Title: Metabolomic Profiles of Scleroderma-PAH are different than idiopathic PAH and associated with worse clinical outcomes

Author list: Mona Alotaibi MD¹, Junzhe Shao², Michael W. Pauciulo MBA^{3,4}, William C. Nichols PhD^{3,4}, Anna R. Hemnes MD⁵, Atul Malhotra MD¹, Nick H. Kim MD¹, Jason X.-J. Yuan MD PhD¹, Timothy Fernandes MD¹, Kim M. Kerr MD¹, Laith Alshawabkeh MD, MSCI⁶, Ankit A. Desai MD⁸, Jeramie D. Watrous PhD⁷, Susan Cheng MD MPH MMSc⁹, Tao Long PhD⁷, Stephen Y. Chan MD PhD¹⁰, Mohit Jain MD PhD⁷

Affiliations:

¹Division of Pulmonary, Critical Care and Sleep Medicine, University of California San Diego, La Jolla, CA, USA

² School of Life Sciences, Peking University, Beijing, China

³Department of Pediatrics, College of Medicine, University of Cincinnati, Cincinnati, OH, USA

⁴Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

⁵Division of Allergy, Pulmonary and Critical Care Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

⁶Division of Cardiovascular Medicine, Sulpizio Cardiovascular Institute, University of California San Diego, La Jolla, CA, USA

⁷Departments of Medicine and Pharmacology, University of California San Diego, La Jolla, CA, USA

⁸Department of Medicine, Indiana University, Indianapolis, IN, USA

⁹Barbra Streisand Women's Heart Center, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA

¹⁰Center for Pulmonary Vascular Biology and Medicine, Pittsburgh Heart, Lung, Blood Vascular Medicine Institute, Division of Cardiology, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Corresponding Author:

Mohit Jain, MD PhD

mjain@health.ucsd.edu

Authors Contributions:

MA, MJ designed the research studies, acquired data, analyzed data, and drafted the manuscript. WCN, MWP, AD and ARH acquired data and drafted the manuscript. NHK, JXJY, TF, KMK, LA and AM designed the research studies and drafted the manuscript. JDW conducted experiments, designed research studies, and analyzed data. TL, SC and JS analyzed data and drafted the manuscript. SYC designed the research studies, analyzed the data and drafted the manuscript.

Conflict of interest statement:

None of the authors have any potential conflicts of interest relative to the study. This work was supported by National Institutes of Health (NIH) grants S10OD020025 and R01ES027595 to M. Jain; K01DK116917 to J.D. Watrous; R01 HL124021, HL105333 to W. Nichols; R01 HL136603 to A. Desai; P01 HL108800 and R01 HL142720 to A.R. Hemnes; HL 122596, HL 138437, and UH2/UH3 TR002073 to S.Y. Chan; R01-HL134168, R01-HL143227, R01-HL142983, and U54-AG065141 to SC. A. Malhotra is funded by NIH. S.Y. Chan was also supported by the American Heart Association Established Investigator Award 18EIA33900027. M. Alotaibi was supported by a postdoctoral fellowship award from the Chest Foundation. A. Malhotra reports income related to medical education from Livanova, Equillium and Corvus. ResMed provided a philanthropic donation to UCSD. K. Kerr received university grant money from Bayer and serve as a consultant for Actelion.

Word Count: 2363

Take Home Message:

Among patients with PAH, those with SSc-PAH suffer disproportionately worse outcomes and disease course. This study represents the most comprehensive analysis of bioactive metabolites profiling comparing two subgroups of PAH. The findings shed light on key differences between SSc-PAH and IPAH that provide important metabolic insight into the disease pathogenesis.

Key words: biomarkers, pulmonary hypertension, scleroderma

Abbreviations list

BMI: Body mass index

EpETE: Epoxyeicosatetraenoic acid

FAHFA: Fatty acid ester of hydroxyl fatty acid

HETE: Hydroxyeicosanoid

IPAH: Idiopathic pulmonary arterial hypertension

LC-MS: Liquid chromatography - mass spectrometry

mRAP: Mean right atrial pressure

MRS: Metabolite risk score

MUFA: Monounsaturated unsaturated fatty acids

PAH: Pulmonary arterial hypertension

PVR: Pulmonary vascular resistance

RHC: Right heart catheterization

SMWD: 6-minute walk distance

SSc: Systemic sclerosis

SSc-PAH: Systemic sclerosis associated pulmonary arterial hypertension

VLCSFA: Very long chain saturated fatty acids

WHO FC: World Health Organization functional class

Abstract

The molecular signature in patients with systemic sclerosis (SSc)-associated pulmonary arterial hypertension (SSc-PAH) relative to idiopathic pulmonary arterial hypertension (IPAH) remain unclear. We hypothesize that patients with SSc-PAH exhibit unfavorable bioactive metabolite derangements compared to IPAH that contribute to their poor prognosis and limited response to therapy. We sought to determine whether circulating bioactive metabolites are differentially altered in SSc-PAH versus IPAH.

Plasma biosamples from 415 patients with SSc-PAH (cases) and 1115 patients with IPAH (controls) were included in the study. Over 700 bioactive metabolites were assayed in plasma samples from independent discovery and validation cohorts using liquid chromatography - mass spectrometry (LC-MS) based approaches. Regression analyses were used to identify metabolites which exhibited differential levels between SSc-PAH and IPAH and associated with disease severity.

From among hundreds of circulating bioactive molecules, twelve metabolites were found to distinguish between SSc-PAH and IPAH, as well as associate with PAH disease severity. SSc-PAH patients had increased levels of fatty acid metabolites including lignoceric acid and nervonic acid, as well as kynurenine, polyamines, eicosanoids/oxylipins and sex hormone metabolites relative to IPAH. In conclusion, SSc-PAH patients are characterized by an unfavorable bioactive metabolic profile that may explain the poor and limited response to therapy. These data provide important metabolic insights into the pathogenesis of SSc-PAH.

Introduction:

Pulmonary arterial hypertension (PAH) is a debilitating disease of the pulmonary circulation leading to elevated pulmonary arterial pressures and pulmonary vascular resistance. The most common subgroups of PAH are idiopathic pulmonary arterial hypertension (IPAH) and systemic sclerosis associated pulmonary arterial hypertension (SSc-PAH).¹ Systemic Sclerosis (SSc) is a complex, immunological disease, characterized by autoimmunity, fibrosis of the skin and internal organs, and small vessel vasculopathy.² PAH is a leading cause of death in patients with SSc with mortality three-fold higher than patients with IPAH.³⁻⁶ Despite the comparable end pathology in SSc-PAH and IPAH, SSc-PAH patients have an impaired response to traditional PAH-targeted therapies and carry a worse prognosis relative to other subgroups of PAH, although they may present with milder hemodynamic impairment.^{3; 7} Proposed factors explaining these striking disparities include more pronounced inflammation⁸, autoimmunity⁸, the nature of the underlying vasculopathy⁹ and the ability of the right ventricle to adapt to the increased afterload.¹⁰ In contrast to IPAH, patients with SSc-PAH have depressed sarcomere function portending worse RV contractility.¹¹ However, little is known about the molecular mechanisms underlying these differences. A clearer understanding of the molecular mechanisms underlying SSc-PAH is critical toward better understanding of the disease pathogenesis, including development of prognostic biomarkers and targeted therapies.

Prior mass-spectrometry methods have identified bioactive molecules in circulation.^{12; 13} These molecules include both endogenous compounds (e.g., amino acids, short peptides, nucleic acids, fatty acids, lipids, amines, carbohydrates) and exogenous

chemicals that are not naturally produced in the body. Their levels provide integrative information on biological functions and define the phenotypes of biological systems in response to genetic or environmental changes. To date, the study of these bioactive metabolites in PAH have revealed changes in key energetic pathways, including abnormal oxidation products as well as elevated levels of circulating acylcarnitine, glutamate, and TCA cycle intermediates.^{14; 15} Despite these early studies, a deeper understanding of the metabolic alterations between subgroups of PAH and more specifically SSc-PAH has been limited to date.

In this study, we hypothesize that patients with SSc-PAH exhibit unfavorable metabolic derangements which are associated with worse clinical outcomes (i.e low 6MWD, mortality, etc.) compared to IPAH that could explain the rapid decline and disease pathogenesis. We compared hundreds of circulating bioactive metabolites between SSc-PAH and IPAH in independent studies and identified selective derangements between these subgroups that also associate with hemodynamic measures.

Methods:

Cohorts and sample collection:

Plasma samples were obtained from patients with PAH enrolled as part of the National Biological Sample and Data Repository for Pulmonary Arterial Hypertension (PAH Biobank) between October 2012 and December 2017 from 37 US centers (www.pahbiobank.org). The PAH biobank samples were divided *a priori* into discovery and validation cohort (1st validation cohort) based on center from which samples were

collected. In the discovery cohort, 310 SSc-PAH patients (cases) and 869 IPAH patients (controls) were included. In the validation cohort (1st validation cohort), 90 SSc-PAH and 213 IPAH patients were included. A second validation cohort (2nd validation cohort) of 15 SSc-PAH patients and 90 IPAH patients was obtained from Vanderbilt Medical Center. Inclusion criteria were confirmed precapillary pulmonary hypertension by right heart catheterization (RHC) ($mPAP \geq 25$ mmHg, $PCWP \leq 15$ mmHg, $PVR > 3WU$). Diagnosis of SSc-PAH was established clinically based on published criteria.¹⁶

Plasma samples were obtained from the antecubital fossa and collected in EDTA vacutainer tubes, immediately put-on ice, centrifuged, and stored at -80°C. World Health Organization functional class (WHO FC), 6-minute walk distance (6MWD) and clinical and hemodynamics data were recorded for all patients, using established criteria. All subjects provided informed consent and local research ethics committees approved the study.

Metabolite Assay

Bioactive metabolites analysis was performed on plasma samples by liquid chromatography - mass spectrometry (LC-MS), using a Vanquish UPLC coupled to high resolution, QExactive orbitrap mass spectrometer (Thermo), similar to as previously described^{12; 17} (details can be found in the online data supplement). Polar metabolites including sugars and organic acids, were assayed using Zic-pHILIC 2.1x150mm 5 μ m column, and small polar, bioactive lipids were measured using a Phenomenex Kinetex C18 column.¹⁷⁻²⁶ Metabolites identified as xenobiotics or detected in <20% of samples were excluded from the analysis, leaving over 700 well-quantified biological metabolites.

Qc/Qa analysis was performed as described in data supplement, and spectral data were extracted as previously described.^{18-20; 22} Data were subsequently normalized using batch median normalization metric with correction for median absolute deviation. Following normalization, metabolite peaks were further compressed for multiple adducts and in source fragments. Normalized, aligned, filtered datasets were subsequently used for statistical analyses, as described below.

Statistical Analysis:

Initial group comparisons between SSc-PAH and IPAH patients were performed using t-test or the Mann-Whitney test for continuous variables and the χ^2 test for categorical variables. Prior to all analyses, metabolite values were natural logarithmically transformed, as needed, and later standardized with mean=0 and SD=1 to facilitate comparisons. Logistic regression analysis was used to determine metabolites that were significantly different between SSc-PAH and IPAH (analysis I). Linear regression was performed between the significant metabolites from analysis I and 6MWD, FC, mRAP, and PVR in both IPAH and SSc-PAH (analysis II). Sensitivity analysis was performed in IPAH only and SSc-PAH only. Sub-analysis adjusting for immunosuppression medication and steroid use was performed in the significant metabolites. All analyses were performed in models adjusting for age, gender, and body mass index (BMI). To determine significance, a Bonferroni corrected *P* value threshold of 0.05 divided by a conservative estimate of the total number of unique small molecules (i.e., $p < 10^{-4}$) was used. Lasso regularized regression model was used to build a prediction model and metabolite risk score (MRS) to select the minimum number of metabolites that distinguish between SSc-

PAH and IPAH. Statistical analysis was performed with R with RStudio and associated packages.²⁷

Results:

Analysis of Study Cohorts:

Baseline demographic, clinical and hemodynamic characteristics and medications for patients enrolled in the study are summarized in **Table 1 and Table e1**. Patients with SSc-PAH were significantly older with female predominance compared to IPAH. At the time of enrollment, SSc-PAH patients had significantly lower mRAP and PVR than IPAH counterparts. Most patients with SSc-PAH were on anti-inflammatory and immunosuppressant therapies at the time of enrollment.

Metabolites distinguishing between SSc-PAH and IPAH:

Circulating levels of 94 bioactive metabolites across broad chemical classes, including bioactive lipids and polar metabolites, distinguished SSc-PAH from IPAH in both discovery and 1st validation cohorts at a 'metabolome wide' statistical threshold of $p < 10^{-4}$ after correction for confounders including age, gender and BMI (**Figure 1**). These metabolites included alterations in fatty acid oxidation and eicosanoids metabolism, steroid hormones, kynurenine pathway, polyamine and pyrimidine pathways (**Figure 1**). Orthogonal partial least squares-discriminant analysis (OPLS-DA) plots showed clear separation between SSc-PAH and IPAH (**Figure 1e**).

Determining Metabolites which most differentiate SSc-PAH from IPAH:

To identify a minimal set of metabolites able to distinguish between SSc-PAH and IPAH, we performed regularized regression analysis. A metabolite risk score (MRS) comprised of 12 metabolites was able to distinguish IPAH from SSc-PAH with discrimination at an AUC of 0.86 (95% CI: 0.82-0.91). This MRS was validated in an independent validation cohort (2nd validation cohort) of 33 IPAH patients and 15 SSc-PAH patients from Vanderbilt Medical Center with an AUC of 0.76 (95% CI: 0.60-0.92) (**Figure 2**). These 12 metabolites were associated with markers of disease severity. 11 of the selected metabolites remained significant after adjusting for immunosuppression medication and steroid use.

Metabolites differentiating SSc-PAH from IPAH and associated with disease severity:

To determine if distinguishing metabolites may contribute to worsening disease prognosis, we next performed association of the 94 metabolite biomarkers with available clinical markers of disease severity and hemodynamic parameters, including 6MWD, FC, mRAP, and PVR. As shown in **Table 2** and **Figure 3**, 30 metabolites were significantly associated with at least one measurement of disease severity ($p < 0.05$ for each metabolite). This includes fatty acid metabolites such as very long chain saturated fatty acids (VLCSCFA) and monounsaturated unsaturated fatty acids (MUFA), several pro-inflammatory eicosanoids, kynurenine, polyamines and sex hormones. Additionally, novel eicosanoids and fatty acid ester of hydroxyl fatty acid (FAHFA) emerged as markers of disease severity and were significantly higher in SSc-PAH.

163

164 **Discussion:**

165 In this report, we provide novel evidence that patients with SSc-PAH have significant
 166 bioactive metabolic alterations compared to those with IPAH. We assayed hundreds of
 167 circulating bioactive metabolites using LC-MS approaches in 415 SSc-PAH patients and
 168 1115 IPAH controls in independent discovery and validation cohorts. We identified a set
 169 of bioactive metabolite biomarkers independently differentiating SSc-PAH from IPAH and
 170 associated with disease severity, after adjusting for age, gender, BMI and medications.
 171 In combination, these biomarkers were able to distinguish SSc-PAH patients from IPAH
 172 with high degree of accuracy. These findings provide molecular insight into the
 173 heterogeneity between PAH subgroups and could explain in part the worse prognosis
 174 and response to therapy in patients with SSc-PAH.

175

176 Our novel findings, which highlight 12 metabolites from among hundreds assayed, are
 177 significant both pathophysiologically and clinically. Pathophysiologically, this work sheds
 178 light on potential different mechanisms between SSc-PAH and IPAH. Clinically, it
 179 suggests that these biomarkers could be explored as potential prognostic tools and
 180 therapeutic targets.

181

182 We observed novel associations between saturated and unsaturated fatty acids that were
 183 significantly higher in patients with SSc-PAH and correlated with worse disease markers.
 184 Very-long-chain saturated fatty acids (VLCSFAs) are group of saturated fatty acids with
 185 a chain length of ≥ 20 carbon atoms. They have distinct functions when compared to long

chain saturated fatty acids, and are involved in liver homeostasis, retinal function, and anti-inflammatory functions³⁰. Little is known about the role of VLCSCFA in relation to pulmonary vascular pathology. In fact, this is the first observation to our knowledge to associate VLCSCFA with PAH. In our study, VLCSCFAs were positively associated with PVR, mRAP and negatively associated with 6MWD. Monounsaturated fatty acids (MUFA), such as nervonic acid, are involved in many physiological processes, including energy metabolism, antioxidant reactions and apoptosis.³¹ In previous reports, nervonic acid was positively associated with greater congestive heart failure, poor performance and increased risk of cardiovascular mortality.^{32; 33}

In our study, patients with SSc-PAH had more significant alterations in several pathways linked to endothelial cells dysfunction and RV dysfunction like kynurenine pathway, polyamines and spermine metabolism, eicosanoids, sex hormones among others, despite having better hemodynamic profile at the time of enrollment. Although some of these markers were described in cardiovascular disease or pulmonary hypertension, this is the first-time showing increase levels in SSc-PAH.

Levels of kynurenine, an immune signaling molecule, correlated with resting pulmonary artery pressure in unexplained dyspnea patients and patients with both mild and severe PH.¹⁵ Plasma levels of polyamines metabolites including N-acetylputrescine was associated with chronic inflammation and previous reports showed increased levels of acetylputrescine in animal models of PAH.³⁴ Interestingly, the administration of polyamines inhibitors in monocrotaline rats prevented the development of pulmonary

hypertension and RV dysfunction³⁵, suggesting this could have therapeutic implications as well in this subgroup specifically. Several pro-inflammatory eicosanoids such as, 15-HETE, 17(18)-EpETE and prostaglandins have been implicated in the pathogenesis of PAH and associated with disease severity.³⁶

Rhodes et al¹⁴ was among the first to use comprehensive LC-MS based metabolomics platform to identify discriminative and prognostic metabolites in PAH. They identified a set of 20 metabolites discriminative between IPAH and healthy and diseased controls. Levels of these metabolites were similar between SSc-PAH and IPAH in our study. Among the prognostic metabolites in IPAH, levels of acetamidobutanoate and acetylputrescine were associated with mortality in the Rhodes study and were significantly elevated in SSc-PAH in our cohorts relative to IPAH. This supports our hypothesis that these metabolites could contribute to the worse outcomes in SSc-PAH.

A few studies have used circulating metabolites to determine potential metabolic pathways altered in scleroderma (with or without PH).^{28; 29} To date however, these studies have been limited by sample size as well as independent validation and did not compare between subgroups of PAH. A comparison between 8 patients with scleroderma without PAH and 10 patients with scleroderma and PAH using nuclear magnetic resonance (NMR) techniques identified increase in glycolysis and altered fatty acid profiles in patients with scleroderma and PAH.²⁹ None of our top differentiating metabolites were measured in this study, mainly related to technical differences (use of NMR vs LC-MS) and small sample size. Thus, a strength of our work is the application of broad plasma

bioactive metabolites analysis in large independent cohorts of IPAH and SSc-PAH. We understand that it is important to compare between scleroderma-PAH and scleroderma without PAH to ascertain that these differences are not due to scleroderma only. However, this is beyond the scope of this work, and we hope to address this in the future. We believe that choosing metabolites that associate with hemodynamic measures and disease severity is suggestive of PH related biomarkers.

Even though scleroderma can be easily distinguishable in most cases clinically from IPAH, our goal was not to develop a diagnostic tool, rather to identify key metabolites distinguishing these two subtypes that could shed light on the pathogenesis between SSc-PAH and IPAH. Our hope is that in the future, this could be explored more for potential therapeutic targets.

This study has several limitations. Importantly, given the study design, adjustment for all potential confounders between IPAH and SSc-PAH remains difficult. It is possible that medications such as anti-inflammatory and immunosuppressants, demographic features including gender, smoking and other comorbidities may contribute to metabolic changes between IPAH and SSc-PAH. We tried our best adjusting for these factors in the analysis. Finally, while metabolite markers were found to distinguish SSc-PAH from IPAH and independently associate with disease severity, establishing a clear *causal* relationship for the role of these metabolic pathways in SSc-PAH will require independent studies in experimental model systems. Despite these acknowledged limitations, we believe that

our findings provide important scientific insight on the metabolic alterations present in SSc-PAH and their potential role in disease pathobiology.

Conclusion:

Our findings suggest that despite the lack of distinctive pathologic features, SSc-PAH is characterized by significant metabolomic alterations compared to IPAH that may contribute to worsening disease and poor response to therapy. Moreover, our study suggests that metabolite levels may distinguish IPAH from SSc-PAH and therefore, different pathways maybe driving the pathogenesis of PAH in these two groups. This observation may lead to much needed novel therapeutic strategies in SSc-PAH patients.

Acknowledgments:

Samples and/or Data from the National Biological Sample and Data Repository for PAH, which receives government support under an investigator-initiated grant (R24 HL105333) awarded by the National Heart Lung and Blood Institute (NHLBI) were used in this study. We thank contributors, including the Pulmonary Hypertension Centers who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible.

Disclosures: S.Y.C. has served as a consultant United Therapeutics; S.Y.C. has held research grants from Actelion and Pfizer. S.Y.C. is a director, officer, and shareholder of Synhale Therapeutics. S.Y.C. has submitted patent applications regarding metabolism in pulmonary hypertension. NHK has served as consultant for Bayer, Janssen, Merck,

278 United Therapeutics and has received lecture fees for Bayer, Janssen. NHK has received
279 research support from Acceleron, Eiger, Gossamer Bio, Lung Biotechnology, SoniVie.

280

281

282

283

284

285

References:

1. Badesch DB, Raskob GE, Elliott CG, Krichman AM, Farber HW, Frost AE, Barst RJ, Benza RL, Liou TG, Turner M et al. 2010. Pulmonary arterial hypertension: Baseline characteristics from the REVEAL registry. *Chest*. 137(2):376-387.
2. Allanore Y, Simms R, Distler O, Trojanowska M, Pope J, Denton CP, Varga J. 2015. Systemic sclerosis. *Nat Rev Dis Primers*. 1:15002.
3. Fisher MR, Mathai SC, Champion HC, Girgis RE, Houston-Harris T, Hummers L, Krishnan JA, Wigley F, Hassoun PM. 2006. Clinical differences between idiopathic and scleroderma-related pulmonary hypertension. *Arthritis Rheum*. 54(9):3043-3050.
4. Chung L, Liu J, Parsons L, Hassoun PM, McGoon M, Badesch DB, Miller DP, Nicolls MR, Zamanian RT. 2010. Characterization of connective tissue disease-associated pulmonary arterial hypertension from reveal: Identifying systemic sclerosis as a unique phenotype. *Chest*. 138(6):1383-1394.
5. Ahmed S, Palevsky HI. 2014. Pulmonary arterial hypertension related to connective tissue disease: A review. *Rheum Dis Clin North Am*. 40(1):103-124.
6. Gashouta MA, Humbert M, Hassoun PM. 2014. Update in systemic sclerosis-associated pulmonary arterial hypertension. *Presse Med*. 43(10 Pt 2):e293-304.
7. Clements PJ, Tan M, McLaughlin VV, Oudiz RJ, Tapson VF, Channick RN, Rubin LJ, Langer A, Investigators PAHQRIP-Q. 2012. The pulmonary arterial hypertension quality enhancement research initiative: Comparison of patients with idiopathic pah to patients with systemic sclerosis-associated pah. *Ann Rheum Dis*. 71(2):249-252.
8. Chaisson NF, Hassoun PM. 2013. Systemic sclerosis-associated pulmonary arterial hypertension. *Chest*. 144(4):1346-1356.
9. Overbeek MJ, Vonk MC, Boonstra A, Voskuyl AE, Vonk-Noordegraaf A, Smit EF, Dijkmans BA, Postmus PE, Mooi WJ, Heijdra Y et al. 2009. Pulmonary arterial hypertension in limited cutaneous systemic sclerosis: A distinctive vasculopathy. *Eur Respir J*. 34(2):371-379.
10. Kelemen BW, Mathai SC, Tedford RJ, Damico RL, Corona-Villalobos C, Kolb TM, Chaisson NF, Harris TH, Zimmerman SL, Kamel IR et al. 2015. Right ventricular remodeling in idiopathic and scleroderma-associated pulmonary arterial hypertension: Two distinct phenotypes. *Pulm Circ*. 5(2):327-334.
11. Hsu S, Kokkonen-Simon KM, Kirk JA, Kolb TM, Damico RL, Mathai SC, Mukherjee M, Shah AA, Wigley FM, Margulies KB et al. 2018. Right ventricular myofilament functional differences in humans with systemic sclerosis-associated versus idiopathic pulmonary arterial hypertension. *Circulation*. 137(22):2360-2370.
12. Watrous JD, Niiranen TJ, Lagerborg KA, Henglin M, Xu YJ, Rong J, Sharma S, Vasan RS, Larson MG, Armando A et al. 2019. Directed non-targeted mass spectrometry and chemical networking for discovery of eicosanoids and related oxylipins. *Cell Chem Biol*. 26(3):433-442.e434.
13. Lagerborg KA, Watrous JD, Cheng S, Jain M. 2019. High-throughput measure of bioactive lipids using non-targeted mass spectrometry. *Methods Mol Biol*. 1862:17-35.

14. Rhodes CJ, Ghataorhe P, Wharton J, Rue-Albrecht KC, Hadinnapola C, Watson G, Bleda M, Haimel M, Coghlan G, Corris PA et al. 2017. Plasma metabolomics implicates modified transfer rnas and altered bioenergetics in the outcomes of pulmonary arterial hypertension. *Circulation*. 135(5):460-475.
15. Lewis GD, Ngo D, Hemnes AR, Farrell L, Domos C, Pappagianopoulos PP, Dhakal BP, Souza A, Shi X, Pugh ME et al. 2016. Metabolic profiling of right ventricular-pulmonary vascular function reveals circulating biomarkers of pulmonary hypertension. *J Am Coll Cardiol*. 67(2):174-189.
16. McLaughlin VV, Archer SL, Badesch DB, Barst RJ, Farber HW, Lindner JR, Mathier MA, McGoon MD, Park MH, Rosenson RS et al. 2009. 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*. 53(17):1573-1619.
17. Kantz ED, Tiwari S, Watrous JD, Cheng S, Jain M. 2019. Deep neural networks for classification of lc-ms spectral peaks. *Anal Chem*. 91(19):12407-12413.
18. Jain M, Nilsson R, Sharma S, Madhusudhan N, Kitami T, Souza AL, Kafri R, Kirschner MW, Clish CB, Mootha VK. 2012. Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science*. 336(6084):1040-1044.
19. Jain M, Ngoy S, Sheth SA, Swanson RA, Rhee EP, Liao R, Clish CB, Mootha VK, Nilsson R. 2014. A systematic survey of lipids across mouse tissues. *Am J Physiol Endocrinol Metab*. 306(8):E854-868.
20. Knoechel B, Roderick JE, Williamson KE, Zhu J, Lohr JG, Cotton MJ, Gillespie SM, Fernandez D, Ku M, Wang H et al. 2014. An epigenetic mechanism of resistance to targeted therapy in t cell acute lymphoblastic leukemia. *Nat Genet*. 46(4):364-370.
21. Nilsson R, Jain M, Madhusudhan N, Sheppard NG, Strittmatter L, Kampf C, Huang J, Asplund A, Mootha VK. 2014. Metabolic enzyme expression highlights a key role for mthfd2 and the mitochondrial folate pathway in cancer. *Nat Commun*. 5:3128.
22. Sharma S, Quintana A, Findlay GM, Mettlen M, Baust B, Jain M, Nilsson R, Rao A, Hogan PG. 2013. An sirna screen for nfat activation identifies septins as coordinators of store-operated ca2+ entry. *Nature*. 499(7457):238-242.
23. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, Lewis GD, Fox CS, Jacques PF, Fernandez C et al. 2011. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 17(4):448-453.
24. Wang TJ, Ngo D, Psychogios N, Dejam A, Larson MG, Vasan RS, Ghorbani A, O'Sullivan J, Cheng S, Rhee EP et al. 2013. 2-aminoadipic acid is a biomarker for diabetes risk. *J Clin Invest*. 123(10):4309-4317.
25. Cheng S, Rhee EP, Larson MG, Lewis GD, McCabe EL, Shen D, Palma MJ, Roberts LD, Dejam A, Souza AL et al. 2012. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation*. 125(18):2222-2231.

26. Ho JE, Larson MG, Vasan RS, Ghorbani A, Cheng S, Rhee EP, Florez JC, Clish CB, Gerszten RE, Wang TJ. 2013. Metabolite profiles during oral glucose challenge. *Diabetes*. 62(8):2689-2698.
27. Team RC. R: A language and environment for statistical computing.
28. Bengtsson AA, Trygg J, Wuttge DM, Sturfelt G, Theander E, Donten M, Moritz T, Sennbro CJ, Torell F, Lood C et al. 2016. Metabolic profiling of systemic lupus erythematosus and comparison with primary sjögren's syndrome and systemic sclerosis. *PLoS One*. 11(7):e0159384.
29. Deidda M, Piras C, Cadeddu Dessalvi C, Locci E, Barberini L, Orofino S, Musu M, Mura MN, Manconi PE, Finco G et al. 2017. Distinctive metabolomic fingerprint in scleroderma patients with pulmonary arterial hypertension. *Int J Cardiol*. 241:401-406.
30. Kihara A. 2012. Very long-chain fatty acids: Elongation, physiology and related disorders. *J Biochem*. 152(5):387-395.
31. Schwingshackl L, Strasser B, Hoffmann G. 2011. Effects of monounsaturated fatty acids on cardiovascular risk factors: A systematic review and meta-analysis. *Ann Nutr Metab*. 59(2-4):176-186.
32. Contaifer D, Buckley LF, Wohlford G, Kumar NG, Morriss JM, Ranasinghe AD, Carbone S, Canada JM, Trankle C, Abbate A et al. 2019. Metabolic modulation predicts heart failure tests performance. *PLoS One*. 14(6):e0218153.
33. Imamura F, Lemaitre RN, King IB, Song X, Steffen LM, Folsom AR, Siscovick DS, Mozaffarian D. 2013. Long-chain monounsaturated fatty acids and incidence of congestive heart failure in 2 prospective cohorts. *Circulation*. 127(14):1512-1521, 1521e1511-1518.
34. Hoet PH, Nemery B. 2000. Polyamines in the lung: Polyamine uptake and polyamine-linked pathological or toxicological conditions. *Am J Physiol Lung Cell Mol Physiol*. 278(3):L417-433.
35. Olson JW, Atkinson JE, Hacker AD, Altieri RJ, Gillespie MN. 1985. Suppression of polyamine biosynthesis prevents monocrotaline-induced pulmonary edema and arterial medial thickening. *Toxicol Appl Pharmacol*. 81(1):91-99.
36. Al-Naamani N, Sagliani KD, Dolnikowski GG, Warburton RR, Toksoz D, Kayyali U, Hill NS, Fanburg BL, Roberts KE, Preston IR. 2016. Plasma 12- and 15-hydroxyeicosanoids are predictors of survival in pulmonary arterial hypertension. *Pulm Circ*. 6(2):224-233.

Figure Legends:

Figure 1: Metabolites distinguishing between SSc-PAH and IPAH.

A. Study Flow chart

Summary of study workflow and data interpretation.

B. Volcano plot of metabolites distinguishing SSc-PAH from IPAH in the discovery and validation cohorts

D. Average metabolite levels in SSc-PAH and IPAH patients for 47 metabolites found to significantly distinguish patients with SSc-PAH from IPAH. Values plotted are log₂ fold change. Negative values indicate metabolites at lower levels in patients with SSc-PAH and positive values indicate metabolites at higher levels in patients with SSc-PAH. FC indicate fold change.

Figure 2: ROC curve

Receiver-operating characteristic curves showing the performance of the model in distinguishing IPAH from SSc-PAH using 12 metabolites. Blue curve represent the discovery cohort and the red curve represent the independent validation cohort (2nd validation cohort). AUC indicate area under the curve, and CI, confidence interval.

Figure 3: Forrest Plot for the 12-MRS metabolites and association with clinical variables

Tables:

Table 1. Patients characteristics

Demographics and clinical features of patients with IPAH and SSc-PAH. Mean±SD, counts or percentages are shown. BMI indicates body mass index; SMWD, six-minute walk distance; mRAP, mean right atrial pressure; PVR, pulmonary vascular resistance.

	Discovery cohort			1 st Validation cohort			2 nd Validation cohort		
	IPAH	SSc-PAH	<i>P</i>	IPAH	SSc-PAH	<i>P</i>	IPAH	SSc-PAH	<i>P</i>
N	864	310		213	91		32	15	
Female (%)	663 (76.7)	267 (86.1)	0.001	166 (77.9)	82 (90.1)	0.019	24 (75.0)	14 (93.3)	0.275
Age (mean (SD))	51.90 (18.47)	64.10 (11.03)	<0.001	53.22 (14.95)	63.74 (10.29)	<0.001	42.09 (17.92)	57.95 (12.72)	0.004
BMI (mean (SD))	30.36 (19.14)	28.24 (11.40)	0.068	30.78 (9.23)	27.32 (8.17)	0.002	30.97 (8.83)	27.08 (4.47)	0.175
Renal Insufficiency (%)	32 (3.7)	28 (9.0)	<0.001	11 (5.2)	7 (7.7)	0.555	NA	NA	
Cirrhosis (%)	13 (1.5)	5 (1.6)	1	2 (0.9)	4 (4.4)	0.125	NA	NA	
Functional Class (%)			0.093			0.51			0.46
I	46 (7.1)	14 (5.7)		5 (3.9)	0 (0.0)		NA	NA	
II	182 (28.3)	81 (33.1)		42 (32.6)	20 (36.4)		5 (16.7)	3 (21.4)	
III	345 (53.6)	135 (55.1)		72 (55.8)	31 (56.4)		22 (73.3)	11 (78.6)	
IV	71 (11.0)	15 (6.1)		10 (7.8)	4 (7.3)		3 (10.0)	0 (0.0)	
SMWD (mean (SD))	353.91 (138.36)	312.44 (118.06)	0.001	348.80 (125.44)	312.30 (130.54)	0.121	293.31 (129.27)	304.42 (78.73)	0.79
mRAP (mean (SD))	9.16 (5.85)	8.33 (5.06)	0.028	8.38 (4.99)	7.48 (4.94)	0.154	8.57 (6.06)	6.93 (4.57)	0.375
PVR (mean (SD))	10.74 (6.84)	8.50 (5.13)	<0.001	12.18 (6.38)	8.48 (4.06)	<0.001	55.76 (213.47)	9.06 (4.62)	0.405
Prostanoids use (%)	415 (48.0)	130 (41.9)	0.075	83 (39.0)	25 (27.5)	0.074	12 (37.5)	4 (26.7)	0.689

Table 2: Metabolites distinguishing SSc-PAH from IPAH and associated with markers of disease severity

Metabolite	Metabolic Pathway	Discovery			Validation		
		<i>P</i>	FC	OR	<i>P</i>	FC	OR
2-deoxy-d-glucose/ d-glucosamine	Amino sugar and nucleotide sugar metabolism	7.00E-07	1.5	1.33	2.00E-02	1.6	1.31
N-Acetylneuraminate	Amino sugar and nucleotide sugar metabolism	2.00E-09	1.4	1.57	1.00E-03	1.34	1.66
17(18) EpETE	Arachidonic acid metabolism	4.00E-05	0.7	0.89	1.00E-02	2.5	0.88
Novel Eic	Arachidonic acid metabolism	1.00E-08	1.5	1.13	3.00E-03	2.2	1.13
Novel Eic 2	Arachidonic acid metabolism	1.00E-06	1.1	1.35	1.00E-03	1.8	1.55
PGE1	Arachidonic acid metabolism	2.00E-05	0.9	0.8	7.00E-03	1	0.84
PGF2a	Arachidonic acid metabolism	4.00E-07	1.3	1.14	2.00E-02	1.3	1.11
N-Acetylputrescine	Arginine and proline metabolism	1.00E-12	2	1.69	3.00E-04	2.9	1.7
4-Acetamidobutanoate	Arginine and proline metabolism	2.00E-14	2.32	1.72	3.00E-05	1.86	2.6
Acetylserine	Cysteine and methionine metabolism	5.00E-06	1.3	1.4	8.00E-03	1.26	1.51
Cystathionine	Cysteine and methionine metabolism	4.00E-05	1.41	1.28	3.00E-02	1.32	1.32
FAHFA	Fatty Acid metabolism	8.00E-08	1.3	1.5	8.00E-03	1.3	1.54
Lignoceric Acid	Fatty Acid metabolism	1.00E-07	1.3	1.5	3.00E-03	1.2	1.57
Nervonic Acid	Fatty Acid metabolism	4.00E-07	1.12	1.5	2.00E-03	1.16	1.58
Tagatose	Galactose metabolism	4.00E-05	1.51	1.38	3.00E-03	1.58	1.63
4-Imidazoleacetate	Histidine metabolism	6.00E-05	1.25	1.35	3.00E-03	1.3	1.58
2-Aminoisobutyrate	Leucine, Isoleucine and Valine Metabolism	1.00E-05	1.22	1.37	3.00E-02	1.25	1.38
Cytosine	Pyrimidine metabolism	8.00E-05	1.4	1.2	1.00E-03	1.64	1.32
Uracil 5-carboxylate	Pyrimidine metabolism	6.00E-12	1.4	1.62	4.00E-04	1.45	1.62
17α-testosterone	Steroid hormones metabolism	1.00E-05	0.8	0.87	1.00E-03	0.7	0.75
17b-Estradiol	Steroid hormones metabolism	1.0E-09	2	1.13	3.00E-03	1.8	1.12
Kynurenine	Tryptophan metabolism	2.00E-13	1.3	1.7	2.00E-03	1.2	1.57
Quinolate	Tryptophan metabolism	1.00E-05	1.55	1.38	2.00E-02	1.2	1.41
Tryptophan	Tryptophan metabolism	1.00E-09	0.8	0.68	4.00E-03	0.8	0.66
Xanthurenate	Tryptophan metabolism	3.00E-06	0.8	0.77	2.00E-02	0.7	0.78
5-Hydroxyindoleacetate	Tryptophan metabolism	2.00E-13	1.7	1.34	6.00E-03	1.27	1.4
N-Acetylasparagine	Alanine, aspartate and glutamate metabolism	7.00E-06	1.5	1.4	9.00E-03	1.4	1.4
Deoxycarnitine	Fatty Acid metabolism	2.00E-05	1.43	1.3	1.00E-02	1.8	1.4
Monoethylmalonate	Fatty Acid metabolism	4.00E-08	1.63	1.4	2.00E-05	1.93	2
Nitrooleate	Fatty Acid metabolism	2.00E-05	0.6	0.9	2.00E-02	0.7	0.9

Figures:

Figure 1.

Figure 2.

Figure 3.