SARCOIDOSIS VASCULITIS AND DIFFUSE LUNG DISEASES 2016; 33; 29-38

Plasma metabolomic profile in fibrosing pulmonary sarcoidosis

Mehdi Mirsaeidi[1], Mohammad Mehdi Banoei[2, 3], Caleb K Nienow[4], Taimur Abassi[1], Anoushirvan Hakim[1], Dean Schraufnagel[1], Brent W. Winston[3], Nadera Sweiss[5], Robert Baughman[6], Joe G.N. Garcia[7], Roberto Machado[1]

[1] Division of Pulmonary, Critical Care, Sleep and Allergy, Department of Medicine, University of Illinois at Chicago, IL, USA, [2] Department of Medical Sciences, University of Calgary, Calgary, Canada, [3] Departments of Critical Care, Medicine and Biochemistry and Molecular Biology, University of Calgary, Calgary, Canada, [4] Mass spectrometry, Metabolomics & Proteomics Facility, University of Illinois at Chicago, IL, USA, [5] Division of Rheumatology, Department of Medicine, University of Illinois at Chicago, IL, USA, [6] Department of Medicine, University of Cincinnati, OH, USA, [7] Department of Medicine, University of Arizona, Tucson, Arizona, USA

Abstract. *Background:* There is no known marker to screen patients with sarcoidosis to determine the risk of progression to pulmonary fibrosis. We aimed to identify potential noninvasive biomarkers for early detection of pulmonary fibrosing sarcoidosis. *Methods:* A case-control study was performed on African Americans with confirmed sarcoidosis included 31 subjects with pulmonary fibrosis vs. 36 without pulmonary fibrosis. Plasma samples were analyzed by liquid chromatography-mass spectrum. Multivariate statistical analysis models were developed in a training set based on 50 age- and sex-matched samples to identify metabolites involved in the discrimination. Principal component analysis and orthogonal partial least squares-discriminant (OPLS) analysis coupled to the most influential variables were used to derive significant metabolic discriminations. *Results:* Of the datasets from 171 feature peaks, 14 features including p-coumaroylagmatine and palmitoylcarnitine showed significant differences between fibrosing and non-fibrosing pulmonary sarcoidosis (p = 0.001). OPLS analysis presented clear separation between two groups with an acceptable goodness of fit (R^2 = 0.522) and predictive power (Q^2=0.322). Discriminating metabolites involved collagen pathway metabolites especially those in the arginine-proline pathway. *Conclusions:* Metabolomics can provide a useful tool to detect pulmonary fibrosis in patients with sarcoidosis. Two discriminating metabolites, p-coumaroylagmatine and palmitoylcarnitine may be potential markers for fibrosing pulmonary sarcoidosis. *(Sarcoidosis Vasc Diffuse Lung Dis 2016; 33: 29-38)*

Key words: sarcoidosis, fibrosis, metabolomics

Introduction

Sarcoidosis is a granulomatous, lymphatic disease of an unknown etiology and exceedingly variable outcome. Its burden is particularly high in African-Americans with a incidence of 39 in 100,000 in the US (1, 2). Its ongoing activity and progression are uncertain and difficult to identify. Sarcoidosis most often involves the lung and disability and death are strongly associated with pulmonary fibrosis (3-5).

Received: 26 March 2015
Accepted after revision: 5 June 2015
Corresponce: Mehdi Mirsaeidi MD, MPH
Section of Pulmonary, Critical Care,
Sleep and Allergy,
Department of Medicine M/C 719,
University of Illinois at Chicago,
840 S. Wood St., Chicago, IL 60612-7323, USA
Tel: +1 240 383 7539
Email: mmirsae@uic.edu, golmeh@yahoo.com

Funding sources: This project is supported by NIH grant 5 T32 HL 82547-7 and a grant from Bernie Mac Sarcoidosis Foundation

The extent of pulmonary fibrosis also varies greatly but may occur to a variable degree in up to 30% of patients (6). Treatment for sarcoidosis is unclear because of the uncertainty of disease progression. Unfortunately, there is no known marker to screen patients with sarcoidosis to determine the risk of progression to pulmonary fibrosis; nor there is a marker of activity of the fibrosing process. Both of which would be extremely useful in caring for these patients.

Metabolomics is an approach to identify and measure small molecules (metabolites) in human fluids. It can provide valuable information about the influence of genetic, epigenetic, and environmental factors on systems biology (7, 8). It is aiding our understanding of the function and malfunction of gene products and the diagnosis and prognosis of diseases. It also can be useful for nutritional and environmental monitoring (9, 10).

Metabolomics can identify biomarkers for the diagnosis and prognosis of conditions such as cardiovascular disorders (11), cancers (12, 13), asthma (14), chronic obstructive pulmonary disease (15), tuberculosis (16), sepsis (17, 18) and HIV (19, 20). Methods that have been used to identify these metabolites including proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy, gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS) (21).

Our study aimed to identify metabolites in patients with sarcoidosis with pulmonary fibrosis that are not present in patients with sarcoidosis without fibrosis in order to try to identify biomarkers of this process. The eventual role of this would be to predict progression to fibrosis and thus poor outcomes and activity of the process to determine a need for intervention.

Methods

This study was approved by an Institutional Review Board of University of Illinois at Chicago (approval number of 20130195001).

Study Design and Patient Data

This was a case-control study on African American subjects with confirmed sarcoidosis. Confirmed sarcoidosis was defined as presence of compatible clinical symptoms and treating physician's diagnosis, and the presence or history of bilateral hilar lymphadenopathy on chest imaging, and characteristic non-necrotizing granulomata in the lung parenchyma or lymph nodes, and the exclusion of all other potential causes (22). Thirty-one subjects with pulmonary fibrosis defined with Scadding score 4 on chest radiograph and evidence of extensive pulmonary fibrosis determined by chest CT scan to occupy more than 25% of the lung field (23) versus 36 subjects without fibrosis (Scadding stage 0 or 1 and no fibrosis in the chest CT scan) were included in the study after giving their written informed consent. Obstructive defect in pulmonary function tests (PFT) was defined per American Thoracic Society and European Respiratory Society (ATS/ERS) guidelines with an FEV1/FVC ratio less than lower limit of normal (24). Obstructive defect was found in 10 out of 31 sarcoidosis subjects with pulmonary fibrosis.

The composite physiological index (CPI) which is a validated weighted index of PFT variables for mortality prediction in sarcoidosis was calculated per Wells and colleagues equation (CPI score = 91.0 − (0.65 x percent predicted DLCO) − (0.53 x percent predicted FVC) + (0.34 x percent predicted FEV1) (23, 25).

Plasma samples were obtained after an overnight fasting from both groups and processed for metabolomics study. Blood samples were taken in the EDTA tubes. The samples centrifuged and plasma aliquoted, frozen at -80' C until use and thawed once. All samples were collected at one center and processed with a single standard system operation.

LC-MS analysis and metabolite concentration profiling

LC–MS analysis was performed using a Shimadzu Ion-Trap (IT)- time of flight (TOF) LCMS coupled with Prominence ultra fast liquid chromatograph (UFLC) pumps (Shimadzu Corporation, Japan). The method has been discussed previously (26). Samples were prepared by protein precipitation with 4X excess ice-cold 8:1:1 methanol, chloroform, water (v:v:v). They were then centrifuged; supernatants were evaporated and reconstituted in 100 μL acetonitrile. Chromatographic separation was performed using a Waters Acquity UPLC BEH amide

column (1.7 μm, 2.1 x 100 mm). The scan range was set to 50-1000 m/z using polarity switching to generate data in a positive and negative mode.

To obtain metabolite concentrations, XCMS online software was used to read and process LC-MS data stored in the mzXML files. Using this software, LC-MS spectra were processed and analyzed for compound detection, providing the relative concentration, retention time, and mass value, as well as univariate statistical analysis (27). XC-MS is an online free software with an open-source license at http://xcmsonline.scripps.edu/.

Metabolites are recognized based by their retention time and unique m/z ratio that exist in the METLIN metabolites database (28). METLIN is a web-based platform used to analyze nontargeted metabolomic data at http://metlin.scripps.edu/.

Pathway analysis

Pathway analysis was performed by using Reactome, an open source peer reviewed pathway database at http://www.reactome.org and http://www.genome.jp/kegg-bin.

Data analysis

For the metabolomic data, the identification of compounds and their intensities and the p values for the differences in the intensities of analyzed ions between case and control groups were generated using the XCMS Online system (Scripps Research Institute, USA).

Multivariate data analysis (MVDA) was performed in a training set based on 50 age- and sex-matched samples to create prediction models to identify metabolites involved in the discrimination. Unsupervised and supervised multivariate statistical analyses, such as principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA), coupled the most influential variables to derive significant metabolic discriminations. PCA analysis as a basis of MVDA was performed to show the intrinsic relationship, differences, and variation between the two cohorts of sarcoidosis patients for LC-MS features. A five-component PCA model was built to summarize the information derived from LC-MS data set from the plasma samples.

OPLS-DA as a supervised method was used to show the separation based on maximizing the covariance between the two cohorts to identify the variables involved in discrimination and measures the predictability of the separation. To verify the OPLS-DA models, three performance parameters CV-ANOVA (p-value), R^2Y and Q^2Y were obtained by multivariate data analysis (29). They are necessary to indicate the reliability and significance of model; R^2Y and Q^2Y are used to evaluate goodness of fit and goodness of prediction, respectively. Multivariate data analysis was performed using SIMCA-P software (Version 13.0, Umetrics AB, Umeå, Sweden). The differences were considered significant when $P< 0.05$.

The prediction set modeling

To obtain sensitivity, specificity and the area under the receiver-operating curve characteristic (AUC), we performed prediction tests. A misclassification table was develpoed for all discriminant analysis models to show the proportion of correctly classified observations in each model. The prediction sets were created from current active models (work sets) by randomly taking 10 samples out from the OPLS-DA model three times and averaged sensitivity, specificity and AUC.

Results

Subject demographic and clinical characteristics

The demographic and clinical characteristics of the enrolled subjects are shown in table 1. Out of 67 subjects, 24 (77%) in the case and 29 (81%) in the control group were females. The mean (SD) ages of patients in the case and control groups were 53.7 (10.9) and 48.7 (10.2) years, respectively (P=0.06). The duration of disease was longer in the subjects with pulmonary fibrosis than the control group, 13 (7.3) vs. 9 (6.7) years respectively, (P=0.024). There was a significant statistical difference in the mean (SD) of CPI score between subjects with and without pulmonary fibrosis 35 (17) vs. 20 (13) in case and control groups respectively (p=0.001).

Table 1. The demographic and clinical characteristics of the enrolled subjects

	Case (N=31)	Control (N=36)	P-value
Female N(%)	24(77)	29(81)	0.753
Age M(SD)(year)	53.7(10.9)	48.7(10.2)	0.06
Asthma	10(32)	9(25)	0.512
BMI M(SD)	32(12)	35(7)	0.229
Duration of disease M(SD)	13(7.3)	9(6.7)	0.024
History of smoking	13(42)	10(27)	0.227
Current smoker	2(6)	3(8)	0.771
HTN	16(52)	17(47)	0.72
Hyperlipidemia	12(39)	5(14)	0.024
Congestive heart failure	3(10)	2(6)	0.527
Pulmonary hypertension	6(19)	1(3)	0.056
Diabetes mellitus	8(26)	11(31)	0.667
Cirrhosis	2(6)	1(3)	0.481
Chronic kidney disease	3(10)	1(3)	0.264
Neurologic disorders	4(13)	5(14)	0.906
GERD	7(23)	11(31)	0.464
Cough	18(58)	11(31)	0.025
Sputum production	7(23)	3(8)	0.115
Dyspnea	25(81)	9(25)	<0.0001
Weight loss	3(10)	1(3)	0.264
Fatigue	10(32)	9(25)	0.512
Eye involvement	9(29)	7(19)	0.361
CPI score mean±SD	35±17	20±13	0.001

Comparing comorbid conditions in two groups only revealed higher rate of hyperlipidemia in the pulmonary fibrosis group (12 (39%) vs. 5 (14%) respectively, P=0.024)

Univariate analysis of LC-MS features

171 features were detected in plasma samples and 10 features had significant difference between two groups. Subsequent, metabolomic analysis showed a significantly higher levels of flabellidine, carnitine, and N-stearoyl serine and lower phenylalanine in subjects with pulmonary fibrosis. Table 2 shows the list of metabolites that were identified in the univariate analysis.

Analysis of metabolic pattern

Based on 171 identified features, a five-component PCA model was obtained from the LC-MS data sets derived from the plasma samples to evaluate the interrelation and grouping between metabolomics data of sarcoidosis subjects with and without fibrosis, including clustering and determining instrumental and biological outliers among the samples. Therefore, the final PCA model was created and interpreted after removing extreme outliers. The PCA score plot of the whole LC-MS data set based on the first, second, and third principal components in 3 dimensional (3D) fashion demonstrated that the fibrosing sarcoidosis group could be distinguished from the non-fibrosing sarcoidosis group (Figure 1). The R^2X (cumulative) value was 0.32 for all first, second

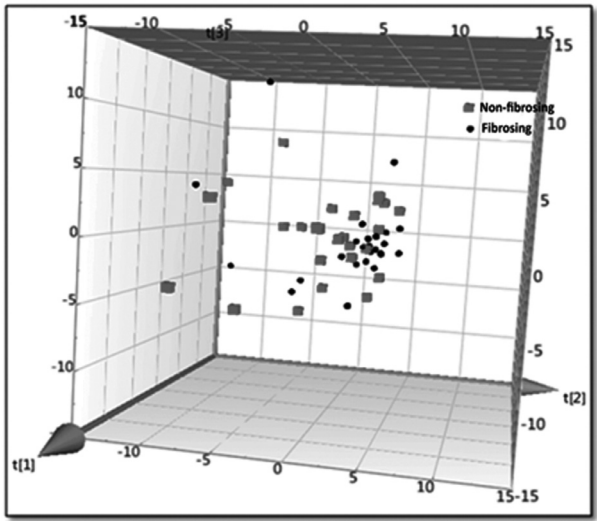

Fig. 1. The unsupervised PCA analysis shows the general separation (clustering) and finds the outliers. T1(X axis) is first principal component that has the greatest variation between the samples and t2 (Y axis) is the second principal component has the second greatest variation, and t3 (Z axis) as the third greatest variation between the samples

Table 2. List of metabolites identified in the metabolomic analysis (online supplemental)

Metabolites	Retention time (min)	m/z	Mass accuracy, ppm	Fold change	Up/down	p-value
Flabellidine	12.06	311.2058	11	1.7	Up	0.00308
Elaidic carnitine	11.32	426.3562	3	1.4	Up	0.00597
Phenylalanine	16.4	166.0865	1	1.7	Down	0.01194
1-Methylestra-1,3,5(10)-triene-3,17beta-diol	12.22	287.2056	17	1.5	Up	0.01201
Nadolol	12.06	310.2024	3	1.4	Up	0.02339
N-stearoyl serine	11.51	394.2932	1	1.7	Up	0.04012
Palmitoyl-L-carnitine	11.4	400.3413	2	1.3	Up	0.04178

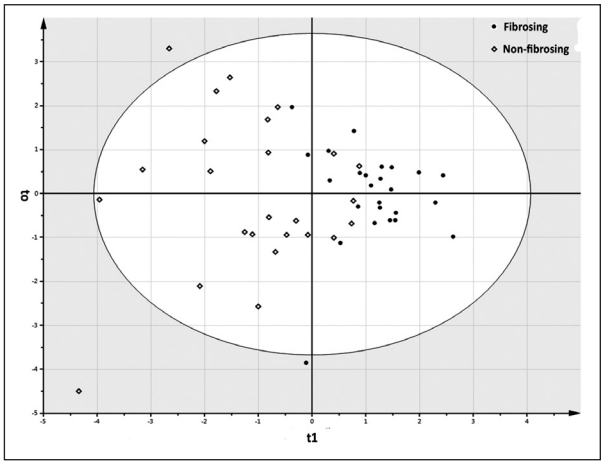

Fig. 2. The supervised OPLS-DA analysis shows the best possible prediction and interpretation of discrimination between case (fibrosing sarcoidosis) and control (non fibrosing sarcoidosis). X axis is as prediction component that shows differences between groups, and Y axis stands for an orthogonal component that shows differences within group. R^2= 0.522, Q^2= 0.322, p= 0.002. R^2 is as goodness of fit, Q^2 is as goodness of prediction and p shows the significance level of the model.

and third components that shows a high variation of X variables (features) among samples (Table 3).

A supervised OPLS-DA analysis was performed to show predictive classifications between two cohorts. By the non-targeted LC-MS metabolomics datasets from 171 feature peaks, 14 features showed significant differences between the two groups in MVDA identified by variable importance in projection (VIP) that computes the influence of X variables (metabolites-features) on observations. The results indicated that a statistically significant separation in metabolic profiles exists between delineated fibrosing from non-fibrosing sarcoidosis (p = 0.001) (Figure 2). OPLS analysis separated the two groups with relative good predictive power (Q^2= 0.322).

As illustrated in coefficient plot (Figure 3), 10 metabolites and 4 features (unidentified) were found as the most meaningful for separation between sarcoidosis with and without pulmonary fibrosis. The model demonstrated higher levels of p-coumaroylagmatine in pulmonary fibrosis subjects as well as carnitine and vitamin D metabolites. The discriminative metabolites show collagen related metabolites, primarily those in the arginine-proline pathway. (On line Supplement figure 1)

The coefficient plot was used to show relative correlation of the most important metabolites with each group that contributed in the separation between two cohorts. Figure 3 depicts 14 features that were increased (n=13) and decreased (n=1) in the fibrosing sarcoidosis in comparison to the non-fibrosing sarcoidosis.

Figure 4 is the observed vs. predicted plot that is created based on the OPLS-DA analysis to verify the discrimination of the two cohorts with showing the

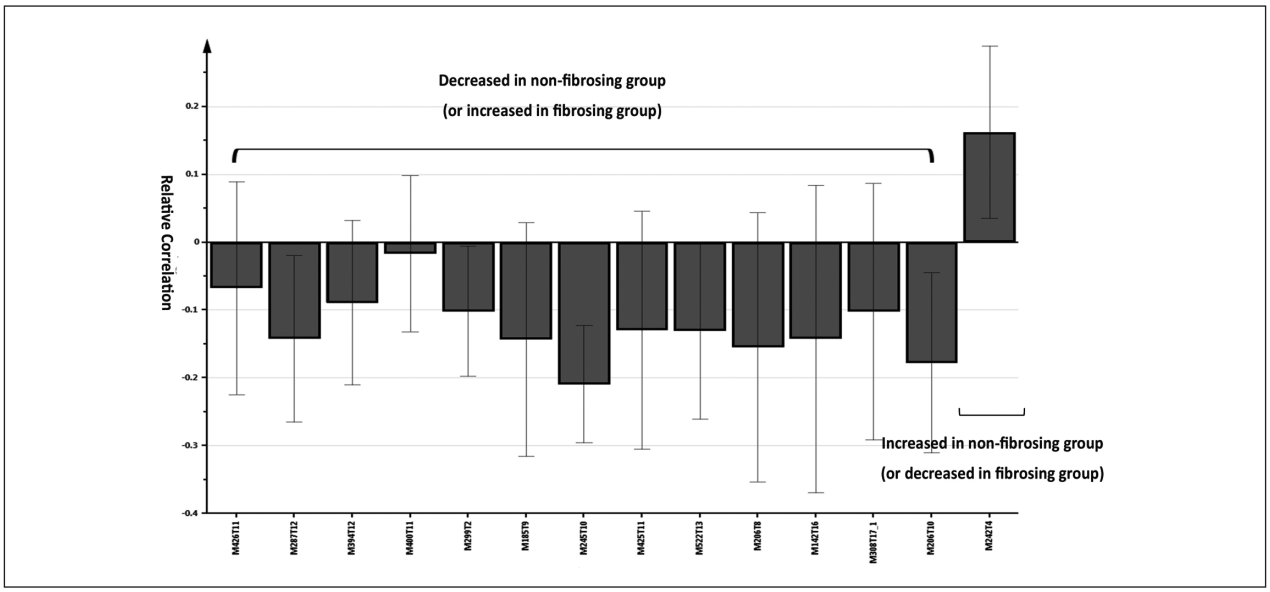

Fig. 3. Coefficient plot refers to scaled and centered metabolites' relative correlation data, with confidence interval derived from jack-knifing. (online supplemental)

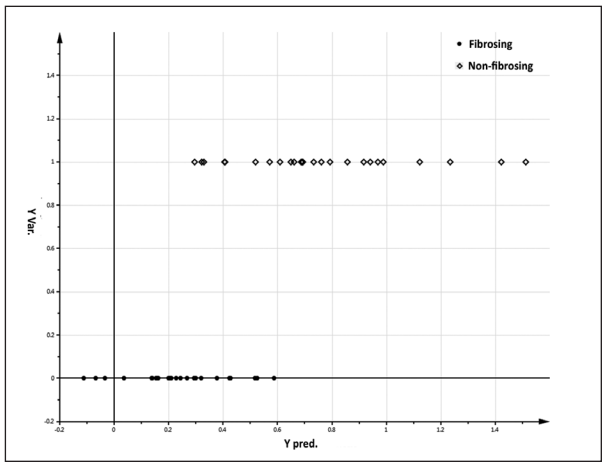

Fig. 4. Observed vs. predicted plot displays the observed values vs. the fitted or predicted values based on the best OPLS-DA model, after the last component. X axis is Y predicted value for each observation and Y axis is for the variable as a value of the observed (original) response. (online supplemental)

predictive ability more transparently. Figure 4 shows that most of the observations in the two cohorts of patients, stay to the right of 0.5 value as threshold to show the separation. Except for some samples that show overlapping, the separation overall is relative good.

Metabolite analysis for subjects with an obstructive defect on PFTs

OPLS-DA analysis produced a predictive model with a good separation (R^2= 0.553, Q^2= 0.422, p= 0.002) to recognize the patients with an obstructive defect on pulmonary function tests (n=11) from those without an obstructive defect (n=14) in the fibrosing sarcoidosis cohort (Figure 5). The observed total variation in the metabolomic profiling between two cohorts amounted to 95% Hotelling T2 confidence interval indicated marked differences between the two groups.

Metabolite analysis for subjects with an CPI score >40

OPLS-DA analysis showed relatively good discrimination (R^2Y=0.476 Q^2Y= 0.381) with statistically significant level (p= 0.0001) between those individuals with CPI score > 40 (n=10) versus those had CPI < 40 (n=31) (Figure 6). The metabolomics profile has a sensitivity in differentiating subjects

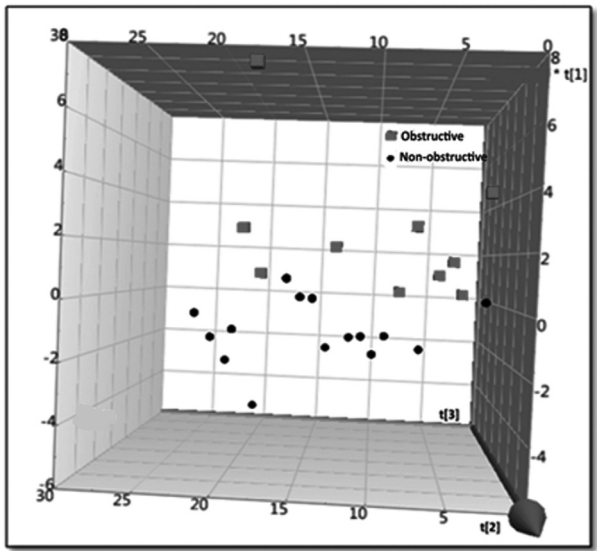

Fig. 5. The supervised OPLS-DA analysis shows a very good discrimination between patients with obstructive defects on PFTs and those with non-obstructive findings in fibrosing sarcoidosis cohorts. (R^2= 0.553, Q^2= 0.422, p= 0.002)

with CPI > 40 of 67%, specificity of 100%, positive predictive value of 100% and negative predictive value of 75%.

The prediction sets results

The prediction set for differentiating subjects with and without pulmonary sarcoidosis using 14 metabolites and features showed a strong predictability (Figure 4 and table 3). Therefore, our proposed model reached a sensitivity of 93%, specificity of 73%, positive predictive value of 79% and negative predictive value of 92%.

The prediction set that metabolic profiling could be strong enough with relatively high predictability and significant (Q^2Y= 0.421, p= 0.002) to separate patients with and without obstructive defect in PFTs and among those with CPI > 40 vs. CPI < 40 score among the fibrosing sarcoidosis cohort (Figure 5, 6, 7 and table 3).

Pathway analysis

Two of the 10 known discriminating metabolites (p-coumaroylagmatine and palmitoylcarnitine) were mapped to two distinct pathways. Palmitoyl-

Table 3. The prediction sets results for fibrosing pulmonary sarcoidosis, obstractive defects among fibrosing pulmonary sarcoidosis and $CPI_{\text{-}40}$

OPLS-DA analysis	R^2Y	Q^2Y	*P* value	Sensitivity	Specificity	PPV	NPV	AUC
Fibrosing vs non-fibrosing sarcoidosis	0.514	0.322	0.002	93%	73%	79%	91%	0.92
Obstructive vs. non-obstructive	0.553	0.421	0.002	85%	100%	100%	98%	1
$CPI_{\text{-}40}$ vs. $CPI_{\text{-}40}$	0.464	0.378	0.0001	67%	100%	100%	75%	1

CPI: Composite physiological index, PPV: positive predictive value, NPV: Negative predictive value, AUC: area under the receiver-operating curve.

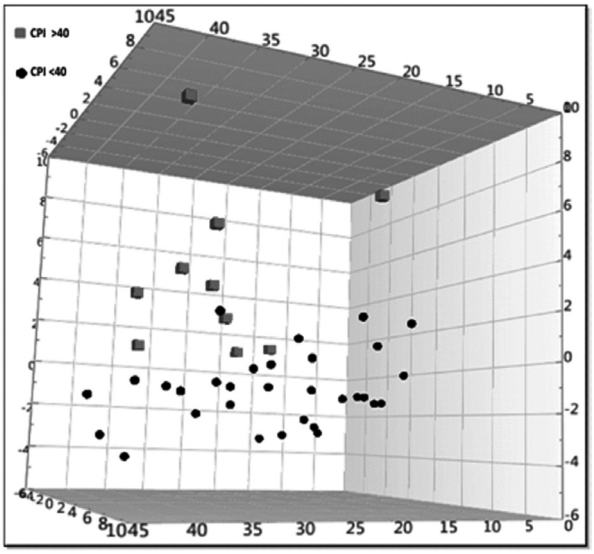

Fig. 6. OPLS-DA analysis showed relatively good discrimination (Q^2Y= 0.381, R^2Y=0.464) with statistically significant level (p= 0.0001) between those individuals with CPI score > 40 (n=10) vs. those had CPI < 40 (n=31)

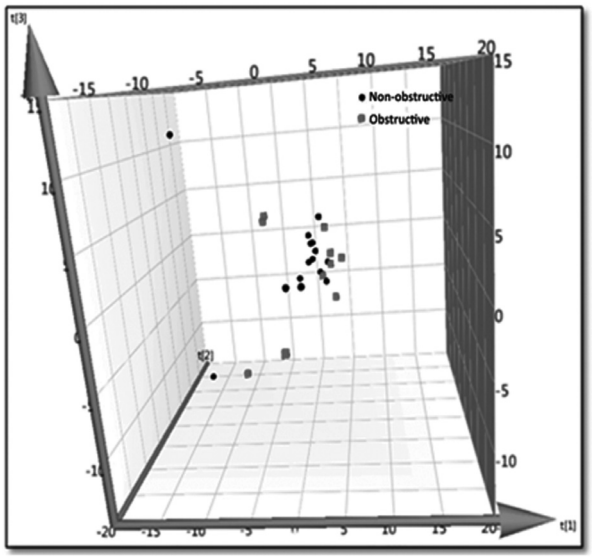

Fig. 7. The unsupervised PCA analysis shows the general separation (grouping) between obstructive and non-obstructive individuals (based on PFTs) in fibrosing sarcoidosis cohorts. T1(X axis) is first principal component that has the greatest variation between the samples and t2 (Y axis) is the second principal component has the second greatest variations, and t3 (Z axis) as the third greatest variation between the samples

CoA is important for the mitochondrial matrix and mitochondrial L-carnitine shuttle. P-coumaroylagmatine is involved in arginine and proline metabolism (30), which is involved in the fibrosis process. On-line Supplemental figures 1 and 2 show these metabolic pathways.

Discussion

Some patients with sarcoidosis have minor non-progressive disease and others progress to disability, respiratory failure, and death. During course of the disease, there is no good way to determine which will progress and which will not (23, 31, 32). Progressive disease may be due to infections, pulmonary hypertension, or worsening fibrosing disease (32). A marker of fibrosis, especially if it could indicate ongoing activity in a manner more sensitive than imaging and pulmonary function testing, could alert clinicians to disease activity and need for treatment. Furthermore, such a marker could help understand the pathophysiology of this disease and be useful for clinical trials.

This study identified markers in sarcoidosis patients that were associated with pulmonary fibrosis. By identifying those patients who are at risk for progressive disease, it may be possible to develop treatment strategies to control or prevent progression.

In this study, we could separate sarcoidosis subjects with pulmonary fibrosis from those without fibrosis by using a plasma metabolomics approach. This 14 metabolite-feature model could be used for assessment of fibrosing activity and thus risk of progression. We also showed that LC-MS was a predic-

tor of obstructive defect in pulmonary function tests in patients with pulmonary fibrosis (Figure 5).

Interestingly, three metabolites in the regression model are associated with arginine and proline metabolism. Collagen molecules are mainly composed of glycine, hydroxyproline, and proline (33). Proline and hydroxyproline form almost 23% of collagen molecules (33), and arginine is a necessary precursor of proline (34). More than 99% of the body's stores of hydroxyproline are found in collagen (35). Given p-coumaroylagmatine is bioactive molecule derived from decarboxylated arginine, a significantly higher p-coumaroylagmatine plasma level is seen in fibrosing pulmonary sarcoidosis individuals (as shown in Table 4 and figure 3) suggest higher collagen metabolism in this group.

The current study found two related carnitine metabolites had significantly higher levels in patients with pulmonary fibrosis. Carnitine has a major role in cellular metabolism (36) through acyl-coA, which is responsible for the transport of cytoplasmic long-chain fatty to the mitochondrial matrix where they are processed for beta oxidation and energy production (37).

Another finding of higher agmatine (p-coumaroylagmatine) and palmitoylcarnitine in the plasma of persons with fibrosing pulmonary sarcoidosis suggests mechanistic pathways. Agmatine (p-coumaroylagmatine) is a bioactive molecule derived from decarboxylated arginine. It is an agonist of the histamine H3 receptor (H3R), which plays an important role in inducing inflammation and allergic reactions in asthma (38, 39).

Palmitoylcarnitine is a surface-active agent with surfactant properties. Palmitoylcarnitine is an intermediate in mitochondrial fatty acid oxidation and, thus, an indicator of mitochondrial respiration (40, 41). Increased palmitoylcarnitine in the fibrosing sarcoidosis would be consistent with alveolar remodeling and leaking of alveolar surfactant to the blood stream. If true, palmitoylcarnitine could be investigated as a new biomarker for alveolar remodeling.

Carnitine can increase collagen and extracellular matrix. Stoppoloni and coworkers showed that carnitine enhances extracellular matrix synthesis in human cartilage (42). Its effect on human lung tissue remains unclear, although it has been claimed that acetyl-L-carnitine could reduce lung injury induced by bleomycin in animal model (43).

The present study found significantly higher levels of vitamin D_3 metabolite (19-hydroxy-10S, 19-dihydrovitamin D_3) in subjects with sarcoidosis and pulmonary fibrosis. Vitamin D_3 and its metabolites levels have been found to be increased in alveolar macrophages in persons with sarcoidosis (44). Higher serum vitamin D levels are linked to greater disease activity and poorer prognosis in sarcoidosis (45). Our vitamin D metabolite findings may be markers of disease activity by themselves and should be explored further.

Our study has several limitations. Our overall and prediction group sample size were not large. We did not study clinical outcomes (not all fibrosis patients do poorly). The study was retrospective and the subjects were not being controlled for in term of their therapies and medications. This study is an

Table 4. List of metabolites identified in multivariate discriminant analysis model

Metabolites	Retention time (min)	m/z	Mass accuracy, ppm	Fold change
p-Coumaroylagmatine	1.800333333	299.1404484	24	1.33187016
Elaidic carnitine	11.3155	426.3562344	3	1.351175314
N-stearoyl serine	11.514	394.2932319	1	1.675419018
Palmitoylcarnitine	11.40333333	400.3413307	1	11.40333333
1-Methylestra-1,3,5(10)-triene-3,17beta-diol	12.21841667	287.2056179	17	1.477334715
19-Hydroxy-10S,19-dihydrovitamin D3	11.40775	425.3435275	10	1.29973358
Pregeijerene	8.7605	185.1292084	4	1.315845762
6-Bromo-2-naphthol	10.48616667	244.9580907	3	1.272196791
2-(9E-octadecenoyl)-sn-glycero-3-phosphocholine	13.20283333	522.3502835	9	1.101795274
1-Hexadecylamine*	4.115	242.2847071	2	1.217777034
M206T10 (unknown feature)	10.482	205.9289544	X	1.157665149
M142T16 (unknown feature)	16.46825	142.1146796	X	1.216470804
M206T10 (unknown feature)	8.424083333	205.9286471	X	1.187330813
M308T17 (unknown feature)	17.00075	307.6882433	X	1.226774939

This metabolite was detected in lower level in the case group

initial step to apply metabolomics to the assessment of pulmonary fibrosis in sarcoidosis. A prospective study is needed to test the value of this information in predicting progression in an independent patient group. The multivariate model has a moderately high negative predictive value but a relatively low positive predictive value.

In conclusion, this study presents a metabolomic approach as a novel tool for assessment of pulmonary fibrosis in sarcoidosis. Our findings show that arginine-proline pathway is increased in patients with fibrosis, which might be expected by the formation and breakdown of collagen. Palmitoylcarnitine is a potential biomarker in fibrosing pulmonary sarcoidosis. A combination of 14 metabolites and features found in this study could be used as biomarkers to detect pulmonary fibrosis in sarcoidosis patients.

Acknowledgement

The authors would like to express their gratitude to the patients who participated in this project. This project is supported by NIH grant 5 T32 HL 82547-7 and a grant from Bernie Mac Sarcoidosis Foundation.

Author Contributions: Conception, hypotheses delineation, and design of the study: M.M. The data analysis and interpretation of such information: M.M., M.M.B, C.K.N., T.A., A.H., Writing the article or substantial involvement in its revision before submission: M.M, M.M.B, B.W.W., D.S., N.J.S, R.B., J.G.N.G., R.M.

References

1. Cragin LA, Laney AS, Lohff CJ, Martin B, Pandiani JA, Blevins LZ. Use of insurance claims data to determine prevalence and confirm a cluster of sarcoidosis cases in Vermont. Public Health Rep 2009; 124 (3): 442-446.
2. Lazarus A. Sarcoidosis: epidemiology, etiology, pathogenesis, and genetics. Dis Mon 2009; 55(11): 649-660.
3. Nunes H, Bouvry D, Soler P, Valeyre D. Sarcoidosis. Orphanet J Rare Dis 2007; 2: 46.
4. Mehdi Mirsaeidi RFM, Dean Schraufnagel, Nadera J Sweiss, Robert Baughman Racial Difference in Sarcoidosis Mortality in the United States. Chest 2014.
5. Sweiss NJ, Lower EE, Mirsaeidi M, Dudek S, Garcia JG, Perkins D, Finn PW, Baughman RP. Rituximab in the treatment of refractory pulmonary sarcoidosis. The European respiratory journal 2014.
6. Lockstone HE, Sanderson S, Kulakova N, Baban D, Leonard A, Kok WL, McGowan S, McMichael AJ, Ho LP. Gene set analysis of lung samples provides insight into pathogenesis of progressive, fibrotic pulmonary sarcoidosis. Am J Respir Crit Care Med 2010; 181 (12): 1367-1375.
7. Heijne WH, Kienhuis AS, van Ommen B, Stierum RH, Groten JP. Systems toxicology: applications of toxicogenomics, transcriptomics, proteomics and metabolomics in toxicology. Expert Rev Proteomics 2005; 2 (5): 767-780.
8. Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. Trends Biotechnol 1998; 16 (9): 373-378.
9. Lindon JC, Nicholson JK, Holmes E, Antti H, Bollard ME, Keun H, Beckonert O, Ebbels TM, Reily MD, Robertson D, Stevens GJ, Luke P, Breau AP, Cantor GH, Bible RH, Niederhauser U, Senn H, Schlotterbeck G, Sidelmann UG, Laursen SM, Tymiak A, Car BD, Lehman-McKeeman L, Colet JM, Loukaci A, Thomas C. Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. Toxicol Appl Pharmacol 2003; 187 (3): 137-146.
10. German JB, Watkins SM, Fay LB. Metabolomics in practice: emerging knowledge to guide future dietetic advice toward individualized health. J Am Diet Assoc 2005; 105 (9): 1425-1432.
11. Ge Y, Wang TJ. Identifying novel biomarkers for cardiovascular disease risk prediction. J Intern Med 2012; 272 (5): 430-439.
12. Davis VW, Schiller DE, Eurich D, Sawyer MB. Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. World J Surg Oncol 2012; 10: 271.
13. Farshidfar F, Weljie AM, Kopciuk K, Buie WD, Maclean A, Dixon E, Sutherland FR, Molckovsky A, Vogel HJ, Bathe OF. Serum metabolomic profile as a means to distinguish stage of colorectal cancer. Genome Med 2012; 4 (5): 42.
14. Ho WE, Xu YJ, Xu F, Cheng C, Peh HY, Tannenbaum SR, Wong WS, Ong CN. Metabolomics reveals altered metabolic pathways in experimental asthma. Am J Respir Cell Mol Biol 2013; 48 (2): 204-211.
15. Ubhi BK, Riley JH, Shaw PA, Lomas DA, Tal-Singer R, MacNee W, Griffin JL, Connor SC. Metabolic profiling detects biomarkers of protein degradation in COPD patients. Eur Respir J 2012; 40 (2): 345-355.
16. Weiner J, 3rd, Parida SK, Maertzdorf J, Black GF, Repsilber D, Telaar A, Mohney RP, Arndt-Sullivan C, Ganoza CA, Fae KC, Walzl G, Kaufmann SH. Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. PLoS ONE 2012; 7 (7): e40221.
17. Mickiewicz B, Vogel HJ, Wong HR, Winston BW. Metabolomics as a Novel Approach for Early Diagnosis of Pediatric Septic Shock and its Mortality. Am J Respir Crit Care Med 2013.
18. Mickiewicz B, Duggan GE, Winston BW, Doig C, Kubes P, Vogel HJ. Metabolic Profiling of Serum Samples by 1H Nuclear Magnetic Resonance Spectroscopy as a Potential Diagnostic Approach for Septic Shock. Crit Care Med 2013.
19. Ghannoum MA, Mukherjee PK, Jurevic RJ, Retuerto M, Brown RE, Sikaroodi M, Webster-Cyriaque J, Gillevet PM. Metabolomics reveals differential levels of oral metabolites in HIV-infected patients: toward novel diagnostic targets. OMICS 2013; 17 (1): 5-15.
20. Philippeos C, Steffens FE, Meyer D. Comparative 1H NMR-based metabonomic analysis of HIV-1 sera. Journal of Biomolecular NMR 2009; 44 (3): 127-137.
21. Ramautar R, Berger R, van der Greef J, Hankemeier T. Human metabolomics: strategies to understand biology. Current opinion in chemical biology 2013; 17 (5): 841-846.
22. Hunninghake GW, Costabel U, Ando M, Baughman R, Cordier JF, du Bois R, Eklund A, Kitaichi M, Lynch J, Rizzato G, Rose C, Selroos O, Semenzato G, Sharma OP. ATS/ERS/WASOG statement on sarcoidosis. American Thoracic Society/European Respiratory Society/World Association of Sarcoidosis and other Granulomatous Disorders. Sarcoidosis, vasculitis, and diffuse lung diseases: official journal

of WASOG/World Association of Sarcoidosis and Other Granulomatous Disorders 1999; 16 (2): 149-173.
23. Walsh SL, Wells AU, Sverzellati N, Keir GJ, Calandriello L, Antoniou KM, Copley SJ, Devaraj A, Maher TM, Renzoni E, Nicholson AG, Hansell DM. An integrated clinicoradiological staging system for pulmonary sarcoidosis: a case-cohort study. The Lancet Respiratory medicine 2014; 2 (2): 123-130.
24. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J, Jensen R, Johnson DC, MacIntyre N, McKay R, Miller MR, Navajas D, Pedersen OF, Wanger J. Interpretative strategies for lung function tests. The European respiratory journal 2005; 26 (5): 948-968.
25. Wells AU, Desai SR, Rubens MB, Goh NS, Cramer D, Nicholson AG, Colby TV, du Bois RM, Hansell DM. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. American journal of respiratory and critical care medicine 2003; 167 (7): 962-969.
26. Gao P, Lu C, Zhang F, Sang P, Yang D, Li X, Kong H, Yin P, Tian J, Lu X, Lu A, Xu G. Integrated GC-MS and LC-MS plasma metabonomics analysis of ankylosing spondylitis. The Analyst 2008; 133 (9): 1214-1220.
27. Tautenhahn R, Patti Gj Fau - Rinehart D, Rinehart D Fau - Siuzdak G, Siuzdak G. XCMS Online: a web-based platform to process untargeted metabolomic data. 2012 (1520-6882 (Electronic)).
28. Tautenhahn R, Patti Gj Fau - Rinehart D, Rinehart D Fau - Siuzdak G, Siuzdak G. XCMS Online: a web-based platform to process untargeted metabolomic data. 2005(1520-6882 (Electronic)).
29. Lennart Eriksson JT, Svante Wold. CV-ANOVA for significance testing of PLS and OPLS® models. Chemometrics 2008; 22 (11-12): 594-600.
30. Haake P, Allen GW. Studies on phosphorylation by phosphoroguanidinates. The mechanism of action of creatine: ATP transphosphorylase (creatine kinase). Proceedings of the National Academy of Sciences of the United States of America 1971; 68 (11): 2691-2693.
31. Baughman RP, Winget DB, Bowen EH, Lower EE. Predicting respiratory failure in sarcoidosis patients. Sarcoidosis, vasculitis, and diffuse lung diseases : official journal of WASOG / World Association of Sarcoidosis and Other Granulomatous Disorders 1997; 14 (2): 154-158.
32. Nardi A, Brillet PY, Letoumelin P, Girard F, Brauner M, Uzunhan Y, Naccache JM, Valeyre D, Nunes H. Stage IV sarcoidosis: comparison of survival with the general population and causes of death. The European respiratory journal 2011; 38 (6): 1368-1373.
33. Barbul A. Proline precursors to sustain Mammalian collagen synthesis. The Journal of nutrition 2008; 138 (10): 2021S-2024S.
34. Wu G, Morris SM, Jr. Arginine metabolism: nitric oxide and beyond. The Biochemical journal 1998; 336 (Pt 1): 1-17.
35. Grant ME, Prockop DJ. The biosynthesis of collagen. 1. The New England journal of medicine 1972; 286 (4): 194-199.
36. Rebouche CJ. Kinetics, pharmacokinetics, and regulation of L-carnitine and acetyl-L-carnitine metabolism. Annals of the New York Academy of Sciences 2004; 1033: 30-41.
37. Kerner J, Hoppel C. Fatty acid import into mitochondria. Biochimica et biophysica acta 2000; 1486 (1): 1-17.
38. Lim HD, Adami M Fau - Guaita E, Guaita E Fau - Werfel T, Werfel T Fau - Smits RA, Smits Ra Fau - de Esch IJP, de Esch Ij Fau - Bakker RA, Bakker Ra Fau - Gutzmer R, Gutzmer R Fau - Coruzzi G, Coruzzi G Fau - Leurs R, Leurs R. Pharmacological characterization of the new histamine H4 receptor agonist VUF 8430; 2009 (1476-5381 (Electronic)).
39. Weltman JK. Update on histamine as a mediator of inflammation. 2000 (1088-5412 (Print)).
40. Sentex E, Laurent A Fau - Martine L, Martine L Fau - Gregoire S, Gregoire S Fau - Rochette L, Rochette L Fau - Demaison L, Demaison L. Calcium- and ADP-magnesium-induced respiratory uncoupling in isolated cardiac mitochondria: influence of cyclosporin A. 1999 (0300-8177 (Print)).
41. Goni FM, Requero Ma Fau - Alonso A, Alonso A. Palmitoylcarnitine, a surface-active metabolite. 1996 (0014-5793 (Print)).
42. Stoppoloni D, Politi L, Dalla Vedova P, Messano M, Koverech A, Scandurra R, Scotto d'Abusco A. L-carnitine enhances extracellular matrix synthesis in human primary chondrocytes. Rheumatology international 2013; 33 (9): 2399-2403.
43. Sayed-Ahmed MM, Mansour HH, Gharib OA, Hafez HF. Acetyl-L-carnitine modulates bleomycin-induced oxidative stress and energy depletion in lung tissues. Journal of the Egyptian National Cancer Institute 2004: 16 (4): 237-243.
44. Adams JS, Singer FR, Gacad MA, Sharma OP, Hayes MJ, Vouros P, Holick MF. Isolation and structural identification of 1,25-dihydroxyvitamin D3 produced by cultured alveolar macrophages in sarcoidosis. The Journal of clinical endocrinology and metabolism 1985: 60 (5): 960-966.
45. Kavathia D, Buckley JD, Rao D, Rybicki B, Burke R. Elevated 1, 25-dihydroxyvitamin D levels are associated with protracted treatment in sarcoidosis. Respiratory medicine 2010: 104 (4): 564-570.