

Expression profiling elucidates a molecular gene signature for pulmonary hypertension in sarcoidosis

Sunit Singla,^{1,a} Tong Zhou,^{2,a} Kamran Javadi,^{1,a} Taimur Abbasi,¹ Nancy Casanova,³ Wei Zhang,⁴ Shwu-Fan Ma,⁵ Michael S. Wade,¹ Imre Noth,⁵ Nadera J. Sweiss,⁶ Joe G. N. Garcia,³ Roberto F. Machado¹

¹Division of Pulmonary, Critical Care, Sleep and Allergy, Department of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA; ²Department of Physiology and Cell Biology, University of Nevada School of Medicine, Reno, Nevada, USA; ³Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Department of Medicine, University of Arizona Health Sciences, Tucson, Arizona, USA; ⁴Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA; ⁵Section of Pulmonary/Critical Care, Department of Medicine, University of Chicago, Chicago, Illinois, USA; ⁶Section of Rheumatology, Department of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA; ^aThese authors contributed equally.

Abstract: Pulmonary hypertension (PH), when it complicates sarcoidosis, carries a poor prognosis, in part because it is difficult to detect early in patients with worsening respiratory symptoms. Pathogenesis of sarcoidosis occurs via incompletely characterized mechanisms that are distinct from the mechanisms of pulmonary vascular remodeling well known to occur in conjunction with other chronic lung diseases. To address the need for a biomarker to aid in early detection as well as the gap in knowledge regarding the mechanisms of PH in sarcoidosis, we used genome-wide peripheral blood gene expression analysis and identified an 18-gene signature capable of distinguishing sarcoidosis patients with PH ($n = 8$), sarcoidosis patients without PH ($n = 17$), and healthy controls ($n = 45$). The discriminative accuracy of this 18-gene signature was 100% in separating sarcoidosis patients with PH from those without it. If validated in a large replicate cohort, this signature could potentially be used as a diagnostic molecular biomarker for sarcoidosis-associated PH.

Keywords: peripheral blood mononuclear cells, gene signature, sarcoidosis, pulmonary hypertension.

Pulm Circ 2016;6(4):465-471. DOI: 10.1086/688316.

Pulmonary hypertension (PH) is a known complication of sarcoidosis, but its prevalence in this patient population has been variably reported and is largely unclear.¹ Historically, it has been detected more commonly in patients with advanced lung disease, with one report documenting its presence in 74% of patients with sarcoidosis awaiting lung transplantation.² Sarcoidosis-associated pulmonary hypertension (SAPH) dramatically decreases functional status and increases oxygen requirements relative to sarcoidosis patients without PH. As reported by Shorr et al.,³ approximately 70% of SAPH patients require assistance with daily functional activities, compared to a small fraction of non-PH sarcoidosis patients. Furthermore, PH is an independent predictor of mortality in patients with sarcoidosis, increasing it by 10-fold among patients listed for lung transplantation.⁴ The overall prognosis of SAPH is a dismal 59% survival at 5 years from diagnosis.⁵

Contributing significantly to both the uncharacterized prevalence of SAPH and its associated poor prognosis are the current challenges faced by clinicians to detect the condition in its early stages.⁶ The presenting symptoms of PH, such as dyspnea, reduced exercise capacity, and fatigue, are clinically indistinguishable from those of advancing pulmonary parenchymal sarcoidosis.⁷ Physical findings of PH are readily detectable only when right heart failure occurs in the late stages of disease.⁸ The diagnosis of SAPH requires invasive right heart catheterization, and there is no reliable noninvasive screening

test for SAPH.⁹ Continuous Doppler flow echocardiography estimating systolic pulmonary arterial pressure from the maximal velocity of the tricuspid regurgitation jet is a noninvasive method routinely used to assess PH in nonsarcoid patients. However, it is known to be less accurate in patients with chronic lung disease, and therefore its ability to perform as a good screening test for SAPH is questionable.¹⁰

There is currently no specific therapy for SAPH.⁶ Many of the usual secondary mechanisms of PH in chronic lung disease have been found to be of minimal significance as contributing factors in the development of SAPH. These include fibrotic destruction of the pulmonary vascular bed,^{2,7} extrinsic compression of the pulmonary vessels by mediastinal lymphadenopathy,¹¹ hypoxic vasoconstriction leading to pulmonary vascular remodeling,² and sarcoid cardiomyopathy.² Increasing evidence points to a sarcoidosis-related mechanism of pulmonary vasculopathy that is distinct from many of the other known causes of PH but is mechanistically uncharacterized.^{12,13}

To address the challenges in early clinical detection of SAPH as well as the current deficit in knowledge regarding its mechanism of pathogenesis, we used gene expression profiling of peripheral blood mononuclear cells (PBMCs) to study cohorts of sarcoidosis patients with and without PH. Whole-genome expression studies have been utilized to identify novel candidate genes for PH by applying an unbiased, genome-based approach.¹⁴ Since gene expression, an intermediate phenotype situated between DNA sequence variation and

Address correspondence to Dr. Roberto F. Machado, 909 South Wolcott Avenue, 3108 COMRB (MC719), Chicago, IL 60612, USA. E-mail: machador@uic.edu.

Submitted March 7, 2016; Accepted June 21, 2016; Electronically published October 10, 2016.

© 2016 by the Pulmonary Vascular Research Institute. All rights reserved. 2045-8932/2016/0604-0006. \$15.00.

cellular, whole-body phenotypes, has a substantial genetic basis,¹⁵⁻¹⁷ profiling gene expression in PBMCs could provide an opportunity to explore gene expression signatures or biomarkers that may be associated with PH in sarcoidosis from both a diagnostic and a pathobiologic standpoint.

METHODS

Inclusion and exclusion criteria

The study was approved by the institutional review board of each collaborating institution, with written informed consent obtained from all subjects, and was performed in accordance with the principles in the Declaration of Helsinki. The inclusion criterion was a diagnosis of sarcoidosis based on the criteria established by the joint international statement on sarcoidosis.¹⁸ Healthy controls were patients without any known chronic medical condition. Among the patients with sarcoidosis, PH was considered to be present in those who had undergone a right heart catheterization for clinical suspicion of PH and who met criteria for PH, as defined by mean pulmonary arterial pressure of ≥ 25 mmHg and a pulmonary artery occlusion pressure of ≤ 15 mmHg. Subjects with cardiac sarcoidosis or neurosarcoidosis and other concurrent systemic inflammatory diseases were excluded.

Functional parameters in test subjects

Subjects with sarcoidosis underwent spirometry in accordance with the joint American Thoracic Society/European Respiratory Society guidelines for standardization of lung function testing.¹⁹ Patients who were clinically suspected to have PH underwent right heart catheterization, as detailed in the American College of Cardiology Foundation/American Heart Association 2009 expert consensus document on PH.²⁰

PBMC preparation

Blood samples were collected from 17 sarcoidosis patients without PH, 8 sarcoidosis patients with PH, and 45 healthy controls (Table 1). PBMCs were isolated from these samples via Ficoll-hypaque density gradient centrifugation, as in Grigoryev et al.²¹

RNA isolation and microarray hybridization

Total RNA was isolated from PBMCs with Trizol LS (Life Technologies), and additional purification was achieved by using the RNeasy purification kit (Qiagen). All RNA sample ($n = 70$) quality was high, without signs of DNA contamination or RNA degradation. Further sample processing (e.g., complementary DNA generation, fragmentation, end labeling, and hybridization to Affymetrix GeneChip Human Exon 1.0 ST arrays) was performed by the Functional Genomics Facility at the University of Chicago and the Core Genomics Facility at University of Illinois at Chicago, per manufacturers' instructions.

Microarray data preprocessing

Expression arrays were analyzed with Affymetrix Power Tools, version 1.12.0 (http://www.affymetrix.com/estore/partners_programs/programs/developer/tools/powertools.affx), as in Zhou et al.²² We used the exper-

Table 1. Clinical characteristics of sarcoidosis patient population

| Characteristic | No PH ($n = 17$) | PH ($n = 8$) |
|-----------------------------------|--------------------|----------------|
| Age, years | 53 \pm 3 | 49 \pm 2 |
| Females, no. | 12 | 4 |
| FVC, % predicted | 43 \pm 9 | 74 \pm 4 |
| ETRA, no. | | 2 |
| PDE-5 inhibitor, no. | | 5 |
| Prostacyclin, no. | | 1 |
| Right atrial pressure, mmHg | | 5 \pm 1 |
| PAP, mmHg | | 37 \pm 4 |
| Cardiac output, min ⁻¹ | | 5.5 \pm 0.4 |
| 6-minute walk distance, m | | 204 \pm 31 |
| BNP, pg/mL | | 127 \pm 57 |

Note: Unless otherwise noted, data are reported as mean \pm SEM. There were no significant differences in age and sex ($P = 0.3$ by t test for age and $P = 0.4$ by χ^2 test for sex), but there were significant differences in pulmonary function ($P = 0.002$ by t test). No PH: patients without pulmonary hypertension; PH: patients with pulmonary hypertension; FVC: forced vital capacity; ETRA: endothelial receptor antagonist; PDE-5: phosphodiesterase 5; PAP: pulmonary arterial pressure; BNP: brain natriuretic protein.

imental probe-masking workflow provided by the Affymetrix Power Tools to filter the probe set (exon-level) intensity files by removing probes that contain known single-nucleotide polymorphisms (SNPs) in the dbSNP database (ver. 129).²³ Overall, of the ~ 1.4 million probe sets on the exon array, $\sim 350,000$ were found to contain at least one probe with an SNP ($\sim 600,000$ probes).²⁴ The resulting probe signal intensities were quantile normalized over all 70 samples. Probe set expression signals were summarized with the robust multiarray average (RMA) algorithm²⁵ and \log_2 transformed with a median polish. We then generated the expression signals of the $\sim 22,000$ transcript clusters (gene level) with the core set (i.e., with RefSeq-supported annotations)²⁶ of exons by taking averages of all annotated probe sets for each transcript cluster. Adjustment for a possible batch effect was conducted by COMBAT (<http://jlab.bu.edu/ComBat/>).²⁷ We consider a transcript cluster to be reliably expressed in these samples if the Affymetrix-implemented DABG (detection above ground)²⁸ P value was less than 0.01 in at least 83% of the samples in either patient group (sarcoidosis patients with or without PH). We removed genes on chromosomes X and Y to avoid the potential confounding factor of sex. We further limited our analysis set to the genes with unique annotations (i.e., transcripts corresponding to unique genes) from the Affymetrix NetAffx website (https://www.affymetrix.com/analysis/index.affx#1_2, accessed March 1, 2012). In total, 12,999 transcript clusters met these criteria and were further analyzed.

Identification of differentially expressed genes

Significance analysis of microarrays (SAM),²⁹ implemented in the *samr* library of the R statistical package,³⁰ was used to compare \log_2 -

transformed gene expression levels between sarcoidosis patients with and those without PH. The false-discovery rate (FDR) was controlled via the q -value method.³¹ Transcripts with a fold change greater than 1.25 and an FDR less than 10% were deemed differentially expressed. We searched for any enriched Kyoto Encyclopedia of Genes and Genomes (KEGG)³² physiological pathways among the differential genes relative to the final analysis set, using the National Institutes of Health Database for Annotation, Visualization and Integrated discovery (DAVID).^{33,34} A P value of <0.05 was used as the cutoff.

Identification of gene signatures for distinguishing sarcoidosis with and without PH

To identify gene signatures useful in the diagnosis of PH in sarcoidosis, a machine-learning algorithm based on a support vector machine (SVM)³⁵ using a linear kernel was applied, in combination with recursive feature elimination (RFE), for generating a predictive model. The decision function of a linear SVM is

$$f(\mathbf{x}) = \text{sign}((\mathbf{w} \cdot \mathbf{x}) + b),$$

where \mathbf{x} is the gene expression vector of a sample, \mathbf{w} is the vector of weights of the features, and b is a scalar offset.³⁶⁻³⁸ The RFE approach recursively reduces the number of genes used in the predictor function by removing those genes with lowest weights and refitting the SVM algorithm with the remaining genes. In the first step, all genes are ranked by the SVM according to their weights. In each of the following steps of the RFE procedure, half of the genes are eliminated from the predictor model, until the gene number is smaller than 20. A 5-fold cross validation (repeated 1,000 times) of the predictive models based on differently sized subsets of genes, as selected by RFE, was performed. After the recursive feature selection steps on each subset, we counted at each level the frequency of the features being selected among all rounds of cross-validation experiments. The most frequently selected features are reported as the final features (signature genes).³⁸ The *e1071* library of the R statistical package³⁰ was used for identifying the gene signature. The SVM linear classifier was trained by the differentially expressed genes between patients with and those without PH. The gene signature that consisted of the most parsimonious number of genes with significant peak prediction accuracy was used in subsequent analyses.

RESULTS

Patient characteristics

PBMC samples were collected from sarcoidosis patients with PH ($n = 8$), sarcoidosis patients without PH ($n = 17$), and healthy controls ($n = 45$). The clinical characteristics of study patients are displayed in Table 1. Significant differences in age and sex were not observed between patients with and those without PH ($P = 0.3$ by t test for age, and $P = 0.4$ by χ^2 test for sex). Not surprisingly, sarcoidosis patients with PH had worse pulmonary function ($P = 0.002$ by t test).

Identification of differentially expressed genes in SAPH

At the specified significance level (fold change > 1.25, FDR < 10%), 275 genes were differentially expressed between the sarcoidosis pa-

tients with and those without PH (Fig. S1; Figs. S1–S3 available online), among which 269 genes were upregulated while 6 genes were downregulated in PH patients. Ten KEGG pathways were significantly enriched among these sarcoidosis PH-associated genes ($P < 0.05$), including Fc gamma R-mediated phagocytosis, cytokine-cytokine receptor interaction, and MAPK signaling pathway (Fig. 1). We also compared the gene expression pattern between the sarcoidosis patients without PH and healthy controls; 263 genes were found to be differentially expressed between these two groups. The top KEGG T cell receptor signaling pathway identified (Fig. S2) was reported in a previous study.²²

Identifying a gene signature for PH in sarcoidosis

The 275 genes differentially expressed between the sarcoidosis patients with and those without PH were used as the initial analysis set for the SVM-based algorithm. Figure S3 shows an overview of the prediction accuracy for gene signatures with different numbers of genes during recursive feature selection (see “Identification of gene signatures for distinguishing sarcoidosis with and without PH” for details). We chose an 18-gene signature (Table 2) as the most parsimonious gene list with the peak prediction accuracy. Among the 18 genes, *HIST1H4C* (encoding histone cluster 1, H4c), *CACHD1* (encoding cache domain containing 1), *STOX1* (encoding storkhead box 1), and *NRCAM* (encoding neuronal cell adhesion molecule) were strongly downregulated in PH patients, whereas the remaining 14 genes were

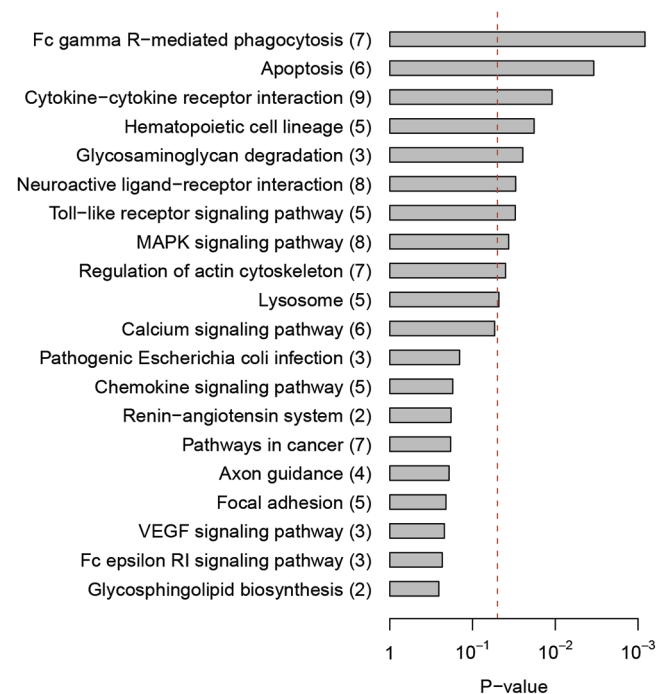


Figure 1. Enriched pathways among the differentially expressed genes between sarcoidosis patients with and those without pulmonary hypertension. The 20 top-ranked Kyoto Encyclopedia of Genes and Genomes pathways are listed. The red line indicates the cutoff of significance (0.05). MAPK: mitogen-activated protein kinase; VEGF: vascular endothelial growth factor.

Table 2. 18-gene signature

| Gene symbol | Gene name | Weight |
|---------------------|--|--------|
| <i>HIST1H4C</i> | Histone cluster 1, H4c | 0.0549 |
| <i>ETHE1</i> | Ethylmalonic encephalopathy 1 | 0.0527 |
| <i>WASF1</i> | WAS protein family, member 1 | 0.0466 |
| <i>RGS19</i> | Regulator of G-protein signaling 19 | 0.0436 |
| <i>CORO1B</i> | Coronin, actin-binding protein, 1B | 0.0364 |
| <i>CACHD1</i> | Cache domain containing 1 | 0.0362 |
| <i>SUMO4</i> | SMT3 suppressor of mif two 3 homolog 4 (<i>S. cerevisiae</i>) | 0.0352 |
| <i>LOC100128751</i> | INM04 | 0.0319 |
| <i>RASSF3</i> | Ras association (RalGDS/AF-6) domain family member 3 | 0.0314 |
| <i>STOX1</i> | Storkhead box 1 | 0.0276 |
| <i>CD177</i> | CD177 molecule | 0.0258 |
| <i>SERPINB8</i> | Serpin peptidase inhibitor, clade B (ovalbumin), member 8 | 0.0256 |
| <i>FXYP6</i> | FXYP domain containing ion transport regulator 6 | 0.0255 |
| <i>NRCAM</i> | Neuronal cell adhesion molecule | 0.0229 |
| <i>RCVRN</i> | Recoverin | 0.0224 |
| <i>BLOC1S1</i> | Biogenesis of lysosomal organelles complex-1, subunit 1 | 0.0214 |
| <i>SH3BGR13</i> | SH3 domain binding glutamic acid-rich protein like 3 | 0.0196 |
| <i>SLC22A16</i> | Solute carrier family 22 (organic cation/carnitine transporter), member 16 | 0.0193 |

upregulated in PH (Fig. 2A). This 18-gene signature distinguished sarcoidosis patients with PH from the ones without PH (Fig. 2B) with 100% discriminative accuracy. We also tested the performance of the 18-gene signature in distinguishing PH patients from all non-PH subjects (including sarcoidosis patients without PH and healthy controls). We found that the discriminative power was also considerably strong, with a classification accuracy of 97.8% (sensitivity: 85.1%, specificity: 99.5%; Fig. 2C).

We mapped the 18-gene signature to a lung gene expression data set, published by Zhao et al.,³⁹ containing 11 controls and 12 PH patients. Among the 14 genes upregulated in PBMCs of our cohort of SAPH patients, only 3 were found to be upregulated in PH lungs (*t* test: $P < 0.05$). Meanwhile, none of the four genes downregulated in PBMCs of SAPH patients were downregulated in lungs of pulmonary arterial hypertension patients. These results suggest that the 18-gene signature is SAPH and PBMC specific.

DISCUSSION

The prevalence of PH in patients with sarcoidosis is unknown, obscuring the predictive value of initial noninvasive diagnostic testing, such as echocardiography.¹ The only described pretest indication of

the potential presence of PH in this population is subjective dyspnea or functional limitation out of proportion to abnormalities on pulmonary function testing, which has been used to identify a subpopulation in which 47% of subjects are found to have PH⁴⁰ but may be difficult to reproduce, given the subjective nature of the assessment. Our study has sought to determine whether peripheral blood gene

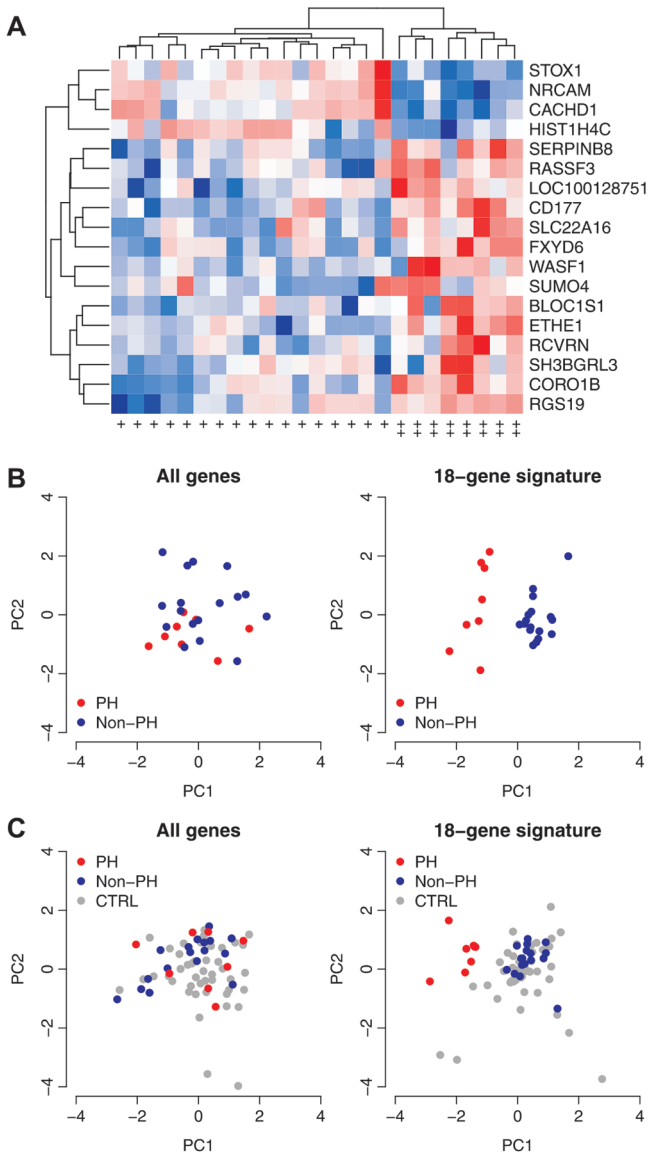


Figure 2. A, Heat map of our 18-gene signature. Red represents increased gene expression, while blue represents gene expression. ++: sarcoidosis patients with PH; +: sarcoidosis patients without PH. B, C, Principal-component analysis (PCA); X-axis: first principal component (PC1); Y-axis: second principal component (PC2); left, PCA on the expression values of all the genes detected by microarray; right, PCA on the expression values of our 18-gene signature. PH: pulmonary hypertension/sarcoidosis patients with PH; Non-PH: sarcoidosis patients without PH. B, PCA on expression of the patients with and without PH. C, PCA on expression of the patients with and without PH and healthy controls. CTRL: healthy controls.

expression profiling, an easily reproducible and objective measurement, can be used to further enrich a sarcoidosis subpopulation for PH, thereby potentially enhancing the predictive value of subsequent diagnostic testing for the disease.

Gene signatures in blood have been shown to serve as a potential diagnostic tool in sarcoidosis.⁴¹ Leveraging whole-genome expression profiles in a cohort of sarcoidosis patients, we identified a gene signature consisting of 18 autosomal genes for differentiating patients with PH from patients without PH. The RFE approach recursively reduces the number of genes that can be used to classify PH and non-PH patients by removing those genes with lowest weights and refitting the SVM algorithm with the remaining genes. The 18-gene signature showed prediction accuracy equivalent to that of other signatures containing a greater number of genes (e.g., 35, 69, 138, or 275), and these larger signatures generally exhibited the same level of classification power as our 18-gene signature. The 18-gene sarcoidosis signature not only correctly distinguished sarcoidosis patients with PH from patients without PH but also successfully differentiated patients with PH from healthy controls. Therefore, the 18-gene signature not only captured differences between sarcoidosis patients with and those without PH but also conveyed information as to differences between PH patients and healthy controls.

We compared our 18-gene signature with other known PBMC gene signatures in PH. None of the genes in our signature overlap with those of a 106-gene PBMC signature separating PH patients from healthy controls described by Bull et al.⁴² Furthermore, Risbano et al.⁴³ elucidated a 9-gene signature that defined scleroderma patients who had PH versus those without PH, and in a recent publication by Desai and colleagues,¹⁴ a 10-gene signature was proposed to identify patients with PH from a cohort of sickle cell disease patients. However, we did not find any overlap between our 18-gene signature gene and either of these etiology-specific PH signatures. Altogether, these comparisons suggest that the mechanisms for PH in sarcoidosis consist of genetic pathways different from those for idiopathic or other secondary etiologies.

While it remains unclear how representative PBMC gene expression changes are of cellular events occurring in the lung, it is very interesting to note that 9 of the 18 genes identified in our study are linked to cellular proliferation,^{44–52} which is one of the central events occurring in pulmonary endothelial cells and/or vascular smooth muscle cells during PH.⁵³ A handful of these are linked to specific pathways or cell types relevant to PH. *ETHE1* and *RASSF3* are mediators of p53-induced apoptosis,^{46,51} suppression of which in pulmonary vascular smooth muscle cells is sufficient to induce PH in rats.⁵⁴ *CORO1B* plays a role in platelet-derived growth factor-mediated smooth muscle cell migration and proliferation,⁴⁹ a pathway that has been studied for its potential as a target for therapy in PH.⁵⁵ *HIST1H4C* has been linked to proliferative pathways involving bone morphogenetic protein (BMP) and activin receptor-like kinase 1 (ALK-1).⁵⁶ BMP and ALK-1 are both known to carry mutations associated with heritable PH syndromes.⁵⁷

Major limitations to our work is the lack of reverse transcription polymerase chain reaction (RT-PCR) validation of the microarray findings and the lack of prospective confirmation in an independent validation cohort. Unfortunately, the amount of sample available

from the original cohort was insufficient to perform RT-PCR validation, and the sample size of SAPH patients is too small to perform an independent validation cohort study at this time. We also recognize that PBMC gene expression changes could be affected by many conditions and factors, not the least of which is the presence of other disease conditions, such as infections. In this small study, we have not collected data regarding indolent coexisting conditions that could play such a confounding role.

In summary, despite several significant limitations, including the relatively small size of our study cohort and the lack of an independent validation cohort, we have identified a molecular gene signature as a potential novel molecular biomarker in the diagnosis of PH in sarcoidosis. This PH gene signature performed with substantial classification accuracy not only between sarcoidosis patients with PH and those without PH but also between patients with PH and healthy controls, and, if validated in a large replicate cohort, may represent a gene signature for PH in sarcoidosis that could be explored as a diagnostic molecular biomarker for the disease.

Source of Support: This work was supported by National Institutes of Health (NIH) grants K23HL098454, 1R01HL111656, and R01 HL127342 to RFM and NIH grants U01 HL112696 and U01HL125208 to JGNG.

Conflict of Interest: None declared.

REFERENCES

1. Cordova FC, D'Alonzo G. Sarcoidosis-associated pulmonary hypertension. *Curr Opin Pulm Med* 2013;19(5):531–537.
2. Shorr AF, Helman DL, Davies DB, Nathan SD. Pulmonary hypertension in advanced sarcoidosis: epidemiology and clinical characteristics. *Eur Respir J* 2005;25(5):783–788.
3. Shorr AF, Davies DB, Nathan SD. Predicting mortality in patients with sarcoidosis awaiting lung transplantation. *Chest* 2003;124(3):922–928.
4. Arcasoy SM, Christie JD, Pochettino A, Rosengard BR, Blumenthal NP, Bavaria JE, Kotloff RM. Characteristics and outcomes of patients with sarcoidosis listed for lung transplantation. *Chest* 2001;120(3):873–880.
5. Nunes H, Humbert M, Capron F, Brauner M, Sitbon O, Battesti J-P, Simonneau G, Valeyre D. Pulmonary hypertension associated with sarcoidosis: mechanisms, haemodynamics and prognosis. *Thorax* 2006;61(1):68–74.
6. Shino MY, Lynch JP III, Fishbein MC, McGraw C, Oyama J, Belperio JA, Saggat R. Sarcoidosis-associated pulmonary hypertension and lung transplantation for sarcoidosis. *Semin Respir Crit Care Med* 2014;35(3):362–371.
7. Sulica R, Teirstein AS, Kakarla S, Nemani N, Behnegar A, Padilla ML. Distinctive clinical, radiographic, and functional characteristics of patients with sarcoidosis-related pulmonary hypertension. *Chest* 2005;128(3):1483–1489.
8. Baughman RP, Sparkman BK, Lower EE. Six-minute walk test and health status assessment in sarcoidosis. *Chest* 2007;132(1):207–213.
9. Galiè N, Humbert M, Vachiéry JL, Gibbs S, Lang I, Torbicki A, Simonneau G, et al. 2015 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension: the Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). *Eur Respir J* 2015;46(4):903–975.
10. Arcasoy SM, Christie JD, Ferrari VA, Sutton MS, Zisman DA, Blumenthal NP, Pochettino A, Kotloff RM. Echocardiographic assessment

- of pulmonary hypertension in patients with advanced lung disease. *Am J Respir Crit Care Med* 2003;167(5):735–740.
11. Handa T, Nagai S, Miki S, Fushimi Y, Ohta K, Mishima M, Izumi T. Incidence of pulmonary hypertension and its clinical relevance in patients with sarcoidosis. *Chest* 2006;129(5):1246–1252.
 12. Rosen Y, Moon S, Huang CT, Gourin A, Lyons HA. Granulomatous pulmonary angiitis in sarcoidosis. *Arch Pathol Lab Med* 1977;101(4):170–174.
 13. Takemura T, Matsui Y, Saiki S, Mikami R. Pulmonary vascular involvement in sarcoidosis: a report of 40 autopsy cases. *Hum Pathol* 1992;23(11):1216–1223.
 14. Desai AA, Zhou T, Ahmad H, Zhang W, Mu W, Trevino S, Wade MS, et al. A novel molecular signature for elevated tricuspid regurgitation velocity in sickle cell disease. *Am J Respir Crit Care Med* 2012;186(4):359–368.
 15. Cheung VG, Jen KY, Weber T, Morley M, Devlin JL, Ewens KG, Spielman RS. Genetics of quantitative variation in human gene expression. *Cold Spring Harb Symp Quant Biol* 2003;68:403–407.
 16. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. *Nature* 2005;437(7063):1365–1369.
 17. Duan S, Huang RS, Zhang W, Bleibel WK, Roe CA, Clark TA, Chen TX, et al. Genetic architecture of transcript-level variation in humans. *Am J Hum Genet* 2008;82(5):1101–1113.
 18. Hunninghake GW, Costabel U, Ando M, Baughman R, Cordier JF, du Bois R, Eklund A, et al. Statement on sarcoidosis: the joint statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) was adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. *Am J Respir Crit Care Med* 1999;160(2):736–755.
 19. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319–338.
 20. McLaughlin VV, Archer SL, Badesch DB, Barst RJ, Farber HW, Lindner JR, Mathier MA, et al. ACCF/AHA 2009 expert consensus document on pulmonary hypertension: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents and the American Heart Association: developed in collaboration with the American College of Chest Physicians; American Thoracic Society, Inc.; and the Pulmonary Hypertension Association. *J Am Coll Cardiol* 2009;53(17):1573–1619.
 21. Grigoryev DN, Mathai SC, Fisher MR, Girgis RE, Zaiman AL, Houstens-Harris T, Cheadle C, et al. Identification of candidate genes in scleroderma-related pulmonary arterial hypertension. *Transl Res* 2008;151(4):197–207.
 22. Zhou T, Zhang W, Sweiss NJ, Chen ES, Moller DR, Knox KS, Ma SF, et al. Peripheral blood gene expression as a novel genomic biomarker in complicated sarcoidosis. *PLoS ONE* 2012;7(9):e44818. doi:10.1371/journal.pone.0044818.
 23. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29(1):308–311.
 24. Duan S, Zhang W, Bleibel WK, Cox NJ, Dolan ME. SNPProbe_1.0: a database for filtering out probes in the Affymetrix GeneChip[®] Human Exon 1.0 ST array potentially affected by SNPs. *Bioinformatics* 2008;24(10):469–470.
 25. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4(2):249–264.
 26. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 2007;35(suppl 1):D61–D65.
 27. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8(1):118–127.
 28. Affymetrix. Exon array background correction. Revision version 1.0. Published September 27, 2005. http://media.affymetrix.com/support/technical/whitepapers/exon_background_correction_whitepaper.pdf.
 29. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001;98(9):5116–5121.
 30. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2005.
 31. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* 2002;99(10):6567–6572.
 32. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res* 2004;32(suppl 1):D277–D280.
 33. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4(1):44–57.
 34. Dennis G Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003;4(5):P3. doi:10.1186/gb-2003-4-5-p3.
 35. Vapnik V. Statistical learning theory. New York: Wiley, 1998.
 36. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn* 2002;46(1–3):389–422.
 37. Thuerigen O, Schneeweiss A, Toedt G, Warnat P, Hahn M, Kramer H, Brors B, et al. Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer. *J Clin Oncol* 2006;24(12):1839–1845.
 38. Zhang X, Lu X, Shi Q, Xu X, Leung HE, Harris LN, Iglehart JD, Miron A, Liu JS, Wong WH. Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data. *BMC Bioinformatics* 2006;7:197. doi:10.1186/1471-2105-7-197.
 39. Zhao Y, Peng J, Lu C, Hsin M, Mura M, Wu L, Chu L, et al. Metabolomic heterogeneity of pulmonary arterial hypertension. *PLoS ONE* 2014;9(2):e88727. doi:10.1371/journal.pone.0088727.
 40. Baughman RP, Engel PJ, Meyer CA, Barrett AB, Lower EE. Pulmonary hypertension in sarcoidosis. *Sarcoidosis Vasculitis Diffus Lung Dis* 2006;23:108–116.
 41. Casanova N, Zhou T, Knox KS, Garcia JG. Identifying novel biomarkers in sarcoidosis using genome-based approaches. *Clin Chest Med* 2015;36(4):621–630.
 42. Bull TM, Coldren CD, Moore M, Sotto-Santiago SM, Pham DV, Nana-Sinkam SP, Voelkel NF, Geraci MW. Gene microarray analysis of peripheral blood cells in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2004;170(8):911–919.
 43. Risbano MG, Meadows CA, Coldren CD, Jenkins TJ, Edwards MG, Collier D, Huber W, et al. Altered immune phenotype in peripheral blood cells of patients with scleroderma-associated pulmonary hypertension. *Clin Transl Sci* 2010;3(5):210–218.
 44. Gao Q, Chen X, Duan H, Wang Z, Feng J, Yang D, Song L, Zhou N, Yan X. FXYD6: a novel therapeutic target toward hepatocellular carcinoma. *Protein Cell* 2014;5(7):532–543.
 45. Dickinson LA, Burnett R, Melander C, Edelson BS, Arora PS, Dervan PB, Gottesfeld JM. Arresting cancer proliferation by small-molecule gene regulation. *Chem Biol* 2004;11(11):1583–1594.
 46. Higashitsuji H, Higashitsuji H, Masuda T, Liu Y, Itoh K, Fujita J. Enhanced deacetylation of p53 by the anti-apoptotic protein HSCO in association with histone deacetylase 1. *J Biol Chem* 2007;282(18):13716–13725.
 47. Zhang J, Zhou S, Tang L, Shen L, Xiao L, Duan Z, Jia L, Cao Y, Mu X. WAVE1 gene silencing via RNA interference reduces ovarian cancer cell invasion, migration and proliferation. *Gynecol Oncol* 2013;130(2):354–361.

48. Tso PH, Yung LY, Wang Y, Wong YH. RGS19 stimulates cell proliferation by deregulating cell cycle control and enhancing Akt signaling. *Cancer Lett* 2011;309(2):199–208.
49. Williams HC, San Martín A, Adamo CM, Seidel-Rogol B, Pounkova L, Datla SR, Lassègue B, Bear JE, Griendling K. Role of coronin 1B in PDGF-induced migration of vascular smooth muscle cells. *Circ Res* 2012;111(1):56–65.
50. He X, Riceberg J, Pulukuri SM, Grossman S, Shinde V, Shah P, Brownell JE, Dick L, Newcomb J, Bence N. Characterization of the loss of SUMO pathway function on cancer cells and tumor proliferation. *PLoS ONE* 2015;10(4):e0123882. doi:10.1371/journal.pone.0123882.
51. Kudo T, Ikeda M, Nishikawa M, Yang Z, Ohno K, Nakagawa K, Hata Y. The RASSF3 candidate tumor suppressor induces apoptosis and G1-S cell-cycle arrest via p53. *Cancer Res* 2012;72(11):2901–2911.
52. Nie X, Zhang K, Wang L, Ou G, Zhu H, Gao WQ. Transcription factor STOX1 regulates proliferation of inner ear epithelial cells via the AKT pathway. *Cell Prolif* 2015;48(2):209–220.
53. Pietra GG, Capron F, Stewart S, Leone O, Humbert M, Robbins IM, Reid LM, Tudor RM. Pathologic assessment of vasculopathies in pulmonary hypertension. *J Am Coll Cardiol* 2004;43(12 suppl):S25–S32.
54. Jacquin S, Rinceval V, Mignotte B, Richard S, Humbert M, Mercier O, Londoño-Vallejo A, Fadel E, Eddahibi S. Inactivation of p53 is sufficient to induce development of pulmonary hypertension in rats. *PLoS ONE* 2015;10(6):e0131940. doi:10.1371/journal.pone.0131940.
55. Schermuly RT, Dony E, Ghofrani HA, Pullamsetti S, Savai R, Roth M, Sydykov A, et al. Reversal of experimental pulmonary hypertension by PDGF inhibition. *J Clin Invest* 2005;115(10):2811–2821.
56. Yoshimatsu Y, Lee YG, Akatsu Y, Taguchi L, Suzuki HI, Cunha SI, Maruyama K, et al. Bone morphogenetic protein-9 inhibits lymphatic vessel formation via activin receptor-like kinase 1 during development and cancer progression. *Proc Natl Acad Sci USA* 2013;110(47):18940–18945.
57. Soubrier F, Chung WK, Machado R, Grünig E, Aldred M, Geraci M, Loyd JE, et al. Genetics and genomics of pulmonary arterial hypertension. *J Am Coll Cardiol* 2013;62(25 suppl):D13–D21.