




SCIENTIFIC REPORTS



OPEN

RNA-sequencing Identifies Novel Pathways in Sarcoidosis Monocytes

Jaya Talreja¹, Pershang Farshi¹, Adnan Alazizi², Francesca Luca ^{2,3}, Roger Pique-Regi ^{2,3} & Lobelia Samavati ^{1,2}

Received: 7 December 2016
Accepted: 25 April 2017
Published online: 02 June 2017

Sarcoidosis is a complex systemic granulomatous disorder of unknown etiology. Genome-wide association studies have not been able to explain a causative role for nucleotide variation in its pathogenesis. The goal of the present study was to identify the gene expression profile and the cellular pathways altered in sarcoidosis monocytes via RNA-sequencing. Peripheral blood monocytes play a role in sarcoidosis inflammation. Therefore, we determined and compared the transcriptional signature of monocytes from peripheral blood from sarcoidosis patients and healthy controls via RNA-sequencing. We found 2,446 differentially expressed (DE) genes between sarcoidosis and healthy control monocytes. Analysis of these DE genes showed enrichment for ribosome, phagocytosis, lysosome, proteasome, oxidative phosphorylation and metabolic pathways. RNA-sequencing identified upregulation of genes involved in phagocytosis and lysosomal pathway in sarcoidosis monocytes, whereas genes involved in proteasome degradation and ribosomal pathways were downregulated. Further studies are needed to investigate the role of specific genes involved in the identified pathways and their possible interaction leading to sarcoidosis pathology.

Sarcoidosis is a multisystem granulomatous disease of unknown etiology characterized by granuloma formation in barrier organs such as lungs, skin, and eyes. It results in significant morbidity and mortality, primarily from respiratory failure in the US¹⁻³. It is thought that exposure to the airborne antigens in a genetically susceptible host leads to disease initiation²⁻⁵. Granulomatous inflammation in sarcoidosis is characterized by activated macrophages as well as CD4⁺ T cell infiltration with predominant Th1 cytokine expression^{6,7}.

Several studies attempted to identify genetic risk of developing sarcoidosis using various techniques such as a candidate gene approach, microarray of candidate genes, and genetic linkage analysis^{3,8,9}. Several genes were linked to sarcoidosis susceptibility including major histocompatibility complex (MHC) and HLA antigens class I and II such as loci HLA-B8 and HLA-DRB1^{3,8,10}. Similar candidate gene approaches additionally identified susceptibility genes involved in antigen processing, antigen presentation, macrophage and T-cell activation, and genes involved in injury repair¹¹⁻¹⁴.

The similarities of sarcoidosis to infectious granulomatous diseases and its association with major histocompatibility loci suggest a major role for inciting microbial triggers^{15,16}. The presence of activated macrophages and the expansion of oligoclonal T cells and B cells suggest an important role of immunity in this disease^{2,17}. Several lines of evidence indicate that aberrant functioning of macrophages, dendritic cells, and monocytes may underlie the Th1 skewness in sarcoidosis^{18,19}. Similarly, previous studies from our laboratory demonstrated that sustained activation of p38 MAPK is associated with a lack of a negative feedback loop through kinase phosphatase (MKP)-1 leading to persistent inflammation in macrophages^{19,20}. Inhibition of inflammatory signaling in macrophages and monocytes led to modulation of activated T cells and their responses to mitogens¹⁹.

Our goal was to unravel the transcriptional signature, and to identify the genes and associated pathways that may be dysregulated in monocytes in sarcoidosis. To this end we interrogated the poly-adenylated fraction of the transcriptome via RNA-sequencing, thus reaching unprecedented depth in the analysis of the gene regulation events that are altered. To our knowledge, this is the first RNA-seq study to compare the transcriptional signature of peripheral blood monocytes from sarcoidosis patients and healthy controls.

¹Department of Internal Medicine, Division of Pulmonary, Critical Care and Sleep Medicine, Wayne State University School of Medicine and Detroit Medical Center, Detroit, MI, 48201, USA. ²Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, 540 E. Canfield, Detroit, MI, 48201, USA. ³Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, MI, 48201, USA. Correspondence and requests for materials should be addressed to L.S. (email: lsamavat@med.wayne.edu)

Characteristic	Patients	Control Subjects
Age (year)	46.11 ± 14.4	48.7 ± 13.4
BMI	29 ± 10.4	28 ± 3.6
Gender, N (%)		
Female	9 (45)	10 (50)
Male	11 (55)	10 (50)
Self-reported Race, N (%)		
African American	20(100)	20 (100)
White	0 (0)	0 (0)
CXR stage, N (%)		
0	0 (0)	NA
1	0 (0)	NA
2	6 (60)	NA
3	4 (40)	NA
Organ Involvements, N (%)		
Neuro-opthalmologic	2(10)	NA
Lung	20 (100)	NA
Skin	2 (10)	NA
Multiorgans	15 (75)	NA
PPD	Negative	Negative
AFB/culture	Negative	NA
IGRA	Negative	NA

Table 1. Subject Demographics Definition of abbreviations: BMI = body mass index, CXR = chest X-ray, NA = not applicable, PPD = purified protein derivative, AFB = acid-fast bacilli, IGRA = interferon gamma release assay.

Results

Differential gene expression and pathway analysis in monocytes. The study included two groups comprising of 20 patients with sarcoidosis and 20 healthy controls. The subject demographics are displayed in Table 1. There was no significant difference in age and BMI between patients and healthy controls ($p > 0.05$). To reduce the effect of potential confounding factors associated with variation in ancestry proportions, we included subjects with the same self-reported race. All subjects were non-smokers, and none was on immune-suppressive medication. All sarcoidosis subjects had lung involvement with chest radiograph stage 2 or 3. We prepared RNA-seq libraries from mRNA isolated from the peripheral blood monocytes obtained from sarcoidosis patients and healthy controls. The schematic study design of RNA-seq library preparation, work flow and analysis are shown in Supplementary Fig. S1. Alignment of the filtered reads to the reference human genome hg19 showed about 90% total aligned reads per RNA-seq library sample.

First, we compared the DE genes between the healthy control and sarcoidosis monocytes using DEseq2. Figure 1a shows the distribution of \log_2 -fold change and expression level for the 8,069 genes expressed in these samples. The DE genes (2,446; FDR < 5%) between sarcoidosis and healthy control monocytes are highlighted in red. Among DE genes, 1,221 were downregulated whereas 1,225 were upregulated in sarcoidosis monocytes. To determine the pathways that are modulated in the sarcoidosis monocytes compared to healthy control monocytes, we performed KEGG pathway analysis using iPathwayGuide and the Gene trail tool on the DE genes (\log_2 -fold change of 0.6 and FDR < 5%). Figure 1b shows the score plot of pathways using impact analysis. Impact analysis uses two types of evidence: i) the over-representation of DE genes in a given pathway and ii) the perturbation of that pathway computed by propagating the measured expression changes across the pathway topology^{21–23}. The pathway topologies consist of genes and their interactions that are obtained from the KEGG database^{24–27}. Significant pathways with FDR < 5% are shown in red and non-significant in black (Fig. 1b). The size of the circle is proportional to the number of genes involved in that pathway. Topological analysis identified 103 pathways significantly enriched in sarcoidosis as compared to healthy controls. The most significant pathways were: ribosome (FDR = 1×10^{-24}), metabolic pathways (FDR = 5.1×10^{-14}), oxidative phosphorylation (FDR = 4.0×10^{-6}), lysosome (FDR = 1.7×10^{-5}), Fc γ R-mediated phagocytosis (FDR = 1.5×10^{-4}), phagosome (FDR = 1.9×10^{-2}), and proteasome (FDR = 1.9×10^{-2}).

Furthermore, pathway analysis identified that genes involved in monocyte activation, inflammation, and innate and adaptive immune responses were highly enriched in sarcoidosis monocytes. Especially, gene expression was increased in sarcoidosis monocytes for various pathways involved in phagosome, lysosome, monocyte activation, and leukocyte migration and trafficking. In contrast, genes involved in proteasome and ribosomal biogenesis were found to be downregulated in sarcoidosis monocytes as compared to control monocytes. Figure 2 shows the heat-map of sarcoidosis DE genes involved in three of the significant (FDR < 5%) pathways: (a) metabolic pathway (most of the genes involved in this pathway were downregulated), (b) phagocytosis (most of the genes involved in this pathway were upregulated), (c) proteasome pathway (genes involved in this pathway were found to be downregulated).

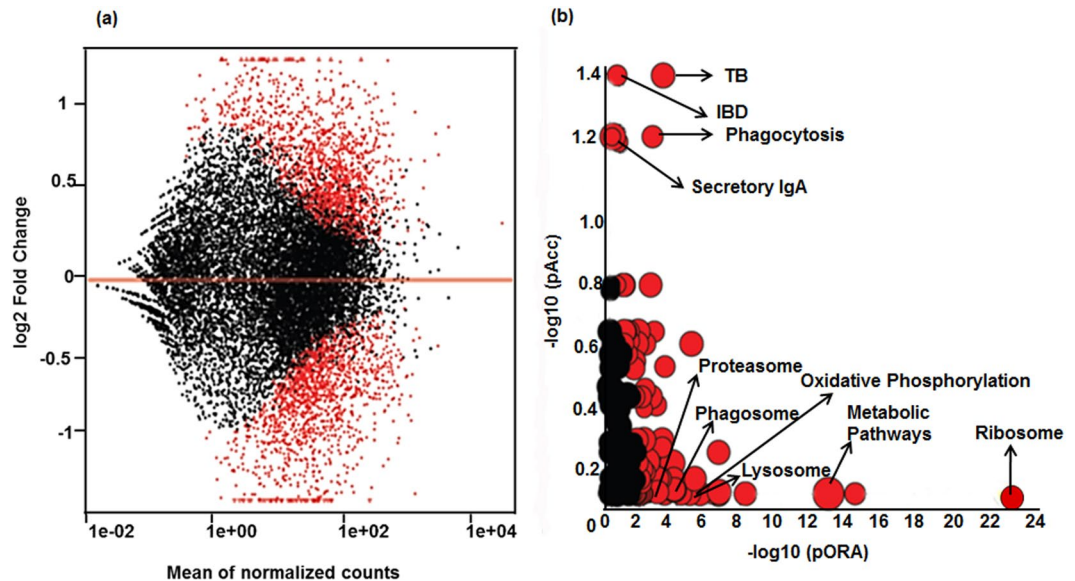


Figure 1. Differential gene expression and pathways between control and sarcoidosis monocytes. **(a)** Scatter plot of the entire gene expression data (healthy versus sarcoidosis monocytes) analyzed by DESeq2 analysis tool, where the \log_2 -fold change (>0.5) of each gene is plotted against the total number of counts recorded for that gene. DE genes (FDR $< 5\%$) are highlighted in red. **(b)** Significant pathways modulated in sarcoidosis monocytes as compared to healthy controls. Pathway analysis was done on the DE genes (\log_2 -fold change > 0.6 and FDR $< 5\%$) using iPathwayGuide analysis tool that uses two types of evidence: the over-representation on the horizontal axis (pORA) and the perturbation on the vertical axis (pAcc). Significant pathways (FDR $< 5\%$) are shown in red, whereas non-significant are in black. The size of the circle is proportional to the number of genes in that pathway.

The top enriched pathway was ribosome pathway (Fig. 3). The genes involved in this pathway were downregulated as compared to control monocytes. Figure 4 shows lysosome pathway that was highly enriched in sarcoidosis monocytes. Most of genes related to lysosomal acid hydrolases and lysosomal membrane proteins were found to be upregulated. Whereas, the genes involved in proteasome pathway were found to be downregulated (Fig. 5). The genes involved in the assembly of regulatory and core particle of proteasome were downregulated. The immunoproteasome gene PA28 γ (also known as PSME3) that encodes the gamma subunit of 11 S regulator of immunoproteasome was downregulated in sarcoidosis monocytes.

Validation of RNA-seq data via qRT-PCR. To validate the RNA-seq data, we evaluated the expression levels of selected DE genes. RNA was isolated from the monocytes obtained from an independent set of sarcoidosis patients ($n = 10$) and healthy controls ($n = 10$) and was subjected to qRT-PCR using selected primers. Our RNA-seq data showed that the phagocytosis, oxidative phosphorylation, ribosomal, and inflammatory pathways are impacted in sarcoidosis. Among DE genes involved in phagocytosis, we chose the gene encoding ATPase H⁺ transporting accessory protein 1 (ATP6AP1) for validation. This gene encodes an essential multitasking protein important for housekeeping acidification of membrane-bound compartments including acidifying endosomes, lysosomes, phagosomes, compartments for uncoupling receptors, and autophagosomes and it is involved in oxidative phosphorylation²⁸. We selected the gene encoding the lysosomal membrane proteins (LAMP)-2, which is integral part of lysosomal membrane²⁹ and was highly expressed in sarcoidosis. Furthermore, we chose the gene encoding cytochrome B-245 β chain that is an integral part of oxidative phosphorylation and its variation has been reported in chronic granulomatous disease³⁰. Serpin family member 1 (SERPINA1) encodes peptidase inhibitor clade 1, which is α -antitrypsin precursor³¹. We also chose two genes from ribosomal pathways for validation. Genes encoding the protein SA of 40S subunit (RPSA), which is involved in ribosomal assembly, and the ribosomal protein L10a (RPL10A), which is the component of 60S subunit were selected for further validation.

We validated selected genes from those pathways using qRT-PCR using primers for the following genes, ATP6AP1, LAMP2, CYBB, SERPINA1, RPSA and RPL10A. The relative gene expression levels for ATP6AP1, CYBB, LAMP2, and SERPINA1 were significantly higher in sarcoidosis monocytes as compared to healthy monocytes (Fig. 6a–d). Additionally, both RPSA and RPL10A genes encoding ribosomal proteins were found to be significantly downregulated in sarcoidosis monocytes as compared to healthy monocytes (Fig. 6e and f). The relative gene expression data are presented as box plots as mean \pm SEM from ten subjects with sarcoidosis monocytes and ten healthy controls monocytes. *Represents a p value < 0.05 and **a p < 0.001 using a paired student *t*-test. These qRT-PCR data confirm in independent sets of patients and controls the RNA-seq data.

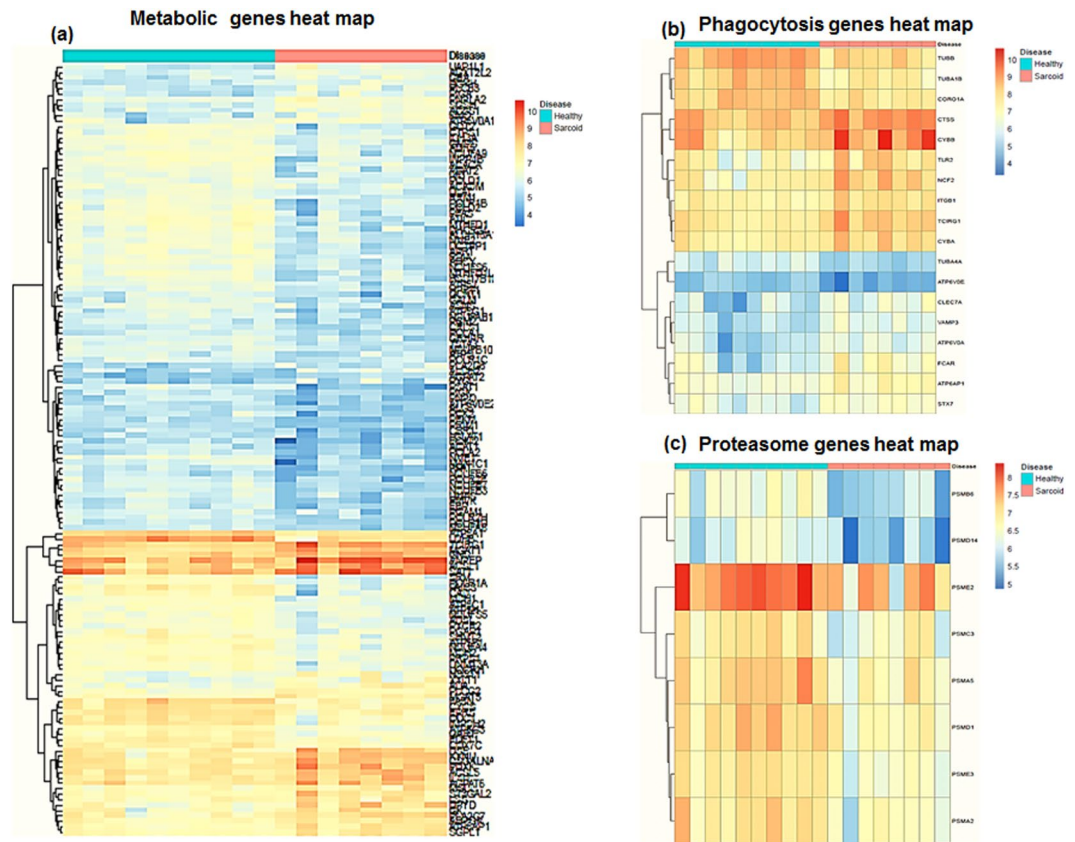


Figure 2. Heat maps of DE genes of three significant pathways in sarcoidosis versus healthy monocytes. **(a)** Heatmap of the 143 genes involved in metabolic pathways between two groups. **(b)** Heatmaps of 19 genes involved in phagocytosis and **(c)** Heatmaps of 9 genes involved in proteasome (\log_2 -fold change > 0.5 & FDR $< 5\%$). Dendrograms according to means identifying genes levels in all three pathways show two distinct clusters. Red shade represents high expression and blue shade represents low expression.

Discussion

It is well accepted that sarcoidosis is a polygenic and multifactorial disease³². Numerous association studies attempted to identify genetic susceptibility in sarcoidosis and various genes were found to increase the risk of developing this disease, among which the chemokine receptors, the tumor necrosis factor (TNF)- α , and several HLA loci and MHC class II antigens^{3,32}. Despite identifying some genes that may contribute to the risk of developing sarcoidosis or modify its severity^{3,33}, none have fully explained the complex nature of this disease. Several studies using heterogeneous samples, including peripheral blood, tissue samples or BALs, attempted to identify a gene signature in this disease^{34–36}. Here, we investigated and compared the transcriptional profile of monocytes, one important player in the formation of multi-nucleated giant cells in sarcoidosis granuloma³⁷. Our aim was to determine the transcriptional responses of peripheral blood monocytes and to unravel the cellular mechanisms responsible for the immunopathogenesis of sarcoidosis.

We found that 2,446 genes are significantly differentially expressed in monocytes between sarcoidosis and healthy controls. Functional analysis of DE genes showed the enrichment for various critical cellular pathways including ribosome and metabolic pathways, the phagosome and phagocytosis, lysosome, oxidative phosphorylation, proteasome, purine metabolism and cellular pathways involved in the inflammatory, innate and adaptive immune response.

One of the highly significant processes was the ribosomal pathway (FDR = 1×10^{-24}). In sarcoidosis monocytes 62 genes related to ribosome were significantly downregulated compared to control monocytes (Fig. 3). In contrast, genes encoding mitochondrial ribosomes were upregulated. This might be due to different regulation between mitochondrial and cytosolic ribosomes. The biogenesis of ribosomes is an extremely energetically expensive cellular process that has long been linked to human health and disease. More recently, it has been shown that ribosome biogenesis is intimately linked to multiple cellular signaling pathways and that defects in ribosome production can lead to a wide variety of human diseases³⁸. This novel pathway previously has not been described in sarcoidosis. Further studies are needed to determine the functional contribution of ribosomal pathways in sarcoidosis pathology.

We found about 50 DE genes that are involved in phagosome and phagocytosis in sarcoidosis monocytes compared to control monocytes. Among these genes, several were related to pattern recognition molecules, inflammatory signaling, leukocyte migration, and trafficking. In particular, genes encoding TLR2, RIPK2, CLEC7A,

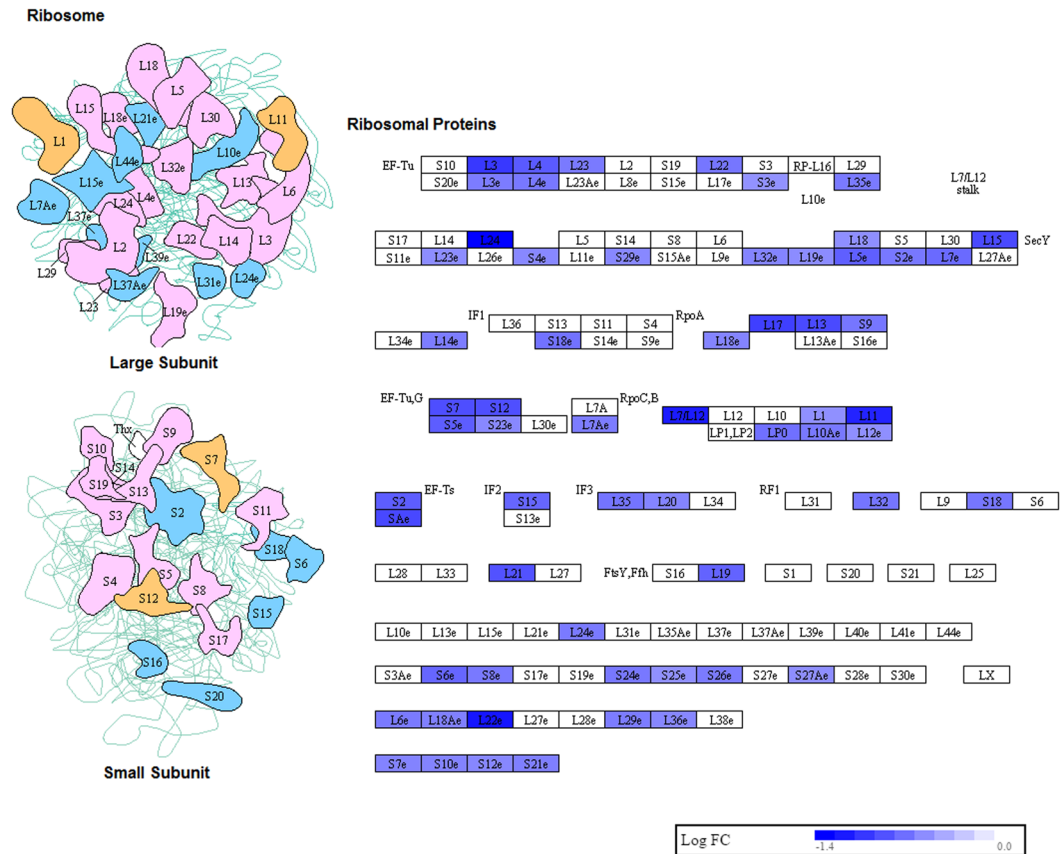


Figure 3. Downregulation of genes involved in ribosomal pathway in sarcoidosis monocytes. Graphic illustration of pathway analysis of DE (\log_2 -fold change with FDR < 0.05) genes related to ribosome in sarcoidosis monocytes. The pathway diagram is overlaid with the computed perturbation of each gene. The perturbation accounts both for the gene's measured fold changes and for accumulated perturbation propagated from any upstream genes (accumulation). The color intensity corresponds to the level of upregulation (red) or downregulation (blue) of the DE genes in sarcoidosis monocytes versus healthy monocytes. Note: A number of genes encoding large and small ribosomal subunits were downregulated.

NLRP3 and MAP2K1 were upregulated. Similarly, we found genes involved in cytokine-cytokine receptor (CCR) signaling, and lipid metabolism that were aberrant in sarcoidosis^{35,39,40}.

Another important pathway related to sarcoidosis pathology is the lysosomal pathway⁵. We identified a large number of genes involved in lysosomal acid hydrolases, lysosomal membrane proteins and various genes related to the function of lysosomes such as the V-ATPase complex that were upregulated in sarcoidosis monocytes as compared to healthy monocytes. Lysosomes are subcellular organelles found in all animal cells that digest cellular debris, damaged organelles and invading microorganisms. Lysosomal membrane proteins maintain the integrity of lysosomes and regulate lysosomal trafficking, fusion and intralysosomal pH. Aberrant lysosomal function in sarcoidosis has long been postulated⁴¹. One conventional immune modulatory therapies in sarcoidosis is Chloroquine⁴² which has been shown to inhibit lysosomal function and modulates autophagy⁴³. Surprisingly, the LAPTM4B gene, which is involved in lysosomal machinery and has been shown to interact with ceramide to facilitate its removal from the late endosomal organelles was downregulated⁴⁴ (Fig. 4). In contrast to genes involved in lysosomal pathways, genes involved in proteasome and immune-proteasome assembly were found to be downregulated in sarcoidosis monocytes (Fig. 5). Proteasomes are macromolecular proteins that are responsible for degradation of intracellular proteins in an ATP dependent manner⁴⁵. Proteasomes degrade regulatory proteins and misfolded or damaged proteins⁴⁵. Along with lysosomal function, proteasomes play a crucial role in protein homeostasis.

We identified several major dysregulated pathways related to metabolism and oxidative phosphorylation in sarcoidosis monocytes. There were 143 differentially expressed genes related to metabolism. Among genes involved in metabolism, we found a predominant downregulation of genes involved in purine metabolism (FDR = 2.9×10^{-7}). A major set of metabolic genes was involved in oxidative phosphorylation and mitochondrial electron transfer chain. Recently, it has been recognized that metabolic reprogramming is a prerequisite for inflammation and immunological responses to diverse stimuli, both in macrophages and lymphocytes⁴⁶⁻⁴⁸. However, most of these studies were done in animal models or isolated macrophages and lymphocytes⁴⁶. In the

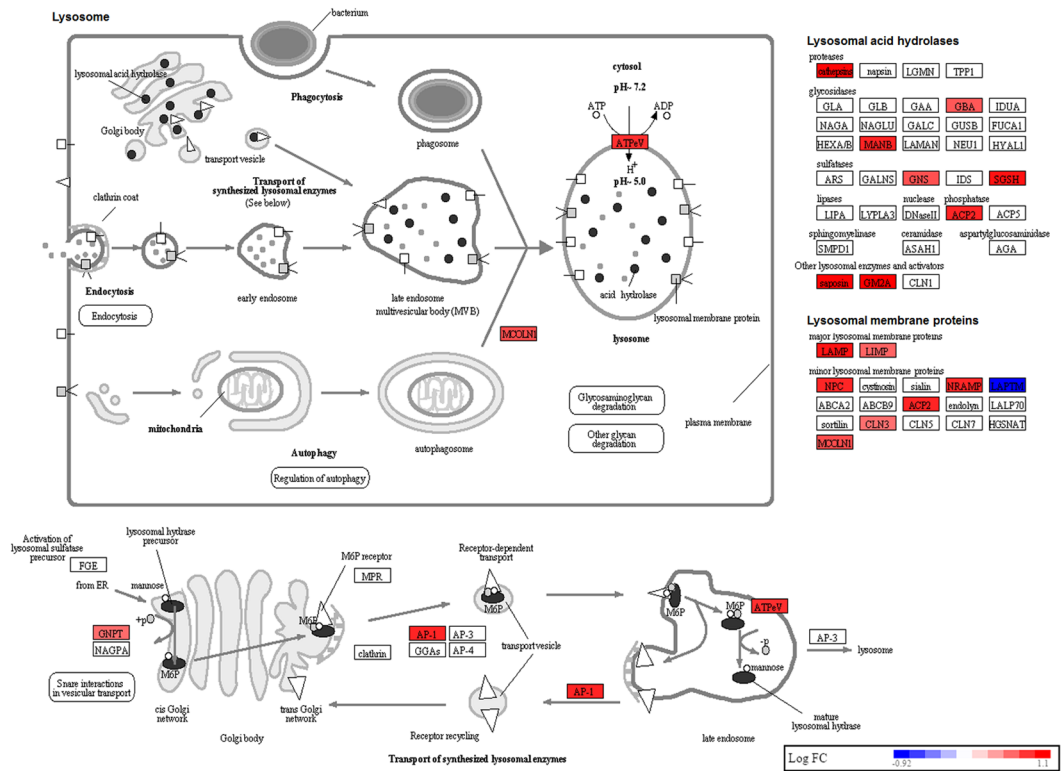


Figure 4. Upregulation of genes involved in lysosomal pathway in sarcoidosis monocytes. Graphic illustration of pathway analysis of DE (\log_2 -fold change with FDR < 0.05) genes related to lysosome in sarcoidosis monocytes. The pathway diagram is overlaid with the computed perturbation of each gene. The perturbation accounts both for the gene’s measured fold changes and for accumulated perturbation propagated from any upstream genes (accumulation). The color intensity corresponds to the level of upregulation (red) or downregulation (blue) of the DE genes in sarcoidosis monocytes versus healthy monocytes. Note: Various genes involved in lysosomal membrane proteins and acid hydrolases were upregulated except LAPT4B gene which was downregulated.

current study, we also found several differentially expressed genes involved in fatty acid and cholesterol metabolism in sarcoidosis monocytes.

We found that various genes involved in cholesterol synthesis and β -oxidation pathways, including acetyl CoA transferases, were lower expressed in sarcoidosis monocytes as compared to healthy monocytes. Other genes involved in lipid metabolisms, including several genes involved in mitochondrial β -oxidation (e.g. various isoforms of Acetyl-Coenzyme A acetyltransferases (ACATs) appear to be unique in sarcoidosis monocytes. No previous studies have identified these genes being associated with the differentiation of human monocytes or macrophages^{49–51}. Previously, in an independent study using ¹H nuclear magnetic resonance (NMR)-based untargeted metabolomic approach, our group showed aberrant metabolism in sarcoidosis⁵². In that study, we found increased levels of several intermediates related to fatty acid metabolism and TCA cycle in sera of sarcoidosis subjects⁵². Our current RNA-seq data confirms our previous metabolic data by identifying specific genes involved in mitochondrial β -oxidation, e.g. ACATs.

Our RNA-sequencing results demonstrate that sarcoidosis monocytes exhibit increased gene expression for phagocytosis and lysosomal genes that may result in the increased uptake of extracellular particles and lysosomal degradation. In contrast, we found a significant down regulation of genes involved in proteasome degradation in sarcoidosis monocytes compared to control monocytes. This may indicate that the accumulation of intracellular proteins or phagocytic degradation products may be the cause of the persistent inflammation in sarcoidosis. Further studies need to delineate in detail the functional role of the identified changes in gene expression in sarcoidosis. This dataset will be useful to develop a model organism for sarcoidosis disease or to investigate potential target drugs.

Methods

Subjects. The Committee for Investigations Involving Human Subjects at Wayne State University approved the protocol for obtaining blood by phlebotomy from control subjects and patients with sarcoidosis. The IRB number for this study is 055208MP4E. All methods were performed in accordance with the relevant guidelines and regulations. Informed consent was obtained from all subjects enrolled for the study. Sarcoidosis diagnosis was based on the ATS/ERS/WASOG statement⁵. The criteria for enrollment in the diseased group were: (i) a compatible clinical/radiographic picture consistent with sarcoidosis, (ii) histologic demonstration of non-caseating

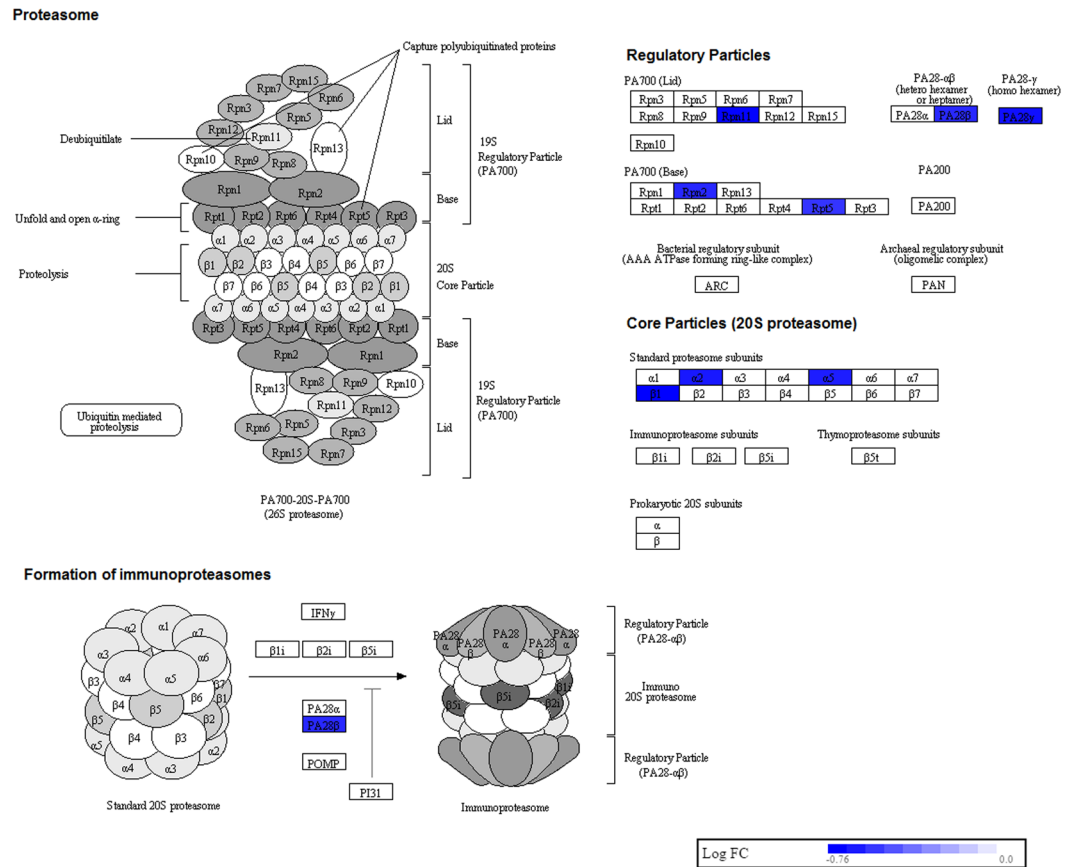


Figure 5. Proteasome in sarcoidosis monocytes. Graphic illustration of pathway analysis of DE (\log_2 -fold change with $FDR < 0.05$) genes related to proteasome and immunoproteasome in sarcoidosis monocytes. The pathway diagram is overlaid with the computed perturbation of each gene. The perturbation accounts both for the gene's measured fold changes and for accumulated perturbation propagated from any upstream genes (accumulation). Note: Several genes involved in regulatory and core particles and immunoproteasomes were downregulated (blue) in sarcoidosis monocytes.

granulomas on the tissue biopsy, and (iii) exclusion of other diseases capable of producing a similar histologic or clinical picture, such as infection by fungus or mycobacteria. Subjects excluded were: (i) smokers, (ii) those receiving immune suppressive medication (defined as corticosteroid alone and/or immune modulatory medications), (iii) those with positive microbial culture in routine laboratory examinations or viral infection; or (iv) those with known hepatitis or HIV infections or any immune suppressive condition. The criteria for enrollment in the control group were: (i) absence of any chronic respiratory diseases, (ii) lifetime nonsmoker, (iii) absence of HIV or hepatitis infection. A total of 20 patients with sarcoidosis and 20 controls participated in this study.

Isolation of PBMCs and purification of monocytes. PBMCs were isolated from heparinized blood using Ficoll-Histopaque (Sigma, St. Louis, MO) density gradient separation method as previously described^{19, 20}. CD14⁺ monocytes were purified from PBMCs by using the MACS monocyte isolation kit (Miltenyl Biotech, San Diego, CA) according to the manufacturer's instructions. The purity of enriched monocytes was evaluated by flow cytometry using fluorochrome-conjugated CD14 antibody; the purity of monocytes was about 95%.

mRNA isolation. Purified monocytes were incubated overnight in endotoxin-free RPMI 1640 medium (HyClone) supplemented with L-glutamine (Life Technologies), penicillin/streptomycin (Life Technologies), and FCS (HyClone) at 37 °C and in air containing 5% CO₂. Cells were washed twice with ice cold PBS and pellets were collected by centrifuging at 2000 rpm. Collected pellets were lysed with Lysis/Binding Buffer (Ambion) and frozen at -80 °C. Poly-adenylated mRNAs were subsequently isolated from thawed lysates using the Dynabeads mRNA Direct Kit (Ambion) following the manufacturer's instructions.

RNA-seq library preparation and sequencing. RNA-seq libraries were prepared using the NEBNext ultradirectional library preparation protocol (New England BioLabs, Ipswich, MA). The individual libraries were quantified using the KAPA real-time PCR system, following the manufacturer instructions and using a custom-made series of standards obtained from serial dilutions of the phi-X DNA⁵³. Individually barcoded RNA-seq libraries were pooled in equimolar quantities. RNA-seq library quality was assessed using an Agilent

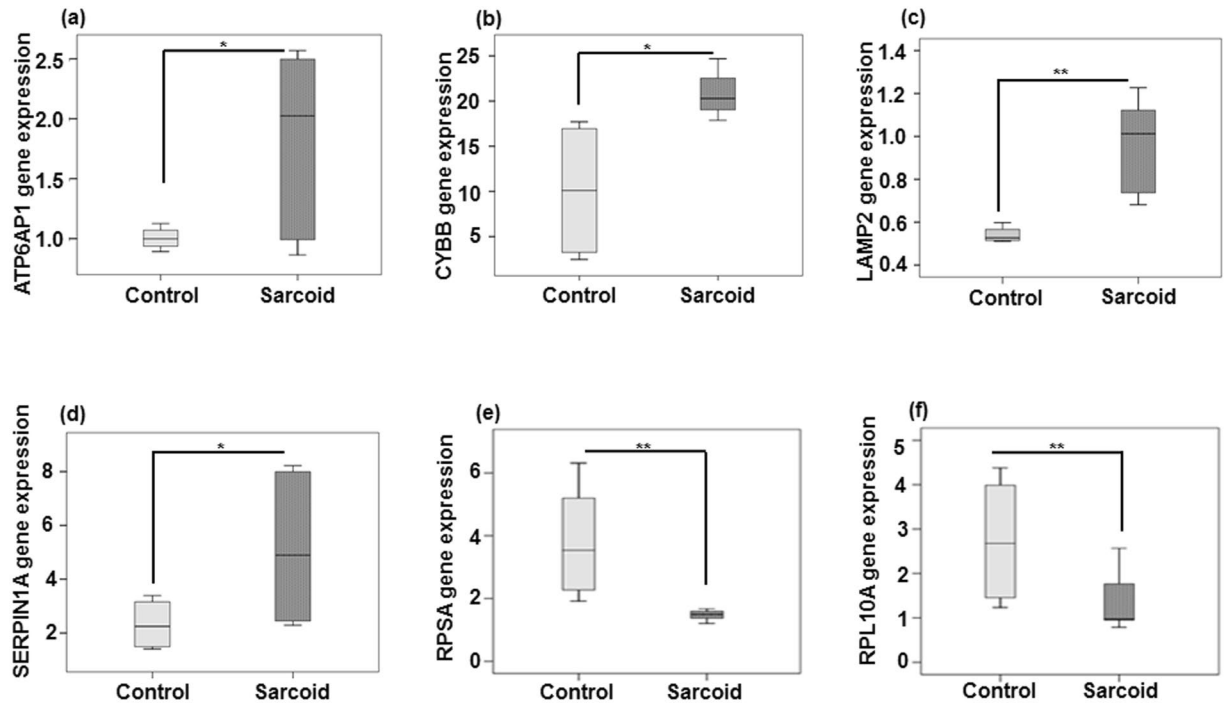


Figure 6. Validation of RNA-seq data by qRT-PCR. Total RNA was extracted from the monocytes from independent sets of 10 sarcoidosis patients and 10 healthy controls. Isolated RNAs were and reverse-transcribed using the Reverse Transcription System. The primers targeted (a) ATP6AP1, (b) CYBB, (c) LAMP2, (d) SERPIN1A, (e) RPSA and (f) RPL10A to amplify cDNA using iQSYBR Green Supermix. Relative mRNA levels were calculated by normalizing to β -actin. Box plots represent the normalized expression level of each gene of monocytes of healthy controls and sarcoidosis monocytes. Data were analyzed using the paired, two-tailed Student's *t* test and the results were expressed as fold change. *Represents a *p* value < 0.05 and **signifies a *p* < 0.001.

Bioanalyzer. A pooled library of 20 samples (10 sarcoidosis monocytes, and 10 healthy monocytes) was sequenced on the Illumina Next-Seq 500 (75 cycles, PE).

RNA-seq data analysis, differential gene expression, and canonical pathway analysis. RNA-seq data were analyzed using the Illumina Basespace RNA express application (app). In this app, the sequencing reads were aligned to the reference human genome hg19 using STAR aligner and differentially expressed (DE) gene data was obtained with the DESeq2 analysis tool^{54,55}. Gene Trail and iPathwayGuide tools were used to identify the cellular pathways and gene ontology categories that were overrepresented based on the list of DE genes (FDR < 5%). The pathways with FDR < 5% were considered to be significant. The type of FDR adjustment used was Benjamini and Hochberg. Heat maps were generated using the DESeq2 package in R-studio.

RNA Extraction and Quantitative Reverse Transcriptase/Real Time-PCR (qRT-PCR). To validate the RNA-seq data, we performed qRT-PCR of selected genes enrolling 10 independent subjects in each study group. RNA was isolated from monocytes of 10 different sarcoidosis patients and 10 healthy controls and qRT-PCR was done as described previously⁵⁶. Relative mRNA expression levels were calculated and normalized to β -actin using delta Ct method. Statistical analysis was done using paired, two-tailed Student's *t*-test. The following primers were used for PCR reactions: β -actin forward; 5'-CATGTACGTTGCTATCCAGGC 3' and reverse; 5' CTCCTTAATGTCACGCACGAT-3', ATP6AP1 forward; 5'-CAAGTCCGAAGATGTCCCATAC-3' and reverse; 5'- CTTGTACGCCACAGAGAAGTT-3'; CYBB forward; 5'-CTGGACAGGAATCTCACCTTTC-3' and reverse; 5'-GCCTCCTTCAGGGTTCTTTATT-3', SerpinA1 forward; 5'-GAAGGTCTGCCAGCTTACAT-3' and reverse; 5'-GCAGCTTCAGTCCCTTTCT-3', LAMP2 forward; 5'-CCAACACTACTGGGATGTTCT-3' and reverse; 5'-GCAGACAAGTATCATTGCCATT-3'; RPSA forward; 5'-TAGGTGGCACCAATCTTGAC-3 and reverse; 5'-GCAGCAAACCTCAGCACAG-3, RPL10A forward, 5'-GCACTGTGACGAGGCTAAG-3' and reverse; 5'-TTGTGTGTGAGCAGGGAAG-3'.

References

- Baughman, R. P. & Lower, E. E. Who dies from sarcoidosis and why? *American journal of respiratory and critical care medicine* **183**, 1446–1447, doi:10.1164/rccm.201103-0409ED (2011).
- Iannuzzi, M. C., Rybicki, B. A. & Teirstein, A. S. Sarcoidosis. *N Engl J Med* **357**, 2153–2165, doi:10.1056/NEJMra071714 (2007).
- Iannuzzi, M. C. & Rybicki, B. A. Genetics of sarcoidosis: candidate genes and genome scans. *Proceedings of the American Thoracic Society* **4**, 108–116, doi:10.1513/pats.200607-141JG (2007).
- Costabel, U. Sarcoidosis: clinical update. *Eur Respir J Suppl* **32**, 56s–68s (2001).

5. Costabel, U. & Hunninghake, G. W. ATS/ERS/WASOG statement on sarcoidosis. Sarcoidosis Statement Committee. American Thoracic Society. European Respiratory Society. World Association for Sarcoidosis and Other Granulomatous Disorders. *Eur Respir J* **14**, 735–737 (1999).
6. Muller-Quernheim, J. Sarcoidosis: immunopathogenetic concepts and their clinical application. *Eur Respir J* **12**, 716–738, doi:10.1183/09031936.98.12030716 (1998).
7. Iida, K. *et al.* Analysis of T cell subsets and beta chemokines in patients with pulmonary sarcoidosis. *Thorax* **52**, 431–437, doi:10.1136/thx.52.5.431 (1997).
8. Rivera, N. V. *et al.* High-Density Genetic Mapping Identifies New Susceptibility Variants in Sarcoidosis Phenotypes and Shows Genomic-driven Phenotypic Differences. **193**, 1008–1022, doi:10.1164/rccm.201507-1372OC (2016).
9. Rybicki, B. *et al.* Genetic linkage analysis of sarcoidosis phenotypes: the sarcoidosis genetic analysis (SAGA) study. *Genes and immunity* **8**, 379–386, doi:10.1038/sj.gene.6364396 (2007).
10. Rossman, M. D. *et al.* HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. *The American Journal of Human Genetics* **73**, 720–735, doi:10.1086/378097 (2003).
11. Pabst, S. *et al.* Toll-like receptor (TLR) 4 polymorphisms are associated with a chronic course of sarcoidosis. *Clinical & Experimental Immunology* **143**, 420–426, doi:10.1111/j.1365-2249.2006.03008 (2006).
12. Morohashi, K., Takada, T., Omori, K., Suzuki, E. & Gejyo, F. Vascular endothelial growth factor gene polymorphisms in Japanese patients with sarcoidosis. *CHEST Journal* **123**, 1520–1526, doi:10.1378/chest.123.5.1520 (2003).
13. Valentonyte, R. *et al.* Study of CC chemokine receptor 2 alleles in sarcoidosis, with emphasis on family-based analysis. *American journal of respiratory and critical care medicine* **171**, 1136–1141, doi:10.1164/rccm.200405-658OC (2005).
14. Kelly, D. M. *et al.* Endotoxin up-regulates interleukin-18: potential role for gram-negative colonization in sarcoidosis. *American journal of respiratory and critical care medicine* **172**, 1299–1307, doi:10.1164/rccm.200411-1594OC (2005).
15. Arai, T., Inoue, Y., Eishi, Y., Yamamoto, S. & Sakatani, M. Propionibacterium acnes in granulomas of a patient with necrotising sarcoid granulomatosis. *Thorax* **63**, 90–91, doi:10.1136/thx.2006.077008 (2008).
16. Drake, W. P. *et al.* Cellular recognition of Mycobacterium tuberculosis ESAT-6 and KatG peptides in systemic sarcoidosis. *Infection and immunity* **75**, 527–530, doi:10.1128/IAI.00732-06 (2007).
17. Fazel, S. B., Howie, S. E., Krajewski, A. S. & Lamb, D. B lymphocyte accumulations in human pulmonary sarcoidosis. *Thorax* **47**, 964–967, doi:10.1136/thx.47.11.964 (1992).
18. Prokop, S., Heppner, F. L., Goebel, H. H. & Stenzel, W. M2 polarized macrophages and giant cells contribute to myofibrosis in neuromuscular sarcoidosis. *The American journal of pathology* **178**, 1279–1286, doi:10.1016/j.ajpath.2010.11.065 (2011).
19. Talreja, J., Talwar, H., Ahmad, N., Rastogi, R. & Samavati, L. Dual Inhibition of Rip2 and IRAK1/4 Regulates IL-1beta and IL-6 in Sarcoidosis Alveolar Macrophages and Peripheral Blood Mononuclear Cells. *J Immunol* **197**, 1368–1378, doi:10.4049/jimmunol.1600258 (2016).
20. Rastogi, R. *et al.* Dysregulation of p38 and MKP-1 in response to NOD1/TLR4 stimulation in sarcoid bronchoalveolar cells. *American journal of respiratory and critical care medicine* **183**, 500–510, doi:10.1164/rccm.201005-0792OC (2011).
21. Draghici, S. *et al.* A systems biology approach for pathway level analysis. *Genome Res* **17**, 1537–1545, doi:10.1101/gr.6202607 (2007).
22. Tarca, A. L., Draghici, S. & Romero, R. Developing classifiers for the detection of cancer using multi-analytes. *Methods Mol Biol* **520**, 259–272, doi:10.1007/978-1-60327-811-9_19 (2009).
23. Khatri, P. *et al.* Onto-Tools: new additions and improvements in 2006. *Nucleic Acids Res* **35**, W206–211, doi:10.1093/nar/gkm327 (2007).
24. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30, doi:10.1093/nar/28.1.27 (2000).
25. Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* **38**, D355–360, doi:10.1093/nar/gkp896 (2010).
26. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**, D109–114, doi:10.1093/nar/gkr988 (2012).
27. Kanehisa, M. *et al.* Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42**, D199–205, doi:10.1093/nar/gkt1076 (2014).
28. Holliday, L. S. Vacuolar H⁺-ATPase: An Essential Multitasking Enzyme in Physiology and Pathophysiology. *New Journal of Science* **2014**, doi:10.1155/2014/675430 (2014).
29. Agrawal, V. *et al.* Altered autophagic flux enhances inflammatory responses during inflammation-induced preterm labor. *Scientific reports* **5**, 9410, doi:10.1038/srep09410 (2015).
30. Defendi, F., Decleva, E., Martel, C., Dri, P. & Stasia, M. J. A novel point mutation in the CYBB gene promoter leading to a rare X minus chronic granulomatous disease variant—impact on the microbicidal activity of neutrophils. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **1792**, 201–210, doi:10.1016/j.bbadis.2009.01.005 (2009).
31. Marques, P. I. *et al.* SERPINA2 is a novel gene with a divergent function from SERPINA1. *PLoS one* **8**, e66889, doi:10.1371/journal.pone.0066889 (2013).
32. Muller-Quernheim, J. *et al.* Genetics of Sarcoidosis. *Clin Chest Med* **29**, 391–414, doi:10.1016/j.ccm.2008.03.007 (2008).
33. Rybicki, B. A. *et al.* The BTNL2 gene and sarcoidosis susceptibility in African Americans and Whites. *The American Journal of Human Genetics* **77**, 491–499, doi:10.1086/444435 (2005).
34. Bloom, C. I. *et al.* Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PLoS one* **8**, e70630, doi:10.1371/journal.pone.0070630 (2013).
35. Zhou, T. *et al.* Peripheral blood gene expression as a novel genomic biomarker in complicated sarcoidosis. *PLoS One* **7**, e44818, doi:10.1371/journal.pone.0044818 (2012).
36. Gharib, S. A. *et al.* Sarcoidosis activates diverse transcriptional programs in bronchoalveolar lavage cells. *Respiratory research* **17**, 93, doi:10.1186/s12931-016-0411-y (2016).
37. Romer, F. K., Christiansen, S. E., Kragballe, K., Herlin, T. & Madsen, M. Studies of peripheral blood monocytes in pulmonary sarcoidosis. *Clin Exp Immunol* **58**, 357–363 (1984).
38. Narla, A. & Ebert, B. L. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**, 3196–3205, doi:10.1182/blood-2009-10-178129 (2010).
39. Maver, A., Medica, I. & Peterlin, B. Search for sarcoidosis candidate genes by integration of data from genomic, transcriptomic and proteomic studies. *Med Sci Monit* **15**, SR22–28 (2009).
40. Su, R. *et al.* Longitudinal analysis of sarcoidosis blood transcriptomic signatures and disease outcomes. *European Respiratory Journal* **44**, 985–993, doi:10.1183/09031936.00039714 (2014).
41. Sheffield, E. A. Pathology of sarcoidosis. *Clinics in chest medicine* **18**, 741–754 (1997).
42. Baltzan, M., Mehta, S., Kirkham, T. H. & Cosio, M. G. Randomized trial of prolonged chloroquine therapy in advanced pulmonary sarcoidosis. *American journal of respiratory and critical care medicine* **160**, 192–197, doi:10.1164/ajrccm.160.1.9809024 (1999).
43. Seguin, S. *et al.* Inhibition of autophagy, lysosome and VCP function impairs stress granule assembly. *Cell Death & Differentiation* **21**, 1838–1851, doi:10.1038/cdd.2014.103 (2014).
44. Blom, T. *et al.* LAPTMB4 facilitates late endosomal ceramide export to control cell death pathways. *Nature chemical biology* **11**, 799–806, doi:10.1038/nchembio.1889 (2015).

45. Jung, T., Catalgol, B. & Grune, T. The proteasomal system. *Molecular aspects of medicine* **30**, 191–296, doi:[10.1016/j.mam.2009.04.001](https://doi.org/10.1016/j.mam.2009.04.001) (2009).
46. Tannahill, G. M. *et al.* Succinate is an inflammatory signal that induces IL-1beta through HIF-1alpha. *Nature* **496**, 238–242, doi:[10.1038/nature11986](https://doi.org/10.1038/nature11986) (2013).
47. Wang, R. *et al.* The transcription factor Myc controls metabolic reprogramming upon T lymphocyte activation. *Immunity* **35**, 871–882, doi:[10.1016/j.immuni.2011.09.021](https://doi.org/10.1016/j.immuni.2011.09.021) (2011).
48. Sinclair, L. V. *et al.* Control of amino-acid transport by antigen receptors coordinates the metabolic reprogramming essential for T cell differentiation. *Nature immunology* **14**, 500–508, doi:[10.1038/ni.2556](https://doi.org/10.1038/ni.2556) (2013).
49. Pinilla-Vera, M. *et al.* Full Spectrum of LPS Activation in Alveolar Macrophages of Healthy Volunteers by Whole Transcriptomic Profiling. *PloS one* **11**, e0159329, doi:[10.1371/journal.pone.0159329](https://doi.org/10.1371/journal.pone.0159329) (2016).
50. Gibbings, S. L. *et al.* Transcriptome analysis highlights the conserved difference between embryonic and postnatal-derived alveolar macrophages. *Blood* **126**, 1357–1366, doi:[10.1182/blood-2015-01-624809](https://doi.org/10.1182/blood-2015-01-624809) (2015).
51. Li, J. *et al.* cDNA microarray analysis reveals fundamental differences in the expression profiles of primary human monocytes, monocyte-derived macrophages, and alveolar macrophages. *J Leukoc Biol* **81**, 328–335, doi:[10.1189/jlb.0206124](https://doi.org/10.1189/jlb.0206124) (2007).
52. Geamanu, A., Gupta, S. V., Bauerfeld, C. & Samavati, L. Metabolomics connects aberrant bioenergetic, transmethylation, and gut microbiota in sarcoidosis. *Metabolomics* **12**, 1–13, doi:[10.1007/s11306-015-0932-2](https://doi.org/10.1007/s11306-015-0932-2) (2016).
53. Moyerbrailean, G. A. *et al.* A high-throughput RNA-seq approach to profile transcriptional responses. *Scientific reports* **5**; doi:[10.1038/srep14976](https://doi.org/10.1038/srep14976) (2015).
54. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21, doi:[10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635) (2013).
55. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106, doi:[10.1186/gb-2010-11-10-r106](https://doi.org/10.1186/gb-2010-11-10-r106) (2010).
56. Samavati, L. *et al.* STAT3 tyrosine phosphorylation is critical for interleukin 1 beta and interleukin-6 production in response to lipopolysaccharide and live bacteria. *Mol Immunol* **46**, 1867–1877, doi:[10.1016/j.molimm.2009.02.018](https://doi.org/10.1016/j.molimm.2009.02.018) (2009).

Acknowledgements

This work was supported by a grant (R01HL113508 (LS) and American lung Association (LS) and as well as the Department of Medicine and the Center for Molecular Medicine and Genetics, Wayne State University School of Medicine (LS).

Author Contributions

J.T. contributed to the study design, conducted the experiments, analyzed the results and in writing the manuscript. P.F. performed qRT-PCR experiments. A.A. conducted the RNA-seq library samples run on Illumina Nextseq500. F.L. and R.P. reviewed the data analysis and the manuscript. L.S. conceived and designed the study, participated in all areas of the research such as patient selection and oversaw patient enrollment, data analysis and writing of the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-02941-4](https://doi.org/10.1038/s41598-017-02941-4)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017