





Article

Genetic Variants Associated with FDNY WTC-Related Sarcoidosis

Krystal L. Cleven^{1,2,*,†}, Kenny Ye^{3,*,†} , Rachel Zeig-Owens^{1,3,4}, Kerry M. Hena⁵,
Cristina Montagna^{6,7,8} , Jidong Shan^{6,8}, H. Dean Hosgood III³, Nadia Jaber⁴,
Michael D. Weiden^{4,5}, Hilary L. Colbeth^{1,4}, David G. Goldfarb^{1,4}, Simon D. Spivack^{1,3,5,6,†} and
David J. Prezant^{1,4,5,†}

¹ Pulmonology Division, Department of Medicine, Montefiore Medical Center, Bronx, NY 10467, USA; rachel.zeig-owens@fdny.nyc.gov (R.Z.-O.); hilary.colbeth@fdny.nyc.gov (H.L.C.); david.goldfarb@fdny.nyc.gov (D.G.G.); simon.spivack@einstein.yu.edu (S.D.S.); david.prezant@fdny.nyc.gov (D.J.P.)

² Pulmonology Division, Department of Medicine, Albert Einstein College of Medicine, Bronx, NY 10467, USA

³ Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA; dean.hosgood@einstein.yu.edu

⁴ Fire Department of the City of New York, Bureau of Health Services, Brooklyn, NY 11201, USA; nadia.jaber@fdny.nyc.gov (N.J.); michael.weiden@nyulangone.org (M.D.W.)

⁵ Pulmonary & Critical Care Division, Department of Medicine, NYU School of Medicine, New York, NY 10016, USA; kerry.hena@nyulangone.org

⁶ Department of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461, USA; cristina.montagna@einstein.yu.edu (C.M.); jidong.shan@einstein.yu.edu (J.S.)

⁷ Department of Pathology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

⁸ Molecular Cytogenetic Core, Albert Einstein College of Medicine, Bronx, NY 10461, USA

* Correspondence: kcleven@montefiore.org (K.L.C.); kenny.ye@einstein.yu.edu (K.Y.); Tel.: +1-718-920-6054 (K.L.C.); +1-718-430-2590 (K.Y.)

† These authors share equal authorship.

Received: 15 February 2019; Accepted: 19 May 2019; Published: 23 May 2019



Abstract: Sarcoidosis is a systemic granulomatous disease of unknown etiology. It may develop in response to an exposure or inflammatory trigger in the background of a genetically primed abnormal immune response. Thus, genetic studies are potentially important to our understanding of the pathogenesis of sarcoidosis. We developed a case-control study which explored the genetic variations between firefighters in the Fire Department of the City of New York (FDNY) with World Trade Center (WTC)-related sarcoidosis and those with WTC exposure, but without sarcoidosis. The loci of fifty-one candidate genes related to granuloma formation, inflammation, immune response, and/or sarcoidosis were sequenced at high density in enhancer/promoter, exonic, and 5' untranslated regions. Seventeen allele variants of human leukocyte antigen (HLA) and non-HLA genes were found to be associated with sarcoidosis, and all were within chromosomes 1 and 6. Our results also suggest an association between extrathoracic involvement and allele variants of HLA and non-HLA genes found not only on chromosomes 1 and 6, but also on chromosomes 16 and 17. We found similarities between genetic variants with WTC-related sarcoidosis and those reported previously in sporadic sarcoidosis cases within the general population. In addition, we identified several allele variants never previously reported in association with sarcoidosis. If confirmed in larger studies with known environmental exposures, these novel findings may provide insight into the gene-environment interactions key to the development of sarcoidosis.

Keywords: sarcoidosis; World Trade Center; 9/11; genetics; firefighters; FDNY

1. Introduction

Sarcoidosis is a multisystem granulomatous disease that presents with pulmonary involvement in 90% of cases and extrapulmonary manifestations, mostly skin, in 10% of patients [1]. The annual age-adjusted incidence in the United States (U.S.) between 1946–2013, in a mostly Caucasian cohort, was 10 per 100,000 (9.4 in males and 10.5 in females) [2]. Between 1985 and 1998, firefighters in the Fire Department of the City of New York (FDNY) had an annual incidence of sarcoidosis somewhat higher than the white male general population at 12.9 per 100,000 and many-fold higher than an internal control group of emergency service workers [3]. The incidence doubled following the World Trade Center (WTC) attacks on September 11, 2001 (9/11), giving an age-adjusted incidence rate of 25 per 100,000 in mostly Caucasian male WTC-exposed FDNY firefighters between 2002–2015 [4].

Since the disorder clusters in families, genetic factors are thought to play an important role in sarcoidosis pathogenesis, presumably through mechanisms of immunologic response to antigenic or other triggering exposures. A recent population-based study suggests a genetic risk factor for sarcoidosis in a large Swedish cohort, where a 3.7-fold increase in risk of sarcoidosis was observed in those with first-degree relatives with the disease. The risk further increased in those with relatives with Lofgren's syndrome or with ≥ 1 relative with sarcoidosis [5]. A case-control etiologic study of sarcoidosis (ACCESS) found a five-fold increase in siblings of those with sarcoidosis [6].

We selected several candidate genes for genetic analysis based on their function encoding proteins involved in granuloma formation, innate and immune recognition, and inflammation. Several have been implicated in genetic studies of sarcoidosis such as the highly polymorphic human leukocyte antigen (HLA) genes. HLA genes encode the major histocompatibility complex (MHC) in humans and are involved in immunologic recognition and regulation. Within various racial groups, allele variations within HLA genes are associated with sarcoidosis. Examples include the HLA-DRB1*1101 allele in both Caucasians and African Americans [7], HLA-DRB1*1501, -*03 and -*0402 specifically in European Americans [7–9], HLA-DRB1*1201 and -*0302 in African-Americans [7,10]. The HLA-DQB1 alleles HLA-DQB1-*0402, and -*0503 are associated in Asian Indians [7] and HLA-DQB1-*0602 in a Dutch cohort [8]. Several other non-HLA genes have been associated with sarcoidosis, such as prostaglandin G/H synthase/cyclooxygenase (PTGS2/COX2), neurogenic locus notch homolog 4 (NOTCH4), nucleotide-binding oligomerization domain-containing protein 2 (NOD2), and Butyrophilin-like 2 (BTNL2) [11–15].

The current case-control study is the first to examine genetic features associated with WTC-related sarcoidosis. Compared to other genetics studies of sarcoidosis, our study uniquely includes a baseline sarcoidosis rate, a known time point of exposure (9/11), and extensive, systematized clinical follow-up for over 14 years. The objective of this study is to explore the genetic characteristics of WTC-associated sarcoidosis and understand its similarities to and differences from sporadic sarcoidosis cases without a clear environmental exposure.

2. Materials and Methods

2.1. Study Population

The source population for the cases and controls was WTC-exposed firefighters. Recruitment of sarcoidosis cases for this study was part of a larger study that explored the clinical course of firefighters with post-9/11 sarcoidosis between the time of diagnosis and 14 years after 9/11 [16]. As outlined in Hena, et al (2018), cases were identified as WTC-exposed firefighters with WTC-related sarcoidosis if they had normal chest imaging prior to 9/11 and then post-9/11 radiographic findings consistent with sarcoidosis [16]. Extra-pulmonary organ involvement was determined from study questionnaires and medical records and was defined according to the World Association of Sarcoidosis and Other Granulomatous Diseases (WASOG) criteria [17].

Controls were randomly selected from the source population and were matched to cases based on race, age on 9/11, smoking history, and WTC-exposure history. WTC-exposure history was determined

by arrival time to the WTC site: arrival group 1 (arriving the morning of 9/11) being the group with the highest dust cloud exposure and arrival group 4 (arriving between 9/13–9/24) having the least dust cloud exposure [18]. Controls also had normal chest radiographs at the time of the study.

The final study population included 55 cases and 100 controls. All cases and controls had peripheral blood drawn in 2015–2016 and completed a questionnaire at the time of blood draw to aid with matching on clinical characteristics and WTC-exposure history. The study was approved by the Montefiore Medical Center/Albert Einstein Institutional Review Board (2014-4291). All study participants provided written informed consent.

2.2. Genotyping

Peripheral blood mononuclear cells were obtained from cases and controls for sequencing analysis. A panel of 51 candidate genes were pre-selected based on their function encoding proteins involved in granuloma formation, innate and immune recognition, inflammation, and sarcoidosis. The loci of targeted candidate genes were sequenced using a custom designed sequencing panel for IonTorrent Ampliseq technology. The custom sequencing panel contained 1193 amplicons spanning all known exons, splice junctions and promoter regions of the genes. Promoter regions were defined as 2 kb upstream from the transcriptional start site. The panel overall coverage was over 90% for 422 exons (Supplementary Table S1). One exception was HLA-A, which had 41% coverage, but is known to be associated with low coverage in prior design assays as well. The sequencing panel was designed for the analysis of amplicons of 125–375 bp using the Ion AmpliSeq technology and was suitable for sequencing using the Ion Proton Sequencer as previously reported [19].

DNA was extracted using the QuickGene 610 system and the whole blood kit (Kurabo #DB-L), and quality verified and quantitated by the NanoDropTM 1000 Spectrophotometer (Thermo Fischer Scientific, Waltham, MA, USA). The Qubit[®] dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA, USA) was used to determine the concentration of each sample. Approximately 20 ng of DNA was used to generate libraries for sequencing. When not in use, samples were stored at -20°C . For each sample, a PCR reaction was set up according to the manufacturer's protocol using 20 ng of genomic DNA, two separate primer pools, and 16 amplification cycles. Following amplification, each sample was treated with FuPa reagent and ligated to a uniquely barcoded adapter to enable sample multiplexing. Libraries were then purified using the 1.5X Agencourt AMPure XP (Beckman Coulter Inc, Brea, CA, USA) kit. Amplification products from each primer pool were quantified individually using the KAPA Library Quantitation kit (Roche, Mannheim, Germany) and then pooled together. Template preparation was performed using the Ion OneTouch 2 system and the Ion PI Template OT2 200 kit v2 (Thermo Fisher Scientific, Waltham, MA, USA), according to the manufacturer's protocol. All libraries were sequenced on the Ion Proton sequencer using the Ion PI chip and the Ion PI Sequencing 200 kit v2 (Thermo Fisher Scientific, Waltham, MA, USA) to generate 200bp single ended sequencing.

For the identification and reporting of germline variants, we used a custom validated bioinformatic pipeline for the identification of SNPs as well as insertions and deletions (INDELs) included in the target panels [20–22]. The sequence reads obtained from the samples were aligned to the human reference genome (hg19-Genome Reference Consortium GRCh37) using the Ion Torrent Suite and then processed by the Torrent Variant Caller for variant calling. The generated BAM files were next imported into the Ion Reporter SoftwareTM (Thermo Fisher Scientific, Waltham, MA, USA) for filtering and annotation.

2.3. Statistical Analysis

2.3.1. Common Variants

We identified 909 variants (SNPs or INDELs) with minor allele frequency (MAF) ≥ 0.05 and p -value for Hardy-Weinberg Equilibrium Test > 0.0001 . The variants were classified by their genomic locations to four categories: exon, intron (likely intron/exon boundaries), upstream, and 5'untranslated region

(5'UTR). For those in an exonic region, they were further classified by their function as synonymous, nonsense, missense, or frameshift (Supplementary Tables S2 and S3). For each of these common variants, Pearson Chi-square test was used to compare the genotype frequencies between the 55 cases and 100 controls that were successfully sequenced (see Supplementary Table S4). To verify the robustness of our results, Fisher's Exact Tests were also performed. For each of the variants, odds ratios were estimated using unconditional logistic regression under the co-dominant model.

We conducted two secondary analyses. First, we explored the interaction between each common variant and degree of WTC exposure, measured as arrival time to the WTC site. For each variant, we applied unconditional logistic regression under the co-dominant model with sarcoidosis as the outcome and the genotypes, WTC exposure group and their interaction term, as the explanatory variables. Likelihood ratio tests are applied to compare two models with and without the interaction term between the genotype and degree of WTC exposure. We also conducted a secondary analysis in which we associate the genetic variants with extrathoracic sarcoidosis. In these analyses, for each variant we applied unconditional logistic regression under the co-dominant model with extrathoracic sarcoidosis as the outcome and the genotypes as the explanatory variable.

2.3.2. Rare Variants

For variants whose MAF < 0.05, for each individual, the total number of rare alleles located in each gene were counted. We then compared the number of rare alleles between the cases and controls for each gene by each genomic location category (exon, intron, upstream, or 5'untranslated region) using two-sample T-tests. Furthermore, we counted rare variants with MAF < 0.01 across all target genes by the above-mentioned categories by genomic location and function, and then performed two-sample t-tests comparing the cases and controls for each of the categories.

3. Results

Final genetics analyses were completed on 55 cases and 100 controls. All but one case (54/55) and one control (99/100) self-identified as Caucasian, the remaining one case and one control self-identified as African American. The cases and controls were well matched, with at least one control for each case (Table 1).

Table 1. Demographics of WTC Study Population.

Demographics	Case Characteristics		Control Characteristics		p-Value
	N	%	N	%	
Total	55	100	100	100	
Arrival group					0.99
Morning of 9/11	12	21.8	22	22	
Afternoon of 9/11	24	43.6	46	46	
Day 9/12	15	27.3	25	25	
Day 9/13–9/24	4	7.3	7	7	
Race					1.00
White	54	98.2	99	99	
African-American	1	1.8	1	1	
Smoking status at time of blood draw					0.66
Never	40	72.7	72	72	
Former	14	25.5	23	23	
Current	1	1.8	5	5	
Age on 9/11, years, median [IQR]	36.8 [32.7–39.0]		35.4 [30.2–39.3]		0.29
Age at sarcoidosis diagnosis, median [IQR]	43.0 [38.2–46.6]				–
Time to diagnosis post 9/11, median [IQR]	6.7 [3.3–9.6]				–

IQR: interquartile range.

Organ involvement for the 55 cases included 47/55 (85.5%) with biopsy proven sarcoidosis. The 8/55 (14.5%) without biopsy results had their imaging and organ involvement reviewed by two

pulmonologists to confirm consistency with sarcoidosis [16]. As outlined in Table 2, there was a decrease in intrathoracic involvement from the time of diagnosis to the time of the blood draw (mean eight-years post diagnosis) from 95% to 47% and an increase in extrathoracic involvement, especially cardiac and joint.

Table 2. Sarcoidosis organ involvement in WTC genetics study population.

Organ Involvement	Study Cases (N = 55)	
	At Diagnosis	Time of Blood Draw ^a
Bone marrow	0	0
Bone/joint	3	8
Calcium	0	0
Cardiac	0	9
Ear/nose/throat	0	1
Extrathoracic lymph nodes	0	0
Eyes	3	3
Intrathoracic involvement by CT Imaging	52	26
Kidney	1	0
Liver	1	1
Muscle	0	0
Nervous system	0	1
Other organs	0	0
Salivary	0	0
Skin	1	1
Spleen	2	2

^a All cases had complete data regarding organ involvement at diagnosis and blood draw with the following exceptions: chest CT scan at diagnosis (N = 52), and at blood draw (N = 53); muscle at diagnosis (N = 54), and at blood draw (N = 54); spleen at diagnosis (N = 54), and at blood draw (N = 53); cardiac at blood draw (N = 53); and ear/nose/throat at diagnosis (N = 51), and at blood draw (N = 53).

Within the 51 candidate genes analyzed, we identified 3619 total number of variants. A total of seventeen common variants were found to be associated with sarcoidosis with Chi-Squared *p*-value < 0.01 (Table 3). All of the 17 variants were within chromosomes 1 and 6. Multiple variants were in HLA genes such as HLA-C, HLA-DRB1, HLA-DQB1, HLA-DPA1, and HLA-DPB1. There was a strong association with sarcoidosis within exonic regions of the HLA-DQB1 gene, represented by rs1049133 and rs1049130, two SNPs 12bp apart (ORs = 2.56 and 1.90, respectively). Two variants upstream from an intronic/exonic border region of HLA-DQB1 were also significantly associated with sarcoidosis, rs4516985 and rs9274614 (ORs = 1.74 and 2.49, respectively). In addition, several genetic variants within or near non-HLA genes were also significantly associated with sarcoidosis: BTNL2, PTGS2/COX2, and PACERR (PTGS2 Antisense NFKB1 Complex-Mediated Expression Regulator RNA). SNPs in a non-coding region of BTNL2, rs2076525 and rs2076524, were significantly associated with sarcoidosis (OR = 1.71). In addition, rs2076523, representing a missense mutation within a BTNL2 coding region was associated (OR = 1.97). Upstream from PTGS2/COX2 gene, rs20417 was also associated with sarcoidosis cases in our cohort (OR = 1.79).

In our secondary analysis, we did not find statistical evidence of an interaction between common variants and the degree of WTC exposure. We also found no statistical significance when we compared the number of rare variants between the cases and controls.

Our results also suggest an association between extrathoracic involvement and genetic variants within several HLA and non-HLA genes: HLA-B, PTGS2/COX2, PACERR, NOTCH4, NOD2, and ITGAE (Integrin Subunit Alpha E) (Table 4). Genetic variants associated with extrathoracic cases were found on chromosomes 1 and 6, similar to the loci associated with all sarcoidosis cases, and were also found on chromosomes 16 and 17. On chromosome 1, rs2066826 represents an intronic region/exonic border of PTGS2/COX2 associated with sarcoidosis, and specifically, extrathoracic disease in this cohort (OR 1.88 and 1.45, respectively). As seen in all sarcoidosis cases with rs20417, another variant only a few

hundred base pairs away and downstream from PACERR, is rs689466, for which the allele C was more common in those with extrathoracic disease. A genetic variant within an HLA-B, rs2276448, was also associated with extrathoracic sarcoidosis. In addition, on chromosome 6, a locus upstream from NOTCH4, rs3134929, was associated with extrathoracic disease. And in chromosome 16 and 17, three variants were associated with extrathoracic involvement: two SNPs within chromosome 16, rs2066843 and rs2066842, both within exonic regions of NOD2; and one within chromosome 17, rs220465, representing a genetic variant within the intronic region/exonic border of ITGAE. As shown in Table 2, cardiac and joint involvement were common, but our cohort size was too small to identify genetic variants associated with any specific extrathoracic organ manifestation. We also found no association between the degree of exposure and extrathoracic organ involvement.

Table 3. Genetic variants most associated with sarcoidosis.

Gene	Position(hg19)	dbSNP	Alleles *	Risk Allele	Chi-sq p-Value	Fisher's p-value	OR **
PTGS2/COX2	chr1:186645927	rs2066826	T/C	C	0.002	0.001	1.88
PTGS2 PACERR †	chr1:186650321	rs20417	G/C	C	0.003	0.001	1.79
HLA-C	chr6:31239681	rs9264669	T/A	A	0.004	0.003	1.75
BTNL2	chr6:32370616	rs2076525	C/T	T	0.006	0.003	1.71
BTNL2	chr6:32370684	rs2076524	G/A	A	0.006	0.003	1.71
BTNL2	chr6:32370835	rs2076523	C/T	T	0.005	0.002	1.97
HLA-DRB1	chr6:32549424	rs112116022	T/C	C	0.003	0.002	5.21
HLA-DQB1	chr6:32629847	rs1049133	G/A	A	0.004	0.003	2.56
HLA-DQB1	chr6:32629859	rs1049130	G/A	A	0.006	0.004	1.90
HLA-DQB1	chr6:32635632	rs4516985	G/A	A	0.004	0.005	1.74
HLA-DQB1	chr6:32635846	rs9274614	G/C	C	0.005	0.004	2.49
HLA-DPA1 HLA-DPB1 †	chr6:33048457	rs386699868 rs1126504 rs386699869	G/C	G	0.007	0.006	1.80 †
HLA-DPA1 HLA-DPB1 †	chr6:33048466	rs1126511 rs1126513	TT/GG	TT	0.007	0.007	1.66 †
HLA-DPA1 HLA-DPB1 †	chr6:33049211	rs928976	T/C	T	0.007	0.008	1.74 †

† indicates that SNP overlaps part of both genes; * listed as alternative allele/reference allele; ** ORs are calculated under the co-dominant model and are relative to the alternative allele unless specified with † which is in relation to the reference allele. This was done to maintain the OR as a comparison between the risk allele and non-risk allele.

Table 4. Genetic variants associated with extrathoracic organ involvement.

Gene	Position(hg19)	dbSNP	Alleles *	Risk Allele	Chi-sq p-Value	Fisher's p-Value	OR **
PTGS2	chr1:186645927	rs2066826	T/C	C	0.009	0.012	1.45
PACERR	chr1:186650751	rs689466	C/T	C	0.006	0.016	2.33 †
HLA-B	chr6:31323020	rs2276448	C/T	T	0.008	0.003	2.13
NOTCH4	chr6:32192107	rs3134929	G/C	G	0.009	0.012	2.39 †
NOD2	chr16:50744624	rs2066842	T/C	T	0.010	0.019	2.00 †
NOD2	chr16:50745199	rs2066843	T/C	T	0.010	0.019	2.21 †
ITGAE	chr17:3637915	rs220465	T/C	T	0.005	0.006	1.08 †

* listed as alternative allele/reference allele; ** ORs are calculated under the co-dominant model and are relative to the alternative allele unless specified with † which is in relation to the reference allele. This was done to maintain the OR as a comparison between the risk allele and non-risk allele.

4. Discussion

Sarcoidosis has been identified in all WTC cohorts (FDNY, general responders other than FDNY, the survivor population and the WTC Registry), the most extensively described being that from FDNY [4,16,23–26]. In our current FDNY case-control study, we examine genetic characteristics in our cases as compared with our WTC-exposed controls without sarcoidosis. Seventeen allele variants of HLA and non-HLA genes were found to be associated with sarcoidosis with p -value < 0.01 and all were within chromosomes 1 and 6. Our results also suggest an association between extrathoracic involvement and allele variants of HLA and non-HLA genes found not only on chromosomes 1 and

6, but also on chromosomes 16 and 17. Comparing our findings to published studies of sporadic sarcoidosis, we found similarities between genetic variants with WTC-related sarcoidosis and those reported in studies of sarcoidosis without known environmental exposures. In addition, we identified several allele variants not previously reported to be in association with sarcoidosis. If confirmed in larger studies with known environmental exposures, these novel findings may provide insight into the gene-environment interactions key to the development of sarcoidosis.

For example, our data are consistent with a prior study showing an association between SNP rs20417, in a non-coding region of the PTGS2/COX2 gene on chromosome 1, and sarcoidosis in a UK and Austrian Caucasian cohort without known environmental exposure [11]. This allele variant, described using prior nomenclature -765C (-765 is the nucleotide base pair position and C is the base pair), was found to be more often present in sarcoidosis subjects with severe pulmonary disease (poor lung function or extensive fibrosis). The authors also created lung fibroblast cell lines that were stimulated by transforming growth factor (TGF)- β 1. They found that the cell lines with the -765C (currently named rs20147) produced little to no prostaglandin E2 (PGE2), an important inhibitor of the proliferation of fibroblasts. The loss of this lung protective inflammatory mediator could explain its relationship to more severe forms of pulmonary sarcoidosis. As eloquently described by Valentonyte and colleagues (2005), allele variations in BTNL2 impact the negative T-cell regulation function of BTNL2 [14]. Thus, the dysregulation of T-cell function from some BTNL2 mutations may explain its relationship to sarcoidosis. Interestingly, Akers and colleagues (2011) determined there was strong linkage disequilibrium between rs9274614, an allele significantly associated with sarcoidosis in our cohort, and several HLA-DQB1*06 alleles [27]. Another study found an association between the HLA-DQB1*0601 allele and cardiac sarcoidosis in a Japanese cohort [28]. Although our sample size was too small to identify specific alleles associated with cardiac sarcoidosis, the known association between the SNP rs9274614 and an allele associated with cardiac sarcoidosis, HLA-DQB1*0601, is one of interest in this cohort given the significant number of participants with cardiac sarcoidosis.

Except as noted above, the genetic variants associated with sarcoidosis in our WTC-exposed cohort have not previously been reported. Several of the SNPs identified in our cohort as associated with WTC-related sarcoidosis have been associated with other diseases in cohorts without known environmental exposures. Perhaps the most frequently investigated is rs20417 in PTGS2/COX2. Studies have identified its association with carotid-calcified plaque in those with diabetes [29], protection against myocardial infarction and stroke [30], as well as asthma [31]. In addition, SNP rs1126513, significantly associated with WTC-related sarcoidosis in our cohort, was found to be associated with ankylosing spondylitis [32].

The main strength of this study was its reliance on a rigorously characterized prospective cohort, with a clearly defined time of exposure (9/11). One limitation of this study is that it was focused on candidate genes already identified in sarcoidosis studies from the general population without WTC exposure. Additionally, the identification of novel SNPs may be partially due to our high-density sequencing which evaluated almost every base pair of these candidate genes. The main limitation, however, is that this study includes a relatively small sample size. Our statistical power was limited to detect only sizable effects. At significance level 0.01, we had 0.8 power to detect an odds ratio of 2.3 at variants with MAF = 0.5 and an odds ratio of 4 at variants with MAF = 0.05; we had 0.2 power to detect an odds ratio of 1.5 at variants with MAF = 0.5 and an odds ratio of 2 at variants with MAF = 0.05. Additional power would be obtained if this study was expanded to include all of the various WTC health program cohorts to allow for both replication of the main effects observed, and to allow for gene-environment analyses. This could potentially allow for more robust causal inferences to be made regarding the etiology of sarcoidosis, and whether there are unique differences in those with extrathoracic disease, particularly cardiac and joint involvement.

5. Conclusions

The incidence of sarcoidosis was increased after the collapse of the WTC towers and this is best demonstrated in the FDNY cohort where there was clear evidence that these cases were new in onset after 9/11 [4,16]. Our findings suggest that the genetic characteristics of WTC-related sarcoidosis are similar to the sporadic cases found in the general population without known environmental exposures. These data suggest the granulomatous disease observed post-9/11 and originally termed “Sarcoid-Like” Granulomatous Disease (SLGD) is better described as WTC-related Sarcoidosis. Our findings, however, suggest that there are several novel allele variants (SNPs) which may be uniquely associated with WTC-related sarcoidosis and with extrathoracic disease. These novel findings may provide some insights into the unique gene-environment interactions key to the development of sarcoidosis in this population, among others. Future studies should focus on increasing the power of our findings by combining sarcoidosis cases from all WTC health program cohorts and comparing them to matched controls with non-WTC associated sarcoidosis.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1660-4601/16/10/1830/s1>, Table S1: Candidate genes and exon coverage, Table S2: Location and function of gene variants associated with sarcoidosis, Table S3: Location and function of gene variants associated with extrathoracic organ involvement, Table S4: Genotype frequencies.

Author Contributions: Conceptualization, K.Y., R.Z.-O., K.M.H., C.M., J.S., S.D.S. and D.J.P.; Data curation, K.L.C., R.Z.-O., K.M.H., C.M., J.S., N.J. and D.G.G.; Formal analysis, K.Y., R.Z.-O. and H.L.C.; Funding acquisition, K.M.H., S.D.S. and D.J.P.; Investigation, K.L.C., K.M.H., N.J. and D.G.G.; Methodology, K.Y., R.Z.-O., K.M.H., C.M., S.D.S. and D.J.P.; Project administration, R.Z.-O., K.M.H., C.M., S.D.S. and D.J.P.; Resources, C.M., S.D.S. and D.J.P.; Software, K.Y., C.M. and J.S.; Supervision, S.D.S. and D.J.P.; Validation, K.L.C. and K.Y.; Visualization, K.L.C. and H.L.C.; Writing—original draft, K.L.C., K.Y., C.M. and S.D.S.; Writing—review & editing, K.L.C., K.Y., H.D.H.III, M.D.W., H.L.C., S.D.S. and D.J.P.

Funding: This research was funded by the National Institute of Occupational Safety and Health (NIOSH), Grant U01-OH010993 and contracts #200-2011-39383, #200-2011-39378, #200-2017-93326 and #200-2017-93426.

Acknowledgments: We would like to acknowledge Thomas K. Aldrich, for his significant contribution to project design, supervision, and funding acquisition and was the original PI on this study. He sadly died of pancreatic cancer while this project was underway. Aldrich’s work was guided by his desire to understand the health impacts of 9/11 in order to improve the lives of those affected. He was a mentor and personal friend to many of us and is deeply missed. We would also like to thank Miao Kevin Shi for technical support in Spivack’s lab.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Baughman, R.P.; Teirstein, A.S.; Judson, M.A.; Rossman, M.D.; Yeager, H., Jr.; Bresnitz, E.A.; DePalo, L.; Hunninghake, G.; Iannuzzi, M.C.; Johns, C.J.; et al. Clinical characteristics of patients in a case control study of sarcoidosis. *Am. J. Respir. Crit. Care Med.* **2001**, *164*, 1885–1889. [[CrossRef](#)] [[PubMed](#)]
2. Ungprasert, P.; Carmona, E.M.; Utz, J.P.; Ryu, J.H.; Crowson, C.S.; Matteson, E.L. Epidemiology of Sarcoidosis 1946–2013: A Population-Based Study. *Mayo Clin. Proc.* **2016**, *91*, 183–188. [[CrossRef](#)]
3. Prezant, D.J.; Dhala, A.; Goldstein, A.; Janus, D.; Ortiz, F.; Aldrich, T.K.; Kelly, K.J. The incidence, prevalence, and severity of sarcoidosis in New York City firefighters. *Chest* **1999**, *116*, 1183–1193. [[CrossRef](#)] [[PubMed](#)]
4. Webber, M.P.; Yip, J.; Zeig-Owens, R.; Moir, W.; Ungprasert, P.; Crowson, C.S.; Hall, C.B.; Jaber, N.; Weiden, M.D.; Matteson, E.L.; et al. Post-9/11 sarcoidosis in WTC-exposed firefighters and emergency medical service workers. *Respir. Med.* **2017**, *132*, 232–237. [[CrossRef](#)]
5. Rossides, M.; Grunewald, J.; Eklund, A.; Kullberg, S.; Di Giuseppe, D.; Askling, J.; Arkema, E.V. Familial aggregation and heritability of sarcoidosis: A Swedish nested case-control study. *Eur. Respir. J.* **2018**, *52*. [[CrossRef](#)] [[PubMed](#)]
6. Rybicki, B.A.; Iannuzzi, M.C.; Frederick, M.M.; Thompson, B.W.; Rossman, M.D.; Bresnitz, E.A.; Terrin, M.L.; Moller, D.R.; Barnard, J.; Baughman, R.P.; et al. Familial aggregation of sarcoidosis. A case-control etiologic study of sarcoidosis (ACCESS). *Am. J. Respir. Crit. Care Med.* **2001**, *164*, 2085–2091. [[CrossRef](#)] [[PubMed](#)]

7. Rossman, M.D.; Thompson, B.; Frederick, M.; Maliarik, M.; Iannuzzi, M.C.; Rybicki, B.A.; Pandey, J.P.; Newman, L.S.; Magira, E.; Beznik-Cizman, B.; et al. HLA-DRB1*1101: A significant risk factor for sarcoidosis in blacks and whites. *Am. J. Hum. Genet.* **2003**, *73*, 720–735. [[CrossRef](#)] [[PubMed](#)]
8. Voorter, C.E.; Drent, M.; van den Berg-Loonen, E.M. Severe pulmonary sarcoidosis is strongly associated with the haplotype HLA-DQB1*0602-DRB1*150101. *Hum. Immunol.* **2005**, *66*, 826–835. [[CrossRef](#)] [[PubMed](#)]
9. Swider, C.; Schnittger, L.; Bogunia-Kubik, K.; Gerdes, J.; Flad, H.; Lange, A.; Seitzer, U. TNF-alpha and HLA-DR genotyping as potential prognostic markers in pulmonary sarcoidosis. *Eur. Cytokine Netw.* **1999**, *10*, 143–146. [[PubMed](#)]
10. Levin, A.M.; Adrianto, I.; Datta, I.; Iannuzzi, M.C.; Trudeau, S.; Li, J.; Drake, W.P.; Montgomery, C.G.; Rybicki, B.A. Association of HLA-DRB1 with Sarcoidosis Susceptibility and Progression in African Americans. *Am. J. Respir. Cell Mol. Biol.* **2015**, *53*, 206–216. [[CrossRef](#)]
11. Hill, M.R.; Papafili, A.; Booth, H.; Lawson, P.; Hubner, M.; Beynon, H.; Read, C.; Lindahl, G.; Marshall, R.P.; McAnulty, R.J.; et al. Functional prostaglandin-endoperoxide synthase 2 polymorphism predicts poor outcome in sarcoidosis. *Am. J. Respir. Crit. Care Med.* **2006**, *174*, 915–922. [[CrossRef](#)]
12. Adrianto, I.; Lin, C.P.; Hale, J.J.; Levin, A.M.; Datta, I.; Parker, R.; Adler, A.; Kelly, J.A.; Kaufman, K.M.; Lessard, C.J.; et al. Genome-wide association study of African and European Americans implicates multiple shared and ethnic specific loci in sarcoidosis susceptibility. *PLoS ONE* **2012**, *7*, e43907. [[CrossRef](#)]
13. Sato, H.; Williams, H.R.; Spagnolo, P.; Abdallah, A.; Ahmad, T.; Orchard, T.R.; Copley, S.J.; Desai, S.R.; Wells, A.U.; du Bois, R.M.; et al. CARD15/NOD2 polymorphisms are associated with severe pulmonary sarcoidosis. *Eur. Respir. J.* **2010**, *35*, 324–330. [[CrossRef](#)]
14. Valentonyte, R.; Hampe, J.; Huse, K.; Rosenstiel, P.; Albrecht, M.; Stenzel, A.; Nagy, M.; Gaede, K.I.; Franke, A.; Haesler, R.; et al. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat. Genet.* **2005**, *37*, 357–364. [[CrossRef](#)]
15. Rybicki, B.A.; Walewski, J.L.; Maliarik, M.J.; Kian, H.; Iannuzzi, M.C. The BTNL2 gene and sarcoidosis susceptibility in African Americans and Whites. *Am. J. Hum. Genet.* **2005**, *77*, 491–499. [[CrossRef](#)]
16. Hena, K.M.; Yip, J.; Jaber, N.; Goldfarb, D.; Fullam, K.; Cleven, K.; Moir, W.; Zeig-Owens, R.; Webber, M.P.; Spevack, D.M.; et al. Clinical Course of Sarcoidosis in World Trade Center-Exposed Firefighters. *Chest* **2018**, *153*, 114–123. [[CrossRef](#)]
17. Judson, M.A.; Costabel, U.; Drent, M.; Wells, A.; Maier, L.; Koth, L.; Shigemitsu, H.; Culver, D.A.; Gelfand, J.; Valeyre, D.; et al. The WASOG Sarcoidosis Organ Assessment Instrument: An update of a previous clinical tool. *Sarcoidosis Vasc. Diffus. Lung Dis. Off. J. WASOG* **2014**, *31*, 19–27.
18. Prezant, D.J.; Weiden, M.; Banauch, G.I.; McGuinness, G.; Rom, W.N.; Aldrich, T.K.; Kelly, K.J. Cough and bronchial responsiveness in firefighters at the World Trade Center site. *New Engl. J. Med.* **2002**, *347*, 806–815. [[CrossRef](#)]
19. Miller, E.M.; Patterson, N.E.; Zechmeister, J.M.; Bejerano-Sagie, M.; Delio, M.; Patel, K.; Ravi, N.; Quispe-Tintaya, W.; Maslov, A.; Simmons, N.; et al. Development and validation of a targeted next generation DNA sequencing panel outperforming whole exome sequencing for the identification of clinically relevant genetic variants. *Oncotarget* **2017**, *8*, 102033–102045. [[CrossRef](#)]
20. Delio, M.; Patel, K.; Maslov, A.; Marion, R.W.; McDonald, T.V.; Cadoff, E.M.; Golden, A.; Grealley, J.M.; Vijg, J.; Morrow, B.; et al. Development of a Targeted Multi-Disorder High-Throughput Sequencing Assay for the Effective Identification of Disease-Causing Variants. *PLoS ONE* **2015**, *10*, e0133742. [[CrossRef](#)]
21. Shastri, A.; Msaouel, P.; Montagna, C.; White, S.; Delio, M.; Patel, K.; Alexis, K.; Strakhan, M.; Elrafei, T.N.; Reed, L.J. Primary Hepatic Small Cell Carcinoma: Two Case Reports, Molecular Characterization and Pooled Analysis of Known Clinical Data. *Anticancer Res.* **2016**, *36*, 271–277.
22. Josephs, K.; Patel, K.; Janson, C.M.; Montagna, C.; McDonald, T.V. Compound heterozygous CASQ2 mutations and long-term course of catecholaminergic polymorphic ventricular tachycardia. *Mol. Genet. Genom. Med.* **2017**, *5*, 788–794. [[CrossRef](#)] [[PubMed](#)]
23. Izbicki, G.; Chavko, R.; Banauch, G.I.; Weiden, M.D.; Berger, K.I.; Aldrich, T.K.; Hall, C.; Kelly, K.J.; Prezant, D.J. World Trade Center “sarcoid-like” granulomatous pulmonary disease in New York City Fire Department rescue workers. *Chest* **2007**, *131*, 1414–1423. [[CrossRef](#)]
24. Crowley, L.E.; Herbert, R.; Moline, J.M.; Wallenstein, S.; Shukla, G.; Schechter, C.; Skloot, G.S.; Udasin, I.; Luft, B.J.; Harrison, D.; et al. “Sarcoid like” granulomatous pulmonary disease in World Trade Center disaster responders. *Am. J. Ind. Med.* **2011**, *54*, 175–184. [[CrossRef](#)] [[PubMed](#)]

25. Parsia, S.S.; Yee, H.; Young, S.; Turetz, M.L.; Marmor, M.; Wilkenfeld, M.; Kazeros, A.; Caplan-Shaw, C.E.; Reibman, J. Characteristics of sarcoidosis in residents and workers exposed to World Trade Center (WTC) dust, gas and fumes presenting for medical care. *Am. J. Respir. Crit. Care Med.* **2010**, *181*, A1740.
26. Jordan, H.T.; Stellman, S.D.; Prezant, D.; Teirstein, A.; Osahan, S.S.; Cone, J.E. Sarcoidosis diagnosed after September 11, 2001, among adults exposed to the World Trade Center disaster. *J. Occup. Environ. Med.* **2011**, *53*, 966–974. [[CrossRef](#)] [[PubMed](#)]
27. Akers, N.K.; Curry, J.D.; Conde, L.; Bracci, P.M.; Smith, M.T.; Skibola, C.F. Association of HLA-DQB1 alleles with risk of follicular lymphoma. *Leuk. Lymphoma* **2011**, *52*, 53–58. [[CrossRef](#)] [[PubMed](#)]
28. Naruse, T.K.; Matsuzawa, Y.; Ota, M.; Katsuyama, Y.; Matsumori, A.; Hara, M.; Nagai, S.; Morimoto, S.; Sasayama, S.; Inoko, H. HLA-DQB1*0601 is primarily associated with the susceptibility to cardiac sarcoidosis. *Tissue Antigens* **2000**, *56*, 52–57. [[CrossRef](#)]
29. Rudock, M.E.; Liu, Y.; Ziegler, J.T.; Allen, S.G.; Lehtinen, A.B.; Freedman, B.I.; Carr, J.J.; Langefeld, C.D.; Bowden, D.W. Association of polymorphisms in cyclooxygenase (COX)-2 with coronary and carotid calcium in the Diabetes Heart Study. *Atherosclerosis* **2009**, *203*, 459–465. [[CrossRef](#)]
30. Cipollone, F.; Toniato, E.; Martinotti, S.; Fazio, M.; Iezzi, A.; Cucurullo, C.; Pini, B.; Ursi, S.; Vitullo, G.; Averna, M.; et al. A polymorphism in the cyclooxygenase 2 gene as an inherited protective factor against myocardial infarction and stroke. *JAMA* **2004**, *291*, 2221–2228. [[CrossRef](#)]
31. Szczeklik, W.; Sanak, M.; Szczeklik, A. Functional effects and gender association of COX-2 gene polymorphism G-765C in bronchial asthma. *J. Allergy Clin. Immunol.* **2004**, *114*, 248–253. [[CrossRef](#)] [[PubMed](#)]
32. Cortes, A.; Pulit, S.L.; Leo, P.J.; Pointon, J.J.; Robinson, P.C.; Weisman, M.H.; Ward, M.; Gensler, L.S.; Zhou, X.; Garchon, H.J.; et al. Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. *Nat. Commun.* **2015**, *6*, 7146. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).