrhosis (PBC), leprosy, and psoriasis (34–39). For CD, AA, and psoriasis marker rs694739 yielded the strongest signal in the associated region. Although this SNP is in high LD with rs479777 according to HapMap data ($r^2 = 0.84$) and is therefore represented in the fine-mapping marker set through LD, we analyzed this variant in the fine-mapping panel D using TaqMan technology and obtained a significant, but still weaker signal than for rs671976 ($P = 3.55 \times 10^{-6}$; OR, 0.80; 95% CI, 0.73–0.88). SNP rs538147, which showed the strongest association in the GWAS of PBC and leprosy, is not in LD with either rs47977 or rs671976, but was tagged by rs516124 in the fine-mapping approach. Because rs516124 obtained a much weaker signal in this analysis than rs671976, we did not follow-up rs538147 in our cohort.

In Silico and Expression Analysis

The associated region harbors eight genes, which could be affected by the causative variants in terms of a disturbed expression regulation. The most significantly associated SNPs, rs671976, rs11231740, rs479777, rs2510066, rs574087, rs499425, and rs647152, are located in regions with an unspecific high regulatory potential according to the UCSC genome browser. An *in silico* analysis that was extended to nine SNPs in high LD ($r^2 > 0.8$) using the NIEHS SNPinfo web server predicted a high regulatory potential for rs2510066 and rs663743 (*see* Table E6). The Genevar database revealed a potential *cis*-regulatory effect of the above mentioned SNPs (except rs479777, which was not included in the database) on the expression of the coiled-coil domain containing 88 (*CCDC88B*) gene in lymphoblastoid cell lines from two female twin pairs (40), with uncorrected *P* values ranging from 2.21×10^{-4} to 3.60×10^{-7}. Furthermore, rs647152 is a nonsynonymous variant located in exon 7 of the *CDCC88B* gene (193 Asp > Glu) that could influence CDCC88B protein function. Besides this, the analysis of an independent set of lymphoblastoid cell lines from the HapMap project using SNPexp suggested rs671976 and rs11231740 as eQTLs for potassium channel, subfamily K, member 4 (*KCNK4*) gene expression ($P = 4.24 \times 10^{-4}$), whereas less significant results were obtained for two peroxiredoxin 5 (*PRDX5*) transcripts ($P = 1.53 \times 10^{-3}$ and 1.34×10^{-3}, respectively). Expression analysis using a commercially available tissue panel detected mRNA of phospholipase C, β 3 (*PLCB3*), BCL2-associated agonist of cell death (*BAD*), G protein-coupled receptor 137 (*GPR137*), *KCNK4*, *PRDX5*, and *CCDC88B* in lymphocytes and in lung tissue, both of which are involved in sarcoidosis pathogenesis (Figure 2). Although *BAD*, *GPR137*, *KCNK4*, and *PRDX5* transcripts were detected in all of the 11 tissues under investigation, *CCDC88B* expression was only shown for leukocytes and spleen, thymus, pancreas, and lung tissue.

Based on the results of the *in silico* and expression analysis, we selected *PLCB3*, *KCNK4*, *PRDX5*, and *CCDC88B* to study differential expression of these genes in BAL samples from healthy individuals and patients with sarcoidosis (n = 12 each, matched for cell composition) using real-time PCR. Relative expression analysis showed that all genes were up-regulated in patients compared to controls (Figure 3). However, statistical analysis revealed significant differences only for *PLCB3*, *KCNK4*, and *PRDX5* expression. Because of the tissue-specific expression of *CCDC88B* and the predicted *cis*-regulatory effect of the associated SNPs on its expression, we selected *CCDC88B* for a more detailed expression analysis as a first step toward a mechanistic explanation of the observed genetic signal. To investigate the predicted *cis*-regulatory effect, we analyzed mRNA expression in BAL panel II, which consists of unselected BAL samples from 27 patients with sarcoidosis, and correlated it with the genotype status at SNP marker rs671976 of the respective patient (Figure 4). Statistical analysis revealed a positive correlation of the *CCDC88B* expression level with the number of rs671976 risk alleles (allele G; *P* value = 0.0187) in this sample. The percentage of alveolar macrophages in BAL panel II varied from 7 to 96% and was not correlated with the rs671976 genotype with a mean percentage of 51% (AA), 53% (AG), and 53% (GG).

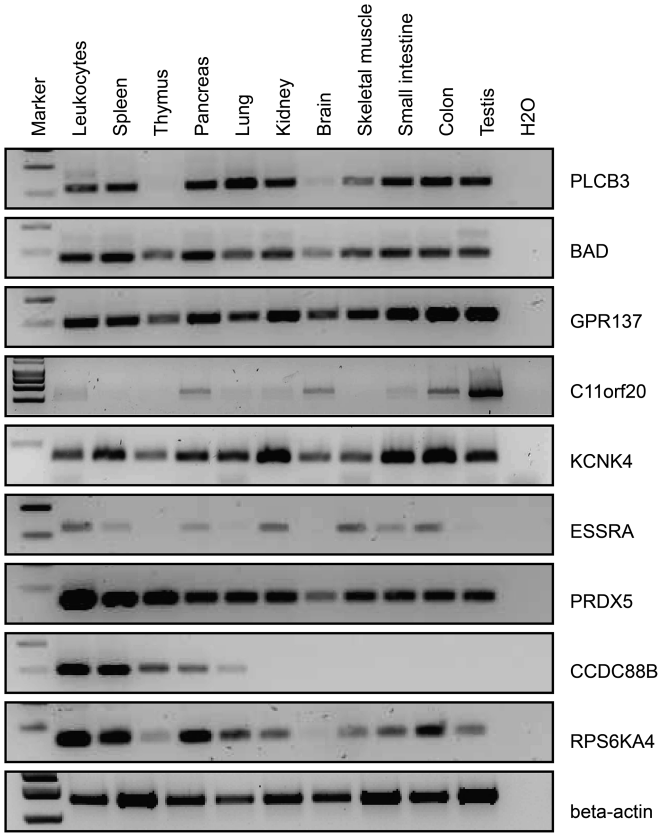


Figure 2. mRNA expression of candidate genes in the associated region on chromosome 11q13.1 in various human tissues and human cell lines.

DISCUSSION

The analysis of the imputed GWAS dataset identified a new risk locus for sarcoidosis with genome-wide significance on chromosome 11q13.1. The association of the respective SNP rs479777 was successfully replicated in four samples from Germany, the Czech Republic, and Sweden, confirming the finding in independent European populations. Interestingly, the locus has been reported before in association with CD, AA, PBC, leprosy, and psoriasis (34, 35, 37–39). It thus adds further evidence to a genetic basis shared by sarcoidosis, CD, leprosy, and psoriasis (12, 33, 38, 41–43), and to our knowledge it is the first report on a genetic overlap of sarcoidosis with AA and PBC outside the HLA region. Although in the GWAS on CD, AA, psoriasis, PBC, and leprosy the strongest association in the region was found for SNPs rs694739 and rs538147, respectively, fine mapping of the associated region in the German sarcoidosis case-control cohort revealed the strongest signal for markers rs479777 and rs671976. Both SNPs represent genetic markers that are in strong LD with the causative SNPs unless functional studies provide evidence for a direct disease-causing effect. For rs671976, the less frequent allele is associated with increased risk to sarcoidosis (OR, 1.31). This fits the widely accepted population genetic assumption that rare

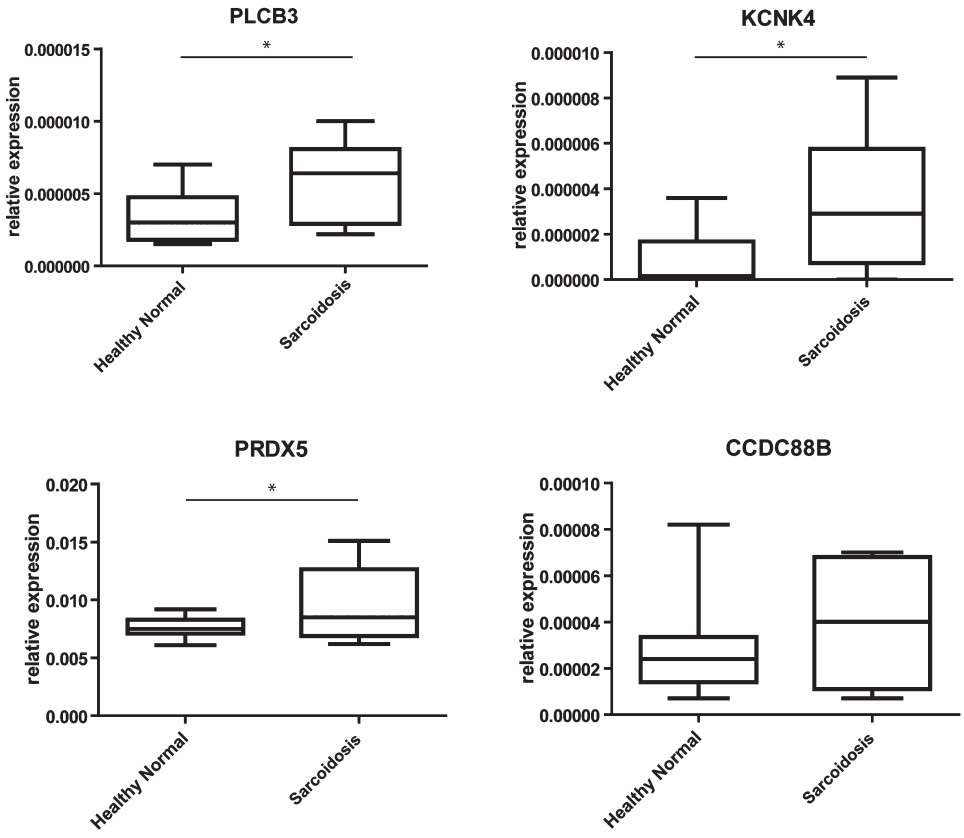


Figure 3. Expression of selected candidate genes in human bronchoalveolar lavage cells from patients with sarcoidosis and unaffected individuals matched by the proportion of alveolar macrophages (bronchoalveolar lavage panel I). *P value <0.05 in the Mann-Whitney U test.

variants are more likely to confer the genetic risk to rare diseases, such as sarcoidosis. In contrast, the less frequent allele of rs479777 reduces the risk for sarcoidosis (OR, 0.76). Therefore, we conclude that in our sarcoidosis sample rs671976 is more likely a causative variant underlying this signal or is in strong LD with the causative SNPs than rs479777. According to HapMap the allele frequencies of rs671976 and rs479777 vary substantially in different populations. Because replication was realized only in European populations, no conclusion can be drawn whether these SNPs confer risk to non-Europeans.

Although *PRDX5*, *ESRRA*, and *RPS6KA4* have been postulated as the most likely underlying risk factors for CD, psoriasis, AA, and leprosy mainly based on genomic proximity, the result of our fine-mapping approach identified the strongest signals at SNP markers rs671976 and rs479777, both of which are predicted eQTLs for *CDCC88B* gene expression. Previously, *CDCC88B* has been shown to be preferentially expressed in endothelial cells and macrophages (44). Here, expression analysis showed that *CDCC88B* expression was restricted to leukocytes and spleen, thymus, pancreas, and lung tissue. In the unselected samples of BAL panel II the level of *CCDC88B* gene expression was significantly correlated with the rs671976 genotype, whereas the percentage of alveolar macrophages in the samples of BAL panel II was not. This observation confirms the predicted *cis*-regulatory effect of rs671976 alleles on *CCDC88B* expression independent of the investigated cellular composition. In BAL panel I, which was matched for cell composition, expression of *CCDC88B* was elevated in diseased individuals compared with control subjects, albeit not significantly. However, this nonsignificant finding does not argue against a true *cis*-regulatory effect of rs671976 on *CCDC88B* expression, because this lack of significance could be caused by the small sample size ($n = 12$ each) or the undefined genetic background of this BAL sample. Altogether the expression analysis suggests *CCDC88B* as an interesting candidate for the underlying risk factor and may serve as a starting point for follow-up functional experiments. The corresponding protein, named GRP78-interacting protein induced by ER-stress (Gipie), is involved in the unfolded protein response and interacts with the lymphocyte-specific protein 1, cathepsin Z, and others (44).

Besides this, *in silico* analysis predicted an influence of the associated SNPs on *KCNK4* and *PRDX5* expression, both of which were detected in any tissue investigated and showed a higher expression in BAL from patients with sarcoidosis compared with those from matched healthy control subjects in BAL panel I. *KCNK4* (also termed TWIK-related arachidonic acid-stimulated K^+ channel) encodes a mechanosensitive two-pore domain K^+ channel (45), which seems to be involved in regulating airway mechanosensing in mouse (46). PRDX5 is a cytoprotective antioxidant enzyme acting against endogenous or exogenous peroxide attacks. Overexpression of the enzyme protects cells against death caused by nitrooxidative stresses (47). However, an impaired capacity to induce oxidative burst is known to cause granuloma in patients suffering from chronic granulomatous disease (48). Hence, a dysregulation of PRDX5 might affect the level of reactive oxygen species produced by macrophages and other phagocytic cells. Whether this scenario might either drive disease pathogenesis caused by granuloma formation or lead to oxidative stress as a hallmark of granulomatous inflammation in sarcoidosis requires further investigation. A further candidate for the underlying gene, the *PLCB3* transcript, was detected in leukocytes and in various tissues including lung tissue and showed elevated levels in sarcoidois BAL compared with control subjects. The phospholipase PCLB3 generates phospholipid second messengers for G-protein coupled receptor signaling and besides other functions, plays a critical role for T lymphocyte chemotaxis (49) and macrophage hypersensitivity to multiple inducers of apoptosis (50). Taking the results from the fine-mapping

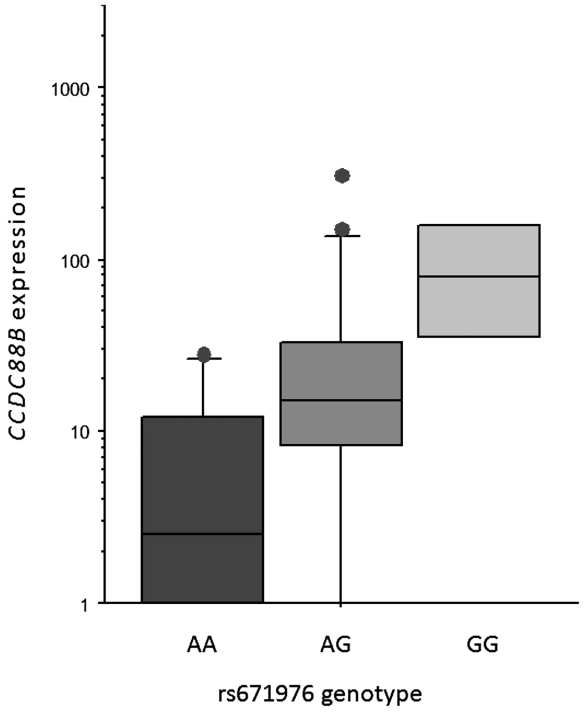


Figure 4. Allele-specific expression of *CCDC88B* mRNA in bronchoalveolar lavage (BAL) panel II. The expression of *CCDC88B* mRNA is positively correlated with the number of rs671976 risk alleles ("G" alleles; Kruskal-Wallis rank test; $P = 0.0187$) in BAL cells from 27 patients with active sarcoidosis (BAL panel II). The relative expression is given as a dimension-free ratio (see METHODS).

experiment, the expression studies, and the available literature, *CCDC88B*, *KCNK4*, *PRDX5*, and *PCLB3* are promising candidates for the underlying risk gene in the associated region. However, further functional studies are needed to explain disease-causing mechanisms and to elucidate the pathogenic mechanisms potentially shared by the various diseases that show a genetic association with this locus.

To account for the clinical heterogeneity of sarcoidosis, we included the analysis of the acute and chronic form of the disease in this GWAS. No subphenotype-specific genetic risk locus was identified with this approach, most likely because of a substantial reduction in sample size at both the screening and validation stage and a thereby reduced statistical power of the subphenotype analyses compared with the overall analysis. Therefore, our study does not argue against the existence of further genetic risk factors that are specific for the acute or chronic course of sarcoidosis, but rather points out the demand for additional genetic analyses of well-diagnosed sarcoidosis cohorts.

In summary, the results presented here provide striking evidence for association of SNP markers at chromosomes 11q13.1 with sarcoidosis in Europeans. Additional genetic analysis of the locus in non-European populations is now warranted to clarify if the novel risk variants predispose patients to sarcoidosis independent of ethnicity, and further experiments are needed to elucidate the role of risk variants on gene function and regulation in the context of sarcoidosis.

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank all patients, families, and physicians for their cooperation. The support of the German Sarcoidosis Patients Organization (Deutsche Sarkoidose-Vereinigung e.V.), the Popgen biobank, and the contributing pulmonologists is gratefully acknowledged. The authors also thank the staff of the Institute of Clinical Molecular Biology (Kiel, Germany) and The Swedish Epidemiological Investigation of Rheumatoid Arthritis group for technical assistance.

References

1. Müller-Quernheim J. Sarcoidosis: immunopathogenetic concepts and their clinical application. *Eur Respir J* 1998;12:716–738.
2. Sverrild A, Backer V, Kyvik KO, Kaprio J, Milman N, Svendsen CB, Thomsen SF. Heredity in sarcoidosis: a registry-based twin study. *Thorax* 2008;63:894–896.
3. Valentonyte R, Hampe J, Huse K, Rosenstiel P, Albrecht M, Stenzel A, Nagy M, Gaede KI, Franke A, Haesler R, *et al.* Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat Genet* 2005;37:357–364.
4. Li Y, Wollnik B, Pabst S, Lennarz M, Rohmann E, Gillissen A, Vetter H, Grohe C. BTNL2 gene variant and sarcoidosis. *Thorax* 2006;61:273–274.
5. Rybicki BA, Walewski JL, Maliarik MJ, Kian H, Iannuzzi MC, Group AR. The BTNL2 gene and sarcoidosis susceptibility in African Americans and whites. *Am J Hum Genet* 2005;77:491–499.
6. Hofmann S, Franke A, Fischer A, Jacobs G, Nothnagel M, Gaede KI, Schürmann M, Müller-Quernheim J, Krawczak M, Rosenstiel P, *et al.* Genome-wide association study identifies ANXA11 as a new susceptibility gene for sarcoidosis. *Nat Genet* 2008;40:1103–1106.
7. Li Y, Pabst S, Kubisch C, Grohe C, Wollnik B. First independent replication study confirms the strong genetic association of ANXA11 with sarcoidosis. *Thorax* 2010;65:939–940.
8. Mrazek F, Stahelova A, Kriegova E, Fillerova R, Zurkova M, Kolek V, Petrek M. Functional variant ANXA11 R230C: true marker of protection and candidate disease modifier in sarcoidosis. *Genes Immun* 2011;12:490–494.
9. Hofmann S, Fischer A, Till A, Muller-Quernheim J, Hasler R, Franke A, Gade KI, Schaarschmidt H, Rosenstiel P, Nebel A, *et al.* A genome-wide association study reveals evidence of association with sarcoidosis at 6p12.1. *Eur Respir J* 2011;38:1127–1135.
10. Müller-Quernheim J, Schürmann M, Hofmann S, Gaede KI, Fischer A, Prasse A, Zissel G, Schreiber S. Genetics of sarcoidosis. *Clin Chest Med* 2008;29:391–414.
11. Statement on Sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and other granulomatous disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee. *Am J Respir Crit Care Med* 1999;160:736–755.
12. Fischer A, Nothnagel M, Franke A, Jacobs G, Saadati HR, Gaede KI, Rosenstiel P, Schurmann M, Muller-Quernheim J, Schreiber S, *et al.* Association of inflammatory bowel disease risk loci with sarcoidosis, and its acute and chronic subphenotypes. *Eur Respir J* 2011;37:610–616.
13. Wijnen PA, Voorter CE, Nelemans PJ, Verschakelen JA, Bekers O, Drent M. Butyrophilin-like 2 in pulmonary sarcoidosis: a factor for susceptibility and progression? *Hum Immunol* 2011;72:342–347.
14. Pabst S, Franken T, Schonau J, Stier S, Nickenig G, Meyer R, Skowasch D, Grohe C. Transforming growth factor {beta} gene polymorphisms in different phenotypes of sarcoidosis. *Eur Respir J* 2011;38:169–175.
15. Fischer A, Valentonyte R, Nebel A, Nothnagel M, Muller-Quernheim J, Schurmann M, Schreiber S. Female-specific association of C–C chemokine receptor 5 gene polymorphisms with Lofgren's syndrome. *J Mol Med* 2008;86:553–561.
16. Spagnolo P, Renzoni EA, Wells AU, Sato H, Grutters JC, Sestini P, Abdallah A, Gramiccioni E, Ruven HJ, du Bois RM, *et al.* C–C chemokine receptor 2 and sarcoidosis: association with Lofgren's syndrome. *Am J Respir Crit Care Med* 2003;168:1162–1166.
17. Grunewald J, Idali F, Kockum I, Seddighzadeh M, Nisell M, Eklund A, Padyukov L. Major histocompatibility complex class II transactivator gene polymorphism: associations with Lofgren's syndrome. *Tissue Antigens* 2010;76:96–101.
18. Kieszko R, Krawczyk P, Chocholska S, Dmoszynska A, Milanowski J. TNF-alpha and TNF-beta gene polymorphisms in Polish patients with sarcoidosis. Connection with the susceptibility and prognosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2010;27:131–137.
19. Pabst S, Golebiewski M, Herms S, Karpushova A, Diaz-Lacava A, Walier M, Zimmer S, Cichon S, Wienker TF, Nothen MM, *et al.* Caspase recruitment domain 15 gene haplotypes in sarcoidosis. *Tissue Antigens* 2011;77:333–337.
20. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, Ronnelid J, Harris HE, Ulfgren AK, Rantapaa-Dahlqvist S, *et al.* A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54:38–46.

21. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904–909.
22. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* 2009;84:210–223.
23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MR, Bender D, Maller J, Sklar P, de Bakker PW, Daly M, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–575.
24. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr Protoc Hum Genet* 2009; Chapter 2: Unit 2.12.
25. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;52:506–516.
26. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265.
27. King DC, Taylor J, Elnitski L, Chiaromonte F, Miller W, Hardison RC. Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences. *Genome Res* 2005;15:1051–1060.
28. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, *et al*. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* 2006;34:D590–D598.
29. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res* 2009;37(Suppl 2):W600–W605.
30. Yang TP, Beazley C, Montgomery SB, Dimas AS, Gutierrez-Arcelus M, Stranger BE, Deloukas P, Dermitzakis ET. Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* 2010;26:2474–2476.
31. Holm K, Melum E, Franke A, Karlsen TH. SNPexp - a web tool for calculating and visualizing correlation between HapMap genotypes and gene expression levels. *BMC Bioinformatics* 2010;11:600.
32. Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999;55:997–1004.
33. Franke A, Fischer A, Nothnagel M, Becker C, Grabe N, Till A, Lu T, Muller-Quernheim J, Wittig M, Hermann A, *et al*. Genome-wide association analysis in sarcoidosis and Crohn's disease unravels a common susceptibility locus on 10p12.2. *Gastroenterology* 2008;135: 1207–1215.
34. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42:1118–1125.
35. Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, Kim H, Singh P, Lee A, Chen WV, *et al*. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature* 2010;466:113–117.
36. Jagielska D, Redler S, Brockschmidt FF, Herold C, Pasternack SM, Garcia Bartels N, Hanneken S, Eigelshoven S, Refke M, Barth S, *et al*. Follow-up study of the first genome-wide association scan in alopecia areata: IL13 and KIAA0350 as susceptibility loci supported with genome-wide significance. *J Invest Dermatol* 2012;132:2192–2197.
37. Mells GF, Floyd JA, Morley KI, Cordell HJ, Franklin CS, Shin SY, Heneghan MA, Neuberger JM, Donaldson PT, Day DB, *et al*. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat Genet* 2011;43:329–332.
38. Zhang F, Liu H, Chen S, Low H, Sun L, Cui Y, Chu T, Li Y, Fu X, Yu Y, *et al*. Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy. *Nat Genet* 2011;43:1247–1251.
39. Ellinghaus D, Ellinghaus E, Nair RP, Stuart PE, Esko T, Metspalu A, Debrus S, Raelson JV, Tejasvi T, Belouchi M, *et al*. Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am J Hum Genet* 2012;90: 636–647.
40. Nica AC, Parts L, Glass D, Nisbet J, Barrett A, Sekowska M, Travers M, Potter S, Grundberg E, Small K, *et al*. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet* 2011;7:e1002003.
41. Kim HS, Choi D, Lim LL, Allada G, Smith JR, Austin CR, Doyle TM, Goodwin KA, Rosenbaum JT, Martin TM. Association of interleukin 23 receptor gene with sarcoidosis. *Dis Markers* 2011;31:17–24.
42. Zhang FR, Huang W, Chen SM, Sun LD, Liu H, Li Y, Cui Y, Yan XX, Yang HT, Yang RD, *et al*. Genomewide association study of leprosy. *N Engl J Med* 2009;361:2609–2618.
43. Zhu KJ, Zhu CY, Shi G, Fan YM. Association of IL23R polymorphisms with psoriasis and psoriatic arthritis: a meta-analysis. *Inflamm Res* (In press).
44. Matsushita E, Asai N, Enomoto A, Kawamoto Y, Kato T, Mii S, Maeda K, Shibata R, Hattori S, Hagikura M, *et al*. Protective role of Gipie, a Girdin family protein, in endoplasmic reticulum stress responses in endothelial cells. *Mol Biol Cell* 2011;22:736–747.
45. Lesage F, Maingret F, Lazdunski M. Cloning and expression of human TRAAK, a polyunsaturated fatty acids-activated and mechano-sensitive K(+) channel. *FEBS Lett* 2000;471:137–140.
46. Lembrechts R, Pintelon I, Schnorbusch K, Timmermans JP, Adriaensen D, Brouns I. Expression of mechanogated two-pore domain potassium channels in mouse lungs: special reference to mechanosensory airway receptors. *Histochem Cell Biol* 2011;136:371–385.
47. Knoops B, Goemaere J, Van der Eecken V, Declercq JP. Peroxiredoxin 5: structure, mechanism, and function of the mammalian atypical 2-Cys peroxiredoxin. *Antioxid Redox Signal* 2011;15:817–829.
48. Baehner RL, Nathan DG. Leukocyte oxidase: defective activity in chronic granulomatous disease. *Science* 1967;155:835–836.
49. Bach TL, Chen Q-M, Kerr WT, Wang Y, Lian L, Choi JK, Wu D, Kazanietz MG, Koretzky GA, Zigmond S, *et al*. Is critical for T cell chemotaxis. *J Immunol* 2007;179:2223–2227.
50. Wang Z, Liu B, Wang P, Dong X, Fernandez-Hernando C, Li Z, Hla T, Claffey K, Smith JD, Wu D. Phospholipase C beta3 deficiency leads to macrophage hypersensitivity to apoptotic induction and reduction of atherosclerosis in mice. *J Clin Invest* 2008;118:195–204.