

Transcriptomic profiles in pulmonary arterial hypertension associate with disease severity and identify novel candidate genes

Casey E. Romanoski^{1,*}, Xinshuai Qi^{2,*}, Shreya Sangam^{2,3}, Rebecca R. Vanderpool², Robert S. Stearman³, Austin Conklin¹, Manuel Gonzalez-Garay², Franz Rischard², Ramon J. Ayon², Jian Wang^{2,4,5}, Tatum Simonson⁵, Aleksandra Babicheva⁵, Yinan Shi³, Haiyang Tang^{2,4}, Ayako Makino^{2,5,6}, Yogendra Kanthi^{7,8}, Mark W. Geraci³, Joe G.N. Garcia^{2,6}, Jason X.-J. Yuan^{2,5,6,#} and Ankit A. Desai^{2,3,#}

¹Department of Cellular and Molecular Medicine, College of Medicine, The University of Arizona, Tucson, AZ, USA; ²Department of Medicine, College of Medicine, The University of Arizona, Tucson, AZ, USA; ³Department of Medicine, Indiana University, Indianapolis, IN, USA; ⁴State Key Laboratory of Respiratory Diseases, Guangzhou Institute of Respiratory Disease, Guangzhou Medical University, Guangzhou, China; ⁵Department of Medicine, University of California, San Diego, La Jolla, CA, USA; ⁶Department of Physiology, College of Medicine, The University of Arizona, Tucson, AZ, USA; ⁷Division of Intramural Research National Heart, Lung and Blood Institute Bethesda, Maryland, USA; ⁸Division of Cardiovascular Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA

Abstract

Using RNAseq, we identified a 61 gene-based circulating transcriptomic profile most correlated with four indices of pulmonary arterial hypertension severity. In an independent dataset, 13/61 (21%) genes were differentially expressed in lung tissues of pulmonary arterial hypertension cases versus controls, highlighting potentially novel candidate genes involved in pulmonary arterial hypertension development.

Keywords

pulmonary arterial hypertension, transcriptome profiling, PBMC, RNAseq

Date received: 16 September 2020; accepted: 3 October 2020

Pulmonary Circulation 2020; 10(4) 1–5

DOI: 10.1177/2045894020968531

Introduction

Group 1 pulmonary arterial hypertension (PAH) is rare and often fatal disease characterized by excessive vasoconstriction and obliterative vascular remodeling contributing to elevated pulmonary vascular resistance (PVR), right ventricular failure, and death.^{1,2} While several studies reported individual genes/pathways associated with severity or disease risk,^{3–5} mechanisms remain unclear. Previously, expression profiling of peripheral blood mononuclear cells (PBMC) isolated from patients with PAH have searched for biomarkers of disease risk, PH Group classification, and have led to insights of disease mechanisms.^{6,7} However, most of these profiles have not been replicated.

Explanations for this shortcoming is likely due, in part, to the dynamic changes in expression profiles over time, profiling in different tissues (e.g., circulating tissue versus target vascular tissue), as well as the quality and heterogeneity of clinical phenotypes. Given the ease of access and the role of inflammation/immunity in PAH, studying PBMC

*Contributed equally as first authors.

#Contributed equally as senior authors.

Corresponding author:

Ankit A. Desai, Indiana University, Krannert Institute of Cardiology, Department of Medicine, 950 W. Walnut Street, R3 Building, Room 408, Indianapolis, IN 46202, USA.

Email: ankdesai@iu.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

© The Author(s) 2020.
Article reuse guidelines:
sagepub.com/journals-permissions
journals.sagepub.com/home/pul



expression profiles remains an appealing method for developing biomarkers for disease progression and prioritizing candidate genes in PAH. In the current study, we developed a single multi-gene classifier based on RNA-seq-based expression profiles from PBMCs derived from patients that associate with four disease severity markers. To gain insight into disease pathology, we next evaluated differential regulation of the gene classifier in target lung tissues available from both PAH patients and non-PAH controls. Taken together, these data represent both possible biomarkers for disease severity as well as robust candidate genes potentially involved with PAH development.

Materials and methods

University of Arizona (UA) cohort

PAH patients receiving care at our University Pulmonary Hypertension clinic between 2012 and 2015 were prospectively recruited in accordance with institutional guidelines and provided informed consent. The cohort was composed of 84 subjects with Group 1 PAH (43 associated, 32 idiopathic, 4 anorexigen/drug-associated, 5 congenital-associated or HIV-associated PAH). For each subject, demographics and multiple measures of clinical severity were collected from four separate clinical tests including right heart catheterization (RHC), transthoracic echocardiogram (TTE), six-minute walk distance (6MWD, m), and brain natriuretic peptide levels (BNP, ng/L), all acquired during their first visit (or as part of their initial collections). Clinical severity was represented by four quantitative traits representing PVR (WU), tricuspid annular plane systolic excursion (TAPSE, cm), BNP, and 6MWD.

RNA processing in UA cohort

PBMCs were stored in RNAtr as described.⁸ In total, approximately 3600 million clusters with paired-end 75 bp reads (~35M cluster per sample) were generated from PBMC-derived RNA. Spearman correlations were performed and adjusted for age, sex, ethnicity, and treatment status (on PAH-specific medication). Only transcripts with $r > |0.2|$ and p -value < 0.05 thresholds with any one of the four clinical variables (TAPSE, BNP, PVR, and 6MWD) were selected for annotation using Ensembl (GRCh38). Transcripts that met this threshold across at least three of four clinical severity traits were further evaluated in the PHBI cohort.

Pulmonary hypertension breakthrough initiative (PHBI) cohort

The PHBI cohort was comprised of 83 lung samples (58 PAH patients/25 control lung failed donors (FD)). Details of these patients have been recently reported⁹ and deposited at NCBI/GEO as accession GSE117261. Patient profiles

including their clinical data are all previously described.¹⁰ The PHBI cohort was analyzed by Affymetrix GeneChip Human Gene 1.0 ST microarray.⁹ Using .cel files, an ANOVA model was used to identify differentially regulated genes between the PAH and FD lung transcriptomes after correcting for sex imbalance in PAH disease population and batch effects (false discovery rate (FDR) q -value < 0.001 , yielding 1140 transcripts).

Results

A total of 84 Group 1 PAH patients were included in the UA cohort (Fig. 1a). The mean age was 59 years and the cohort was composed of 75% women, consistent with the sexual dimorphism of PAH. All four independent clinical measures (PVR, BNP, TAPSE, and 6MWD) were characterized in the cohort to provide multiple quantitative diagnostic measures that together bolster stratifications of PAH severity.

To identify genes whose expression is indicative of PAH severity, we first correlated the normalized expression values per individual with individual PAH traits. Using a Spearman correlation coefficient threshold of > 0.2 , 61 transcripts demonstrated correlation with at least three of the four PAH severity measures (Fig. 1b and c). Zinc finger protein 841 (*ZNF841*) RNA levels were uniquely correlated with all four PAH measured traits ($p < 0.05$). Lower expression levels of *ZNF841* from PBMCs correlated with higher PVR and BNP levels as well as lower TAPSE and 6MWD results. We then confirmed that *ZNF841* was down-regulated in human pulmonary artery endothelial cells (HPAECs, data not shown) isolated from patients with PAH versus those from controls. In combination with PBMC data, these data suggest *ZNF841* deficiency may have a causal role in PAH risk and development.

In an independent cohort, we tested whether any of the 61 genes identified were differentially expressed in whole lung tissues between those collected from patients with PAH and those from non-PAH controls. Beyond discovery of putative genomic biomarkers of PAH severity, the PHBI cohort prioritized gene sets that may be involved in the development or in advanced disease. Based on available microarray data, we identified that 13 of the 61 gene set were differentially expressed (Fig. 1d; FDR $< 10\%$).

We also evaluated association of gene expression levels of individual transcripts of the top 13 genes filtered from the Replication cohort against two available measures of severity—PVR and a quantitative measure of percent lung inflammation. Specifically, *CLASP2* (cytoplasmic linker associated protein 2) and *CSNK1E* (casein kinase 1 epsilon) displayed significant correlation and consistent directionality as the Discovery cohort with PVR (Fig. 1e and f, $r > 0.2$, $p < 0.05$). Given their reproducibility across two cohorts from two tissue sources, these data suggest that *CLASP2* and *CSNK1E* may represent novel candidate genes that modify PAH severity and possibly involved in inflammatory signaling in PAH lungs.

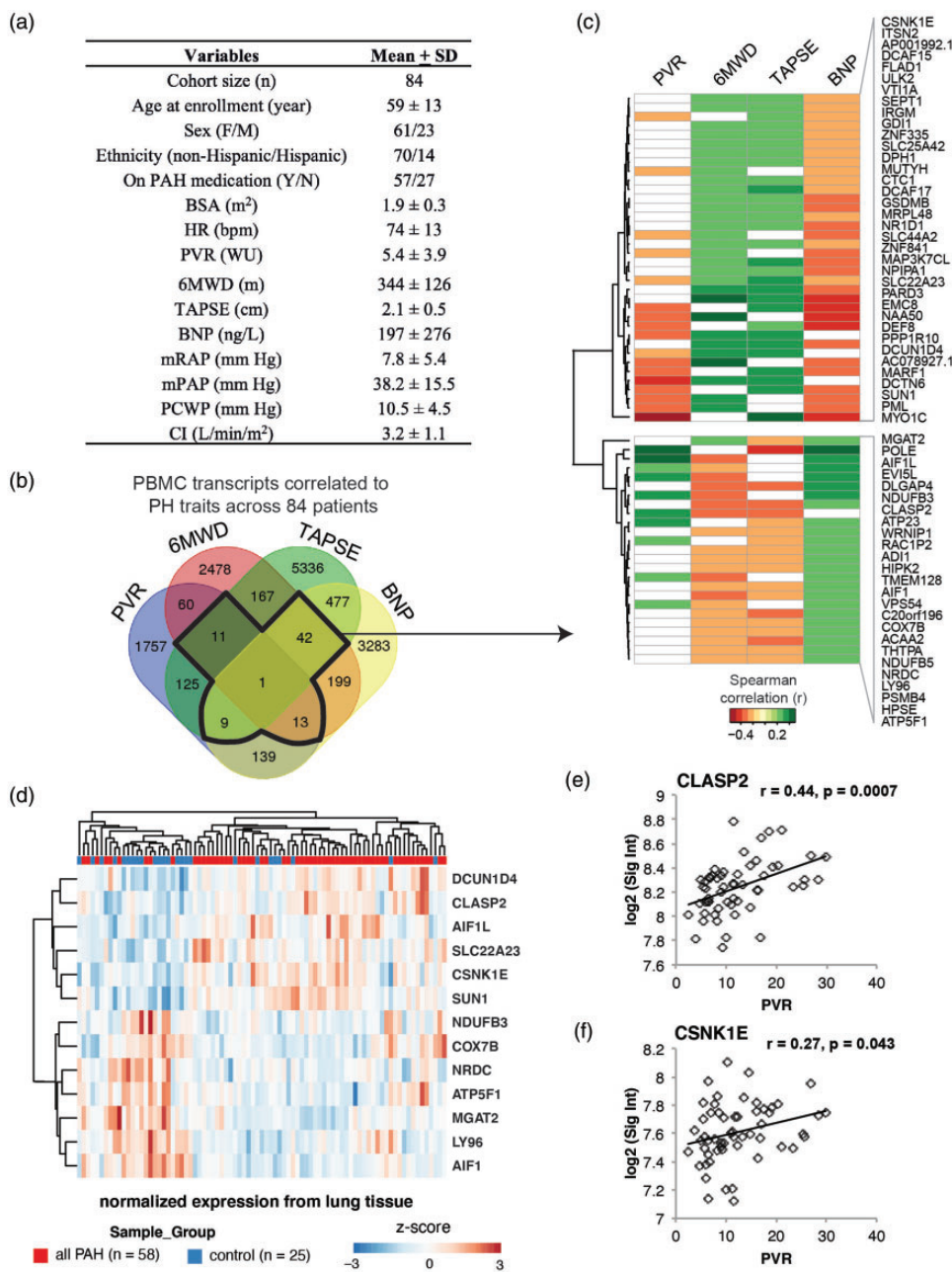


Figure 1. 61-gene signature associated with PAH severity and risk. (a) Table depicts demographic and clinical characteristics of the UA cohort. (b) Venn diagram for the numbers of transcripts shared among four correlation analyses with correlation coefficient $r \geq 0.2$ ($p < 0.05$). A single transcript (*ZNF841*) shows significant correlation with all four clinical parameters. (c) The heat map depicts correlation coefficients for the 61 genes correlated with PVR, BNP, TAPSE, or 6MWD ($r > 0.2$, $p < 0.05$). (d) The heatmap depicts 13 of the 61 genes from the original classifier, which is significantly differentially regulated between lungs of patients with PAH ($n = 58$) versus non-PAH controls ($n = 25$). (e and f) Two of the 13 filtered genes in lungs were also significantly correlated with PVR in the Replication cohort (*CLASP2*, *CSNK1E*). BSA: body surface area; HR: heart rate; PVR: pulmonary vascular resistance; WU: wood units; 6MWD: six-minute walk distance; TAPSE: tricuspid annular plane systolic excursion; BNP: brain natriuretic peptide; mRAP: mean right atrial pressure; mPAP: mean pulmonary artery pressure; PCWP: pulmonary capillary wedge pressure; CI: cardiac index; PA saturation: pulmonary artery saturation. Descriptive statistics are presented as mean \pm SD.

Discussion

We report novel RNAseq profiles from PBMCs that associate with PAH disease severity. For the first time, we integrate expression profiles associated with multi-modality

assessment of severity to generate a single multi-gene classifier. In an independent testing cohort, a subset (~21%) of these genes consistently mirrored expression profiles in target lung tissues from patients with and without PAH

with moderate correlation to disease severity. This latter observation likely reflects the heterogeneity of whole lung tissue samples and provides initial perspectives into the ability of the circulating PBMC transcriptome to provide insights into the expression of disease-causing candidate genes in target tissues. Replication in larger cohorts will validate the robustness of potentially causal biomarkers, highlighting promising targets for novel therapies.

Our analysis also highlights *ZNF841*, *CLASP2*, and *CSNK1E* as some of the top genes associated with selected markers of disease severity. These genes were not observed to be differentially regulated in a recent meta-analysis¹¹ of circulating PBMCs from patients and non-PAH controls. Part of this likely stems from differences in the utility of microarray (used in the meta-analysis) versus RNAseq (current work) technologies as well as the focus of their analyses on disease risk versus disease severity in the current work. Furthermore, while the function *ZNF841* is unknown, other zinc fingers are reported to play a role in lung diseases such as asthma.¹² Zinc fingers are structurally diverse and are present among proteins that perform a broad range of functions in various cellular processes, such as replication and repair, metabolism, cell proliferation, and apoptosis.¹³ *CLASP2* belongs to a family of microtubule plus-end tracking proteins that localizes to the distal ends of microtubules and regulates microtubule dynamics.¹⁴ *CLASP2* functions in various microtubule-dependent processes, including cell division, cytoskeletal remodeling for cell migration, podosome regulation, and stabilization of adherens junctions.^{15,16} *CLASP2* is also involved in epithelial-mesenchymal transition and progression in cancer.¹⁷ And finally, *CSNK1* is a group of ubiquitous serine/threonine kinases that are involved in normal cellular functions and several pathological conditions, such as DNA repair, cell cycle progression, and apoptosis.¹⁸ Recent studies indicate that *CSNK1E* expression has a role in human cancers.^{19,20} Based on these reports, all three genes are involved in hyperproliferative cellular processes, which is highly relevant to vascular remodeling observed in PAH, and may represent novel biomarkers of PAH severity.

Limitations of this study include a small sample size reflecting low severity of disease and heterogeneity in PAH subtypes. We do not have the power to evaluate differences in patients with advanced PAH or in homogenous cases such as idiopathic PAH alone due to incrementally smaller sample sizes during subgroup analysis. We also present a cross-sectional view of gene expression patterns and not serial patterns over time, which may better reflect progression of disease. While there are many phenotypes in PAH that provide a measure of severity, we chose one representative and well-studied variable from each diagnostic test (i.e., PVR from RHC) for association analyses, limiting interpretation to other markers of severity.

In summary, the current study demonstrates the utility of PBMC transcriptomes to identify molecular classifiers of

disease risk and severity in PAH. They further highlight how PBMCs may shed insights into both target tissue expression profiles as well as possible roles of inflammation in the molecular pathogenesis of PAH.

Scientific Knowledge on the Subject

Pulmonary arterial hypertension is a complex disease with multiple molecular mediators that contribute to disease predilection and severity. While a number of studies have reported individual genes/pathways associated with severity or disease risk, mechanisms remain unclear.

What This Study Adds to the Field

This study has identified novel transcripts that are potential valuable biomarkers of both pulmonary arterial hypertension severity assessed via multi-modality testing. Importantly, a subset of these gene expression profiles may play a pathological role in disease development.

Author Contributions

CER: contributed to the concept of this study, data collection, drafting of the manuscript, and data analysis; XQ: contributed to the concept of this study, data collection, drafting of the manuscript, and data analysis; SS: contributed to the data collection and drafting of the manuscript; RRV: contributed to the data collection; RSS: contributed to the data collection and analysis and drafting of the manuscript; AC: contributed to the data analysis; MGG: contributed to the data analysis; FR: contributed to the data collection; RJA: contributed to the data collection; JW: contributed to the drafting of the manuscript; TS: contributed to the drafting of the manuscript; AB: contributed to the data collection and analysis; YS: contributed to the data collection; HT: contributed to the data collection and analysis and drafting of the manuscript; AM: contributed to the data collection and drafting of the manuscript; YK: contributed to the drafting of the manuscript; MWG: contributed to the data collection and drafting of the manuscript; JGG: contributed to the data collection and drafting of the manuscript; JXJ: contributed to the design of this study, data collection, drafting of the manuscript, and data analysis; AAD: contributed to the concept and design of this study, data collection, drafting of the manuscript, and data analysis.

Conflict of interest

The author(s) declare that there is no conflict of interest.

Funding

This work was supported in part by the grants from the National Heart, Lung and Blood Institute of the National Institutes of Health [R01HL136603 (AAD), R35HL135807 (JY), U01HL125208 (FR/JY), R00HL123485 (CER), R01HL147187 (CER), R24HL123767 (MWG), K08HL131993 and R01HL150392 (YK)] the Intramural Research Program of the NIH and NHLBI (YK), NHLBI T32 HL007249-44 (AC), ARCS Scholar Award (Phoenix Chapter) (AC), ABRC/ADHS18-198871 (RRV) and funding from the Cardiovascular Medical Research

and Education Fund (MWG) and. We also want to acknowledge Dr. Zarema Arbieva and the Core Genomics Laboratory at the University of Illinois at Chicago for processing RNAseq samples.

References

- Desai AA and Machado RF. Diagnostic and therapeutic algorithm for pulmonary arterial hypertension. *Pulm Circ* 2011; 1: 122–124.
- McLaughlin VV, Archer SL, Badesch DB, 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. *Circulation* 2009; 119: 2250–2294.
- Makino A, Firth AL and Yuan JX. Endothelial and smooth muscle cell ion channels in pulmonary vasoconstriction and vascular remodeling. *Compr Physiol* 2011; 1: 1555–1602.
- Rhodes CJ, Batai K, Bleda M, et al. Genetic determinants of risk in pulmonary arterial hypertension: international case-control studies and meta-analysis. *Lancet Resp Med* 2018; 7: 227–238.
- Tang H, Wu K, Wang J, et al. Pathogenic role of mTORC1 and mTORC2 in pulmonary hypertension. *JACC Basic Transl Sci* 2018; 3: 744–762.
- Desai AA, Zhou T, Ahmad H, et al. A novel molecular signature for elevated tricuspid regurgitation velocity in sickle cell disease. *Am J Respir Crit Care Med* 2012; 186: 359–368.
- Grigoryev DN, Mathai SC, Fisher MR, et al. Identification of candidate genes in scleroderma-related pulmonary arterial hypertension. *Transl Res* 2008; 151: 197–207.
- Desai AA, Lei Z, Bahroos N, et al. Association of circulating transcriptomic profiles with mortality in sickle cell disease. *Blood* 2017; 129: 3009–3016.
- Stacher E, Graham BB, Hunt JM, et al. Modern age pathology of pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2012; 186: 261–272.
- Stearman RS, Bui QM, Speyer G, et al. Systems analysis of the human pulmonary arterial hypertension lung transcriptome. *Am J Respir Cell Mol Biol* 2019; 60: 637–649.
- Elinoff JM, Mazer AJ, Cai R, et al. Meta-analysis of blood genome-wide expression profiling studies in pulmonary arterial hypertension. *Am J Physiol Lung Cell Mol Physiol* 2020; 318: L98–L111.
- Wu AC, Himes BE, Lasky-Su J, et al. Inhaled corticosteroid treatment modulates ZNF432 gene variant's effect on bronchodilator response in asthmatics. *J Allergy Clin Immunol* 2014; 133: 723.e3–728.e3.
- Krishna SS, Majumdar I and Grishin NV. Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res* 2003; 31: 532–550.
- Lawrence EJ, Arpag G, Norris SR, et al. Human CLASP2 specifically regulates microtubule catastrophe and rescue. *Mol Biol Cell* 2018; 29: 1168–1177.
- Shahbazi MN, Megias D, Epifano C, et al. CLASP2 interacts with p120-catenin and governs microtubule dynamics at adherens junctions. *J Cell Biol* 2013; 203: 1043–1061.
- Efimova N, Grimaldi A, Bachmann A, et al. Podosome-regulating kinesin KIF1C translocates to the cell periphery in a CLASP-dependent manner. *J Cell Sci* 2014; 127: 5179–5188.
- Zhu B, Qi L, Liu S, et al. CLASP2 is involved in the EMT and early progression after transurethral resection of the bladder tumor. *BMC Cancer* 2017; 17(1): 105.
- Knippschild U, Krüger M, Richter J, et al. The CK1 family: contribution to cellular stress response and its role in carcinogenesis. *Front Oncol* 2014; 4: 96.
- Varghese RT, Young S, Pham L, et al. Casein kinase 1 epsilon regulates glioblastoma cell survival. *Sci Rep* 2018; 8: 13621.
- Babron MC, Kazma R, Gaborieau V, et al. Genetic variants in DNA repair pathways and risk of upper aerodigestive tract cancers: combined analysis of data from two genome-wide association studies in European populations. *Carcinogenesis* 2014; 35: 1523–1527.