SARCOIDOSIS VASCULITIS AND DIFFUSE LUNG DISEASES 2022; 39 (4); e2022040 DOI: 10.36141/svdld.v39i4.13407

Metagenomic sequencing of the bronchoalveolar lavage extracellular virome and cellular transcriptome of sarcoidosis patients does not detect rubella virus

Emma L. Keeler[1], Milica Vukmirovic[5], Xiting Yan[5], Kristin Gulino[6], Elodie Ghedin[6,7], Naftali Kaminski[5], Kathleen E. Sullivan[4,8], Frederic D. Bushman[1], Ronald G. Collman[1,2], Misha Rosenbach[3]

Departments of [1]Microbiology, [2]Medicine, [3]Dermatology and [4]Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; [5]Department of Medicine, Yale University School of Medicine, New Haven, CT, USA; [6]Department of Biology, Center for Genomics and Systems Biology, New York University, New York, NY, USA; [7]Laboratory of Parasitic Diseases, NIAID, NIH, Washington, DC, USA; [8]Children's Hospital of Philadelphia, Philadelphia, PA, USA

Abstract. *Background:* Sarcoidosis is a multisystem granulomatous inflammatory disease of unclear etiology that involves the lung, skin and other organs, with an unknown antigenic trigger. Recently, evidence has been found in both immune deficient and immune competent patients for rubella virus in cutaneous granulomas. These granulomatous lesions share overlapping features with cutaneous sarcoidosis, raising the question of rubella virus in sarcoidosis. *Objective:* To investigate the presence of rubella virus in sarcoidosis lung samples. *Methods:* We employed metagenomic sequencing to interrogate extracellular virome preparations and cellular transcriptomes from bronchoalveolar lavage (BAL) of 209 sarcoidosis patients for rubella virus sequences. *Results:* We found no evidence for rubella virus genomes in acellular fluid or rubella virus gene expression in BAL cells of sarcoidosis patients. *Conclusions:* These findings argue against rubella virus infection or persistence within the lung at time of sampling as a sarcoidosis trigger.

Key words: Virome, Transcriptome, Sarcoidosis, Rubella

Introduction

Sarcoidosis is a granulomatous multisystem disease with frequent involvement of lungs, skin, lymph nodes, eyes and other tissues. The pathogenesis of sarcoidosis involves aberrant T cell and macrophage immune responses, but the trigger for dysregulated immunity remains unknown. Infectious agents have been long suspected as initiators of the inflammatory response including mycobacteria and propionibacteria, but definitive evidence is lacking (1-3). Recently, rubella virus genetic sequences and/or proteins have been reported in immunocompromised patients with cutaneous granulomatous dermatitis, particularly in the setting of common variable immunodeficiency (CVID) (4-12). More recently, rubella virus was reported in skin granulomas of immunocompetent adults (13). Given the fact that many clinical and pathological features of cutaneous granulomatous dermatitis overlap with those of sarcoidosis, we sought to determine whether rubella virus sequences were present in lung specimens of sarcoidosis patients.

Received: 30 June 2022
Accepted after revision: 16 November 2022
Correspondence: Ronald Collman, MD
522 Johnson Pavilion,
University of Pennsylvania School of Medicine,
Philadelphia PA 19104, USA
E-mail: collmanr@pennmedicine.upenn.edu
Misha Rosenbach, MD
3600 Spruce Street,
University of Pennsylvania School of Medicine,
Philadelphia PA 19104, USA
E-mail: Misha.Rosenbach@pennmedicine.upenn.edu

Methods

Subjects, samples and sequencing

Bronchoalveolar lavage (BAL) specimens were obtained from well-characterized patients with sarcoidosis enrolled in the Genomic Research in Alpha1 Disease and Sarcoidosis (GRADS) study as described (14, 15). The cohort was 53% female with a mean±SD age of 51.6±10.2 years and by design GRADS recruited individuals across different phenotypes. Approximately 7% had acute sarcoidosis, 36% had Scadding II or III disease, 14% had Scadding IV, and 11% had multiorgan involvement; additional details of the cohort are provided in (14, 15). Participants were recruited at GRADS study sites based on an established diagnosis of sarcoidosis or high degree of clinical suspicion that was confirmed prior to inclusion in the study. All subjects provided informed consent to participate in the study.

BAL was separated into cellular and acellular components by centrifugation (700g x 10 minutes). BAL cell transcriptome cDNA libraries were generated from poly-A-selected RNA, sequenced on the Ion Torrent platform (Thermo Fisher, Waltham, MA, USA) and processed as previously described (15). Data are publicly available on the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo/) under the accession number GSE109516. For the acellular fraction that would contain virions, virus-like particles (VLPs) were isolated and purified from 5 mL of BAL and lavage saline controls. Nucleic acids were extracted and amplified using a sequence-independent single primer amplification method and prepared for sequencing with the NEB Next Ultra 2 Library Preparation Kit (New England Biolabs, Ipswich, MA, USA) and sequenced on the Illumina NextSeq500 (Illumina Inc., San Diego, CA, USA) with 2x150 bp reads. Data are publicly available for download from the NCBI with BioProject accession number PRJNA646239.

Rubella virus sequence analysis

To investigate the presence of rubella virus sequences in BAL cell transcriptome data, we interrogated 209 BAL datasets that passed sequence quality metrics. Thirty rubella virus genomes from GenBank (NCBI) were used as local alignment targets (Table 1). Alignments were performed using the hisss pipeline (https://github.com/louiejtaylor/hisss) (16), which uses Bowtie 2 (option, –very-sensitive-local) (17) to align reads to target genomes, SAMtools (18) and BEDtools (19) to calculate the coverage of the reads to the target genomes, and ggplot2 in R (version 3.2.3) to visualize the alignments. Positive identification was prospectively defined as ≥0.10 fraction coverage (≥10% coverage) to any of the rubella virus sequences. As a positive control, we aligned reads from rubella virus sequencing project PRJNA70479 and confirmed that our pipeline yielded robust virus identification. For acellular virion-encapsidated

Table 1. Query rubella virus genomes

GenBank accession	RubV_strain
JN635281	RVi/NJ.USA/61/1a
JN635282	RVi/Brooklyn.NY.USA/98/1B CRS
JN635283	RVi/LA.CA.USA/91/1C
JN635284	RVi/TX.USA/34.98/1C
JN635285	RVi/CA.USA/88/1D
JN635286	RVi/BarHarbor.ME.USA/43.08/1E
JN635287	RVi/Springfield.MA.USA/49.98/1E
JN635288	RVi/Pullman.WA.USA/30.08/1E
JN635289	RVi/Boston.MA.USA/13.07/1G
JN635290	RVi/Lebanon.NH.USA/3.05/1G CRS
JN635291	RVi/Daly City.CA.USA/97/1j CRS
JN635292	RVi/Kalamazoo.MI.USA/4.07/2B
JN635293	RVi/Seattle.WA.USA/16.00/2B
JN635294	RVi/LA.CA.USA/45.08/2B CRS
JN635295	RVi/Eagan.MN.USA/13.09/2B
JN635296	RVi/Bismarck.ND.USA/23.08/2B
M15240	F-therin
AF435865	Ulrike
AB047330	TO-336
AB222609	RVi/Matsue.JPN/68
DQ085339	Cordoba_Argentina_1988
DQ085341	Anima_Mexico_1997
DQ388281	JC2 NZL91
DQ085343	6423_ITALY_1997
DQ388280	GUZ GER92
AY258322	BR1
DQ085342	Anam5_Korea_1996
DQ085338	I-11_Israel_1968
DQ388279	C4 RUS67

datasets, sequencing reads were analyzed with Taxonomer (20), which uses the UniProt90 viral database to assign taxonomy.

Results

A total of 209 sarcoidosis BAL cell transcriptome datasets (with a median of 30,598,036 reads per sample) were queried. Only three short reads (one each from three separate BAL samples) mapped to rubella virus sequences. The mapped reads revealed fractional coverage of the rubella virus genomes of 0.002971 (0.2971% coverage), 0.002971 (0.2971% coverage), and 0.002868 (0.2868% coverage), respectively, far below the 0.1 fraction (10% coverage) threshold prespecified for positive virus identification. While rubella virus does express subgenomic transcripts (21) and may replicate to only very low levels in infected cells, this low coverage argues against the authentic detection of rubella virus gene expression in BAL cells.

Within the 115 acellular sarcoidosis BAL samples subject to virome analysis (with an average of 851,000 paired-end reads per sample), only a single read from one patient aligned with rubella virus.

Thus, we found no strong evidence for rubella virus in sarcoidosis lung samples.

Discussion

A microbial trigger for sarcoidosis has been long suspected due to pathological features of granulomatous inflammation and evidence for aberrant immune activation. The granulomatous dermatitis seen in patients with CVID can clinically and histologically resemble cutaneous sarcoidosis (5). The identification of rubella virus proteins and nucleic acids in granulomatous dermatitis in individuals with immune deficiency (4-12) and more recently immunocompetent patients (13) raises the question of whether this virus might be present in sarcoidosis as well. In addition, patients with CVID can also develop interstitial lung disease (22) and patients with sarcoidosis occasionally present with lymphopenia, suggesting additional features of clinical overlap. However, our study provides strong evidence against the presence of extracellular rubella virions or intracellular rubella virus transcription in lung lavage, arguing against active infection and virus replication at the time of sampling.

A limitation of our study is that we did not sample granulomatous tissue from biopsied lung or lymph node, which could conceivably harbor sequences not present in BAL. Perhaps most important, these findings do not exclude that this agent or other potential antigenic triggers might have been present at some earlier time but no longer persist by the time of sampling. Indeed, most of the patients from whom rubella virus was identified in cutaneous granulomatous disease were immunodeficient, which might contribute to viral persistence beyond an initial trigger event. It thus may be fruitful to investigate the presence of immune responses to rubella in sarcoidosis, as has been pursued for mycobacterial agents (23, 24), which may persist beyond the time of disease initiation and microbial elimination.

Acknowledgement: We thank patients who provide samples and investigators in the GRADS study who collected samples. This work was supported by NIH grants R33-HL137063 and U01-HL112712 (RGC, FDB) and by the NHLBI U01-HL112707 Genomic Research in AAT-Deficiency and Sarcoidosis (GRADS) Genomics and Informatics Center. This work was also supported in part by the Division of Intramural Research (DIR) of the NIAID/NIH (EG).

Conflicts of Interest: Each author declares that he or she has no commercial associations (e.g. consultancies, stock ownership, equity interest, patent/licensing arrangement etc.) that might pose a conflict of interest in connection with the submitted article.

Communication: This article was discussed with Dr. Baughman on 29-June-2022

References

1. Chen ES, Moller DR. Etiologic role of infectious agents. Semin Respir Crit Care Med. 2014;35(3):285-95. Epub 2014/07/10. doi: 10.1055/s-0034-1376859. PubMed PMID: 25007081.
2. Beijer E, Veltkamp M, Meek B, Moller DR. Etiology and Immunopathogenesis of Sarcoidosis: Novel Insights. Semin Respir Crit Care Med. 2017;38(4):404-16. Epub 2017/07/28. doi: 10.1055/s-0037-1603087. PubMed PMID: 28750457.
3. Clarke EL, Lauder AP, Hofstaedter CE, et al. Microbial Lineages in Sarcoidosis. A Metagenomic Analysis Tailored for Low-Microbial Content Samples. Am J Respir Crit Care Med. 2018;197(2):225-34. Epub 2017/08/29. doi: 10.1164/rccm.201705-0891OC. PubMed PMID: 28846439; PubMed Central PMCID: PMC5768903.
4. Perelygina L, Faisthalab R, Abernathy E, et al. Rubella Virus Infected Macrophages and Neutrophils Define Patterns of Granulomatous Inflammation in Inborn and Acquired Errors of Immunity. Frontiers in immunology. 2021;12:796065. Epub 2022/01/11. doi: 10.3389/fimmu.2021.796065. PubMed PMID: 35003119; PubMed Central PMCID: PMC8728873.
5. Shields BE, Perelygina L, Samimi S, et al. Granulomatous Dermatitis Associated With Rubella Virus Infection in an Adult With

Immunodeficiency. JAMA dermatology. 2021;157(7):842-7. Epub 2021/05/27. doi: 10.1001/jamadermatol.2021.1577. PubMed PMID: 34037685; PubMed Central PMCID: PMC8156178 Karger Inc and serving as a paid member of the advisory board for Kyowa Kirin. Dr Rosenbach reported receiving personal fees from Merck Consulting: Drug Reactions, grants from Processa Research; consulting fees from necrobiosis lipoidica trial planning, Janssen Pharmaceuticals, and Eli Lilly outside the submitted work. No other disclosures were reported.

6. Perelygina L, Icenogle J, Sullivan KE. Rubella virus-associated chronic inflammation in primary immunodeficiency diseases. Current opinion in allergy and clinical immunology. 2020;20(6):574-81. Epub 2020/10/13. doi: 10.1097/aci.0000000000000694. PubMed PMID: 33044342; PubMed Central PMCID: PMC7730704.
7. Dhossche J, Johnson L, White K, et al. Cutaneous Granulomatous Disease With Presence of Rubella Virus in Lesions. JAMA dermatology. 2019;155(7):859-61. Epub 2019/06/06. doi: 10.1001/jamadermatol.2019.0814. PubMed PMID: 31166586.
8. Leclerc-Mercier S, Moshous D, Neven B, et al. Cutaneous granulomas with primary immunodeficiency in children: a report of 17 new patients and a review of the literature. Journal of the European Academy of Dermatology and Venereology : JEADV. 2019;33(7):1412-20. Epub 2019/03/15. doi: 10.1111/jdv.15568. PubMed PMID: 30869812.
9. Buchbinder D, Hauck F, Albert MH, et al. Rubella Virus-Associated Cutaneous Granulomatous Disease: a Unique Complication in Immune-Deficient Patients, Not Limited to DNA Repair Disorders. J Clin Immunol. 2019;39(1):81-9. Epub 2019/01/05. doi: 10.1007/s10875-018-0581-0. PubMed PMID: 30607663; PubMed Central PMCID: PMC7739844.
10. Bodemer C, Sauvage V, Mahlaoui N, et al. Live rubella virus vaccine long-term persistence as an antigenic trigger of cutaneous granulomas in patients with primary immunodeficiency. Clin Microbiol Infect. 2014;20(10):O656-63. Epub 2014/01/31. doi: 10.1111/1469-0691.12573. PubMed PMID: 24476349.
11. Perelygina L, Plotkin S, Russo P, et al. Rubella persistence in epidermal keratinocytes and granuloma M2 macrophages in patients with primary immunodeficiencies. J Allergy Clin Immunol. 2016;138(5):1436-9.e11. Epub 2016/09/11. doi: 10.1016/j.jaci.2016.06.030. PubMed PMID: 27613149; PubMed Central PMCID: PMC5392721.
12. Perelygina L, Chen MH, Suppiah S, et al. Infectious vaccine-derived rubella viruses emerge, persist, and evolve in cutaneous granulomas of children with primary immunodeficiencies. PLoS Pathog. 2019;15(10):e1008080. Epub 2019/10/29. doi: 10.1371/journal.ppat.1008080. PubMed PMID: 31658304; PubMed Central PMCID: PMC6837625 Andrey Zharkikh is an employee of Myriad Genetics. This company works in cancer genetics and diagnostics without any relation to pathogen studies and treatments. This does not alter our adherence to all PLOS Pathogens policies on sharing data and materials.
13. Wanat KA, Perelygina L, Chen MH, et al. Association of Persistent Rubella Virus With Idiopathic Skin Granulomas in Clinically Immunocompetent Adults. JAMA dermatology. 2022;158(6):626-33. Epub 2022/03/27. doi: 10.1001/jamadermatol.2022.0828. PubMed PMID: 35338705; PubMed Central PMCID: PMC8957700 outside the submitted work. Dr Sullivan reported consulting fees from the Immune Deficiency Foundation and Enzyvant and being an editor for Elsevier. Dr Rosenbach reported personal fees from Merck, Novartis, Eli Lilly, AbbVie, and Processa outside the submitted work. No other disclosures were reported.
14. Moller DR, Koth LL, Maier LA, et al. Rationale and Design of the Genomic Research in Alpha-1 Antitrypsin Deficiency and Sarcoidosis (GRADS) Study. Sarcoidosis Protocol. Annals of the American Thoracic Society. 2015;12(10):1561-71. Epub 2015/07/21. doi: 10.1513/AnnalsATS.201503-172OT. PubMed PMID: 26193069; PubMed Central PMCID: PMC4627423.
15. Vukmirovic M, Yan X, Gibson KF, et al. Transcriptomics of bronchoalveolar lavage cells identifies new molecular endotypes of sarcoidosis. Eur Respir J. 2021;58(6). Epub 2021/06/05. doi: 10.1183/13993003.02950-2020. PubMed PMID: 34083402.
16. Abbas AA, Taylor LJ, Dothard MI, et al. Redondoviridae, a Family of Small, Circular DNA Viruses of the Human Oro-Respiratory Tract Associated with Periodontitis and Critical Illness. Cell Host Microbe. 2019;25(5):719-29.e4. Epub 2019/05/10. doi: 10.1016/j.chom.2019.04.001. PubMed PMID: 31071295; PubMed Central PMCID: PMC6510254.
17. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9. Epub 2012/03/06. doi: 10.1038/nmeth.1923. PubMed PMID: 22388286; PubMed Central PMCID: PMC3322381.
18. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. Epub 2009/06/10. doi: 10.1093/bioinformatics/btp352. PubMed PMID: 19505943; PubMed Central PMCID: PMC2723002.
19. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2. Epub 2010/01/30. doi: 10.1093/bioinformatics/btq033. PubMed PMID: 20110278; PubMed Central PMCID: PMC2832824.
20. Flygare S, Simmon K, Miller C, et al. Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. Genome Biol. 2016;17(1):111. Epub 2016/05/27. doi: 10.1186/s13059-016-0969-1. PubMed PMID: 27224977; PubMed Central PMCID: PMC4880956.
21. Oker-Blom C, Ulmanen I, Kääriäinen L, Pettersson RF. Rubella virus 40S genome RNA specifies a 24S subgenomic mRNA that codes for a precursor to structural proteins. J Virol. 1984;49(2):403-8. Epub 1984/02/01. doi: 10.1128/jvi.49.2.403-408.1984. PubMed PMID: 6694262; PubMed Central PMCID: PMC255479.
22. Mooney D, Edgar D, Einarsson G, Downey D, Elborn S, Tunney M. Chronic lung disease in common variable immune deficiency (CVID): A pathophysiological role for microbial and non-B cell immune factors. Critical reviews in microbiology. 2017;43(4):508-19. Epub 2017/01/11. doi: 10.1080/1040841x.2016.1268568. PubMed PMID: 28068853.
23. Wikén M, Ostadkarampour M, Eklund A, et al. Antigen-specific multifunctional T-cells in sarcoidosis patients with Lofgren's syndrome. Eur Respir J. 2012;40(1):110-21. Epub 2011/12/21. doi: 10.1183/09031936.00166110. PubMed PMID: 22183480.
24. Richmond BW, Ploetze K, Isom J, et al. Sarcoidosis Th17 cells are ESAT-6 antigen specific but demonstrate reduced IFN-γ expression. J Clin Immunol. 2013;33(2):446-55. Epub 2012/10/18. doi: 10.1007/s10875-012-9817-6. PubMed PMID: 23073617; PubMed Central PMCID: PMC3850764.