Title
Integrative Multiomics and Systems Biology of Pulmonary Arterial Hypertension

Author
Hong, Jason

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Integrative Multiomics and Systems Biology of Pulmonary Arterial Hypertension

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular, Cellular, and Integrative Physiology

by

Jason Hong

2021

ABSTRACT OF THE DISSERTATION

Integrative Multiomics and Systems Biology of Pulmonary Arterial Hypertension

by

Jason Hong

Doctor of Philosophy in Molecular, Cellular, and Integrative Physiology

University of California, Los Angeles, 2021

Professor Xia Yang, Co-Chair

Professor Mansoureh Eghbali, Co-Chair

Pulmonary arterial hypertension (PAH) is a lung disease characterized by narrowing of the pulmonary arteries causing hemodynamic resistance which eventually leads to right heart failure and death. Current therapies mainly act through vasodilation but none reverse the underlying vascular remodeling characteristic of PAH. A deeper understanding of the molecular and cellular mechanisms of PAH is needed to bridge this translational gap. The goal of this dissertation is to investigate the transcriptional alterations in PAH lungs using integrative

multiomics to identify and prioritize candidate genes, pathways, and cell types implicated in PAH.

First, we identified reprogramming of genes and pathways in various cell types in the lungs of two commonly used rat models of PAH, namely Sugen-hypoxia (SuHx) and monocrotaline (MCT), using single-cell RNA sequencing (scRNAseq). We found that genes dysregulated in SuHx nonclassical monocytes were significantly enriched for PAH-associated genes and GWAS variants. We further identified candidate drugs predicted to reverse the dysregulated gene programs. This study revealed the distinct and shared reprogramming of genes and pathways in two commonly used PAH models for the first time at single-cell resolution and demonstrated their relevance to human PAH and utility for drug repositioning.

Next, we dissected the human PAH lung transcriptome at the tissue level using an innovative network and systems biology methods on a well-powered RNA sequencing dataset of human PAH lungs. We discovered many DEGs and pathways in human PAH lungs at the tissue level, and through integration with clinical data and PAH GWAS, our network analysis revealed co-expressed gene modules that are not only associated with PAH diagnosis and severity, but also risk of PAH implicating their causal role in PAH pathogenesis. Key driver analysis utilizing a comprehensive gene regulatory network of the human lung identified and prioritized candidate genes. Furthermore, we integrated the tissue-level networks with scRNAseq to uncover the specific cell types mediating the tissue-level gene programs.

Overall, this integrative multiomics and systems biology study revealed and prioritized the dysregulation of many genes, pathways, and cell types in the lungs of PAH animal models and patients, thereby opening new avenues for therapeutic targeting.

The dissertation of Jason Hong is approved.

Steven Bensinger

Kristina Bostrom

Song Li

Mansoureh Eghbali, Committee Co-Chair

Xia Yang, Committee Co-Chair

University of California, Los Angeles

2021

For my family, friends, and patients

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Chapter 2

ACKNOWLEDGEMENTS

First, I would like to thank my thesis advisors, Xia Yang and Mansoureh Eghbali for their outstanding mentorship throughout my PhD. They provided tremendous support and guidance without which I would not have been able to develop as a scientist. They gave me the freedom to pursue my research interests but also selflessly invested so much of their time and energy to ensure my academic success. I am forever grateful to them.

I would also like to thank my thesis committee, Steven Bensinger, Kristina Bostrom, and Song Li. I am very fortunate to have had such exceptional scientists guiding me not only in my research, but also my career as a junior physician-scientist. They provided such invaluable perspective and insights that have contributed immensely to my scientific development.

I am also indebted to Alan Fogelman, Linda Demer, Mitch Wong, Tamer Sallam, and the rest of the STAR leadership for giving me the opportunity to pursue a PhD through the STAR program at UCLA. Their guidance and support have been instrumental in my growth as a physician-scientist.

I am also grateful for the support of the pulmonary division at UCLA, especially Steven Dubinett and Tisha Wang. They showed faith by investing in me and giving me the protected time needed for my career development first as a fellow and now as a faculty member.

I would also like to thank my clinical mentors, Rajan Saggar and Richard Channick, who trained me in caring for patients with pulmonary arterial hypertension and welcomed me into the pulmonary vascular disease program.

I am also grateful to the Molecular, Cellular, and Integrative Physiology PhD program at UCLA, especially Mark Frye, Yesenia Rayos, and the outstanding faculty and students, for

giving me such an enriching experience in pursuing my PhD. I have also met some great peers and colleagues within the program, who have been encouraging and provided me with advice in one way or another.

I would also like to express my deepest gratitude to the current and former members in the Yang and Eghbali labs who made my experience in research at UCLA so fun and enjoyable. They are like family to me.

I am also grateful to our collaborators at PHBI, Bob Stearman and Mark Geraci, who allowed us to take part in analyzing their RNA sequencing data. I also thank David-Alexandre Trégouët and the French PAH group, Chris Rhodes and the UK PAH group as well as other stakeholders in the US for kindly providing the PAH GWAS data.

I also would like to thank all the patients and their families who trusted me with the privilege of taking care of them. They are my inspiration and motivation for pursuing research in pulmonary arterial hypertension.

Last but not least, I would like to thank my family for their steadfast support and encouragement throughout all these years of training.

Chapter 2 is reprinted with permission of the American Thoracic Society.

Cite: Hong, Jason, Douglas Arneson, Soban Umar, Gregoire Ruffenach, Christine M. Cunningham, In Sook Ahn, Graciel Diamante, May Bhetraratana, John F. Park, Emma Said, Caroline Huynh, Trixie Le, Lejla Medzikovic, Marc Humbert, Florent Soubrier, David Montani, Barbara Girerd, David-Alexandre Trégouët, Richard Channick, Rajan Saggar, Mansoureh Eghbali, Xia Yang. 2021. Single-cell Study of Two Rat Models of Pulmonary Arterial

Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning. 203(8):1006-1022.

The American Journal of Respiratory and Critical Care Medicine is an official journal of the American Thoracic Society.

This research was supported in part by UCLA Pulmonary and Critical Care Medicine T32 grant (T32HL072752), American Lung Association Catalyst Award (CA-675591), NIH Loan Repayment Program (L30 HL154170), and UCLA Specialty Training and Advanced Research (STAR) program.

VITA

Jason Hong

EDUCATION

2001 – 2005	University of California, Berkeley, Berkeley, California Bachelor of Science (Honors) – Bioengineering
2007 – 2012	Loma Linda University School of Medicine, Loma Linda, California Doctor of Medicine
2010 – 2011	UC Berkeley School of Public Health, Berkeley, California Master of Public Health – Epidemiology

EMPLOYMENT

2012 – 2015	Resident Physician (Internal Medicine), Department of Medicine Columbia University Medical Center; New York, NY
2014 – 2017	Fellow Physician, Division of Pulmonary & Critical Care Medicine UCLA Health System; Los Angeles, CA
2015 – 2016	Clinical Instructor, Department of Medicine University of Washington School of Medicine; Seattle, WA
2019 – current	Clinical Instructor, Division of Pulmonary & Critical Care Medicine David Geffen School of Medicine at University of California, Los Angeles

PUBLICATIONS

ORCID Bibliography: https://orcid.org/0000-0002-5062-8926
NCBI Bibliography: https://www.ncbi.nlm.nih.gov/myncbi/jason.hong.1/bibliography/public/

‡Corresponding author

1. Heath L, Conway S, Jones L, Semrau K, Nakamura K, Walter J, Decker W, **Hong J**, Chen T, Heil M, Sinkala M, Kankasa C, Thea D, Kuhn L, Mullins J, Aldrovandi G. Restriction of HIV-1 genotypes in breast milk does not account for the population transmission genetic bottleneck that occurs following transmission. *PLoS ONE*. 2010 Apr 20;5(4):e10213. https://doi.org/10.1371/journal.pone.0010213.

2. Ruffenach G, Umar S, Vaillancourt M, Cao N, **Hong J**, Sarji S, Moazeni S, Cunningham C, Ardehali A, Reddy ST, Saggar R, Fishbein G, Eghbali M. Histological hallmarks and role of Slug/PIP axis in pulmonary hypertension secondary to pulmonary fibrosis. *EMBO Mol Med***.** 2019 Sep;11(9):e10061. https://doi.org/10.15252/emmm.201810061.

3. Bhetraratana M, Orozco LD, **Hong J**, Diamante G, Majid S, Bennett BJ, Ahn IS, Yang X, Lusis AJ, Araujo JA. Diesel exhaust particles dysregulate multiple immunological pathways in murine macrophages: Lessons from microarray and scRNA-seq technologies. *Arch Biochem Biophys***.** 2019 Dec 15;678:108116. https://doi.org/10.1016/j.abb.2019.108116.

4. Zhang G, Byun HR, Ying Z, Zhao Y, **Hong J**, Shu L, Gomez-Pinilla F, Yang X. Differential Metabolic and Multi-tissue Transcriptomic Responses to Fructose Consumption among Genetically Diverse Mice. *Biochim Biophys Acta Mol Basis Dis***.** 2020 Jan 1;1866(1):165569. https://doi.org/10.1016/j.bbadis.2019.165569.

5. Umar S, Ruffenach G, Moazeni S, Vaillancourt M, **Hong J**, Cunningham CM, Cao N, Navab S, Sarji S, Lee L, Li M, Fishbein G, Ardehali A, Navab M, Reddy ST, Eghbali M. Involvement of Low Density Lipoprotein

Receptor in the Pathogenesis of Pulmonary Hypertension. *J Am Heart Assoc*. 2020 Jan 21;9(2):e012063. https://doi.org/10.1161/JAHA.119.012063.

6. Liu W, Venugopal S, Majid S, Ahn IS, Diamante G, **Hong J**, Yang X, Chandler SH. Single-cell RNA-seq analysis of the brainstem of mutant SOD1 mice reveals perturbed cell types and pathways of amyotrophic lateral sclerosis. *Neurobiol Dis***.** 2020 Jul;141:104877. https://doi.org/10.1016/j.nbd.2020.104877.

7. Medzikovic L, Cunningham CM, Li M, Amjedi M, **Hong J**, Ruffenach G, Eghbali M. Sex differences underlying preexisting cardiovascular disease and cardiovascular injury in COVID-19. *J Mol Cell Cardiol*. 2020 Aug 22;148:25–33. https://doi.org/10.1016/j.yjmcc.2020.08.007.

8. Ruffenach G, O'Connor E, Vaillancourt M, **Hong J**, Cao N, Sarji S, Moazeni S, Papesh J, Grijalva V, Cunningham CM, Shu L, Chattopadhyay A, Tiwari S, Mercier O, Perros F, Umar S, Yang X, Gomes AV, Fogelman AM, Reddy ST, Eghbali M. Oral 15-Hydroxyeicosatetraenoic Acid Induces Pulmonary Hypertension in Mice by Triggering T Cell-Dependent Endothelial Cell Apoptosis. *Hypertension*. 2020 Sep;76(3):985-996. https://doi.org/10.1161/hypertensionaha.120.14697.

9. Ruffenach G, **Hong J**, Vaillancourt M, Medzikovic L, Eghbali M. Pulmonary hypertension secondary to pulmonary fibrosis: clinical data, histopathology and molecular insights. *Respir Res*. 2020 Nov 18;21:303. https://doi.org/10.1186/s12931-020-01570-2.

10. **Hong J**‡, Arneson D, Umar S, Ruffenach G, Cunningham CM, Ahn IS, Diamante G, Bhetraratana M, Park JF, Said E, Huynh C, Le T, Medzikovic L, Humbert M, Soubrier F, Montani D, Girerd B, Trégouët D-A, Channick R, Saggar R, Eghbali M, Yang X. Single-cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning. *Am J Respir Crit Care Med*. 2021 Apr 15;203(8):1006-1022. https://doi.org/10.1164/rccm.202006-2169oc. Editorial: https://doi.org/10.1164/rccm.202010_-3889ED.

11. Park JF, Clark VR, Banerjee S, **Hong J**, Razee A, Williams T, Fishbein G, Saddic L, Umar S. Transcriptomic Analysis of Right Ventricular Remodeling in Two Rat Models of Pulmonary Hypertension: Identification and Validation of EMT in Human Right Ventricular Failure. *Circ Heart Fail*. 2021 Feb 5. https://doi.org/10.1161/circheartfailure.120.007058.

12. Saggar R, Giri PC, Chunqin D, Johnson D, McCloy MK, Lloyd L, Faisal S, **Hong J**, Channick RN, SS, Lynch III JP, Belperio JA., Weigt SS, Ramsey A, Ross DJ, Sayah D, Shino M, Derhovanessian A, Sherman A, Saggar R. Significance of Autoimmune Disease in Severe Pulmonary Hypertension Complicating Extensive Pulmonary Fibrosis: A Prospective Cohort Study. *Pulm Circ*. 2021 May 2;11(2):20458940211011329. https://doi.org/10.1177/20458940211011329.

PRESENTATIONS

1. **Hong J**, Arneson D, Ahn IS, Umar S, Diamente G, Bhetraratana M, Vaillancourt M, Cunningham C, Eghbali M, Yang X. Lung Single-Cell Transcriptomics of Two Animal Models of Pulmonary Arterial Hypertension Reveals Vulnerable Cell Types and Pathways. Pulmonary Vascular Research Institute (PVRI) Annual World Congress on Pulmonary Vascular Disease, Barcelona, Spain; 2019.

2. **Hong J**, Arneson D, Ahn IS, Umar S, Diamente G, Bhetraratana M, Cunningham C, Eghbali M, Yang X. Lung Single-Cell Transcriptomics of Two Animal Models of Pulmonary Arterial Hypertension Reveals Relevant Cell Types and Pathways. Am J Respir Crit Care Med 2019;199:A5867.

3. **Hong J**, Arneson D, Umar S, Ruffenach G, Ahn IS, Diamente G, Bhetraratana M, Cunningham C, Eghbali M, Yang X. Integrative single-cell omics to uncover mechanistic insights and drug repurposing candidates for pulmonary arterial hypertension. Pulmonary Vascular Research Institute (PVRI) Annual World Congress on Pulmonary Vascular Disease, Lima, Peru; 2020.

CHAPTER 1

Introduction

Pulmonary arterial hypertension (PAH) is an important clinical problem

PAH is characterized by narrowing of the pulmonary arteries causing hemodynamic resistance which eventually leads to right heart failure and death. Current therapies mainly act through vasodilation but none reverse the underlying vascular remodeling characteristic of PAH[1]. Therefore, PAH patients continue to suffer from a poor quality of life and face a grim long-term prognosis[2]. To make significant gains in improving quality of life and survival, new therapies targeting pathways central to disease pathogenesis are needed to counteract the underlying disease process. However, many novel drugs while effective in PAH models have failed to translate to the bedside[3]. A deeper understanding of the molecular and cellular mechanisms of PAH is needed to bridge this translational gap.

Current state and gaps in PAH omics

Over the past decade, data-driven multiomics have emerged as powerful tools to uncover novel genes and pathways involved in PAH pathogenesis. Transcriptomic studies of PAH lungs have uncovered genes and pathways differentially expressed in PAH[4,5]. However, whether such findings are robust and causal in disease pathogenesis remain unknown since lung samples are usually from limited numbers of advanced stage PAH patients and experimental follow-up of findings is often lacking. Genetic studies including genome-wide association studies (GWAS) have identified only a limited number of causal genes in PAH leaving greater than 80% of cases genetically unexplained[6–9]. In addition to causal genes, the culprit cell types are also critically important to understand the molecular underpinnings of PAH. Besides pulmonary vascular cells, many other lung cell types including various immune cells have been implicated in PAH by prior studies[10–14]. A human PAH scRNAseq study was also recently published[15] but was limited to

three diseased lungs and a focused analysis of a select few cell types. A more comprehensive and integrated examination of the altered genes, pathways, and cell types in PAH lungs is needed to advance our understanding of PAH mechanisms and help inform translational research efforts.

The dissertation

This dissertation investigates the transcriptional alterations in the lungs of PAH animal models and patients using integrative multiomics to identify and prioritize candidate genes, pathways, and cell types implicated in PAH pathobiology.

Chapter 2 is a reprint of "Single-Cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning", which was originally published in *Am J Respir Crit Care Med*. In this study, we used single-cell RNA sequencing (scRNAseq) to determine and prioritize dysregulated genes, pathways, and cell types in lungs of PAH rat models to assess relevance to human PAH and identify drug repositioning candidates. We identified distinct changes in genes and pathways in numerous cell types in Sugen-hypoxia (SuHx) and monocrotaline (MCT) lungs. We also found that genes altered in SuHx nonclassical monocytes were significantly enriched for PAH-associated genes and genetic variants, and candidate drugs predicted to reverse the changes were identified. Our study revealed the distinct and shared dysregulation of genes and pathways in two commonly used PAH models for the first time at single-cell resolution and demonstrated their relevance to human PAH and utility for drug repositioning.

Chapter 3 dissects the PAH lung transcriptional landscape at the tissue level using a large RNA sequencing (RNAseq) dataset of human PAH and control lungs. We identified many differentially expressed genes and pathways at the tissue level in human PAH lungs. Through

integration with clinical data and PAH GWAS, our network analysis revealed modules of co-expressed genes that are not only associated with PAH, but likely causal for disease severity and disease risk. Furthermore, we integrated the tissue-level networks with scRNAseq to uncover the specific cell types mediating the tissue-level gene programs. Thus, our findings implicate novel genes, pathways, and cell types in PAH pathobiology.

Chapter 4 is a concluding summary of the dissertation and covers future directions for integrative multiomics in PAH.

References

1. Harvey LD, Chan SY. Emerging Metabolic Therapies in Pulmonary Arterial Hypertension. *J Clin Med* 2017;6:.

2. Benza RL, Miller DP, Barst RJ, Badesch DB, Frost AE, McGoon MD. An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry. *Chest* 2012;142:448–456.

3. Bonnet S, Provencher S, Guignabert C, Perros F, Boucherat O, Schermuly RT, Hassoun PM, Rabinovitch M, Nicolls MR, Humbert M. Translating Research into Improved Patient Care in Pulmonary Arterial Hypertension. *American Journal of Respiratory and Critical Care Medicine* 2017;195:583–595.

4. Hoffmann J, Wilhelm J, Olschewski A, Kwapiszewska G. Microarray analysis in pulmonary hypertension. *European Respiratory Journal* 2016;48:229–241.

5. Stearman RS, Bui QM, Speyer G, Handen A, Cornelius AR, Graham BB, Kim S, Mickler EA, Tuder RM, Chan SY, Geraci MW. Systems Analysis of the Human Pulmonary Arterial Hypertension Lung Transcriptome. *Am J Respir Cell Mol Biol* 2019;60:637–649.

6. Rhodes CJ, Batai K, Bleda M, Haimel M, Southgate L, Germain M, Pauciulo MW, Hadinnapola C, Aman J, Girerd B, Arora A, Knight J, Hanscombe KB, Karnes JH, Kaakinen M, Gall H, Ulrich A, Harbaum L, Cebola I, Ferrer J, Lutz K, Swietlik EM, Ahmad F, Amouyel P, Archer SL, Argula R, Austin ED, Badesch D, Bakshi S, *et al.* Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir Med* 2019;7:227–238.

7. Germain M, Eyries M, Montani D, Poirier O, Girerd B, Dorfmüller P, Coulet F, Nadaud S, Maugenre S, Guignabert C, Carpentier W, Vonk-Noordegraaf A, Lévy M, Chaouat A, Lambert J-C, Bertrand M, Dupuy A-M, Letenneur L, Lathrop M, Amouyel P, Ravel TJL de, Delcroix M, Austin ED, Robbins IM, Hemnes AR, Loyd JE, Berman-Rosenzweig E, Barst RJ, Chung WK, *et al.* Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. *Nature Genetics* 2013;45:518–521.

8. Zhu N, Pauciulo MW, Welch CL, Lutz KA, Coleman AW, Gonzaga-Jauregui C, Wang J, Grimes JM, Martin LJ, He H, Hirsch R, White RJ, Simon M, Badesch D, Rosenzweig E, Burger C, Chakinala M, Thenappan T, Elliott G, Simms R, Farber H, Frantz R, Elwing J, Hill N, Ivy D, Klinger J, Nathan S, Oudiz R, Robbins I, *et al.* Novel risk genes and mechanisms implicated by exome sequencing of 2572 individuals with pulmonary arterial hypertension. *Genome Medicine* 2019;11:69.

9. Gräf S, Haimel M, Bleda M, Hadinnapola C, Southgate L, Li W, Hodgson J, Liu B, Salmon RM, Southwood M, Machado RD, Martin JM, Treacy CM, Yates K, Daugherty LC, Shamardina O, Whitehorn D, Holden S, Aldred M, Bogaard HJ, Church C, Coghlan G, Condliffe R, Corris PA, Danesino C, Eyries M, Gall H, Ghio S, Ghofrani H-A, *et al.* Identification of rare

sequence variation underlying heritable pulmonary arterial hypertension. *Nature Communications* 2018;9:.

10. Hong J, Arneson D, Umar S, Ruffenach G, Cunningham CM, Ahn IS, Diamante G, Bhetraratana M, Park JF, Said E, Huynh C, Le T, Medzikovic L, Humbert M, Soubrier F, Montani D, Girerd B, Trégouët D-A, Channick R, Saggar R, Eghbali M, Yang X. Single-cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning. *Am J Respir Crit Care Med* 2020;doi:10.1164/rccm.202006-2169OC.

11. Huertas A, Phan C, Bordenave J, Tu L, Thuillet R, Le Hiress M, Avouac J, Tamura Y, Allanore Y, Jovan R, Sitbon O, Guignabert C, Humbert M. Regulatory T Cell Dysfunction in Idiopathic, Heritable and Connective Tissue-Associated Pulmonary Arterial Hypertension. *Chest* 2016;149:1482–1493.

12. Perros F, Dorfmüller P, Souza R, Durand-Gasselin I, Mussot S, Mazmanian M, Hervé P, Emilie D, Simonneau G, Humbert M. Dendritic cell recruitment in lesions of human and experimental pulmonary hypertension. *Eur Respir J* 2007;29:462–468.

13. Florentin J, Coppin E, Vasamsetti SB, Zhao J, Tai Y-Y, Tang Y, Zhang Y, Watson A, Sembrat J, Rojas M, Vargas SO, Chan SY, Dutta P. Inflammatory Macrophage Expansion in Pulmonary Hypertension Depends upon Mobilization of Blood-Borne Monocytes. *The Journal of Immunology* 2018;200:3612–3625.

14. Hoffmann J, Yin J, Kukucka M, Yin N, Saarikko I, Sterner-Kock A, Fujii H, Leong-Poi H, Kuppe H, Schermuly RT, Kuebler WM. Mast cells promote lung vascular remodelling in pulmonary hypertension. *European Respiratory Journal* 2011;37:1400–1410.

15. Saygin D, Tabib T, Bittar HET, Valenzi E, Sembrat J, Chan SY, Rojas M, Lafyatis R. Transcriptional profiling of lung cell populations in idiopathic pulmonary arterial hypertension. *Pulm Circ* 2020;10:.

CHAPTER 2

Single-cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning

ORIGINAL ARTICLE

Single-Cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning

Jason Hong[1], Douglas Arneson[2], Soban Umar[3], Gregoire Ruffenach[3], Christine M. Cunningham[3], In Sook Ahn[2], Graciel Diamante[2], May Bhetraratana[4], John F. Park[3], Emma Said[3], Caroline Huynh[2], Trixie Le[3], Lejla Medzikovic[3], Marc Humbert[5], Florent Soubrier[6], David Montani[5], Barbara Girerd[5], David-Alexandre Trégouët[7], Richard Channick[1], Rajan Saggar[1], Mansoureh Eghbali[3]*, and Xia Yang[2]*

[1]Division of Pulmonary and Critical Care Medicine, [2]Department of Integrative Biology and Physiology, [3]Department of Anesthesiology and Perioperative Medicine, and [4]Division of Cardiology, University of California Los Angeles, Los Angeles, California; [5]Department of Respiratory and Intensive Care Medicine, Bicêtre Hospital, University of Paris-Saclay, National Institute of Health and Medical Research Joint Research Unit S 999, Public Assistance Hospitals of Paris, Le Kremlin-Bicêtre, France; [6]Institut Hospitalo–Universitaire Cardiométabolisme et Nutrition, Paris, France; and [7]Bordeaux Population Health Research Center, University of Bordeaux, National Institute of Health and Medical Research Joint Research Unit 1219, Bordeaux, France

ORCID ID: 0000-0001-8036-3079 (S.U.).

Abstract

Rationale: The cellular and molecular landscape and translational value of commonly used models of pulmonary arterial hypertension (PAH) are poorly understood. Single-cell transcriptomics can enhance molecular understanding of preclinical models and facilitate their rational use and interpretation.

Objectives: To determine and prioritize dysregulated genes, pathways, and cell types in lungs of PAH rat models to assess relevance to human PAH and identify drug repositioning candidates.

Methods: Single-cell RNA sequencing was performed on the lungs of monocrotaline (MCT), Sugen-hypoxia (SuHx), and control rats to identify altered genes and cell types, followed by validation using flow-sorted cells, RNA *in situ* hybridization, and immunofluorescence. Relevance to human PAH was assessed by histology of lungs from patients and via integration with human PAH genetic loci and known disease genes. Candidate drugs were predicted using Connectivity Map.

Measurements and Main Results: Distinct changes in genes and pathways in numerous cell types were identified in SuHx and MCT lungs. Widespread upregulation of NF-κB signaling and downregulation of IFN signaling was observed across cell types. SuHx nonclassical monocytes and MCT conventional dendritic cells showed particularly strong NF-κB pathway activation. Genes altered in SuHx nonclassical monocytes were significantly enriched for PAH-associated genes and genetic variants, and candidate drugs predicted to reverse the changes were identified. An open-access online platform was developed to share single-cell data and drug candidates (http://mergeomics.research.idre.ucla.edu/PVDSingleCell/).

Conclusions: Our study revealed the distinct and shared dysregulation of genes and pathways in two commonly used PAH models for the first time at single-cell resolution and demonstrated their relevance to human PAH and utility for drug repositioning.

Keywords: pulmonary hypertension; single-cell RNA sequencing; drug repurposing; monocrotaline; Sugen-hypoxia

(Received in original form June 7, 2020; accepted in final form October 6, 2020)

*Co–senior authors.

Supported by NIH/NHLBI grant T32 HL072752 (J.H.), an American Lung Association Catalyst Award (J.H.), NIH/NHLBI grant R01HL147883 (X.Y.), and NIH/NHLBI grants R01HL129051 and R01HL147586 (M.E.).

Author Contributions: J.H., D.A., M.E., and X.Y. contributed to the conception and design of the research and interpretation of the data. J.H. conducted or helped with all experiments, analyzed the data, made the figures, and wrote the manuscript. S.U. performed the animal experiments. G.R. and C.M.C. contributed to staining. G.R. contributed to the bulk RNA sequencing experiment. I.S.A. and G.D. performed the single-cell experiments. C.M.C. and M.B. contributed to lung dissociation experiments. C.H. analyzed drug data. C.H. and E.S. analyzed images. T.L. and E.S. prepared lung sections for staining. J.F.P. analyzed echo data. M.H., F.S., D.M., B.G., and D.-A.T. provided the genome-wide association study data. X.Y., M.E., L.M., R.C., and R.S. provided intellectual input.

Correspondence and requests for reprints should be addressed to Jason Hong, M.D., Division of Pulmonary and Critical Care Medicine, David Geffen School of Medicine at University of California Los Angeles, 200 UCLA Medical Plaza, Suite 365-B, Box 951693, Los Angeles, CA 90095. E-mail: jasonhong@mednet.ucla.edu.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Am J Respir Crit Care Med Vol 203, Iss 8, pp 1006–1022, Apr 15, 2021

Originally Published in Press as DOI: 10.1164/rccm.202006-2169OC on October 6, 2020
Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: The cellular and molecular landscape and translational value of commonly used models of pulmonary arterial hypertension (PAH) are poorly understood. Single-cell transcriptomics can enhance molecular understanding of preclinical models and facilitate their rational use and interpretation.

What This Study Adds to the Field: Our study revealed the distinct and shared dysregulation of genes and pathways in two commonly used PAH models for the first time at single-cell resolution and demonstrated their relevance to human PAH and utility for drug repositioning.

Despite advances in the management of pulmonary arterial hypertension (PAH), it remains an incurable and progressive disease characterized by severe pulmonary vascular remodeling, poor quality of life, and guarded long-term prognosis (1). Notably, current therapies focus on relieving symptoms and do not reverse vascular remodeling, the key pathological feature of PAH. The lack of therapies targeting underlying mechanisms in PAH may be in part because of our limited understanding of the pathogenic cell types and their specific molecular pathways. It has been increasingly recognized that in addition to pulmonary vascular cells, other cell types in the lung, including various immune-cell populations, may also play an important role in PAH and other pulmonary diseases (2–7). However, to our knowledge, a comprehensive evaluation to systematically compare these various cell types has not been undertaken in the most widely used preclinical models of PAH, namely the monocrotaline (MCT) and Sugen-hypoxia (SuHx) rat models. Given that numerous novel therapies tested in PAH animal models have not translated to the bedside (8), a more comprehensive understanding of the cellular and molecular landscape of these models is needed to unravel mechanistic insights and enhance the ability of preclinical work to predict drug efficacy in humans (9).

In this study, we performed lung single-cell RNA expression (scRNA-seq) in MCT and SuHx rats to investigate altered cell types, genes, and pathways and further integrated the findings with human genetics to assess human relevance. We then identified potential drug-repurposing candidates through computational screening of drug transcriptional profiles against the dysregulated transcriptional programs revealed by scRNA-seq. Lastly, to facilitate dissemination of the data and findings, we offer an open-access online platform for the wider research community (http://mergeomics.research.idre.ucla.edu/PVDSingleCell/). Some of the results of these studies have been previously reported in the form of an abstract (10).

Methods

The main methods are below with additional details provided in an online supplement.

Animals

Adult male Sprague-Dawley rats (250–350 g) were used for all animal experiments, which were approved by the University of California, Los Angeles, Animal Research Committee. For the SuHx model, rats were injected subcutaneously with Sugen 5416 (20 mg/kg) followed by being kept in hypoxia at 10% O_2 for 21 days and then by being kept in normoxia for 14 days. For the MCT model, rats were injected subcutaneously with MCT (60 mg/kg) followed by being kept in normoxia for 28 days. Age-matched control rats were kept in normoxia for 28 days. Echocardiography and right heart catheterization were performed. Lungs were then harvested and enzymatically dissociated into single-cell suspensions, which was followed by scRNA-seq (11) ($n = 6$/group).

scRNA-seq Analysis

Expression data was normalized, filtered, and clustered using the Seurat R package (R Foundation for Statistical Computing) (12). Cell types were identified on the basis of known cell-type marker genes. Cell-type proportions were quantified and compared between PAH models and control animals, as previously described (13). Global transcriptomic shifts between groups were assessed using a Euclidian distance method (14). Differentially expressed genes (DEGs) were determined for each cell type between control and either SuHx or MCT rats using MAST (Model-based Analysis of Single-Cell Transcriptomics) (15). To annotate DEGs for biological pathways or PAH relevance, gene-set enrichment analysis was performed using hallmark pathways from the Molecular Signature Database (16) as well as using human PAH-associated gene sets obtained from DisGeNET (17) and the Comparative Toxicogenomics Database (18).

scRNA-seq Validation in Rat and Human Lung Tissues

The identities of select cell types were validated using bulk RNA-seq on cells purified by fluorescence-activated cell sorting (FACS) from the lungs of an additional set of rats ($n = 4$/group). Select scRNA-seq DEGs from SuHx and MCT rats were validated by RNA *in situ* hybridization (ISH) and immunofluorescence using rat lung sections ($n = 5$–6 rats/group). The same DEGs were further evaluated by RNA ISH on human lung sections from patients with PAH compared with control patients ($n = 7$–8 subjects/group).

Integration of Rat PAH Single-Cell DEGs with Human PAH Genome-Wide Association Study

To evaluate the relevance of the rat DEGs to human PAH, we assessed the human orthologs of rat DEGs for enrichment of genetic variants associated with PAH from a human genome-wide association study (GWAS) (19) using marker set enrichment analysis in the Mergeomics R package (20).

Identification of Drugs Predicted to Reverse Rat Disease Signatures Using Connectivity Map

Signatures of MCT and SuHx DEGs for each cell type were queried against the full Connectivity Map (CMap) (21) database of compound expression signatures induced in human cell lines to prioritize those with highly matching or opposing signatures. Pattern-matching algorithms scored each reference perturbagen profile for the direction and strength of enrichment with query scRNA-seq DEG signatures. Perturbagens with positive or negative connectivity scores have similar or opposite signatures to that of the query (i.e., genes that are increased in the scRNA-seq DEG query are decreased by the perturbagen or vice versa).

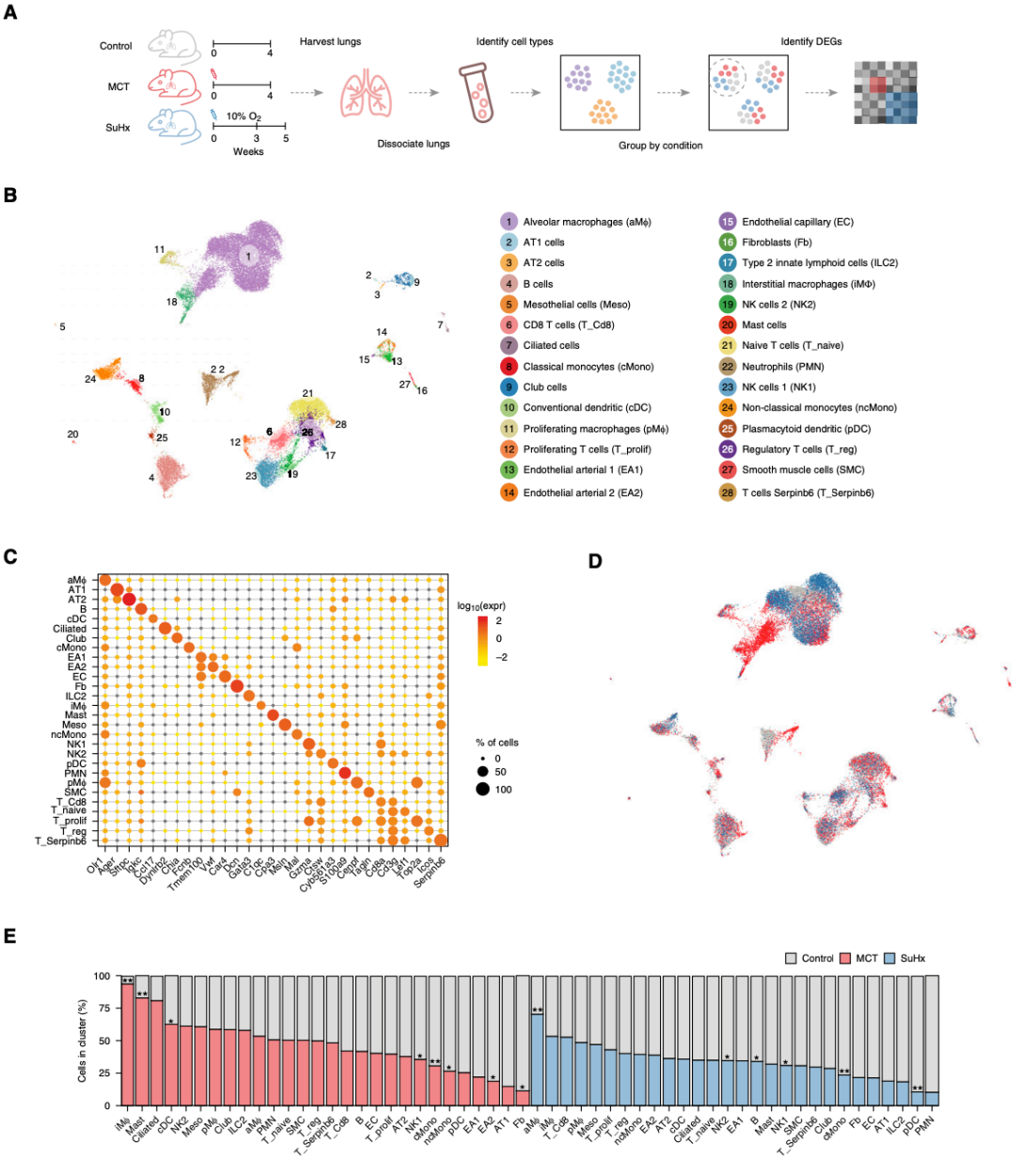


Figure 1. Single-cell RNA sequencing identifies diverse lung cell types in rat models of pulmonary arterial hypertension. (*A*) Schematic of study design for single-cell RNA sequencing analysis on the lungs of monocrotaline (MCT), Sugen-hypoxia (SuHx), and control rats ($n = 6$/group). (*B*) Uniform manifold approximation and projection plot showing lung cells from 18 rats with clusters labeled by cell type. (*C*) Dot plot highlighting $\log_{10}$ average expression of select marker genes used to identify cell clusters. The dot size corresponds to the percentage of cells expressing a gene in a given cluster. (*D*) Uniform manifold approximation and projection plot showing lung cells colored by disease condition: MCT in red, SuHx in blue, and control in gray ($n = 6$/group). (*E*) Bar table showing relative contributions of cells from disease models (MCT in red and SuHx in blue) versus the control model (gray) within each cell-type

Data Availability

The scRNA-seq data set and lists of cell type–specific marker genes and disease DEGs are available online at http://mergeomics.research.idre.ucla.edu/PVDSingleCell/CellBrowser/. Connectivity scores of the entire panel of perturbagens from the CMap analysis are available at http://mergeomics.research.idre.ucla.edu/PVDSingleCell/CMap/.

Results

scRNA-seq Identifies Diverse Cell Populations in the Rat Lung

The PAH phenotype in MCT and SuHx rats was confirmed by echocardiography (*see* Figures E1A, E1B, and E2B–E2L in the online supplement), immunohistochemistry (Figure E1C), and right heart catheterization (Figure E2A). The scRNA-seq of 18 lungs (6/group) profiled 33,392 cells (Figure 1A) after quality control (Figures E3A–E3D, E4A, and E4B), with even representation of groups (Figures E4C and E4D). After clustering on the basis of transcriptomic similarity, we identified 28 distinct cell types expressing established markers for epithelial, stromal, lymphoid, and myeloid cell populations and rare populations, including conventional dendritic cells (cDCs) and regulatory T cells (Tregs) (Figures 1B, 1C, and E5). Batch correction did not further optimize clustering and cell-type identification (Figure E4E) (12). Each cluster included cells from each group (Figures 1D and 1E). Compared with control animals, we observed a significant increase in the normalized cell fractions of interstitial macrophages (iMΦs) in MCT rats and alveolar macrophages (aMΦs) in SuHx rats (Figure 1E).

FACS and Bulk RNA-seq Validate scRNA-seq Cell-Type Identities and Proportions

To validate the rare lung cell types identified from scRNA-seq, namely cDCs and Tregs, and their corresponding gene signatures, we performed bulk RNA-seq on FACS-purified cells and subsequent deconvolution using our scRNA-seq signatures as a reference (Figure 2A). We used canonical markers to isolate cDCs ($CD64^-$ $CD11b/c^+$, $RT1B^+$) and Tregs ($CD4^+$, $CD25^+$, $CD278^+$) by FACS (Figures 2B and 2C). Deconvolution of FACS-purified transcriptomes showed strong enrichment for the correct cell types as identified by scRNA-seq, thus validating the accuracy of scRNA-seq cell signatures (Figure 2D). Furthermore, FACS-determined relative cell proportions between disease models and the control model showed a pattern similar to that from scRNA-seq (Figures 2E–2H). Specifically, both scRNA-seq and FACS showed significantly increased cDCs in MCT rats, but not in SuHx rats, and Tregs did not change in either model compared with the control model.

scRNA-seq Reveals DEGs with Cell-Type Specificity in PAH Models

A total of 4,724 and 2,324 DEGs were identified in MCT and SuHx rats (false discovery rate < 0.05), respectively, across 17 cell types (Figure 3A). There were 1,511 DEGs common in both models, of which 921 were regulated in the same direction. aMΦs, the largest cell cluster, had the most DEGs, likely due to high statistical power. We also assessed changes on a transcriptome scale within each cell type using a Euclidean distance-based approach that is less influenced by cluster size (14) (Figure 3B). Despite MCT rats having more DEGs, aMΦs and nonclassical monocytes (ncMonos) from the SuHx model demonstrated the strongest global transcriptomic shifts from the control model.

A closer examination of DEGs revealed genes whose differential expression was model and cell-type specific (Figure 3C). In total, there were 2,088 and 574 DEGs specific to one cell type in MCT and SuHx rats, respectively. For example, *Il6st*, which encodes a signal transducer that mediates IL-6 signaling, was upregulated exclusively in a subpopulation of endothelial arterial type 1 (EA1) cells from SuHx; *Il6* was specifically upregulated in SuHx ncMonos and MCT neutrophils, suggesting model-specific differences in IL-6 signaling. *Gpr15* was exclusively upregulated in SuHx Tregs and encodes an orphan G protein–linked receptor implicated in Treg homing (22).

Furthermore, we identified 19 and 8 DEGs that were differentially expressed in the same direction in at least five cell types in either MCT or SuHx rats, among which 6 (*Nfkbia*, *Scgb1a1*, *Ifi27*, *Slfn3*, *Mt-cox3*, and *AY172581.24*) were altered across various immune cells in both models (Figure 3D). For example, *Ifi27*, which encodes IFNα–inducible protein 27 and plays a role in apoptosis and vascular response to injury (23, 24), was downregulated across cell types in both models and in human PAH lungs (Figure E6) (25).

scRNA-seq Reveals Pathways with Cell-Type Specificity in PAH Models

Pathway enrichment of DEGs revealed cell type–specific dysregulation of many pathways (Figure 4A). The most distinct difference between models was a strong downregulation of IFN signaling across multiple cell types in the MCT model that in the SuHx model was weaker or in the opposite direction (Figure 4B). The relevance of IFN downregulation to human PAH was demonstrated in EA1 cells as an example (Figures 4C and 4D). The most notable commonality between models was a widespread upregulation of TNFα/NF-κB signaling across cell types, most notably in SuHx ncMonos (Figure 4E and 4F).

Validation of Select DEGs by RNA ISH and Immunofluorescence

Given the importance of ncMonos, suggested in our analyses above, we validated a DEG from ncMonos by RNA ISH on both rat and human lung sections (Table E1). We defined ncMonos as positive for both CD16 and Mal. Mal is the top marker for ncMonos in our scRNA-seq (Figure 1C) and is involved in the MyD88 pathway, important in human lung ncMonos (26). We validated the upregulation of *Ccrl2*, a top SuHx ncMono DEG, encoding a chemokine receptor–like protein whose function is unknown but is upregulated during monocyte-to-MΦ differentiation (Figure 5A) (27). We also validated the upregulation of *Fabp4*, a top

Figure 1. (*Continued*). cluster. The cell-type cluster referred to on the *y*-axis is defined as the total number of cells of a cell type from the control model and either the MCT or SuHx model (but not both models). A significant increase in p … relative to the control model. Wilcoxon rank-sum test: $^*P < 0.05$ and $^{**}P < 0.01$

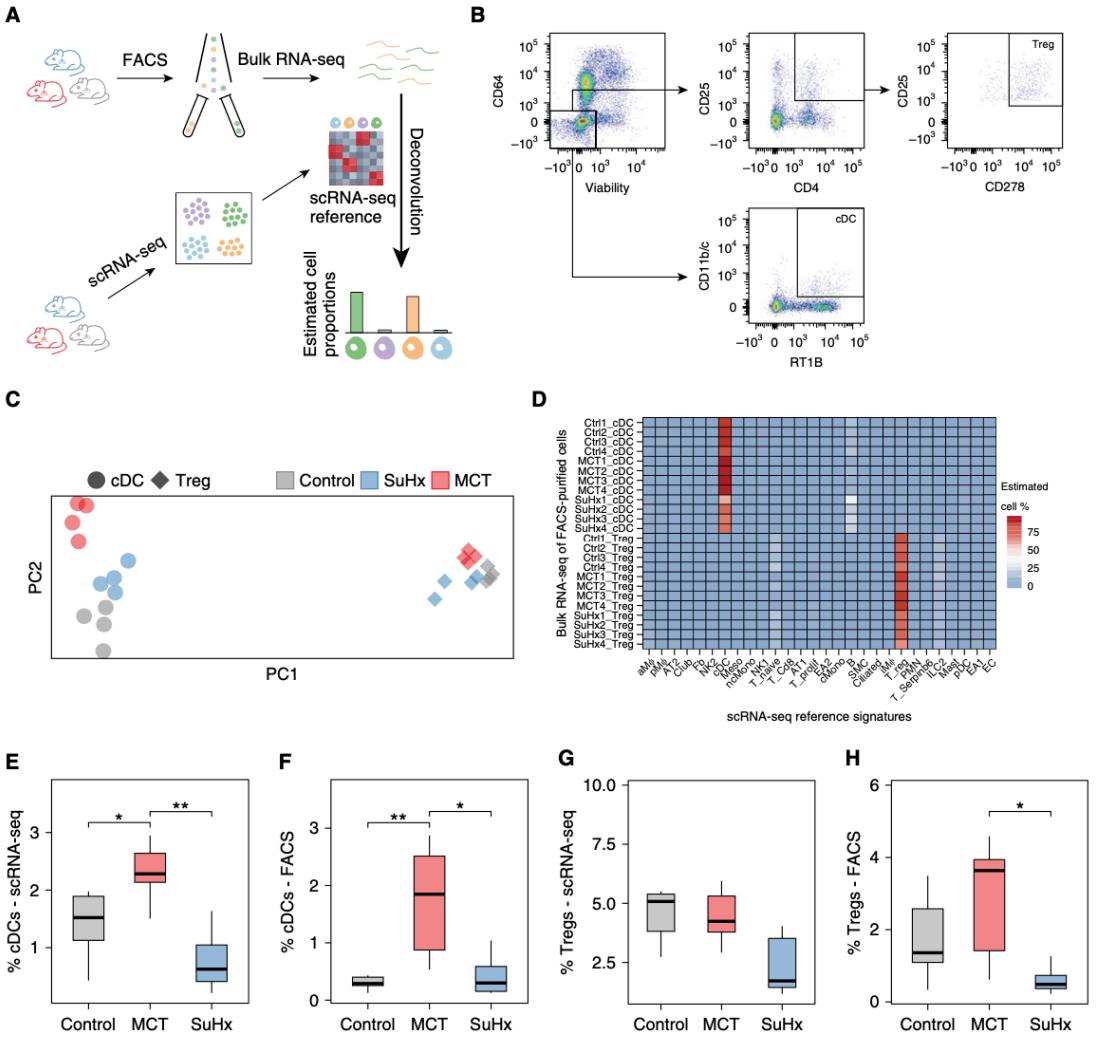


Figure 2. Fluorescence-activated cell sorting (FACS) and bulk RNA sequencing (RNA-seq) validate single-cell RNA-seq (scRNA-seq) cell-type identities and proportions. (*A*) Schematic showing the study design for deconvolution. The lungs from a separate set of Sugen-hypoxia (SuHx) (blue), monocrotaline (MCT) (red), and control (gray) rats ($n = 4$/group) underwent FACS using canonical cell-surface markers for two specific cell types, conventional dendritic cells (cDCs) and regulatory T cells (Tregs), after which bulk RNA-seq was performed on sorted cells to validate the cell-type identities in our scRNA-seq data using cell-type deconvolution of FACS-purified transcriptomes with CIBERSORTx (Stanford University). (*B*) Gating strategy for the isolation of $CD64^{-}$, $CD11b/c^{+}$, and $RT1B^{+}$ cDCs and $CD4^{+}$, $CD25^{+}$, and $CD278^{+}$ Tregs after gating for singlets and live cells. (*C*) Principal component (PC) analysis plot showing the clustering of bulk RNA-seq transcriptomes based on cell type (cDCs as circles and Tregs as diamonds) and disease condition. The first and second PCs (PC1 and PC2) explained 82% and 3% of the variance, respectively. (*D*) Heatmap showing deconvolution results using CIBERSORTx, in which relative proportions of cell types were estimated in each bulk RNA-seq sample of flow-sorted cells using a gene-expression signature for each cell type derived from scRNA-seq. High specificity for the correct cell types as identified by scRNA-seq was noted (83 ± 11% for cDCs; 77 ± 6% for Tregs). (*E–H*) FACS-determined relative cell-type proportions between disease models and the control model (*F* and *H*) showed a similar pattern to that of our scRNA-seq results (*E* and *G*). A significant increase in cDCs was noted in MCT compared with control rats. Furthermore, both methods consistently showed no significant changes in the number of cDCs in the SuHx model compared with the control model. The number of Tregs was also unchanged us ith the control model. Wilcoxon rank-sum test: $^{*}P < 0.05$ and $^{**}P < 0.01$. FACS = fluorescence-activated cell sorter.

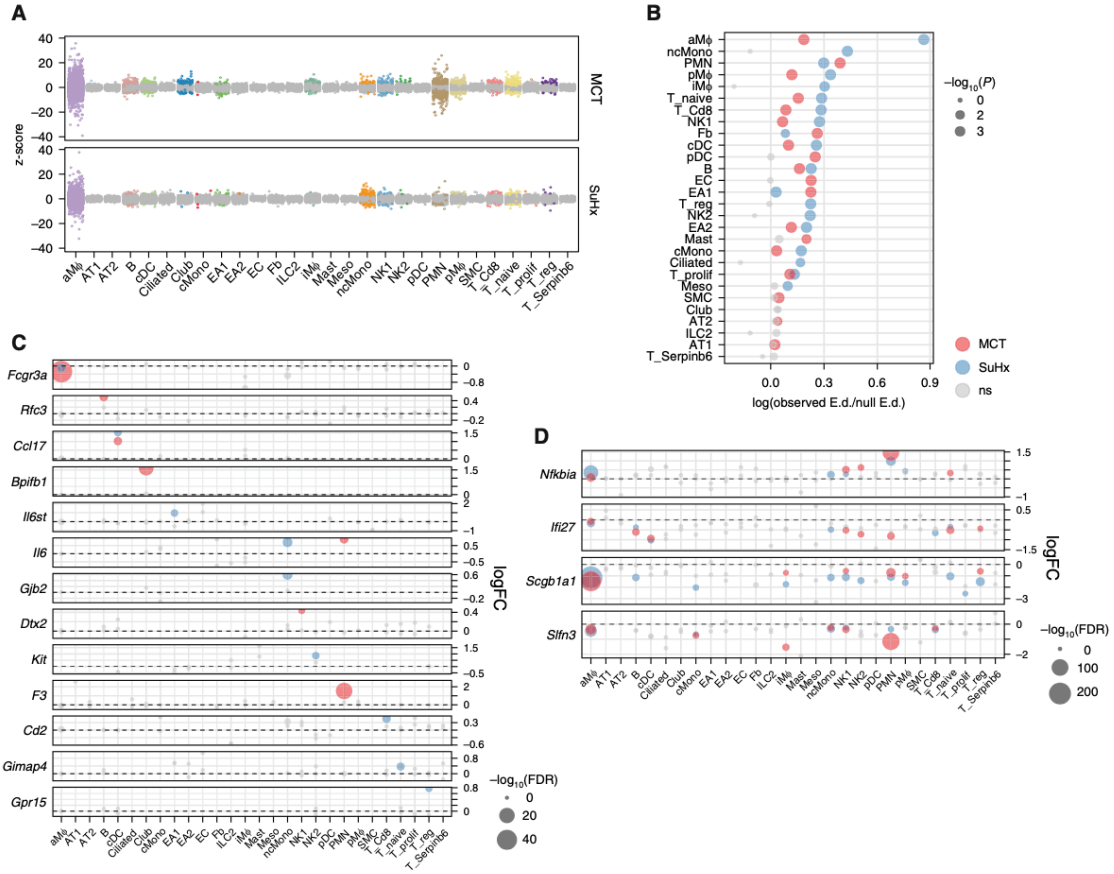


Figure 3. Single-cell RNA sequencing reveals differentially expressed genes (DEGs) in individual cell types of pulmonary arterial hypertension models. (*A*) Jitter plot showing changes in gene expression for each cell type due to monocrotaline (MCT) (top) or Sugen-hypoxia (SuHx) (bottom) conditions compared with the control condition. Each dot represents the differential expression MAST (Model-based Analysis of Single-Cell Transcriptomics) *z*-score of a gene. Dots indicating a false discovery rate (FDR) < 0.05 are in color. The gray dots indicate values that were not significant (ns). (*B*) Dot plot showing shifts in gene expression on a whole-transcriptome scale within each cell type for MCT (red) and SuHx (blue) models compared with the control model using a Euclidean distance (E.d.)-based statistical approach as previously described (14). The *x*-axis shows the log ratio of observed-to-null E.d. The alveolar macrophages and nonclassical monocytes from the SuHx model demonstrated the strongest global shifts in gene expression from the control model. (*C*) Dot plot comparing DEGs across cell types and disease models shows genes whose differential expression was specific to a disease model and a particular cell type. For example, *Gpr15*, which encodes an orphan G protein–linked receptor believed to be important in regulatory T cell (Treg) homing (22), was exclusively upregulated in Tregs from SuHx rats. (*D*) Dot plot showing DEGs consistent across immune-cell types. For instance, *Ifi27*, which encodes IFNα-inducible protein 27 and plays a role in apoptosis and vascular response to injury (23, 24), was downregulated across cell types in both models. (*C* and *D*) The horizontal dashed line for each gene represents zero logFC. (*B–D*) Gray dots indicate values tha: were ns, and the size of the dots corresponds to $-\log_{10}(P)$ values (*B*) and $-\log_{10}$(FDR) values (*C* and *D*). logFC = log fold change.

MCT aMΦ DEG, encoding a fatty acid–binding protein involved in lipid metabolism and inflammation (Figure 5B) (28). We further demonstrate similar upregulation of both proteins by immunofluorescence in rat lungs (Figures 5C and 5D).

Integrative Analysis of Rat scRNA-seq DEGs with Human PAH Genes Supports the Relevance to Humans

We curated genes implicated in PAH from DisGeNET and the Comparative Toxicogenomics Database (Figure 6A) and demonstrated that the top pathways enriched in these public gene sets (Figure 6B) were also highly enriched in our rat scRNA-seq (Figures 4A and 4E). When directly testing rat DEGs for enrichment of these PAH genes, we noted a marked upregulation in SuHx ncMonos in particular (Figures 6C–6E).

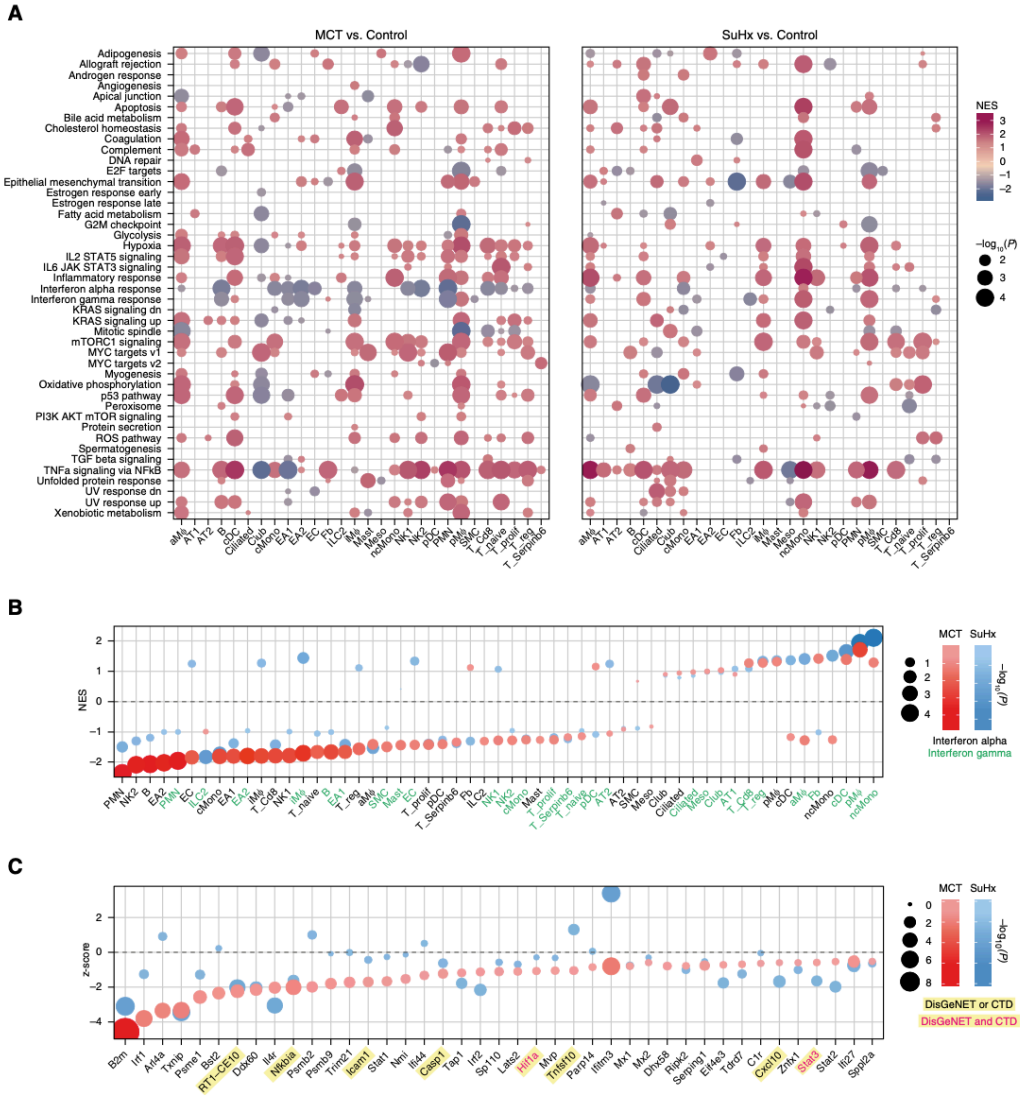


Figure 4. Single-cell RNA sequencing reveals pathways in individual cell types of pulmonary arterial hypertension models. (*A*) Heatmap showing cell type–specific pathway enrichment of gene signatures of monocrotaline (MCT) (left) and Sugen-hypoxia (SuHx) (right) models compared with the control model using gene-set enrichment analysis (GSEA) ($P < 0.05$) and hallmark pathways from the Molecular Signatures Database on the *y*-axis. The dot size corresponds to $-\log_{10}(P)$, and color represents the normalized enrichment score (NES) from GSEA, indicating upregulation (red) or downregulation (blue). TNFα/NF-κB signaling was significantly upregulated across many cell types in both disease models. (*B*) Dot plot showing NES of IFNα (black text) and IFNγ (green text) response pathways across cell types in the MCT (red) and SuHx (blue) models, in which the size and color tint of dots represent strength of $-\log_{10}(P)$ values. A strong downregulation of IFN pathways was seen across cell types in the MCT model. (*C*) Dot plot showing MAST (Model-based Analysis of Single-Cell Transcriptomics) *z*-scores of leading-edge genes accounting for the MCT EA1 downregulation of IFNγ response as determined by GSEA from the MCT (red) and SuHx (blue) models, in which the size and color tint of dots represent the strength of $-\log_{10}(P)$ values. Gene labels highlighted in yellow represent human pulmonary arterial hypertension–associated genes from either (black text) or both (red text) of the Comparative Toxicogenomics Database and DisGeNET databases. (*D*) Boxplots showing RNA expression of human orthologs of select IFN leading-edge genes shown in *C* derived from a public microarray (Gene Expression Omnibus series 70456) in which primary human pulmonary arterial endothelial cells were transfected with control (gray) or BMPR2 (purple)

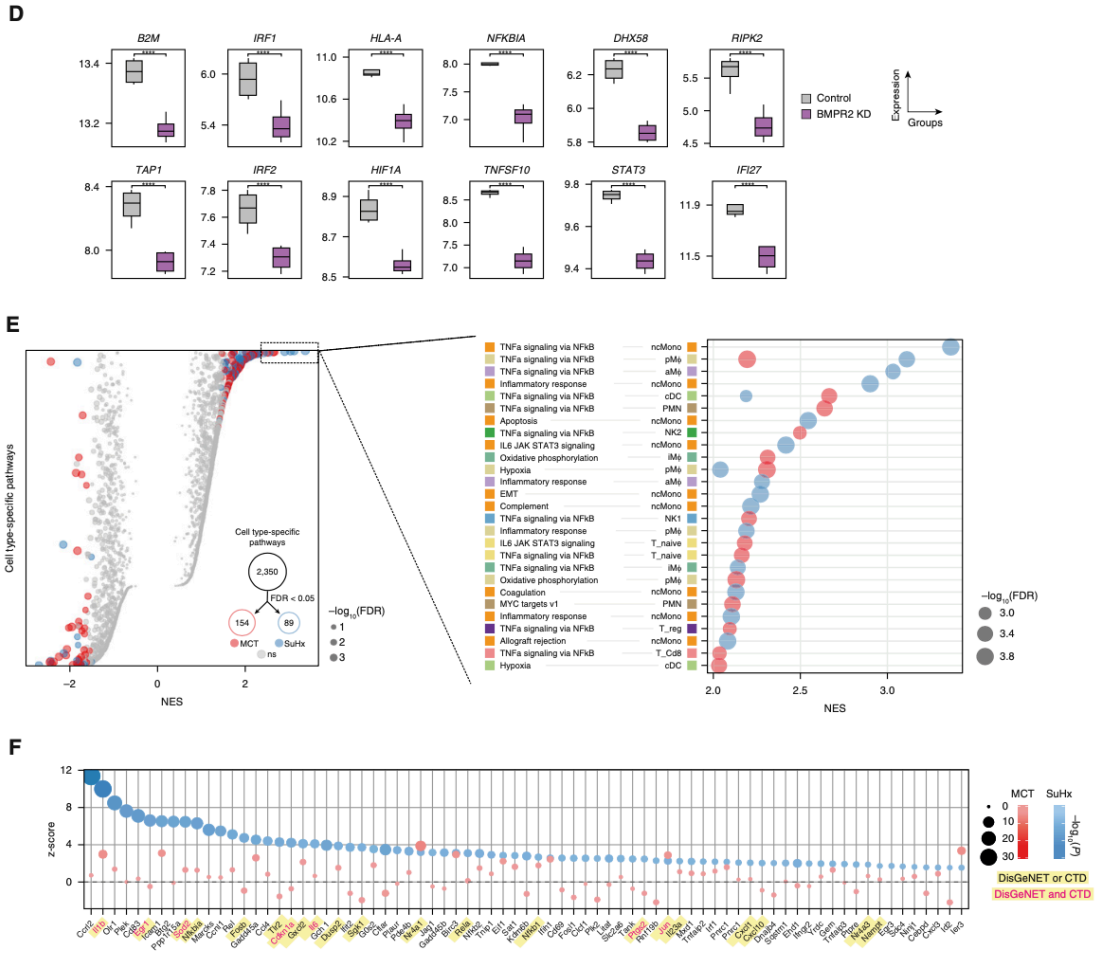


Figure 4. (*Continued*). siRNA ($n = 4$/group from 4 donors). *P* values were determined by using the limma R package: ****False discovery rate (FDR) < 0.05. (*E*) Dot plots showing all (left) and top 30 (right) cell type–specific pathways in descending order on the *y*-axis by NES in which positive scores indicate upregulation. Red (MCT) and blue (SuHx) dots met the FDR < 0.05 criterion, and gray dots were not significant (ns). The dot size indicates the strength of the FDR. The number of significant cell type–specific rat signatures by disease model is shown in the lower right (FDR < 0.05). In the left plot, dots on opposite sides of an NES of 0 for a given row represent opposite directionalities of cell type–specific enrichment of MCT and SuHx models. Many more cell type–specific pathways were significant in the MCT model compared with the SuHx model, but TNFα/NF-κB signaling in SuHx nonclassical monocytes (ncMonos) was the most prominently upregulated pathway overall (right). Cell-type colors correspond to those as labeled in Figure 1B. (*F*) Dot plot showing MAST *z*-scores of leading-edge genes accounting for the SuHx ncMono upregulation of TNFα/NF-κB signaling with figure legend as described in *C*. CTD = Comparative Toxicogenomics Database; KD = knockdown; NK = natural killer.

We further integrated rat scRNA-seq with a PAH GWAS using Mergeomics to assess human relevance (Figures 6F and 6G). We found significant enrichment for GWAS signals among DEGs in both models from a number of immune cells of both myeloid and lymphoid origins, supporting that DEGs from our rat models are relevant to PAH pathogenesis in humans (Figure 6H and Tables E2–E4).

scRNA-seq Uncovers Perturbations in Vascular Cell Types Relevant to Human PAH

Given the importance of pulmonary vascular cell types to PAH pathogenesis, we provide a closer examination of the endothelial arterial subpopulations of EA1 and EA2 cells, as well as of smooth muscle cells (SMCs) and fibroblasts (Figures 7A and 7B). We show model- and cell type–specific alterations of many established and unknown genes in PAH (Figure 7C) and a distinct pathway

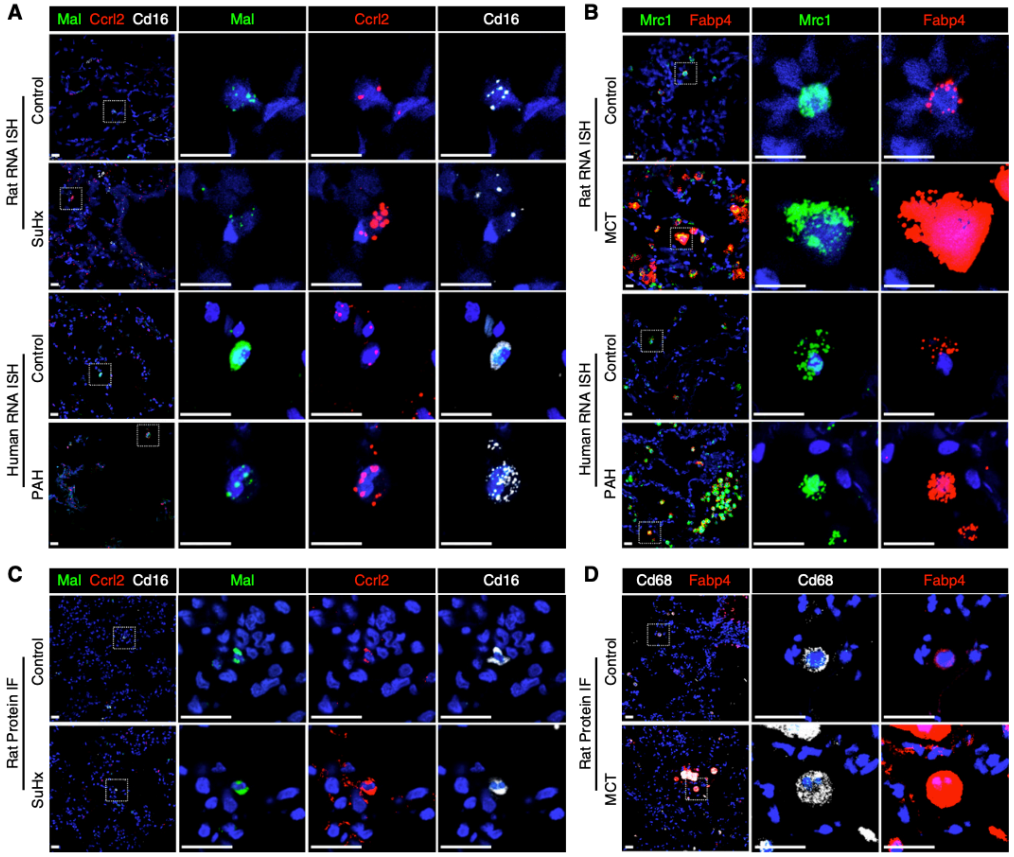


Figure 5. RNA ISH and immunofluorescence (IF) validate select differentially expressed genes. (*A* and *B*) The upregulation of *Ccrl2* (red) in Sugen-hypoxia (SuHx) nonclassical monocytes (ncMonos) from single-cell RNA sequencing (scRNA-seq) was observed by RNAscope (Advanced Cell Diagnostics) in SuHx rats and patients with PAH (*A*) and by IF in SuHx rats (*B*). ncMonos were defined as cells positive for both Cd16 (white) and Mal (green). We chose Mal for double-labeling because it was the top marker gene specific for ncMonos in our data. (*C* and *D*) The upregulation of *Fabp4* (red) in MCT alveolar macrophages (aMΦs) from scRNA-seq was demonstrated, in which aMΦs were defined as cells positive for *Mrc1* (green) for rat and human RNAscope (*C*) or Cd68 (white) for rat IF (*D*). Both *Mrc1* and *Cd68* are canonical markers and were cell type–specific markers for aMΦs. The cell nuclei are labeled with DAPI (blue). Scale bars, 20 μm. ISH = *in situ* hybridization; MCT = monocrotaline; PAH = pulmonary arterial hypertension.

dysregulation (Figure 7D). We also highlight select DEGs that are similarly altered in public human cell type–specific data sets (Figure 7E). For example, *Bmpr2*, the most well-studied PAH gene, was downregulated in MCT EA1 cells but not in EA2 cells, whereas downregulation was not observed in SuHx vascular cells, which is consistent with a recent study (29). *Cst3*, encoding cystatin C, was upregulated in MCT and SuHx EA1 cells. Furthermore, Cst3 serum levels correlated with right ve

mortality in patients with PAH, despite unknown mechanism in PAH (30, 31).

Integration of Rat DEGs with CMap Identifies Potential Candidate Drugs for Repositioning

To investigate the utility of scRNA-seq DEGs to identify the therapeutic potential of existing drugs for PAH, we screened all cell-specific transcriptional signatures against the molecular profiles of thousands of pharmacologic perturbagens tested in human cell lines from CMap, including 5 PAH drugs, 26 drugs tested in PAH clinical trials, and 15 drugs that have shown efficacy in PAH animal models (Figures 8A and 8B). Comparisons of drug connectivity profiles revealed interesting patterns. For example, bosentan, an endothelin receptor antagonist currently used in patients with PAH, and tacrolimus, a calcineurin inhibitor for which a phase 2 clinical trial for PAH was completed, had very similar connectivity profiles across cell types and disease models, suggesting converging pathways that are likely due to similar activation of BMP signaling (32, 33). In contrast, distinct connectivity

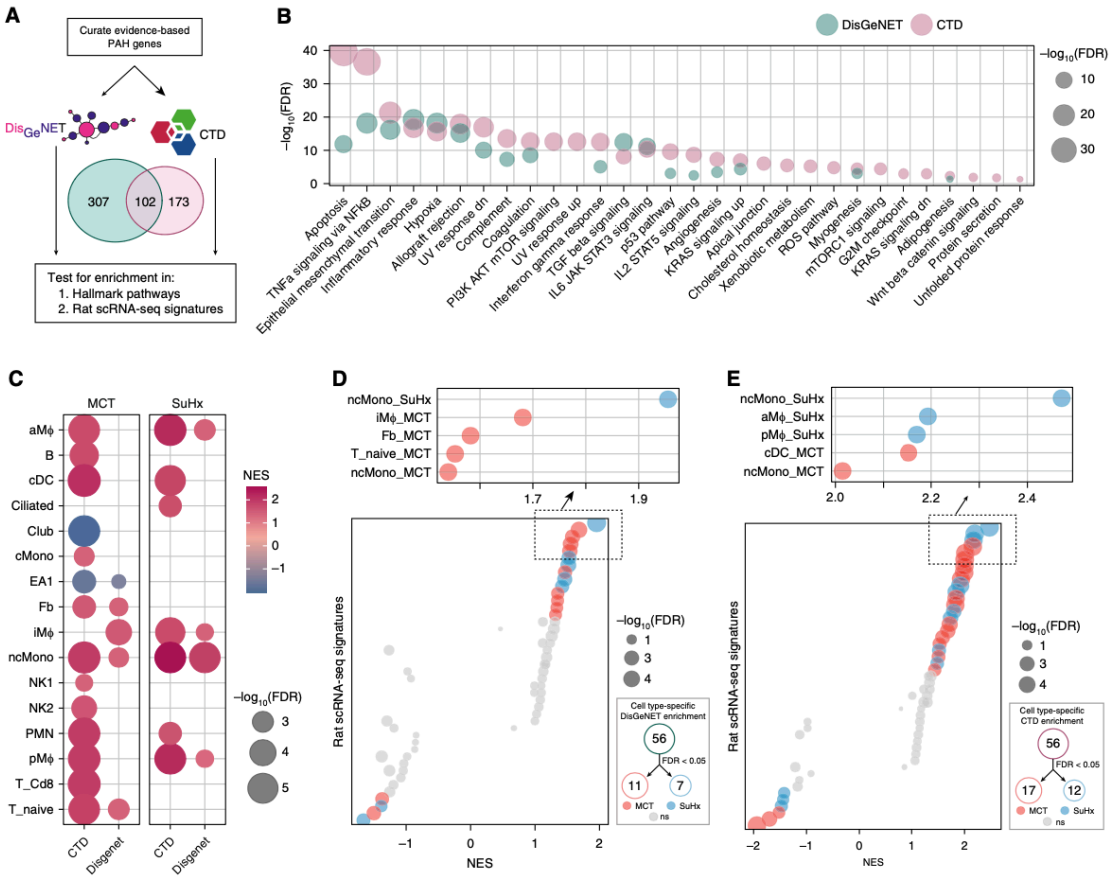


Figure 6. Integrative analysis of rat single-cell RNA sequencing (scRNA-seq) differentially expressed genes (DEGs) with human pulmonary arterial hypertension (PAH) genetics points to the relevance of the DEGs to human PAH. (*A*) Schematic of analytical approach whereby genes implicated in human PAH were curated from DisGeNET (409 genes) and Comparative Toxicogenomics Database (CTD) (275 genes), of which 102 genes were shared between the databases. These genes were then tested for enrichment in (*B*) Molecular Signatures Database hallmark pathways and (*C–E*) rat scRNA-seq signatures. (*B*) Dot plot showing pathways significantly enriched for PAH genes by hypergeometric test (false discovery rate [FDR] < 0.05), whether from the DisGeNET database (green) or the CTD database (purple). The dot size represents the $-\log_{10}$(FDR). These gene sets were highly enriched for known or implicated PAH pathways, such as apoptosis, NF-κB signaling, and endothelial-to-mesenchymal transition and were similar overall to those altered in Sugen-hypoxia (SuHx) and monocrotaline (MCT) rat lung scRNA-seq (Figure 4). (*C*) Heatmap showing most highly significant (FDR < 0.01) enrichment of PAH genes in MCT (left) and SuHx (right) cell type–specific signatures using gene-set enrichment analysis, in which red indicates upregulation and blue indicates downregulation. The dot size represents $-\log_{10}$(FDR). Significant upregulation of PAH genes was noted in myeloid cell types in both models, and in nonclassical monocytes (ncMonos) in particular. (*D* and *E*) Dot plots showing all (left) and top 5 (upper right) cell type–specific rat signature enrichment for PAH genes from the DisGeNET (*D*) and CTD (*E*) databases. The red (MCT) and blue (SuHx) dots indicate meeting the FDR < 0.05 criterion, and gray dots indicate values that were not significant (ns). The dot size represents $-\log_{10}$(FDR). A number of significant cell type–specific rat signatures by disease model are shown in the lower right (FDR < 0.05). In the left-sided plots, dots on opposite sides of an NES of 0 for a given row represent opposite directionalities of cell type–specific enrichment of MCT and SuHx models. The SuHx ncMonos DEGs were most highly enriched for PAH genes comparing both models. For the MCT model, DEGs from iMΦs and cDCs demonstrated the strongest enrichment for PAH genes. (*F*) Schematic of analytical approach for human PAH genome-wide association study (GWAS) integration. Human orthologs of rat scRNA-seq DEGs were assessed for enrichment of genetic variants associated with PAH in human GWAS to further assess human relevance of the rat signatures. GWAS SNPs were filtered by keeping the top 50% by *P* value strength and LD $r^2 < 0.5$, after which SNPs were mapped to genes by integrating with lung-specific expression quantitative loci (eQTLs) curated from public databases. DEGs within each cell type ($P < 0.01$ to include DEGs from rare cell types with low statistical power) were then tested for enrichment of these GWAS-integrated expression SNPs (eSNPs). The GWAS *P* values of each eSNP set (by cell type and disease model) were then compared against that of eSNPs generated from random gene sets to assess the significance of enrichment for stronger GWAS association *P* values using a modified chi-square statistic. (*G*) Manhattan plot showing $-\log_{10}(P)$ values of 39,263 eSNPs used for rat

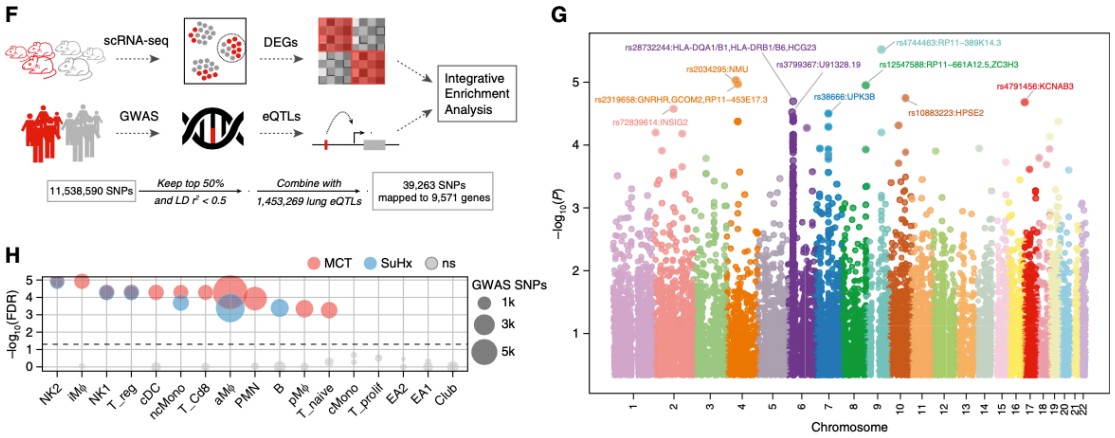


Figure 6. (*Continued*). integration as described in *F*. SNPs are ordered along the *x*-axis according to their chromosomal location; the colors represent different chromosomes. Select eSNPs are labeled with their reference SNP identifiers and the corresponding genes they regulate. (*H*) Dot plot showing rat cell type–specific DEGs enriched for PAH-associated genetic variants from human GWAS. Red (MCT) and blue (SuHx) dots indicate an FDR < 0.05, and gray dots indicate values that were ns. The horizontal dashed line corresponds to an FDR = 0.05. The dot size is proportional to the number of enriched GWAS SNPs in the thousands. Significant enrichment for human PAH GWAS signals among DEGs in both rat disease models was noted from a number of immune-cell types of both myeloid and lymphoid origin, supporting that the rat disease signatures are relevant to PAH pathogenesis in humans. LD = linkage disequilibrium; NES = normalized enrichment score.

profiles were observed among drugs from the same class, such as with treprostinil and iloprost (Figure 8B), both of which are synthetic analogs of prostacyclin but have opposing effects on a preclinical model of lung cancer (34), likely due to different off-target effects mediated by differential engagement of other eicosanoid receptors. Our analysis also predicted that sildenafil reverses MCT's SMC signature most enriched for genes related to epithelial-to-mesenchymal transition (EMT), consistent with prior studies investigating sildenafil's effect on pulmonary arterial SMCs (PASMCs) from humans (35–37) and MCT rats (38). Upregulated EMT genes in our MCT SMC signature were also found to be downregulated in a toxicogenomic microarray of sildenafil-exposed rat hepatocytes (Figures E7A–E7C) (39).

We next examined drugs predicted to reverse the transcriptional signature of SuHx ncMonos, given that these cells were most highly enriched for genes and pathways relevant to PAH. Out of 2,429 compounds screened, the drug with the strongest reversal of SuHx DEGs in ncMonos was treprostinil, one of the most effective PAH-targeted therapies currently in use (Figure 8E). Two other top drugs s recently shown to attenuate or prevent PAH in animal models: tipifarnib, a farnesyltransferase inhibitor currently undergoing clinical trials for various cancers (40), and memantine, an NMDA receptor antagonist used to treat Alzheimer's dementia (41). Our analysis also revealed novel candidates for repurposing, such as palonosetron, a serotonin-receptor antagonist used to treat chemotherapy-induced nausea, and guaifenesin, an expectorant used to loosen airway mucus via unknown mechanisms (Figure 8C).

Dissemination of scRNA-seq and Drug Repositioning Results

We implemented a web server to enable interactive browsing of the entire scRNA-seq data set, as well as searching and downloading of cell type–specific marker genes, disease signatures, and drug predictions (Figure 9).

Discussion

In this study, we uncover comprehensive cellular landscapes of altered genes and pathways at single-cell resolution in the o widely used animal models of PAH. Comparison of landscapes between SuHx and MCT models reveals critical similarities and differences in key cell types. We integrated rat scRNA-seq with human data to determine human relevance and with pharmacotranscriptomic data to identify potential drug-repurposing candidates. Finally, we offer an open-access platform for the wider research community to access the data and findings.

The mechanisms leading to the PAH phenotype in MCT and SuHx rat models are believed to differ in that MCT is associated with endothelial toxicity and marked lung inflammation, whereas SuHx is characterized by angioproliferative pulmonary vascular disease (42). SuHx is overall believed to recapitulate PAH more closely than MCT, such as in the formation of neointimal plexiform lesions, the pathological hallmark of PAH. However, human relevance of these models is likely more nuanced at the cellular and molecular level. The lack of in-depth understanding of lung cell types and their respective pathway alterations in PAH limits our ability to rationally leverage these models in translational science. In this study, we found cell type–specific upregulation of known PAH pathways across multiple cell types in both MCT and SuHx models, most notably involving NF-κB signaling.

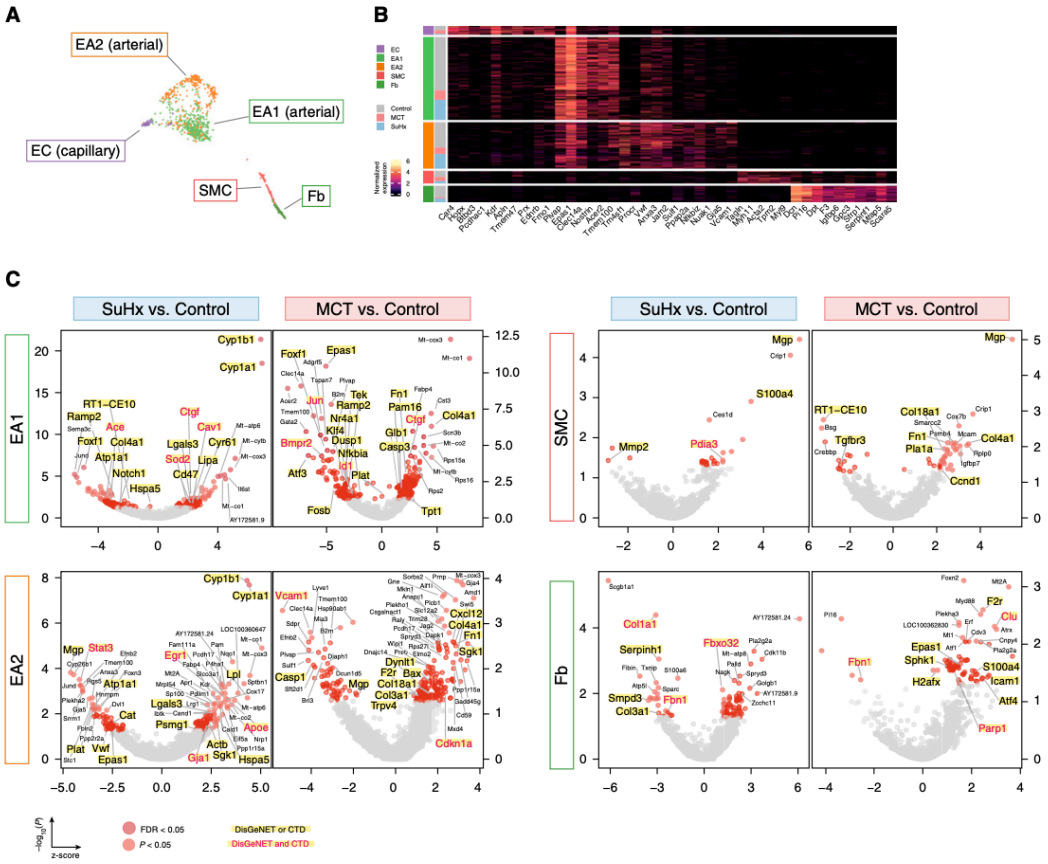


Figure 7. Single-cell RNA sequencing uncovers perturbations in lung vascular cell types relevant to human pulmonary arterial hypertension (PAH). (*A*) Uniform manifold approximation and projection plot showing vascular cells from 18 rat lungs with clusters labeled by cell type. (*B*) Heatmap showing normalized expression of top marker genes used to identify the vascular cell types, in which each row is an individual cell. Shown to the left are the condition and cell type to which each cell belongs. (*C*) Volcano plots showing differentially expressed genes (DEGs) within vascular cell types for the Sugen-hypoxia (SuHx) or monocrotaline (MCT) models versus the control model, in which the *x*-axis represents MAST (Model-based Analysis of Single-Cell Transcriptomics) *z*-scores and the *y*-axis indicates $-\log_{10}(P)$. Significant upregulated ($z > 0$) or downregulated ($z < 0$) genes are shown as red ($P < 0.05$) or dark red (false discovery rate [FDR] < 0.05) dots. DEGs ($P < 0.05$) labeled and highlighted in yellow represent human PAH-associated genes from either (black text) or both (red text) of the CTD and DisGeNET databases. Otherwise, DEGs are labeled with their gene names if the FDR < 0.05 (endothelial arterial type 1 cell [EA1]) or $P < 0.01$ (EA2, SMC, Fb). (*D*) Dot plots showing the top five upregulated and top five downregulated pathways within vascular cell types as determined by gene-set enrichment analysis. Colored dots in red (MCT) or blue (SuHx) indicate significant values ($P < 0.05$), whereas gray dots represent values that were not significant (ns). The dot size corresponds to the $-\log_{10}(P)$ value. (*E*) Box plots showing expression of select DEGs in rat lung vascular cell types with similar changes shown side by side in human orthologs from public cell type–specific expression data sets: *BMPR2*: Gene Expression Omnibus series (GSE) 126262, primary PAECs from two patients with PAH with BMPR2 mutations versus nine unused donor controls; *FOXF1*: GSE126262, primary PAECs from four male patients with PAH versus five male unused donor controls; *CST3*, *STAT3, SGK1* and *AMD1*: GSE70456, four BMPR2 siRNA–transfected versus four control siRNA–transfected primary human PAECs from four donors; *MGP*, *MMP2*, *CCND1*, *F2R*, *FBN1*, and *EPAS1*: GSE2559, primary human PASMCs from two patients with PAH versus two normal subjects ($n = 4$ vs. 3, respectively, including BMP2-treated vs. untreated). *P* values from RNA sequencing (GSE126262) were determined by using DESeq2, whereas those from microarray (GSE70456 and GSE2559) were determined by using R limma: $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and ****FDR < 0.05. CTD = Comparative Toxicogenomics Database; Fb = fibroblast; KD = knockdown; NES = normalized enrichment score; PAEC = pulmonary artery endothelial cell; PASMC = pulmonary arterial SMC; SMC = smooth muscle cell.

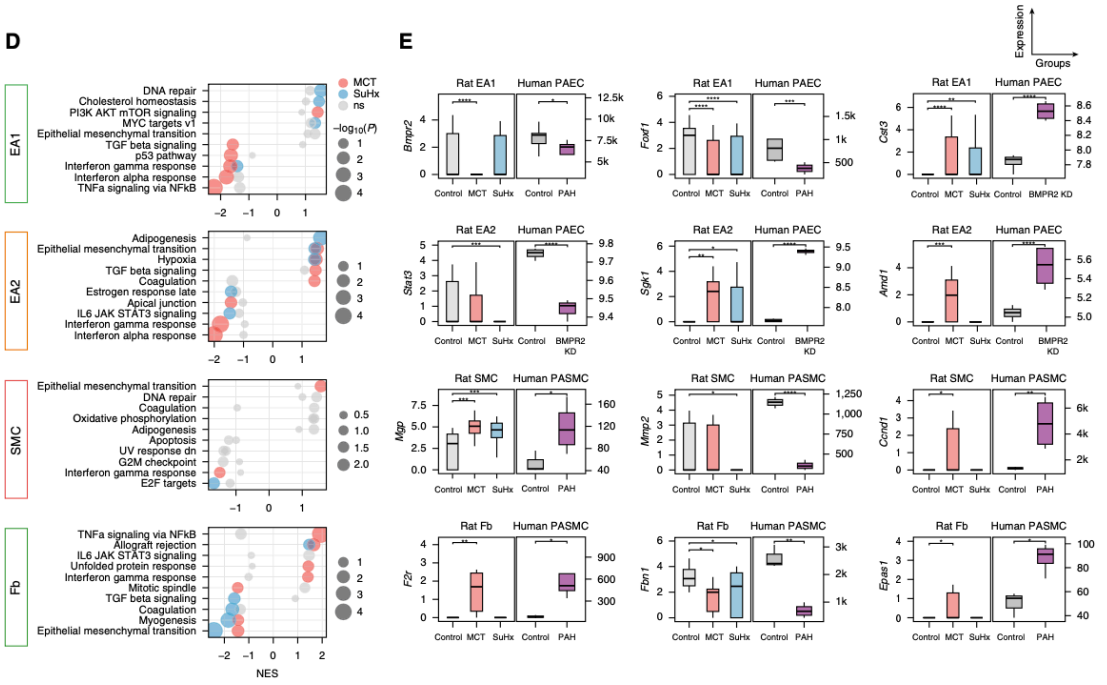


Figure 7. (*Continued*).

Although prior studies have demonstrated NF-κB's importance in PAH (43, 44), our study systemically resolves and implicates understudied cell types that most strongly mediate this critical pathway: SuHx ncMonos and MCT cDCs. Similar to cDCs in human PAH (45), cDCs in rat PAH were increased in MCT (but not SuHx) lungs compared with control lungs, as determined by scRNA-seq and FACS. In addition, iMΦs were also increased in MCT lungs but not in SuHx lungs, which is concordant with prior studies in MCT rats and human PAH (46). MCT iMΦ and cDC transcriptional signatures were significantly enriched for PAH GWAS single-nucleotide polymorphisms and known PAH genes, whereas those of SuHx signatures were not (Figures 6D, 6E, and 6H). Therefore, for the investigation of cell types such as cDCs or iMΦs, the MCT model may recapitulate human PAH better than the SuHx model does.

A rather unexpected finding from pathway analysis was the widespread downregulation of IFNα and IFNγ si cells, which was most notable in the MCT model. In endothelial cells, many of the genes accounting for the decrease in IFN signaling in our study were also downregulated in BMPR2-silenced human pulmonary arterial endothelial cells. A harmful downregulation of this pathway is supported by prior data showing that exogenous IFNα decreased proliferation in human pulmonary arterial endothelial cells and PASMCs *in vitro* and reversed PAH in animal models (47). However, other studies suggest that excess IFN signaling may contribute to PAH (48, 49). Further research is needed to dissect the nuanced role of this pathway in PAH.

In addition to resolving PAH-relevant cell types and pathways, scRNA-seq revealed many altered genes with cell-type and model specificity. *Ccrl2*, which has not been previously implicated in PAH, was the top upregulated gene contributing to the strong NF-κB pathway enrichment and transcriptional signature in SuHx ncMonos; we confirmed *Ccrl2*'s upregulation in AH ncMonos by lung histology. *Fabp4*, a fatty acid–binding protein that we found to be highly upregulated in MCT and human PAH aMΦs, is a transcriptional target of HIF-1α and has been implicated in regulating inflammatory cytokines and NF-κB signaling in aMΦs (28, 50). Further investigation of the role and mechanisms of altered genes, such as *Ccrl2* in ncMonos and *Fabp4* in aMΦs, is warranted.

Many of the DEGs identified in our study could be simply correlated with disease rather than pathogenic drivers of PAH. Given that genetic risk signals carry causal information, we integrated rat scRNA-seq with a human GWAS to infer causality of cell-type transcriptional programs in human PAH. Transcriptional signatures of ncMonos from SuHx and MCT models, along with numerous other myeloid and lymphoid cells, including iMΦs and cDCs from MCT rats and Tregs from MCT and SuHx rats, were significantly enriched for genetic variants from a human PAH GWAS, suggesting that these cell types and their corresponding pathways may play a causal role in PAH. The relevance of ncMonos is supported by

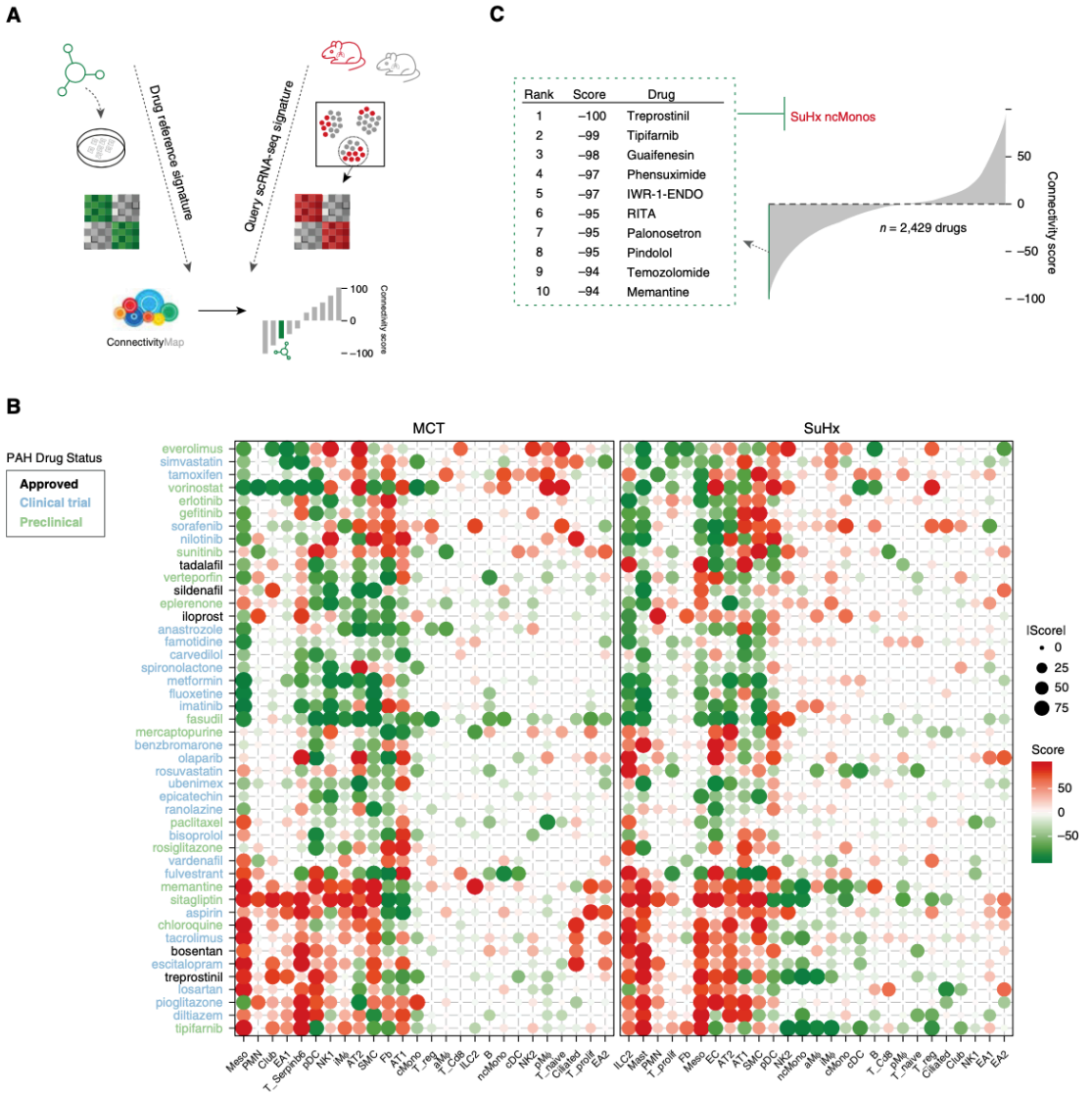


Figure 8. Integration of rat differentially expressed genes (DEGs) with Connectivity Map identifies potential candidate drugs for repositioning. (*A*) Schematic of analytical approach whereby signatures of rat DEGs ($P < 0.01$ to include DEGs from rare cell types with low statistical power) for each cell type for both Sugen-hypoxia (SuHx) and monocrotaline (MCT) models against the control model were queried against the full Connectivity Map database of compound and genetic perturbational expression signatures induced in human cell lines. The pattern-matching algorithms scored each reference perturbagen profile for the direction and strength of enrichment with the query single-cell RNA sequencing (scRNA-seq) DEG signature. Perturbagens with strongly positive connectivity scores have highly similar signatures to that of the query, whereas those perturbagens with strongly negative scores have signatures that strongly oppose that of the query (i.e., genes that are upregulated in the scRNA-seq DEG query are downregulated by treatment with the perturbagen or vice versa). (*B*) Heatmap showing connectivity scores of rat scRNA-seq DEGs to drugs approved for use in patients with pulmonary arterial hypertension (PAH) (black), drugs currently or previously in PAH clinical trials (blue), and preclinical drugs with demonstrated efficacy in PAH animal models (green). The size of dots corresponds to absolute values of the connectivity score. The PAH-related drugs showed distinct matching patterns to cell type–specific PAH rat signatures. For example, bosentan and tacrolimus had very similar connectivity profiles across cell types and disease models, although they come from different classes of drugs. (*C*) The top 10 drugs with the n st

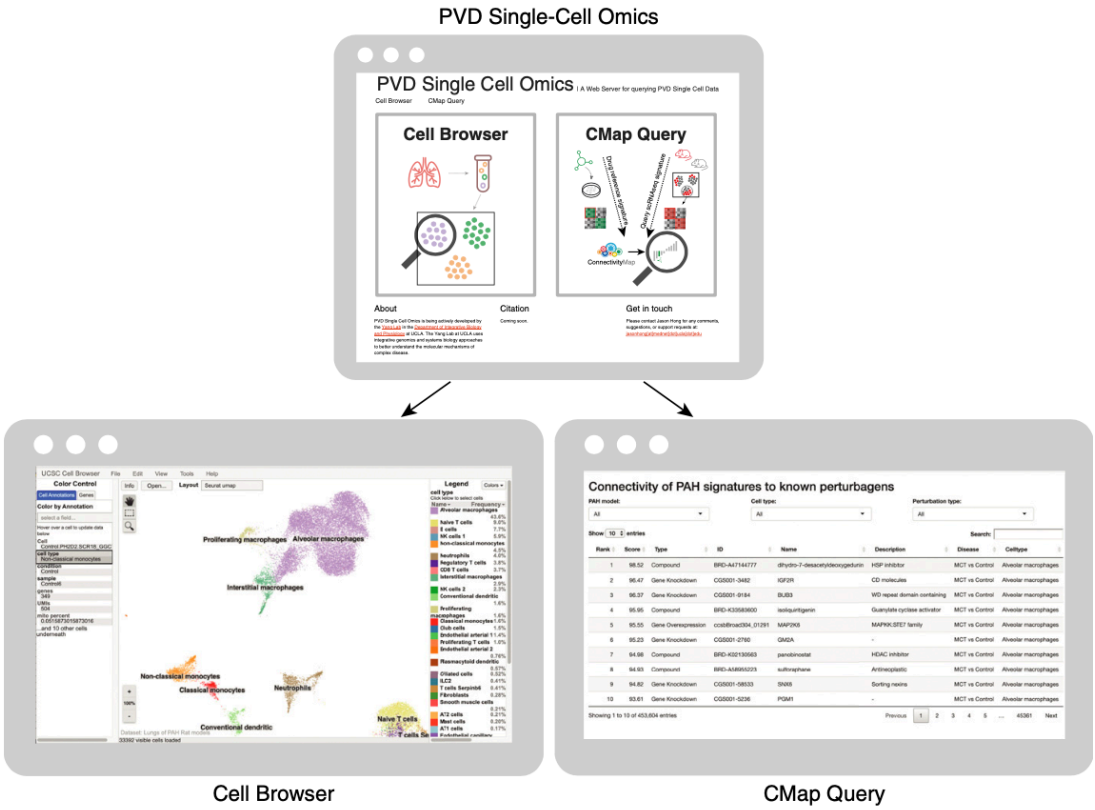


Figure 9. The PVD Single-Cell Omics website offers an open-access online platform. The entire rat single-cell RNA sequencing gene-expression data set and lists of cell type–specific marker genes and disease differentially expressed genes are available online in the form of an interactive cell browser at http://mergeomics.research.idre.ucla.edu/PVDSingleCell/CellBrowser/. Connectivity scores of the entire panel of perturbagens from the CMap analysis are available to query at http://mergeomics.research.idre.ucla.edu/PVDSingleCell/CMap/. CMap = Connectivity Map; PAH = pulmonary arterial hypertension; PVD = pulmonary vascular disease.

increased ncMonos in patients with PAH (46), and its deficiency in HIF-1α leads to impaired maturation of iMΦs and disease attenuation in hypoxic mice (51). Blocking CX3CL1–CX3CR1 signaling, which is important for ncMono survival (52), decreased lung iMΦs and attenuated vascular remodeling in rodent models (46). Our unbiased comparative study further prioritizes ncMonos as a particularly important cell type in PAH pathogenesis.

To gain further translational insights, we leveraged a wealth of pharmaco transcriptomic data to query PAH signatures identified from all lung cell types in both models. Supporting this approach, our analysis predicted sildenafil's action in reversing MCT's SMC signature most enriched for EMT genes, consistent with prior studies investigating sildenafil's effect on human PASMCs (35–37). Thus, sildenafil may have a therapeutic effect in PAH beyond pulmonary vasodilation. Indeed, targeted *in vivo* delivery of sildenafil to PASMCs was recently shown to inhibit vascular remodeling and improve survival in MCT rats (38). Further validating our approach, treprostinil, one of the most effective PAH-targeted therapies currently in use for patients with PAH, was the top drug predicted to reverse the SuHx ncMono disease signature that was most enriched for NF-κB signaling. Although treprostinil's primary clinical effect in PAH is believed to be pulmonary vasodilation via the prostacyclin pathway, an additional therapeutic mechanism may be attenuation of NF-κB signaling, based on our results and previous human and murine studies (53–56).

Figure 8. (*Continued*). strongly reverse the transcriptional signature of SuHx nonclassical monocytes (ncMonos), are shown (out of 2,429 compounds screened). The drugs predicted against SuHx ncMonos were of particular interest, given the strong upregulation of both NF-κB signaling and human PAH genes. The drug with the most negative connectivity score was treprostinil, one of the most effective PAH-targeted therapies currently in use in patients wi

Our analysis also predicted other top drugs that may reverse the SuHx ncMono signature for repositioning in PAH. Among these, palonosetron, a 5-HT3 receptor antagonist used for chemotherapy-induced nausea, may have therapeutic potential in PAH through inhibiting the upregulation in serotonin signaling, known to occur in PAH pathobiology, and/or through inhibiting NF-κB, the top pathway implicated in SuHx ncMonos. Indeed, another 5-HT3 receptor antagonist, tropisetron, modulates NF-κB in a rat model of type 2 diabetes (57). Another top prediction was guaifenesin, an expectorant which acts to loosen airway mucus through unknown mechanisms. Guaifenesin may act as a NMDA receptor antagonist targeting the dysregulation in glutamate–NMDA receptor signaling in PAH (41, 58) or may act through modulating pathways like TNFα/NF-κB signaling that are upregulated in SuHx ncMonos, as revealed in our analysis and suggested in a prior human study (59). Although guaifenesin is a common over-the-counter medicine used by patients with PAH on an as-needed basis for mucus relief, it may not be available at the doses and frequency needed to treat PAH. Such drug predictions warrant further investigation as potential repurposing candidates for PAH.

The overall strengths of our study include offering the first single-cell resolution landscape of two widely used rat models of PAH; comprehensive comparative and integrative omics analysis to prioritize cell types, genes, and pathways relevant to human PAH; high-throughput computational screening to identify potential drug repositioning candidates for future testing; and an open-access resource for the wider PAH research community. Select key findings were also experimentally validated using alternative methods, such as FACS with bulk RNA-seq, ISH, and immunofluorescence. There are also limitations to our study that underscore the need to further improve on current single-cell approaches, particularly the need to enzymatically dissociate the heterogenous cell populations that may be particularly fragile or tightly embedded in the extracellular matrix (3). Our lung dissociation yielded fewer vascular cells overall than immune cells, which limited statistical power within these cell clusters and potentially favored larger immune clusters in our comparative analyses. Alternative methods such as single-nucleus RNA-seq may mitigate such limitations. Despite this limitation, separate analysis of the captured vascular cells still provided valuable insights. Furthermore, our study unravels numerous genes, pathways, and drugs that warrant experimental and functional testing in future studies.

In conclusion, our scRNA-seq study of SuHx and MCT rat lungs dissects the distinct and shared dysregulation of gene-expression programs and pathway activation in individual cell populations, elucidates their relevance to human PAH pathobiology and drug repositioning, and will help guide the rational use of these preclinical models in future translational studies in PAH. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank Dr. Ramin Salehirad and Dr. Min Zhou for their input on the flow cytometry experiments, Dr. Zeyneb Kurt for her guidance on Mergeomics, Dr. Min Li for her assistance with the lung dissociation protocol, and Dr. Gregory Fishbein for his help in identifying control human lung sections.

References

1. Benza RL, Miller DP, Barst RJ, Badesch DB, Frost AE, McGoon MD. An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL registry. *Chest* 2012;142:448–456.
2. Thenappan T, Ormiston ML, Ryan JJ, Archer SL. Pulmonary arterial hypertension: pathogenesis and clinical management. *BMJ* 2018;360: j5492.
3. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, *et al*. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;199:1517–1536.
4. Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, *et al*. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* 2018;24:1277–1289.
5. Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, *et al*. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med* 2019;25:1153–1163.
6. Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, *et al*. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 2018;560:319–324.
7. Saygin D, Tabib T, Bittar HET, Valenzi E, Sembrat J, Chan SY, *et al*. Transcriptional profiling of lung cell populations in idiopathic pulmonary arterial hypertension. *Pulm Circ* 2020;10:1–15.
8. Bonnet S, Provencher S, Guignabert C, Perros F, Boucherat O, Schermuly RT, *et al*. Translating research into improved patient care in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2017;195:583–595.
9. Denayer T, Stöhr T, Van Roy M. Animal models in translational medicine: validation and prediction. *New Horiz Transl Med* 2014;2:5–11.
10. Hong J, Arneson D, Ahn IS, Umar S, Diamente G, Bhetraratana M, *et al*. Lung single-cell transcriptomics of two animal models of pulmonary arterial hypertension reveals vulnerable cell types and pathways. *Am J Respir Crit Care Med* 2020;201:A5867.
11. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, *et al*. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015;161:1202–1214.
12. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–420.
13. Schirmer L, Velmeshev D, Holmqvist S, Kaufmann M, Werneburg S, Jung D, *et al*. Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* 2019;573:75–82.
14. Arneson D, Zhang G, Ying Z, Zhuang Y, Byun HR, Ahn IS, *et al*. Single cell molecular alterations reveal target cells and pathways of concussive brain injury. *Nat Commun* 2018;9:3894.
15. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, *et al*. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 2015;16:278.
16. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417–425.
17. Piñero J, Bravo À, Queralt-Rosinach N, Gutiérrez-Sacristán A, Deu-Pons J, Centeno E, *et al*. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res* 2017;45:D833–D839.
18. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J, *et al*. The comparative toxicogenomics database: update 2019. *Nucleic Acids Res* 2019;47:D948–D954.
19. Germain M, Eyries M, Montani D, Poirier O, Girerd B, Dorfmüller P, *et al*. Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. *Nat Genet* 2013;45: 518–521.
20. Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am J Hum Genet* 2010;86:581–591.

21. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, *et al*. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;313:1929–1935.
22. Kim SV, Xiang WV, Kwak C, Yang Y, Lin XW, Ota M, *et al*. GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. *Science* 2013;340:1456–1459.
23. Rosebeck S, Leaman DW. Mitochondrial localization and pro-apoptotic effects of the interferon-inducible protein ISG12a. *Apoptosis* 2008; 13:562–572.
24. Papac-Milicevic N, Breuss JM, Zaujec J, Ryban L, Plyushch T, Wagner GA, *et al*. The interferon stimulated gene 12 inactivates vasculoprotective functions of NR4A nuclear receptors. *Circ Res* 2012;110:e50–e63.
25. Mura M, Cecchini MJ, Joseph M, Granton JT. Osteopontin lung gene expression is a marker of disease severity in pulmonary arterial hypertension. *Respirology* 2019;24:1104–1110.
26. Zheng Z, Chiu S, Akbarpour M, Sun H, Reyfman PA, Anekalla KR, *et al*. Donor pulmonary intravascular nonclassical monocytes recruit recipient neutrophils and mediate primary lung allograft dysfunction. *Sci Transl Med* 2017;9:eaal4508.
27. Regan-Komito D, Valaris S, Kapellos TS, Recio C, Taylor L, Greaves DR, *et al*. Absence of the non-signalling chemerin receptor CCRL2 exacerbates acute inflammatory responses in vivo. *Front Immunol* 2017;8:1621.
28. Makowski L, Brittingham KC, Reynolds JM, Suttles J, Hotamisligil GS. The fatty acid-binding protein, aP2, coordinates macrophage cholesterol trafficking and inflammatory activity. Macrophage expression of aP2 impacts peroxisome proliferator-activated receptor gamma and IkappaB kinase activities. *J Biol Chem* 2005;280:12888–12895.
29. Happé C, Kurakula K, Sun X-Q, da Silva Goncalves Bos D, Rol N, Guignabert C, *et al*. The BMP receptor 2 in pulmonary arterial hypertension: when and where the animal model matches the patient. *Cells* 2020;9:1422.
30. Fenster BE, Lasalvia L, Schroeder JD, Smyser J, Silveira LJ, Buckner JK, *et al*. Cystatin C: a potential biomarker for pulmonary arterial hypertension. *Respirology* 2014;19:583–589.
31. Blok IM, van Riel ACMJ, Schuuring MJ, de Bruin-Bon RHACM, van Dijk APJ, Hoendermis ES, *et al*. The role of cystatin C as a biomarker for prognosis in pulmonary arterial hypertension due to congenital heart disease. *Int J Cardiol* 2016;209:242–247.
32. Maruyama H, Dewachter C, Sakai S, Belhaj A, Rondelet B, Remmelink M, *et al*. Bosentan reverses the hypoxia-induced downregulation of the bone morphogenetic protein signaling in pulmonary artery smooth muscle cells. *Life Sci* 2016;159:111–115.
33. Spiekerkoetter E, Tian X, Cai J, Hopper RK, Sudheendra D, Li CG, *et al*. FK506 activates BMPR2, rescues endothelial dysfunction, and reverses pulmonary hypertension. *J Clin Invest* 2013;123:3600–3613.
34. Geraci MW. Targeting the prostacyclin/peroxisome proliferator-activated receptor gamma axis in lung cancer chemoprevention. *Trans Am Clin Climatol Assoc* 2018;129:48–55.
35. Tantini B, Manes A, Fiumana E, Pignatti C, Guarnieri C, Zannoli R, *et al*. Antiproliferative effect of sildenafil on human pulmonary artery smooth muscle cells. *Basic Res Cardiol* 2005;100:131–138.
36. Wharton J, Strange JW, Møller GMO, Growcott EJ, Ren X, Franklyn AP, *et al*. Antiproliferative effects of phosphodiesterase type 5 inhibition in human pulmonary artery cells. *Am J Respir Crit Care Med* 2005; 172:105–113.
37. Milara J, Escrivá J, Ortiz JL, Juan G, Artigues E, Morcillo E, *et al*. Vascular effects of sildenafil in patients with pulmonary fibrosis and pulmonary hypertension: an ex vivo/in vitro study. *Eur Respir J* 2016;47:1737–1749.
38. Li B, He W, Ye L, Zhu Y, Tian Y, Chen L, *et al*. Targeted delivery of sildenafil for inhibiting pulmonary vascular remodeling. *Hypertension* 2019;73:703–711.
39. Svoboda DL, Saddler T, Auerbach SS. An overview of national toxicology program's toxicogenomic applications: drugMatrix and ToxFX. In: Hong H, editor. Advances in computational toxicology: methodologies and applications in regulatory science Cham, Switzerland: Springer International Publishing; 2019. pp. 141–157.
40. Duluc L, Ahmetaj-Shala B, Mitchell J, Abdul-Salam VB, Mahomed AS, Aldabbous L, *et al*. Tipifarnib prevents development of hypoxia-induced pulmonary hypertension. *Cardiovasc Res* 2017;113:276–287.
41. Dumas SJ, Bru-Mercier G, Courboulin A, Quatredeniers M, Rücker-Martin C, Antigny F, *et al*. NMDA-type glutamate receptor activation promotes vascular remodeling and pulmonary arterial hypertension. *Circulation* 2018;137:2371–2389.
42. Stenmark KR, Meyrick B, Galie N, Mooi WJ, McMurtry IF. Animal models of pulmonary arterial hypertension: the hope for etiological discovery and pharmacological cure. *Am J Physiol Lung Cell Mol Physiol* 2009;297:L1013–L1032.
43. Farkas D, Alhussaini AA, Kraskauskas D, Kraskauskiene V, Cool CD, Nicolls MR, *et al*. Nuclear factor κB inhibition reduces lung vascular lumen obliteration in severe pulmonary hypertension in rats. *Am J Respir Cell Mol Biol* 2014;51:413–425.
44. Hosokawa S, Haraguchi G, Sasaki A, Arai H, Muto S, Itai A, *et al*. Pathophysiological roles of nuclear factor kappaB (NF-kB) in pulmonary arterial hypertension: effects of synthetic selective NF-kB inhibitor IMD-0354. *Cardiovasc Res* 2013;99:35–43.
45. Marsh LM, Jandl K, Grünig G, Foris V, Bashir M, Ghanim B, *et al*. The inflammatory cell landscape in the lungs of patients with idiopathic pulmonary arterial hypertension. *Eur Respir J* 2018;51: 1701214.
46. Florentin J, Coppin E, Vasamsetti SB, Zhao J, Tai Y-Y, Tang Y, *et al*. Inflammatory macrophage expansion in pulmonary hypertension depends upon mobilization of blood-borne monocytes. *J Immunol* 2018;200:3612–3625.
47. Bauer EM, Zheng H, Lotze MT, Bauer PM. Recombinant human interferon alpha 2b prevents and reverses experimental pulmonary hypertension. *PLoS One* 2014;9:e96720.
48. George PM, Oliver E, Dorfmuller P, Dubois OD, Reed DM, Kirkby NS, *et al*. Evidence for the involvement of type I interferon in pulmonary arterial hypertension. *Circ Res* 2014;114: 677–688.
49. Savale L, Sattler C, Günther S, Montani D, Chaumais M-C, Perrin S, *et al*. Pulmonary arterial hypertension in patients treated with interferon. *Eur Respir J* 2014;44:1627–1634.
50. Hu B, Guo Y, Garbacz WG, Jiang M, Xu M, Huang H, *et al*. Fatty acid binding protein-4 (FABP4) is a hypoxia inducible gene that sensitizes mice to liver ischemia/reperfusion injury. *J Hepatol* 2015;63:855–862.
51. Yu YA, Malakhau Y, Yu CA, Phelan SJ, Cumming RI, Kan MJ, *et al*. Nonclassical monocytes sense hypoxia, regulate pulmonary vascular remodeling, and promote pulmonary hypertension. *J Immunol* 2020;204:1474–1485.
52. Narasimhan PB, Marcovecchio P, Hamers AAJ, Hedrick CC. Nonclassical monocytes in health and disease. *Annu Rev Immunol* 2019;37:439–456.
53. Raychaudhuri B, Malur A, Bonfield TL, Abraham S, Schilz RJ, Farver CF, *et al*. The prostacyclin analogue treprostinil blocks NFkappaB nuclear translocation in human alveolar macrophages. *J Biol Chem* 2002;277:33344–33348.
54. Watzinger K, Tancevski I, Sonnweber T, Löffler-Ragg J. Antiinflammatory properties of PAH drugs. *Eur Respir J* 2016;48: PA1818.
55. Yeh C-H, Kuo C-H, Yang S-N, Huang M-Y, Wu H-C, Wang H-P, *et al*. Prostaglandin I2 analogs suppress tumor necrosis factor α production and the maturation of human monocyte-derived dendritic cells. *J Investig Med* 2011;59:1109–1115.
56. Aronoff DM, Peres CM, Serezani CH, Ballinger MN, Carstens JK, Coleman N, *et al*. Synthetic prostacyclin analogs differentially regulate macrophage function via distinct analog-receptor binding specificities. *J Immunol* 2007;178:1628–1634.
57. Ali Mohamed R, Galal O, Refaat Mohammed A, Salah El-Abhar H. Tropisetron modulates peripheral and central serotonin/insulin levels via insulin and nuclear factor kappa B/receptor for advanced glycation end products signalling to regulate type-2 diabetes in rats. *RSC Advances* 2018;8:11908–11920.
58. Keshavarz M, Showraki A, Emamghoreishi M. Anticonvulsant effect of guaifenesin against pentylenetetrazol-induced seizure in mice. *Iran J Med Sci* 2013;38:116–121.
59. Zhang Z, Cherryholmes G, Mao A, Marek C, Longmate J, Kalos M, *et al*. High plasma levels of MCP-1 and eotaxin provide evidence for an immunological basis of fibromyalgia. *Exp Biol Med (Maywood)* 2008; 233:1171–1180.

Single-cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning

Jason Hong, Douglas Arneson, Soban Umar, Gregoire Ruffenach, Christine M. Cunningham, In Sook Ahn, Graciel Diamante, May Bhetraratana, John F. Park, Emma Said, Caroline Huynh, Trixie Le, Lejla Medzikovic, Marc Humbert, Florent Soubrier, David Montani, Barbara Girerd, David-Alexandre Trégouët, Richard Channick, Rajan Saggar, Mansoureh Eghbali, Xia Yang

Online Data Supplement

Table E1: Clinical characteristics of patients whose lung tissue sections underwent RNA *in situ* hybridization. *History of methamphetamine exposure. PTX = pneumothorax; NSCLC = non-small cell lung cancer; FD = failed donor; nd = not done; FFPE = formalin-fixed paraffin-embedded; OCT = optimal cutting temperature; mPAP = mean pulmonary artery pressure (mm Hg); PVR = pulmonary vascular resistance (Wood units).

Diagnosis	*Age*	*Sex*	*mPAP*	*PVR*	*Prep*
PTX	20	M	nd	nd	FFPE
PTX	21	M	nd	nd	FFPE
PTX	34	M	nd	nd	FFPE
NSCLC	66-70	M	nd	nd	FFPE
NSCLC	71-75	M	nd	nd	FFPE
NSCLC	71-75	M	nd	nd	FFPE
FD	21-25	M	nd	nd	OCT
IPAH	24	F	65	9.7	OCT
IPAH	39	F	51	15.8	OCT
IPAH	25	F	66	21	OCT
IPAH	57	F	51	10.8	OCT
IPAH	52	F	69	9.3	OCT
PAH*	49	F	65	9.3	FFPE
PAH*	48	M	70	10.5	FFPE
IPAH	50	M	82	16.1	FFPE

Table E2: Integration of MCT iMΦ signature with PAH GWAS.

DEG (ortholog)	*# SNPs mapped*	*Top SNP mapped*	*-log10(P) of top SNP*
HLA-A	584	rs2517695	3.46
ZSCAN21	4	**rs7811201**	3.07
PLD3	8	rs41322049	2.9
FCGR2C	9	rs12142755	2.6
CFD	6	rs55939330	2.44
AP2S1	2	**rs6509308**	2.22
LYZ	12	rs315145	2.04
USP7	1	**rs254951**	2.03
CTSS	9	rs7511649	1.98
TREM2	1	**rs169531**	1.95
FCGR3A	1	rs6672453	1.85
THEMIS2	1	**rs3766399**	1.84
ITGB2	5	rs2006271	1.55
GNGT2	1	**rs4794029**	1.51
CTSB	6	rs2409833	1.29
API5	1	rs12224644	1.27
FTL	1	rs6509408	1.23
LDHA	4	rs10766474	1.22
MRC1	5	rs4348791	1.21
PLEK	2	rs17035378	1.18
CAPG	4	rs6886	1.11
HOPX	6	**rs4864600**	1.09
LILRA4	2	rs11084321	1.08
CSTB	2	rs11089098	1.07
GRAMD1B	1	rs17671708	1.07
SKIL	1	rs12147896	1.06
RALBP1	1	rs16955240	1.04
TNIP3	1	**rs6818199**	1.04
NDUFB2	1	rs2364397	1.03
NAA16	2	rs17198746	1.03
FAM20C	1	rs7457923	0.99
SRRM2	1	**rs2240145**	0.95
ABCA1	1	rs4149339	0.94
MRPS33	4	**rs38736**	0.88
TPR	1	rs3131550	0.84
MGST2	1	**rs6835548**	0.84

IL18	6	rs4592454	0.82
GNB2	3	rs41280986	0.79
TSPO	3	rs25245	0.73
RPS27A	1	**rs2941577**	0.69
UQCRQ	1	rs2288387	0.66
PRPF38B	2	rs41279714	0.64
ATP6V0E1	1	rs1643600	0.59
MX1	4	rs17000900	0.55
MKKS	1	rs6039931	0.54
RPL32	1	**rs3773307**	0.52
ANPEP	2	rs4932250	0.52
PRDX1	3	**rs3219492**	0.45
TMEM14C	1	rs1045911	0.45
ADPRH	2	**rs4687867**	0.4
C1orf21	1	rs6696816	0.39
ZBED5	1	rs12417274	0.39
GLUL	1	rs1325744	0.38
CCDC102A	1	**rs7185200**	0.36
NDUFC1	1	rs56290907	0.35
COPB2	1	**rs211584**	0.34
LTA4H	1	rs34579	0.34
HPGDS	1	**rs2048299**	0.32
GOLM1	1	rs11141239	0.32
APRT	1	rs12934088	0.32
PSAP	1	rs4747203	0.31

Table E3: Integration of MCT cDC signature with PAH GWAS.

DEG (ortholog)	***# SNPs mapped***	***Top SNP mapped***	***$-log_{10}$(P) of top SNP***
HLA-A	584	rs2517695	3.46
PCM1	6	**rs379066**	3.07
NUP107	9	rs2546521	2.51
USE1	3	**rs2014686**	2.27
CPNE8	7	rs58670318	2.11
RPS10	2	rs12662905	2.04
ATP8B2	5	rs9787014	2.03
RPSA	10	rs62243409	1.93
TMF1	1	rs4365633	1.91

CD200	5	rs2399422	1.68
SLFN5	10	rs1871894	1.61
ANXA5	6	rs13137821	1.47
PPIA	4	rs4720481	1.42
RPS16	1	rs11881477	1.4
STOM	1	rs2289069	1.4
PKIB	2	**rs1132630**	1.29
CCM2	4	rs3812261	1.29
FTL	1	**rs6509408**	1.23
LDHA	4	rs10766474	1.22
GM2A	4	**rs153450**	1.09
LILRA4	2	rs11084321	1.08
NUDT9	1	rs28406846	1.04
SEPT2	9	rs10933538	1.01
PSMB5	2	rs11543947	1
NIPBL	1	rs11741514	1
CCND2	1	**rs1049606**	0.97
LUC7L3	1	rs58836800	0.97
BIRC3	1	**rs7127583**	0.91
HDLBP	5	rs4675973	0.85
LAYN	2	rs11602450	0.85
FNTA	3	rs7462864	0.83
ZFAND2A	6	**rs1880296**	0.82
CALM1	2	rs2110114	0.82
NCL	1	rs10445829	0.81
CCL17	2	rs8323	0.8
VRK1	1	**rs8019921**	0.77
MMD	1	rs888206	0.76
SF3B2	2	rs72938506	0.76
PIGK	2	rs1253224	0.75
CISD2	1	rs10020659	0.72
MTCH2	1	rs35624992	0.67
PAPSS1	4	**rs2522431**	0.64
NR4A2	1	rs9646870	0.64
HERPUD1	1	rs11863728	0.63
LITAF	1	rs72781049	0.62
AIF1	1	**rs1077394**	0.61
EID1	1	rs7169052	0.57
SNRNP70	2	**rs2379085**	0.52
RPL32	1	rs3773307	0.52

NCOA6	1	rs2378260	0.49
POLDIP3	1	rs137037	0.48
PSTPIP1	1	rs12148612	0.46
NAP1L1	1	rs11829167	0.44
ZNF385A	1	rs34094446	0.43
RAMP1	1	rs3754701	0.42
RPS8	1	rs59366549	0.42
CST3	1	rs34792920	0.4
UBC	1	**rs7971677**	0.39
CD63	1	rs772254	0.38
MYLIP	1	**rs2056937**	0.36
SERPINB1	1	rs1262184	0.35
CSNK2A2	1	**rs6499938**	0.35
SLC25A28	1	rs11190439	0.32
EEF1A1	1	**rs561930**	0.32

Table E4: Integration of SuHx ncMono signature with PAH GWAS.

DEG (ortholog)	***# SNPs mapped***	***Top SNP mapped***	***-log10(P) of top SNP***
HLA-A	584	rs2517695	3.46
LYZ	12	**rs315145**	2.04
CTSS	9	rs7511649	1.98
HAGH	22	**rs3760038**	1.93
HEXA	4	rs11636289	1.89
CD274	3	**rs6476982**	1.88
DDX5	2	rs7218583	1.86
FCGR3A	1	**rs6672453**	1.85
GCH1	1	rs28703642	1.71
IL17RA	8	**rs4819962**	1.64
S1PR4	1	rs10409051	1.63
CSNK1G3	1	**rs1579036**	1.61
SLFN5	10	rs1871894	1.61
TMBIM4	2	rs55638130	1.58
NAAA	19	rs35882473	1.46
GNG10	14	rs10981012	1.46
LIG4	1	rs1931336	1.44
STOM	1	**rs2289069**	1.4
TLR2	3	rs1439167	1.38

MRC1	5	rs4348791	1.21
UQCRH	4	rs2295464	1.2
ASCC3	8	**rs9485222**	1.2
PLEK	2	rs17035378	1.18
MYBPH	2	**rs946263**	1.17
ARHGDIB	2	rs1558789	1.17
RAC2	1	**rs4821758**	1.16
N4BP2L1	1	rs12427676	1.14
FAM111A	3	rs56239606	1.07
CLEC4D	5	rs74059866	1.07
SKIL	1	rs12147896	1.06
CCL3	1	rs1634525	1.05
CLEC7A	1	rs11053628	1.05
ST20	10	rs8037808	1.02
GJB2	2	**rs1967781**	1.01
LUC7L3	1	rs58836800	0.97
UPF2	1	rs10906040	0.92
BIRC3	1	rs7127583	0.91
CST7	8	**rs227591**	0.9
NR3C1	4	rs10041520	0.88
ATF4	1	**rs2009999**	0.85
IVNS1ABP	3	rs10911729	0.82
DDX39B	9	rs62395355	0.82
OSTF1	1	rs17623732	0.82
CALM1	2	**rs2110114**	0.82
JAG1	5	rs6040439	0.76
RNASET2	4	**rs3093018**	0.74
RPL36AL	7	rs17121520	0.73
PSIP1	1	**rs3087653**	0.69
LIMK1	4	rs150866	0.69
FAM49A	2	rs17643147	0.68
ITM2B	3	rs9567953	0.65
HMHA1	1	rs4147934	0.65
SCP2	1	rs11206019	0.64
TES	1	**rs7794106**	0.63
GBP5	1	rs4556332	0.63
LITAF	1	rs72781049	0.62
RPS14	2	rs13184111	0.59
CD55	4	rs12117380	0.58
FCGRT	1	rs34400029	0.55

E17

SP110	1	rs3948464	0.54
NFKBIB	1	rs3136642	0.52
OPTN	2	rs599557	0.51
CXCL16	2	rs7221871	0.49
CCL4	1	rs712043	0.47
HNRNPM	1	rs12977861	0.46
PLAUR	1	rs4760	0.43
DSTN	1	rs6111572	0.39
SIK1	1	rs229353	0.39
NLRP3	1	rs10925027	0.39
BCAT1	2	rs2353483	0.38
BMP2K	1	rs13122868	0.37
PLK2	1	rs1558367	0.36
GPR19	1	rs11610142	0.36
TSPAN13	1	rs3807509	0.34
SSU72	1	rs3766169	0.33
CLEC4E	1	rs7299116	0.33

Table E5: Antibodies used for FACS and IF.

Target	*Supplier*	*Catalog #*	*Conjugate*	*Host*	*Clonality*	*Application*	*Dilution*
Cd64	Sino Biological	80016-R005-P	PE	Rabbit	Mono	FACS	1:400
Cd278	eBioscience	62-9949-82	Super Bright 436	Armenian hamster	Mono	FACS	1:50
Cd4	eBioscience	63-0040-82	Super Bright 600	Mouse	Mono	FACS	1:50
RT1.B	Invitrogen	MA5-17432	FITC	Mouse	Mono	FACS	1:400
Cd25	eBioscience	46-0390-82	PerCP-eFluor 710	Mouse	Mono	FACS	1:800
Cd11b/c	Invitrogen	MA5-17507	APC	Mouse	Mono	FACS	1:4000
Ccrl2	Thermo Fisher	PA575492	n/a	Rabbit	Poly	IF	1:100
Cd16	Novus	NBP2-42228AF647	Alexa Fluor 647	Mouse	Mono	IF	1:100

Mal	Bioss	bs-4693R-A488	Alexa Fluor 488	Rabbit	Poly	IF	1:100
Fabp4	**Santa Cruz**	**sc-271529 AF594**	**Alexa Fluor 594**	**Mouse**	**Mono**	**IF**	**1:100**
Cd68	Bio-Rad	MCA341A647	Alexa Fluor 647	Mouse	Mono	IF	1:100
alpha-SMA	**Novus**	**NBP2-34522AF647**	**Alexa Fluor 647**	**Mouse**	**Mono**	**IF**	**1:500**
Rabbit IgG	Invitrogen	A11012	Alexa Fluor 594	Goat	Poly	IF	1:500

Table E6: Leading-edge genes of pathways $P < 0.05$.

Leading-edge genes (human orthologs) as determined by GSEA accounting for the enrichment signal of all pathways that met a cutoff of $P < 0.05$.

Table E7: Leading-edge genes of pathways FDR < 0.05.

Leading-edge genes (human orthologs) as determined by GSEA accounting for the enrichment signal of all pathways that met a cutoff of FDR < 0.05.

Supplementary Figure Legends

Figure E1: Phenotypic characterization of PAH rat models. (A) Echocardiography images in the parasternal short-axis view showing enlarged right ventricles and "D" shaped left ventricles in end-systole in both SuHx and MCT rats as compared to control. Scale bar, 10 mm. (B) Pulsed-wave doppler of the right ventricular outflow tract showing mid-systolic flow deceleration and notching indicative of elevation of pulmonary artery pressure and pulmonary vascular resistance. The x- and y-axes represent time and velocity, respectively. Horizontal and vertical scale bars, 0.1 seconds and 400 mm/s, respectively. Vertical (C) Immunofluorescence staining showing increased α-smooth muscle actin (SMA, white) in the vessel walls of MCT and SuHx compared to control, suggesting increased vascular remodeling. RV = right ventricle; LV = left ventricle.

Figure E2: Hemodynamic assessment of PAH rat models. (A-L) Box plots showing (A) RV systolic pressure (RVSP), (B) Fulton index of RV hypertrophy (RVH), (C) RV fractional area change (RV FAC), (D) RV internal diameter (RVID), (E) pulmonary artery acceleration time (PAT), (F) ratio of PAT and pulmonary ejection time (PAT/PET), (G) stroke volume (SV), (H) cardiac output (CO), (I) cardiac index (CI), (J) total pulmonary vascular resistance index (TPVRI), (K) heart rate (HR), and (L) LV systolic pressure (LVSP). Compared to control rats, MCT and SuHx rats showed increased RVSP, RVH, RVID, and TPVRI and decreased RV FAC, PAT, PAT/PET, SV, CO, and CI. Open circular dots represent animals that underwent FACS.

Filled triangle dots represent animals that underwent scRNA-seq. Wilcoxon rank-sum test: *$P <$ 0.05, **$P \leq 0.01$, ***$P \leq 0.001$, and $P \leq 0.0001$. RV = right ventricle; LV = left ventricle; S = septum; ns = not significant.

Figure E3: scRNA-seq quality control. (A-D) Violin plots for each animal used for scRNA-seq showing (A) number of transcripts detected per cell, (B) number of genes detected per cell, (C) fraction of mitochondrial genes detected per cell, and (D) collapse rate for each cell calculated as the total number of transcripts detected for a given cell divided by the number of unique transcripts detected for that cell. Collapse rate is an indicator of how deeply a library is sequenced.

Figure E4: scRNA-seq UMAP visualizations. (A-E) UMAP plots showing cells colored by (A) number of genes detected, (B) number of transcripts detected, (C) animal of origin, (D) disease condition, and (E) lung dissociation protocol. (D) The use of different cell dissociation protocols helps resolve potential biases in tissue dissociation and recover a more comprehensive atlas of lung cell types. Batch correction methods such as canonical correlation analysis (CCA) did not further optimize clustering and cell type identification and thus were not used for downstream analyses.

Figure E5: scRNA-seq cell type markers. UMAP plots showing cells colored by expression level of cell type marker genes for all cell types identified encompassing those of myeloid, lymphoid, mesothelial, and epithelial origin.

Figure E6: *Ifi27* downregulated in human PAH lungs. Box plot showing downregulation of *IFI27* in the lungs of PAH patients (n = 15) versus control lungs (n = 11) from GEO dataset GSE11343. Expression values are the original submitter-supplied data to GEO. Wilcoxon rank-sum test.

Figure E7: Sildenafil predicted to reverse MCT's SMC signature. (A) Schematic of sildenafil's predicted reversal of the MCT SMC transcriptional signature (based on a connectivity score of -95). This prediction infers that sildenafil would inhibit the top upregulated pathway enriched in MCT's SMC signature which was EMT. (B) Dot plot showing upregulation of leading-edge genes accounting for the MCT SMC upregulation of EMT as determined by GSEA. Red and blue dots indicate MCT and SuHx, respectively. Size and color tint of dots represent strength of $-\log_{10}(P)$ values. (C) Dot plot showing predominant downregulation of the same EMT genes in (B) from a public microarray of sildenafil-exposed vs. unexposed primary rat hepatocytes (n = 3 and 61 biological replicates, respectively). Data was accessed at ToxicoDB (26) and generated by the U.S. National Toxicology Program for the toxicogenomic database DrugMatrix (27). Of the 23 genes in (B), 2 genes were not in this microarray and thus are not shown (*Lamc1 and Ctgf*). P values were determined by limma. EMT = epithelial-to-mesenchymal transition; log2FC = $\log_2$(fold change).

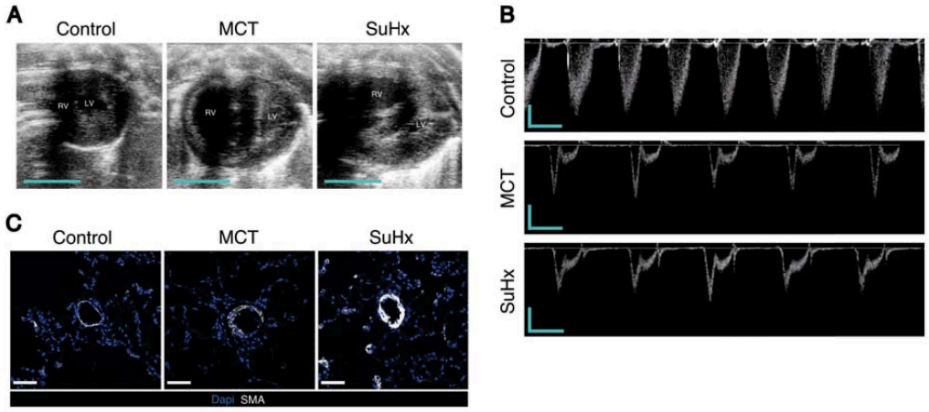
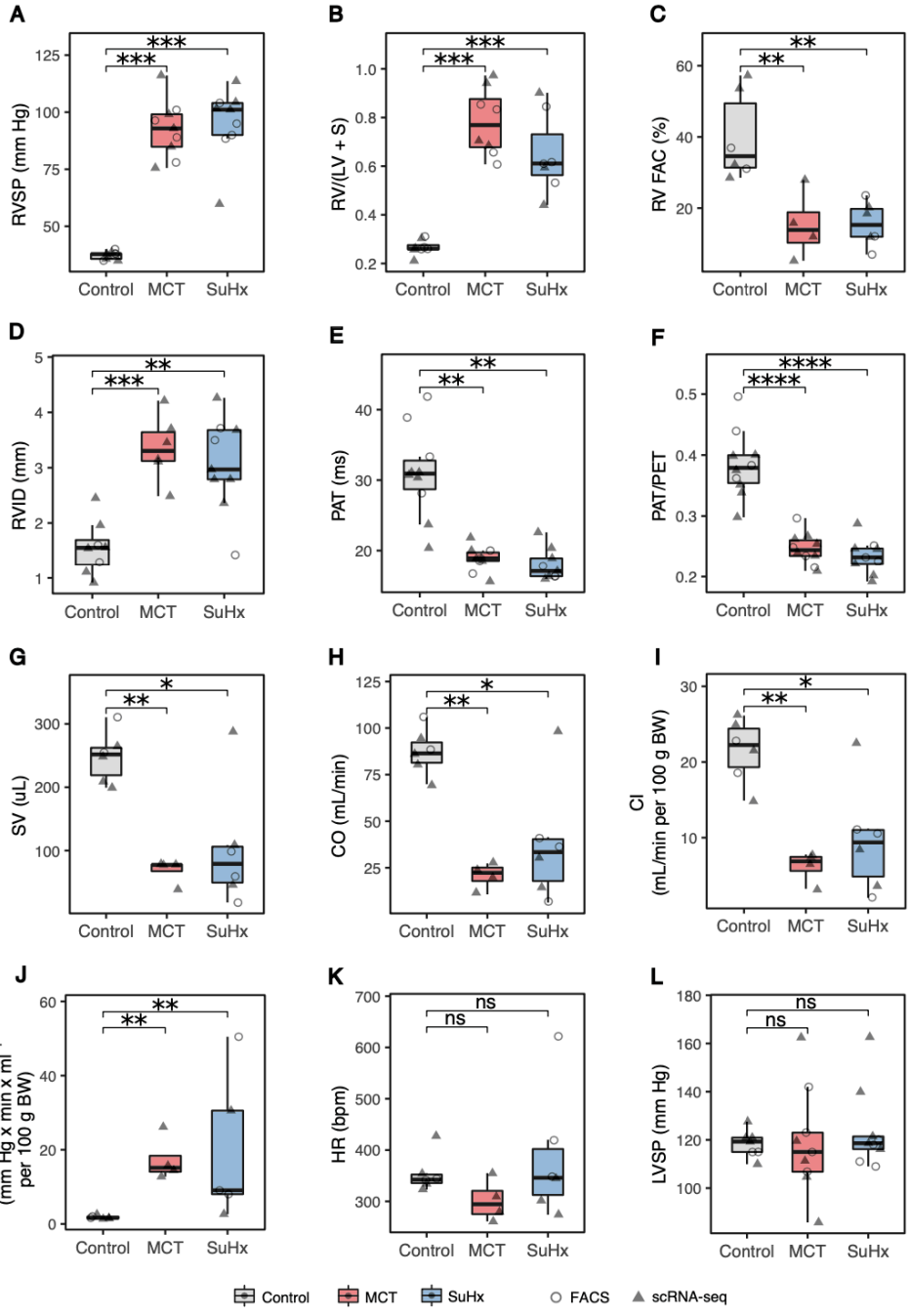

A
Control
MCT
SuHx
RV
LV
B
Control
MCT
SuHx
C
Control
MCT
SuHx
Dapi SMA

Figure E1


A
RVSP (mm Hg)
125
100
75
50
***
***
Control
MCT
SuHx
B
RV/(LV + S)
1.0
0.8
0.6
0.4
0.2
***
***
Control
MCT
SuHx
C
RV FAC (%)
60
40
20
**
**
Control
MCT
SuHx
D
RVID (mm)
5
4
3
2
1
**
***
Control
MCT
SuHx
E
PAT (ms)
40
30
20
**
**
Control
MCT
SuHx
F
PAT/PET
0.5
0.4
0.3
0.2
****
****
Control
MCT
SuHx
G
SV (uL)
300
200
100
*
**
Control
MCT
SuHx
H
CO (mL/min)
125
100
75
50
25
*
**
Control
MCT
SuHx
I
CI
(mL/min per 100 g BW)
30
20
10
*
**
Control
MCT
SuHx
J
TPVRI
(mm Hg x min x ml-1
per 100 g BW)
60
40
20
0
**
**
Control
MCT
SuHx
K
HR (bpm)
700
600
500
400
300
ns
ns
Control
MCT
SuHx
L
LVSP (mm Hg)
180
160
140
120
100
ns
ns
Control
MCT
SuHx
Control
MCT
SuHx
FACS
scRNA-seq

E28

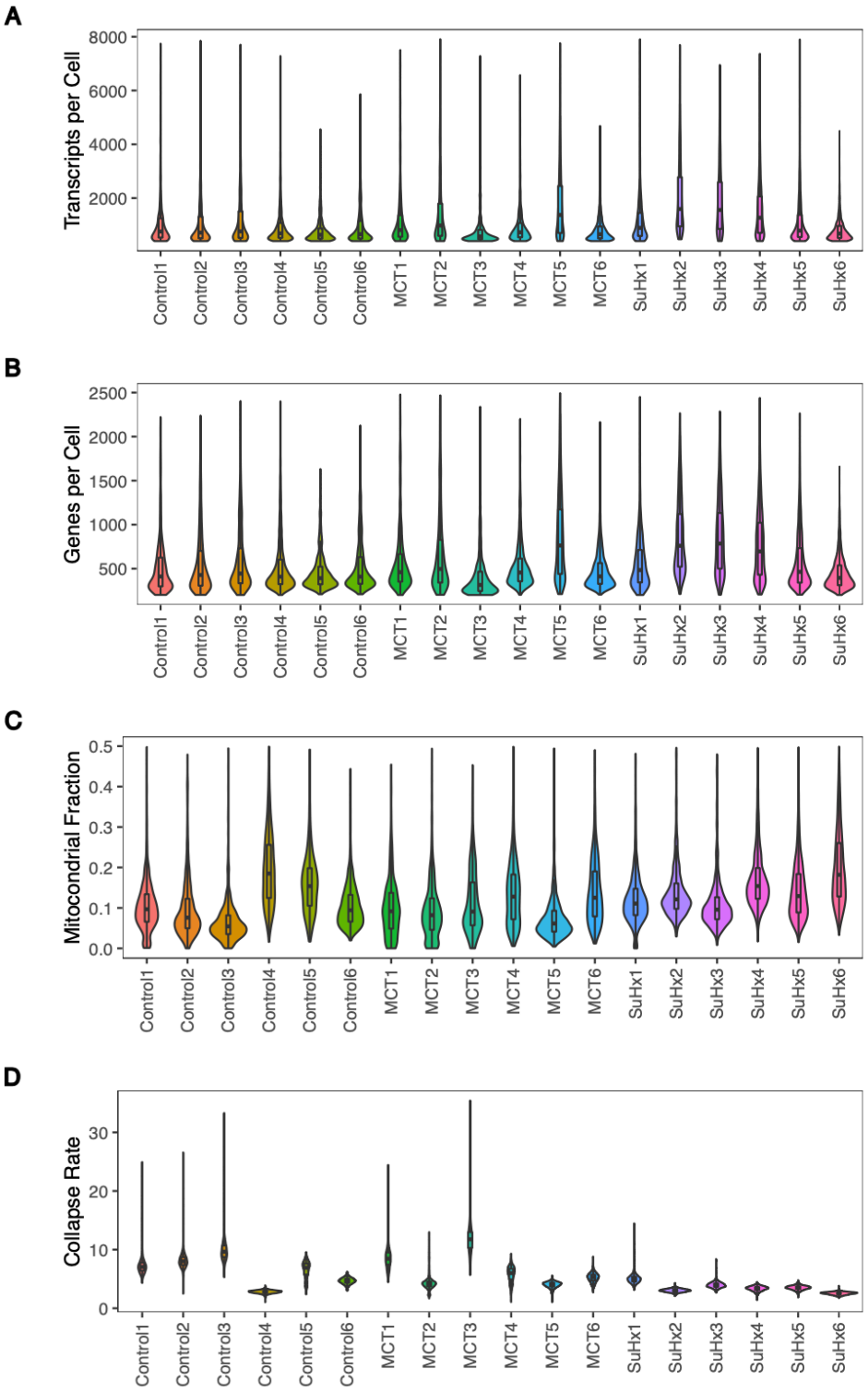
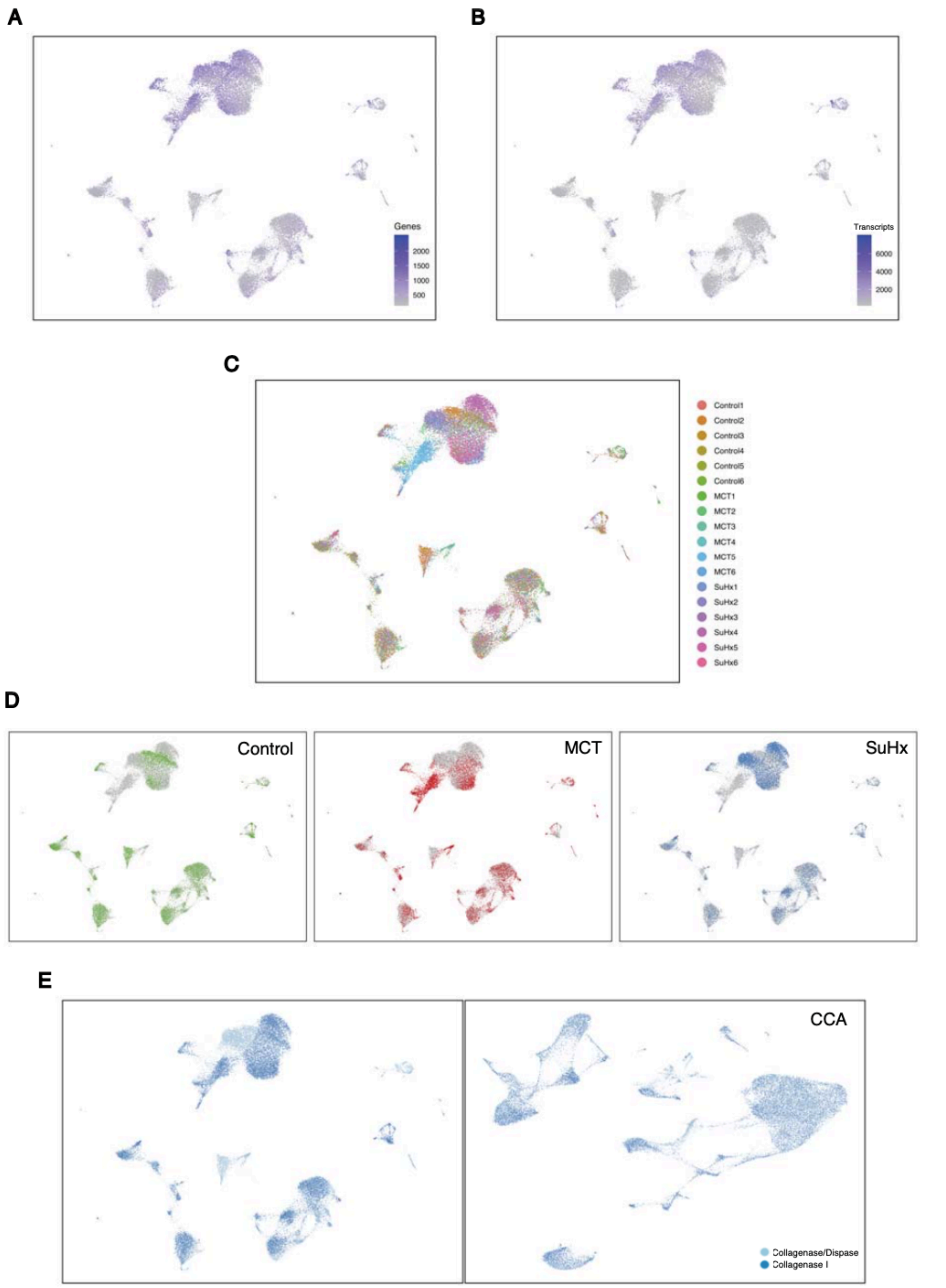


Figure E3

E29


A
Genes
2000
1500
1000
500
B
Transcripts
6000
4000
2000
C
Control1
Control2
Control3
Control4
Control5
Control6
MCT1
MCT2
MCT3
MCT4
MCT5
MCT6
SuHx1
SuHx2
SuHx3
SuHx4
SuHx5
SuHx6
D
Control
MCT
SuHx
E
CCA
Collagenase/Dispase
Collagenase I

E30


aMΦ: Mrc1
iMΦ: C1qb
pMΦ: Cenpf
cDC: Havcr2
pDC: Irf8
cMono: Fcnb
ncMono: Nr4a1
PMN: S100a9
T cells: Cd3g
T_naive: Lef1
T_reg: Il2ra
T_Serpinb6: Serpinb6
T_Cd8: Cd8a
T_prolif: Top2a
NK1: Prf1
NK2: Xcl1
ILC2: Gata3
B cells: Igkc
AT1: Ager
AT2: Sftpc
Ciliated: Dynlrb2
Club: Scgb3a2
Mesothelial: Msln
EA1: Plvap
EA2: Vwf
EC: Car4
Fb: Dcn
SMC: Acta2
Myeloid
Lymphoid
Epithelial
Stromal

E31

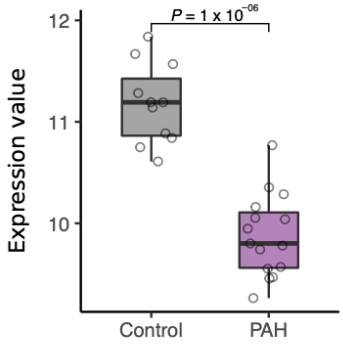


Figure E6

E32

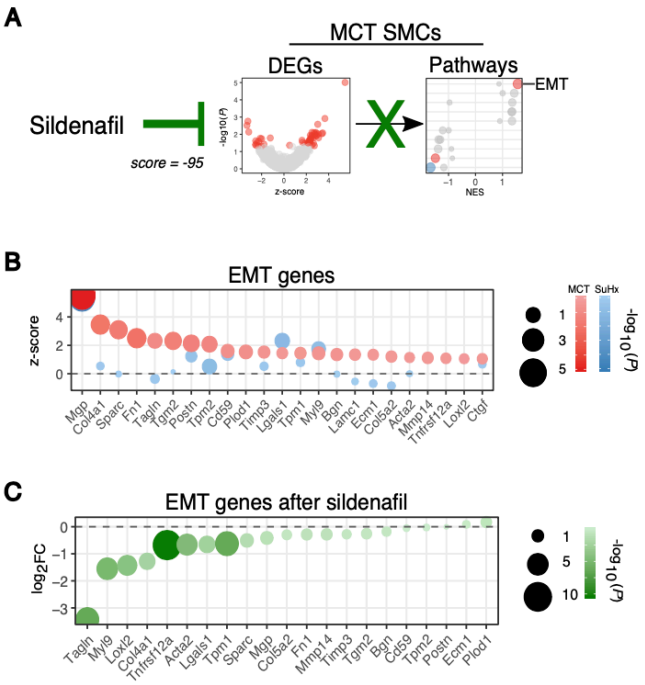


Figure E7

Cite: Hong, Jason, Douglas Arneson, Soban Umar, Gregoire Ruffenach, Christine M. Cunningham, In Sook Ahn, Graciel Diamante, May Bhetraratana, John F. Park, Emma Said, Caroline Huynh, Trixie Le, Lejla Medzikovic, Marc Humbert, Florent Soubrier, David Montani, Barbara Girerd, David-Alexandre Trégouët, Richard Channick, Rajan Saggar, Mansoureh Eghbali, Xia Yang. 2021. Single-cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning. 203(8):1006-1022.

The American Journal of Respiratory and Critical Care Medicine is an official journal of the American Thoracic Society.

CHAPTER 3

Dysregulated Genes and Pathways in the Lungs of Pulmonary Arterial Hypertension Patients

Abstract

The pathological hallmark of PAH occurs inside the lungs where the pulmonary arteries undergo irreversible remodeling. Therefore, studying the changes in gene expression that occur inside the lungs of PAH patients is critical to advance our understanding of underlying disease mechanisms. Well-powered RNAseq combined with systems analysis at the tissue-level can capture the collective and coordinated activity of a diversity of cell types that make up a complex tissue, insights that may otherwise be lost when analyzing gene expression in individual cells. In this study, we dissected the human PAH lung transcriptional landscape at the tissue level using an innovative network and systems biology methods on a large RNAseq dataset of human PAH lungs. We discovered many dysregulated genes and pathways in human PAH lungs at the tissue level, and through integration with clinical data and PAH GWAS, our network analysis revealed modules of co-expressed genes that are not only associated with PAH diagnosis and severity, but also risk of PAH implicating their causal role in PAH pathogenesis. Key driver analysis utilizing a comprehensive gene-gene regulatory network of the human lung identified and prioritized candidate genes. Furthermore, we integrated the tissue-level networks with scRNAseq to uncover the specific cell types mediating the tissue-level gene programs. Thus, our findings implicate novel genes, pathways, and cell types in PAH pathobiology.

Introduction

The pathological hallmark of PAH occurs inside the lungs where the pulmonary arteries undergo irreversible remodeling. Therefore, studying the changes in gene expression that occur inside the lungs of PAH patients is critical to advance our understanding of underlying disease mechanisms. Advances in sequencing have enabled more expansive molecular profiling of

tissues which combined with rapidly evolving bioinformatic methods have led to major biological discoveries. However, such technological and methodological advances remain relatively unexploited in studying the lungs of PAH patients, who do not routinely undergo lung biopsies and rarely undergo lung transplantations. Thus, relatively small sample sizes have limited prior molecular studies of PAH lungs.

The Pulmonary Hypertension Breakthrough Initiative (PHBI) is a consortium of 13 academic centers that maintain a biorepository of lung specimens from PAH patients undergoing lung transplant and from donors to serve as control. A recent study by our collaborators at PHBI analyzed the transcriptomes of 58 PAH lungs[1], the largest to date, but was limited by use of microarray which requires a predefined probe set to detect gene expression. To our knowledge, PAH lung studies have yet to leverage RNA sequencing (RNAseq) which is more comprehensive and sensitive than microarray profiling. While single-cell RNA sequencing (scRNAseq) is now available, well-powered bulk RNAseq of a large sample size combined with state-of-the-art systems biology and multi-omic integration at the tissue level can capture the collective and coordinated activity of a diversity of cell types that make up a complex tissue, insights that may otherwise be lost when analyzing gene expression in individual cells. In this chapter, we employ bulk RNAseq leveraging a large biorepository of PAH and healthy control lungs to perform an in-depth and well-powered investigation into the tissue-level alterations of genes and pathways. We further integrate bulk RNAseq with scRNAseq to deconvolute the contribution of individual cell types and cell-cell interactions in PAH lungs.

Materials and Methods

Bulk RNAseq dataset and quality control

Through collaboration with Mark Geraci, MD and Bob Stearman, PhD at PHBI, we obtained bulk RNAseq from explanted lungs of 96 patients with pulmonary hypertension (94 with WHO Group 1 PAH) and 52 failed donor (FD) lungs which did not have an appropriate recipient but still met physiologic standards to serve as control[1]. We mapped raw reads to the reference human genome (hg19) and assembled transcripts with HISAT2 and StringTie[2], respectively. We then performed hierarchical clustering and principal components analysis (PCA, **Fig. 3**) to identify potential outliers for removal and technical batch effects for correction using ComBat[3]. We focused on WHO Group 1 PAH (n=94) in downstream analyses given the likely significant differences in the underlying pathophysiology compared to WHO Group 4 (n=2).

Differentially expressed genes (DEGs)

We determined DEGs between PAH and control using DESeq2[4]. While the analysis will focus on WHO Group 1 PAH versus control, we will also perform pairwise analysis to assess for differences between PAH subgroups including idiopathic (IPAH, n=41), pulmonary veno-occlusive disease (n=7), familial (n=8), connective tissue disease (n=11), congenital heart disease (n=19), and drug-induced (n=8).

Co-expression networks

We performed Weighted Gene Co-expression Network Analysis (WGCNA)[5] to identify modules of co-expressed genes likely controlled by the same transcriptional regulatory program, functionally related, or members of the same pathway[6].

Clinical integration

The PHBI subjects were deeply phenotyped enabling us to correlate the first principal component (PC) of WGCNA modules with clinical data such as disease status (i.e. PAH vs control), mean pulmonary artery pressure (mPAP), pulmonary vascular resistance (PVR), cardiac output, six-minute walk test (6MWT), degree of hypoxemia (i.e. PaO_2/FiO_2 ratio), prostacyclin use, pulmonary function tests (i.e. DLCO) and PAH subgroup.

Pathway enrichment of DEGs and coexpression modules

We evaluated the functional role of DEGs and WGCNA modules by testing for enrichment in known biological pathways using Gene Set Enrichment Analysis (GSEA)[7].

PAH GWAS integration

To assess for causal molecular links between risk of PAH and advanced-stage PAH lungs, we tested DEGs and WGCNA modules for enrichment in single-nucleotide polymorphisms (SNPs) from PAH genome-wide association studies (GWAS) using Mergeomics[8], a computational pipeline developed by our lab that integrates multi-omics data for mechanistic discoveries. To evaluate for consistency across methods and across GWAS cohorts, we a) implemented two other GWAS enrichment approaches: MAGMA[9] and GSA-SNP2[10] and b) utilized four independent GWAS cohorts as well as their meta-analysis totaling 2,085 PAH and 9,659 control subjects[11] obtained from our collaborator Christopher Rhodes, PhD with approval from key stakeholders in UK, US and France. We mapped GWAS SNPs to genes using two approaches: a) chromosomal proximity to transcription start sites (20 kilobases) and b) eQTLs derived from 1626 human lungs[12,13].

Bayesian networks (BNs)

A complementary approach to co-expression networks are BNs which model gene-gene regulation and can flexibly incorporate various types of information such as genetic causality, expression quantitative trait loci (eQTL), and transcription factor (TF) binding to model directional gene regulatory networks[14–20]. The combination of these two network types is particularly powerful in uncovering novel biological insights, as co-expression networks offer a more global view of gene co-regulation and BNs provide granular and directional regulatory relationships. Therefore, we routinely construct and utilize these complementary networks based on multi-omics data[14–16,21–32]. We constructed and took the union of three distinct human lung BNs derived from 148 PHBI, 577 GTEx[12], and 1,343 GSE23546[13] samples to build a comprehensive gene regulatory network of the human lung. BNs were constructed using Reconstructing Integrative Molecular Bayesian Network (RIMBANet)[19]. For this method, 1000 networks were generated from different random seed genes using continuous and discrete expression data as well as cis eQTL and transcription factor data as priors. Then, the final network was obtained by taking a consensus network from the 1000 randomly generated networks whereby only edges that passed a probability of >30% across the 1000 BNs were kept. We projected WGCNA modules associated with PAH onto this regulatory network and performed weighted Key Driver Analysis to identify hub genes of PAH[8]. We then constructed and performed differential network analysis on separate PAH and control networks using DyNet[33] to find genes and subnetworks most rewired in PAH.

Statistical power

A minimum of 5 samples per group is needed for adequate power to detect DEGs based on a variety of RNAseq datasets[34] and 100 total samples are needed for network analyses[18]. To maximize power, we included 94 WHO Group 1 PAH and 52 FD lungs.

Results

Bulk RNAseq samples met quality control for downstream analyses

Prior to performing integrative multiomics and network analyses (**Figure 3.1**), RNAseq samples were checked for quality control. Principal components analysis (PCA) of bulk RNAseq samples showed that the majority of samples clustered together. When colored by disease status, PCA revealed global differences between PAH and FD samples as a whole (**Figure 3.2**). Hierarchal clustering revealed two outliers, a 55 year-old Hispanic female FD and a 13 year-old White male with PAH, which we removed for downstream analysis (**Figure 3.3**).

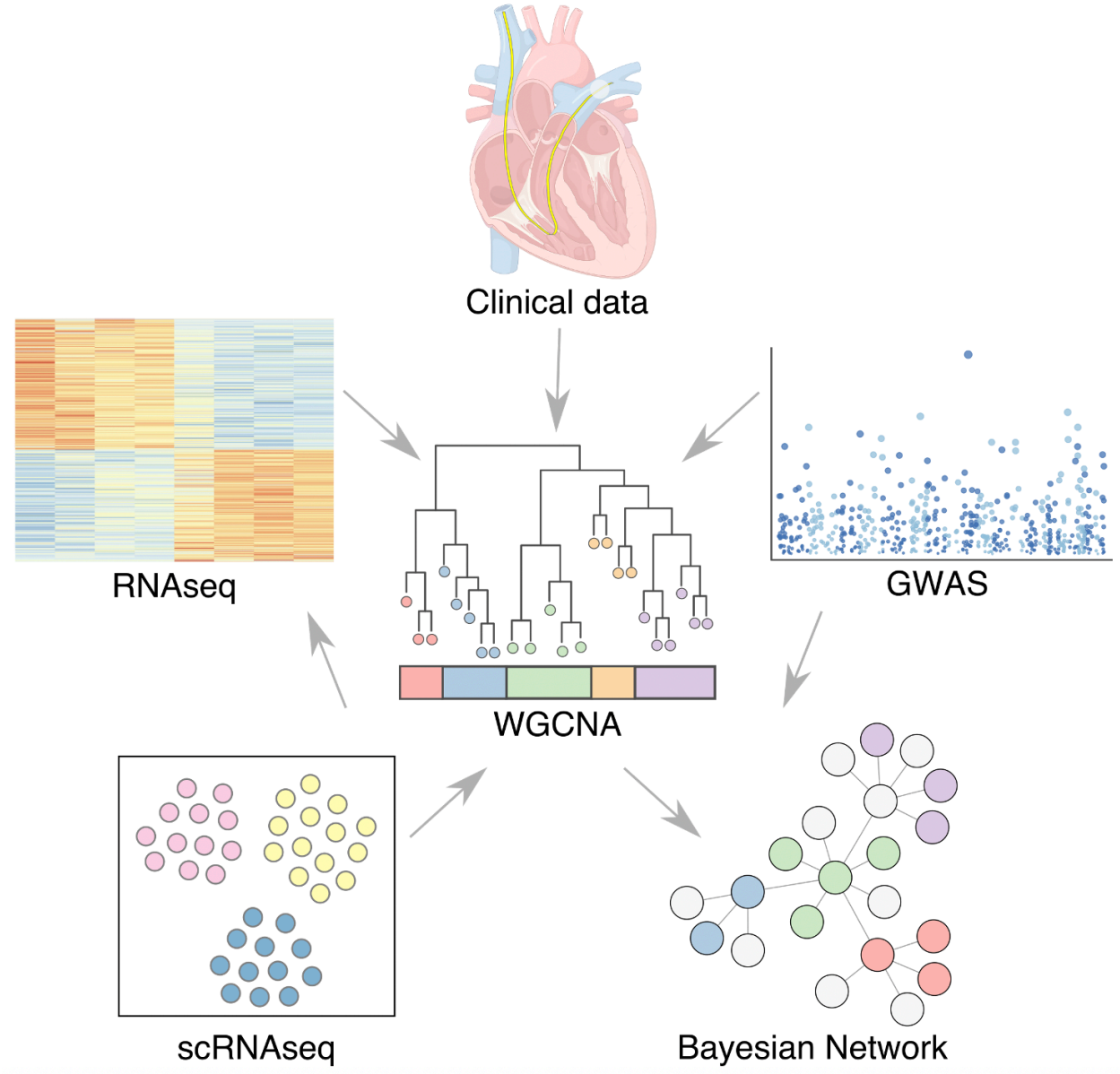


Figure 3.1. Overall study design for integrative multiomics and network analyses of bulk RNAseq.

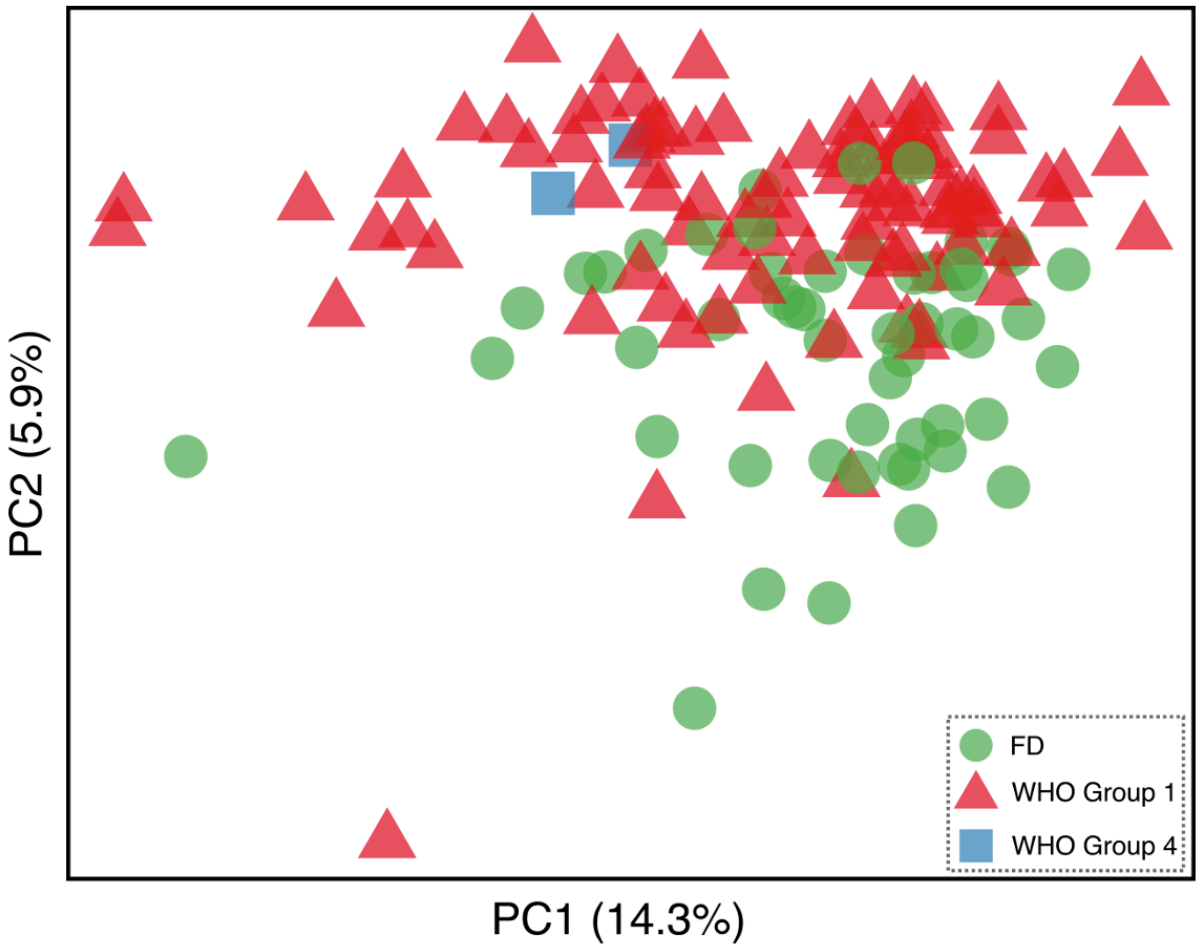


Figure 3.2. Lung RNAseq samples cluster by disease status. PCA of 148 lung RNAseq samples.

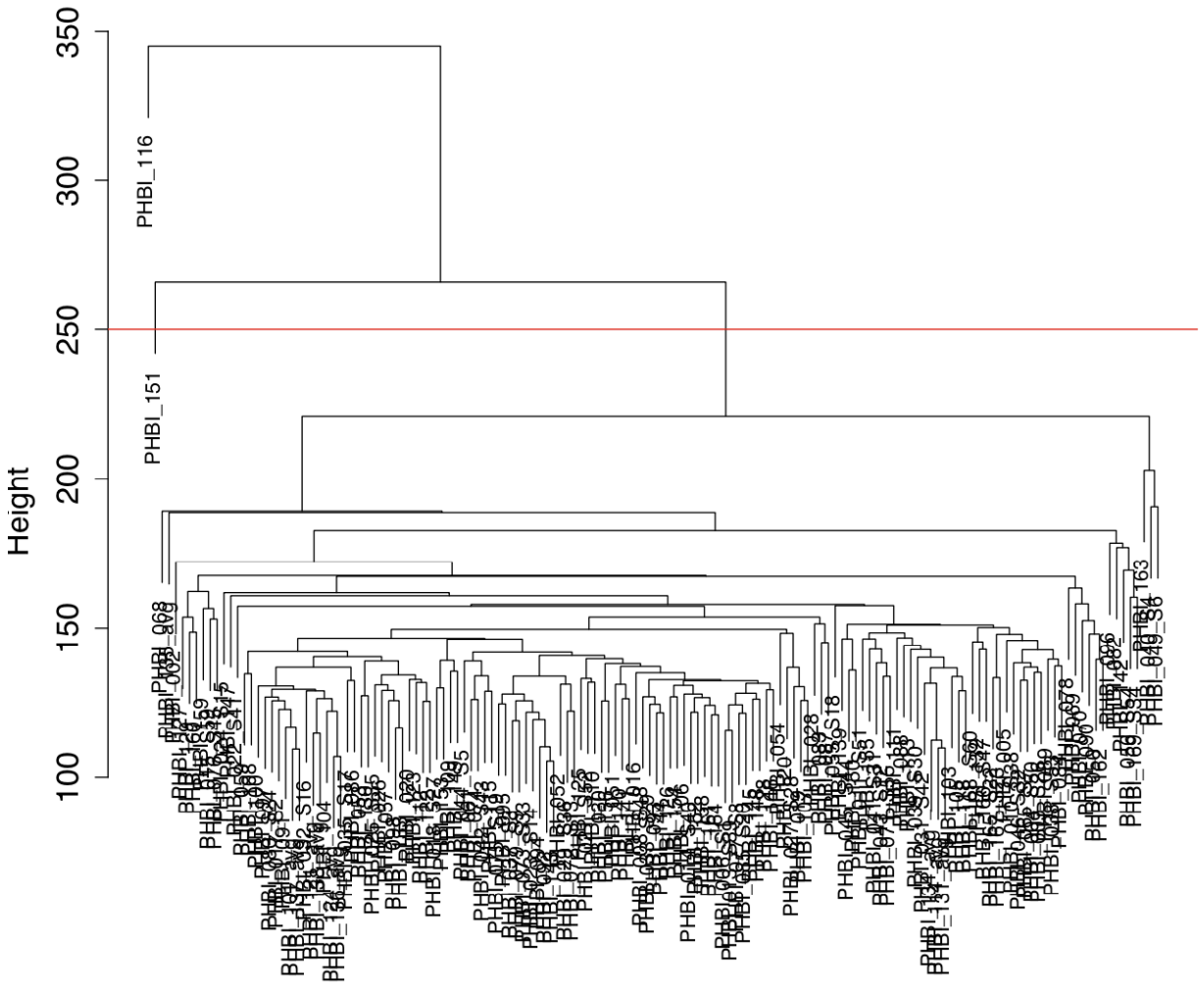


Figure 3.3. Two lung RNAseq sample outliers identified by hierarchal clustering. Dendrogram showing hierarchal clustering of 148 lung RNAseq samples. Horizontal red line represents the heigh above which outliers were removed.

Many genes are altered in PAH lungs

Differential expression analysis comparing PAH and FD lungs revealed 2719 upregulated and 2534 downregulated in PAH with a total of 5253 dysregulated genes (FDR < 0.05) (**Figure 3.4**). Top upregulated genes included *HBA2*, *HBB*, *LAMP5*, *HBA1*, and *MFAP4*. Top downregulated genes included *SIGLEC10*, *PI3*, *SAA2*, *SLC36A1*, and *ALPP*.

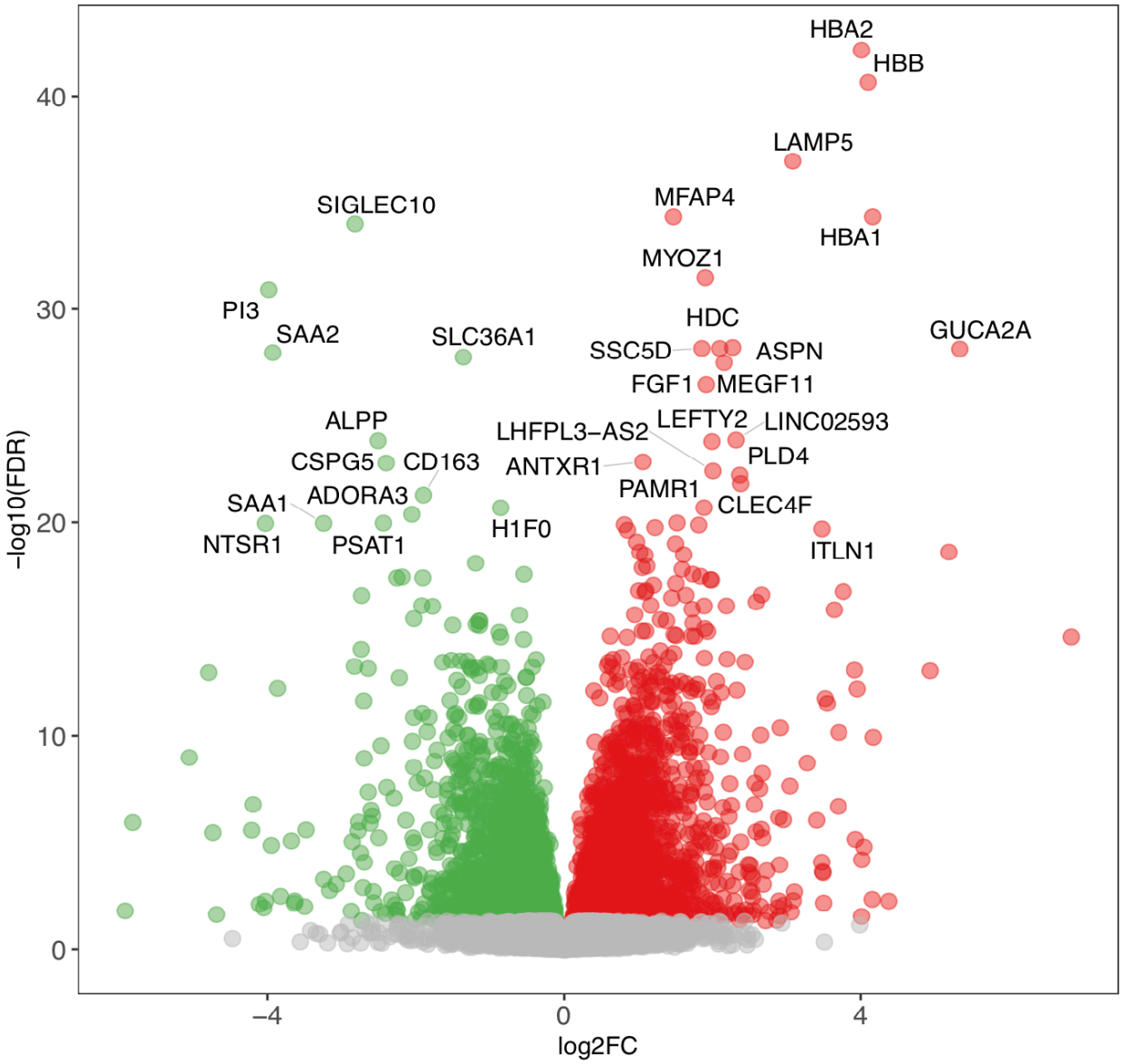


Figure 3.4. Differential gene expression analysis shows many genes dysregulated in PAH lungs. Volcano plot showing genes differentially expressed where red (upregulated) or green (downregulated) dots correspond to genes with FDR < 0.05 and grey dots represent genes not statistically significant. Select top upregulated and downregulated genes are labeled. FDR = false discovery rate; FC = fold change.

Many pathways are altered in PAH lungs

Pathway enrichment analysis revealed that many pathways are dysregulated in PAH lungs as compared to FD (**Figure 3.5**). Top upregulated pathways included epithelial mesenchymal transition, hedgehog signaling, and apical junction. Top downregulated pathways included mTORC1 signaling, G2M checkpoint, and MYC targets.

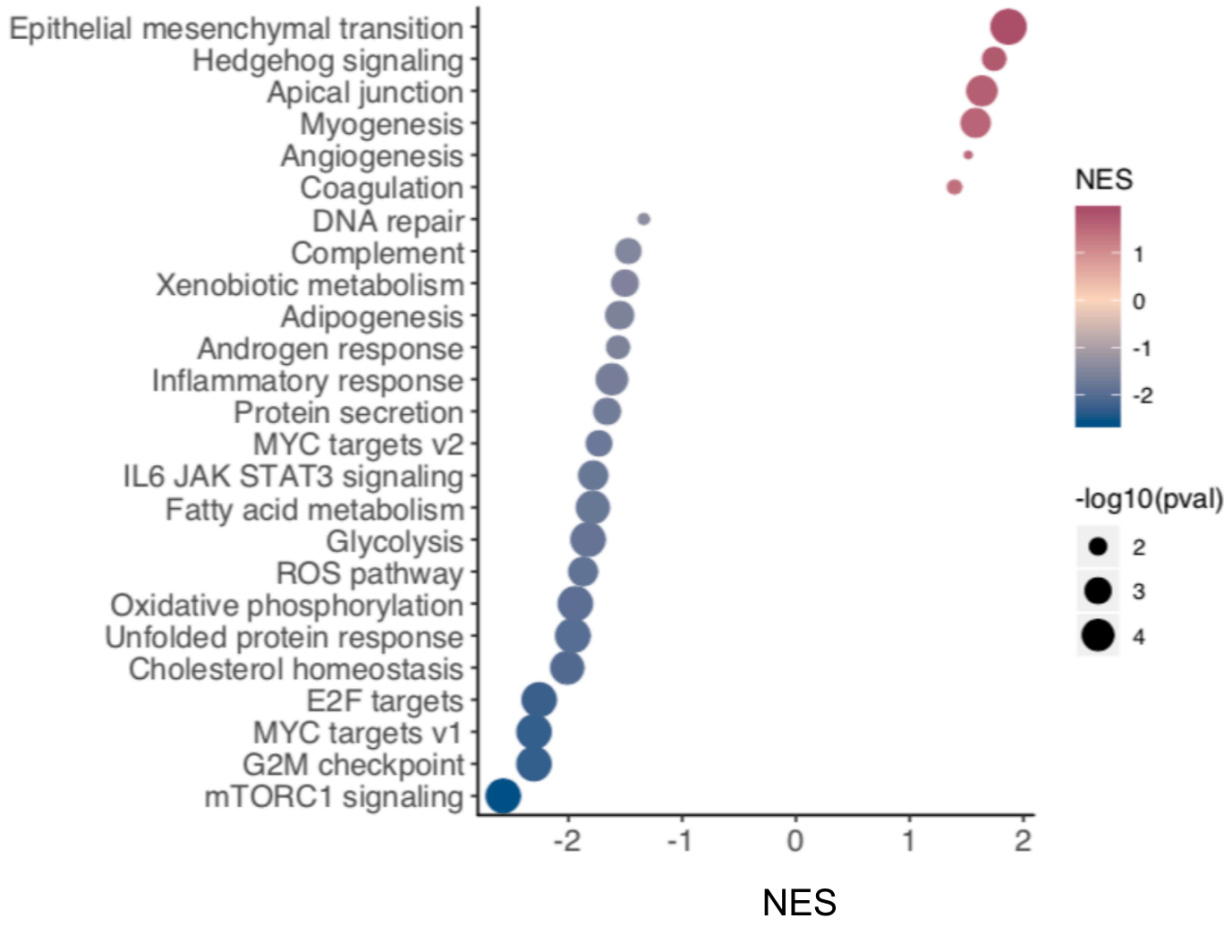


Figure 3.5. Many pathways are dysregulated in PAH lungs. Red dots indicate upregulated pathways. Blue dots indicate downregulated pathways. Larger dots indicate stronger statistical significance. NES = normalized enrichment score derived from GSEA; pval = p value.

Distinct modules of co-expressed genes identified in PAH lungs

WGCNA analysis revealed 20 distinct co-expression modules with the largest being the turquoise module at 5391 genes and the smallest being the orange module at 42 genes, with a median size across modules of 140.5 genes (**Figure 3.6**). Pathway enrichment analysis for each WGCNA module as determined by GSEA identified many pathways up- and downregulated

(**Figure 3.7**). The top upregulated pathway was TNF alpha signaling via NFkB in the lightcyan module and the top downregulated pathway was MYC targets v1 in the black module.

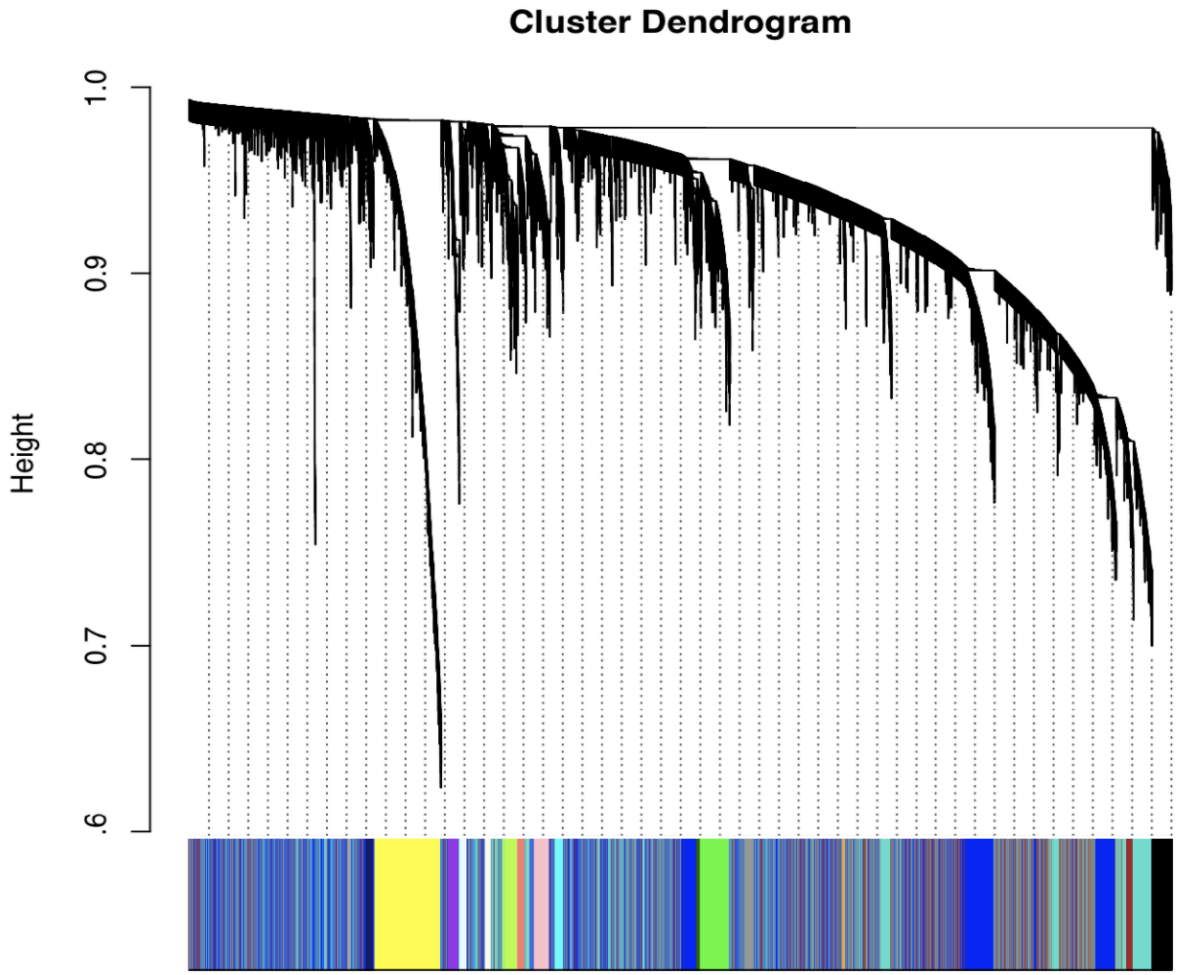


Figure 3.6. Distinct modules of co-expressed genes identified in PAH lungs. Dendrogram showing clustering of genes, with dissimilarity based on topological overlap as determined by WGCNA analysis. Module color assignments for each gene is also shown.

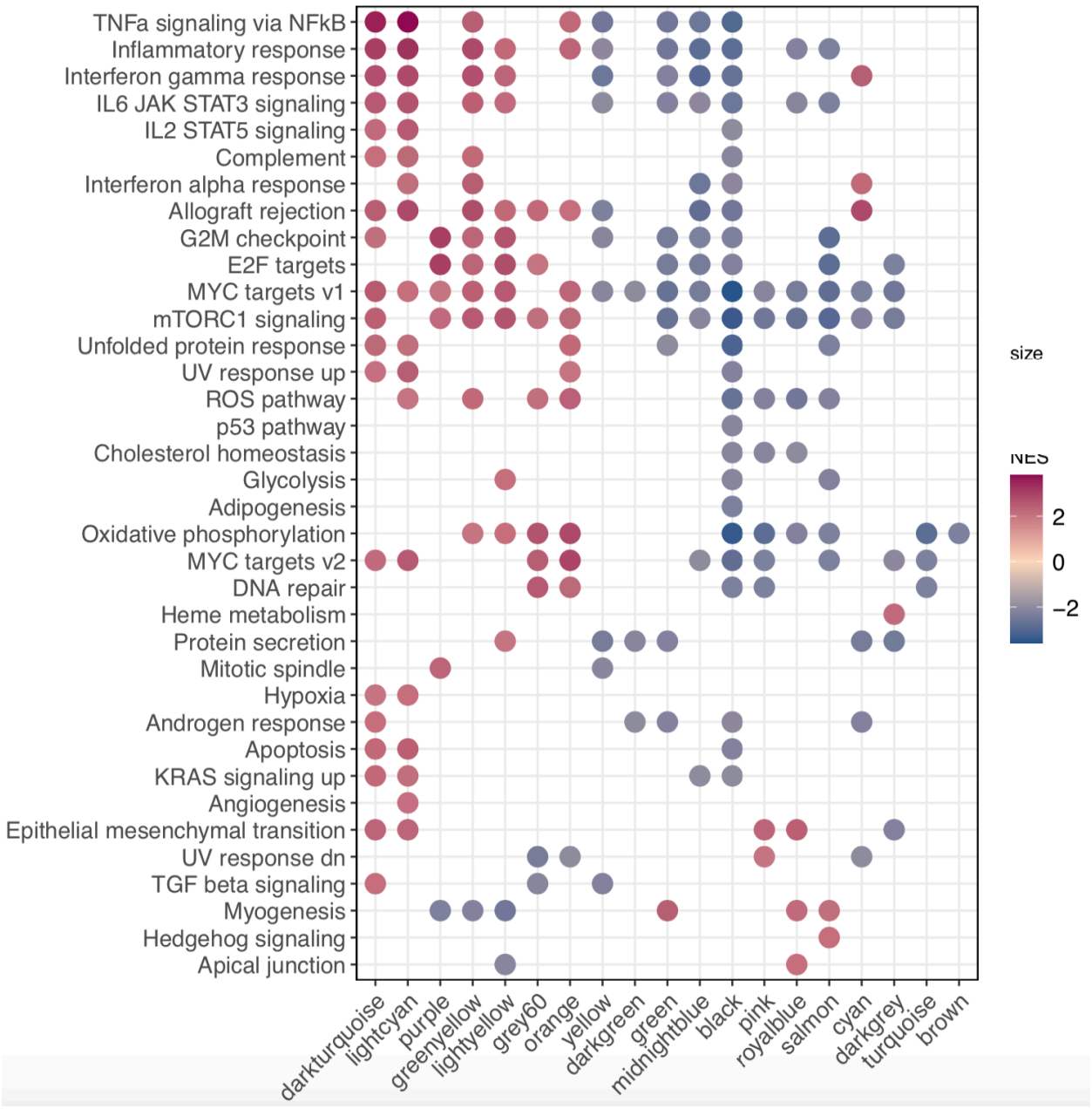


Figure 3.7. Various pathways are involved in lung modules of co-expressed genes. Dots represent statistically significant pathways enriched in WGCNA modules with FDR < 0.01 that also had the strongest NES scores (>2 or <2) where red signifies upregulation and blue signifies downregulation. NES = normalized enrichment score.

WGCNA lung modules associated with PAH diagnosis and severity

Correlation analysis of the first PC of WGCNA modules with clinical data revealed that the pink, royalblue, salmon, and black modules positively correlate with both PAH diagnosis and PAH severity (mPAP and PVR) (**Figure 3.8**). The pink module was also negatively correlated with cardiac output (CO) which also suggests correlation with more severe PAH. Pink and royalblue were also positively correlated with FVC/DLCO suggesting worse diffusion impairment relative to forced vital capacity, and thus more severe disease. Salmon and royalblue were negatively correlated with PaO2/FiO2 suggesting more severe hypoxemia and thus more severe disease. When repeating the same analysis but with the 5253 PAH DEGs as its own module, only PAH diagnosis was correlated with the DEG module. The other clinical traits related to disease severity were not correlated with the DEG module. GSEA analysis of PAH vs FD DEGs with WGCNA modules as the gene sets showed that the modules that were positively or negatively correlated with PAH diagnosis were also positively or negatively enriched for PAH DEGs by GSEA. Out of all the modules, the pink module had the strongest positive enrichment for PAH DEGs.

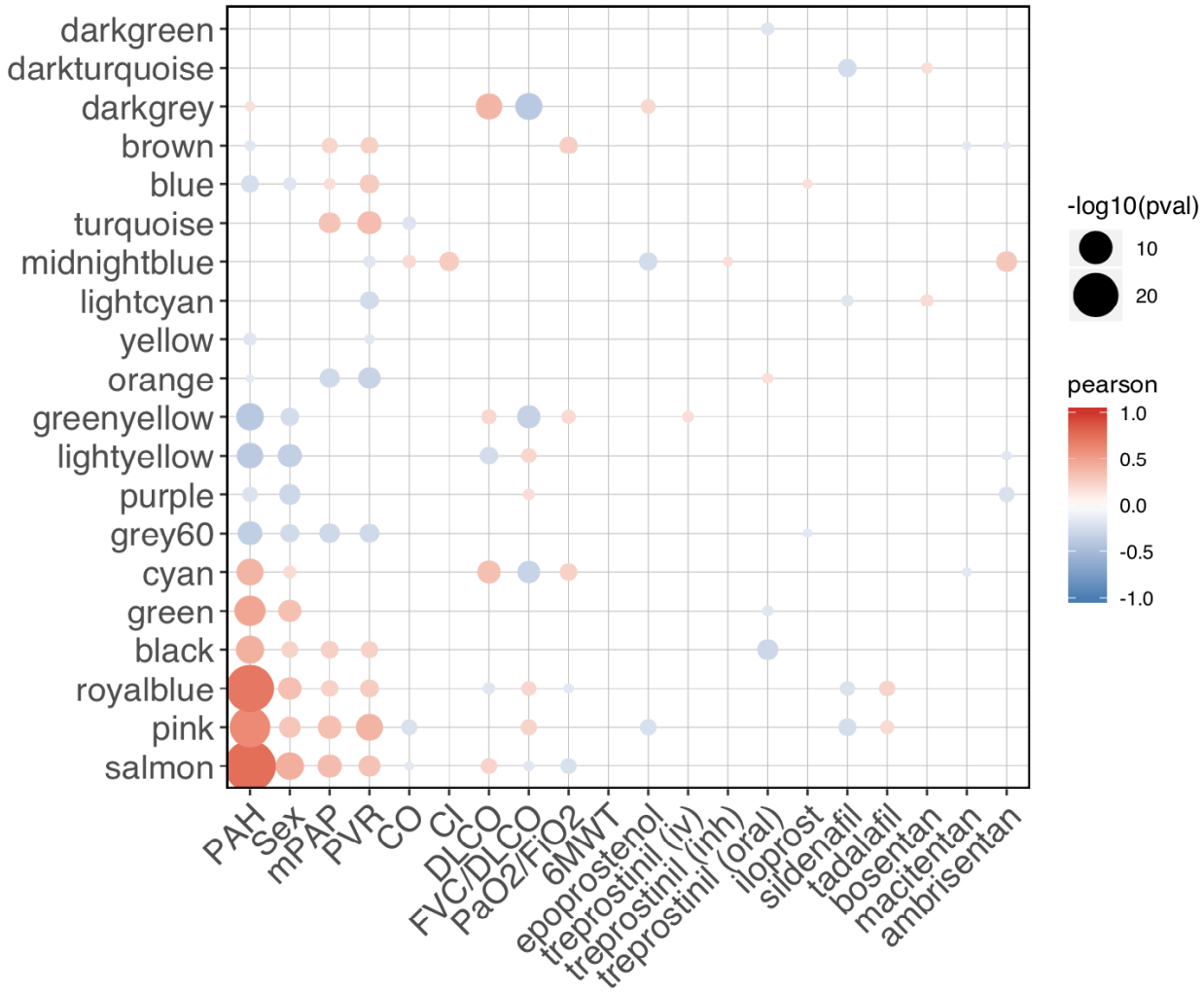


Figure 3.8. WGCNA lung modules associated with PAH diagnosis and severity. Dots represent statistically significant Pearson correlation between clinical data and the first PC of WGCNA modules (P value < 0.05). Red represents positive correlation and blue represents negative correlation. Dots with larger size denote stronger statistical significance. Pval = p value.

WGCNA lung modules associated with PAH genetic risk

To infer causal roles of the PAH-associated DEGs and WGCNA modules, we assessed the enrichment of DEGs and modules for PAH GWAS signals using three distinct approaches. Mergeomics Marker Set Enrichment Analysis (MSEA) revealed that the black module was

significantly enriched for PAH GWAS in multiple cohorts (PHAAR, PAHB, and GWAS meta-analysis of all 4 cohorts) using two different approaches to mapping SNPs to genes (20 kilobase chromosomal proximity to transcription start site and lung eQTLs). The pink module was also significantly enriched for PAH GWAS using two other GWAS enrichment methods (MAGMA and GSA-SNP2) across multiple different GWAS cohorts (PHAAR, PAHB, and BHFPAH). Using GSA-SNP2, the PAH-associated DEG module totaling 5253 genes showed statistically significant enrichment in GWAS signals in the BHFPAH cohort and GWAS meta-analysis, as well as relatively high enrichment in the other cohorts (though not statistically significant after multiple testing correction). PAH-associated DEGs were not enriched in GWAS signals using the MAGMA approach and were not included in the Mergeomics analysis (default maximum module size of 500 genes was used).

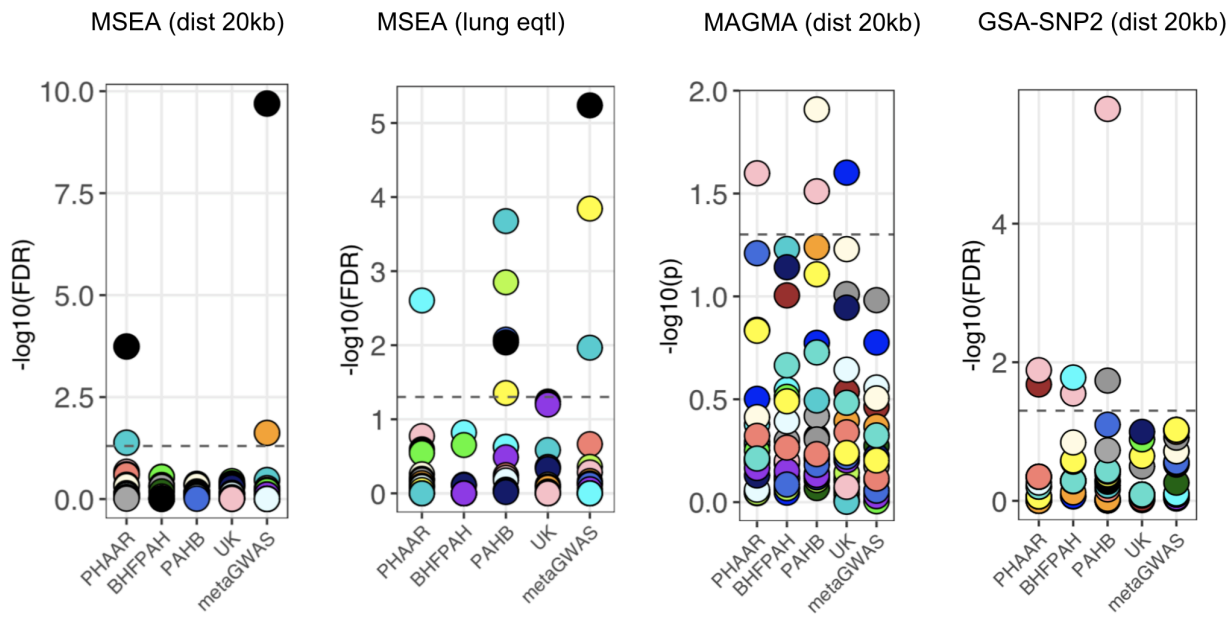


Figure 3.9. WGCNA lung modules associated with PAH risk. Dot plots showing GWAS enrichment of WGNCA modules across GWAS cohorts and GWAS enrichment methods. Colors represent corresponding WGCNA color modules. Horizontal dashed lines represent statistical significance threshold. "Dist 20kb" signifies SNPs were mapped to genes if within 20 kilobases of the transcription start site. "Lung eqtl" signifies that SNPs were mapped to genes if they were also lung eQTLs to the corresponding gene.

Bayesian network analysis reveals key genes implicated in PAH

To identify potential regulators of the PAH-associated DEGs and coexpression modules, we performed weighted key driver analysis (wKDA) using Bayesian networks (BNs). We first constructed and took the union of three distinct human lung BNs derived from 148 PHBI, 577 GTEx[12], and 1,343 GSE23546[13] samples to build a comprehensive gene regulatory network of the human lung consisting of 22,444 nodes and 62,151 edges (**Figure 3.10**). We projected PAH-associated DEGs and WGCNA modules onto this regulatory network and performed wKDA to

identify hub genes of PAH[8] (**Figure 3.11**). We also projected known human PAH-associated gene sets obtained from DisGeNET [35] and Comparative Toxicogenomics Database [36] onto the lung BN to see how they would localize with respect to WGCNA modules.

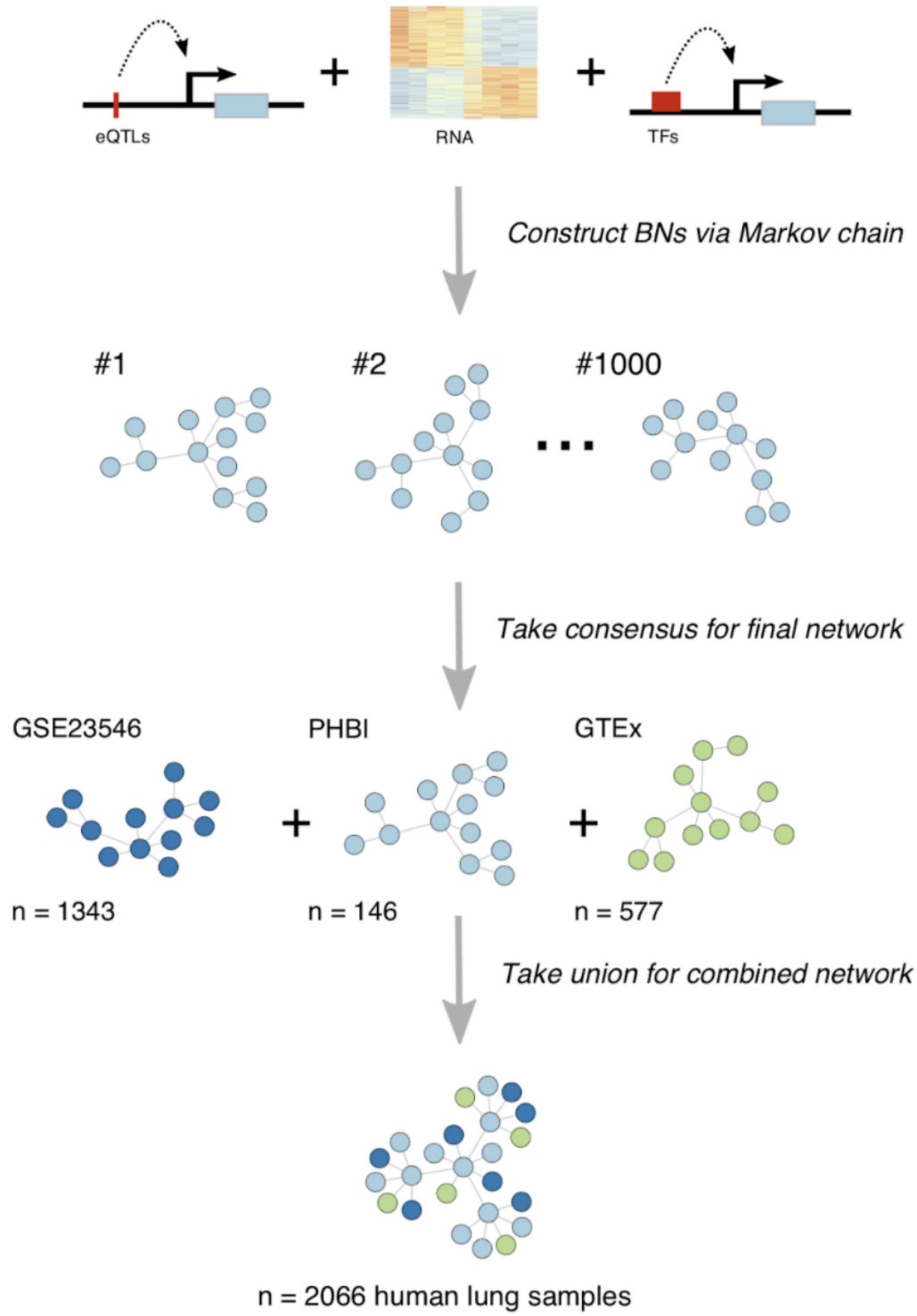


Figure 3.10. Schematic of Bayesian gene-gene regulatory network construction for the human lung.

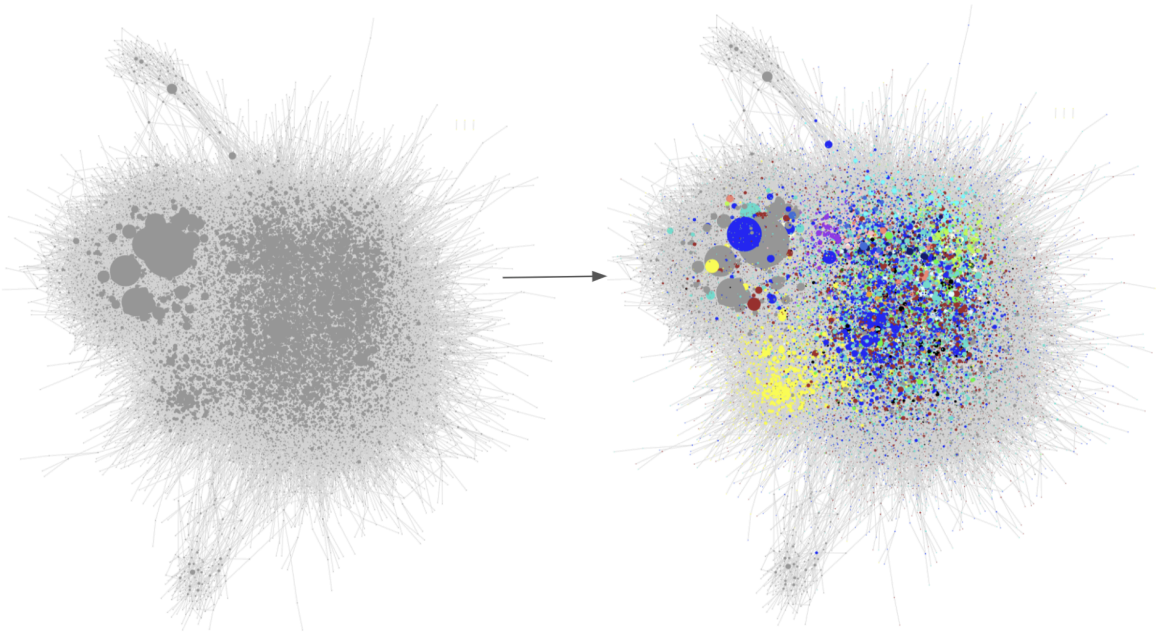


Figure 3.11. WGCNA modules projected onto Bayesian gene-gene regulatory network of the human lung.

Focusing our attention on modules associated with PAH diagnosis, severity, and risk, we found *PDE7B* to be the most central gene in the pink subnetwork by KDA (**Figure 3.12**). *PDE7B* encodes the phosphodiesterase 7B protein which hydrolyzes the second messenger cAMP. Many known PAH genes were also connected to pink subnetwork including *BMPR2*, the most well-established causal PAH gene. *PDE7B* was also a key driver gene of PAH-associated DEGs.

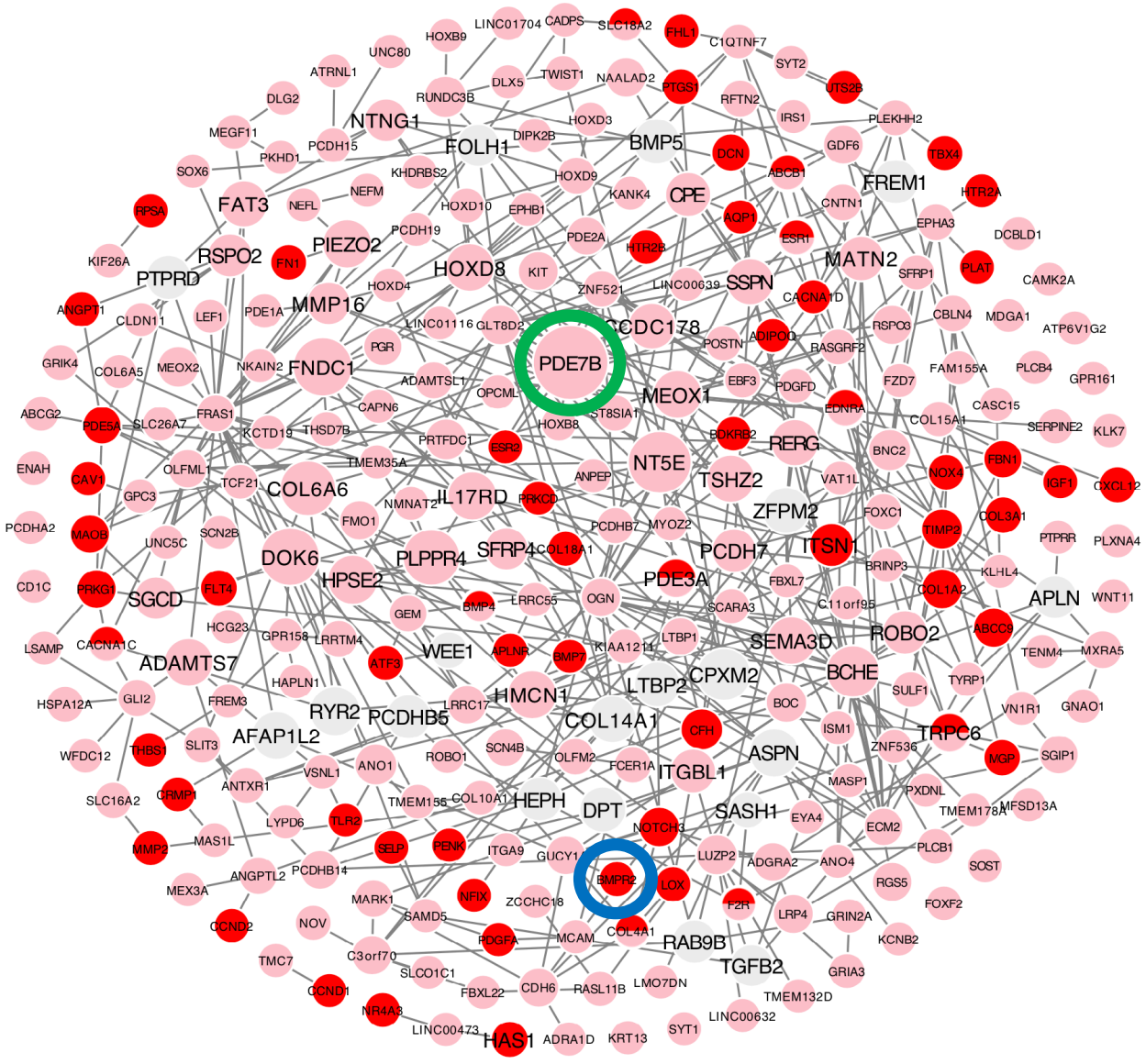


Figure 3.12. PDE7B as the top key driver gene in the pink subnetwork. Pink subnetwork of the lung BN where pink nodes represent genes in the pink module, grey nodes represent genes not in the pink module but connected to the pink subnetwork, red nodes represent known PAH genes connected to the pink subnetwork, and red and pink nodes represent known PAH genes that are also members of the pink module. Larger size nodes correspond to stronger statistical strength of the gene as a key driver of the pink subnetwork. *PDE7B* is highlighted by a green circle and *BMPR2* is highlighted by a blue circle.

We then constructed and performed differential network analysis on separate PAH and control BN using DyNet[33] to find genes and subnetworks most rewired in PAH. We found that the most rewired node whose expression is also upregulated in PAH lungs was MELTF-AS1, a long non-coding RNA increasingly recognized as an important pathogenic driver in many cancers[37–39] but never before implicated in PAH (**Figure 3.13** and **Figure 3.14**). In KDA using the lung BN, MELTF-AS1 was also a central hub gene of the black WGCNA module which is associated with PAH diagnosis, severity, and PAH.

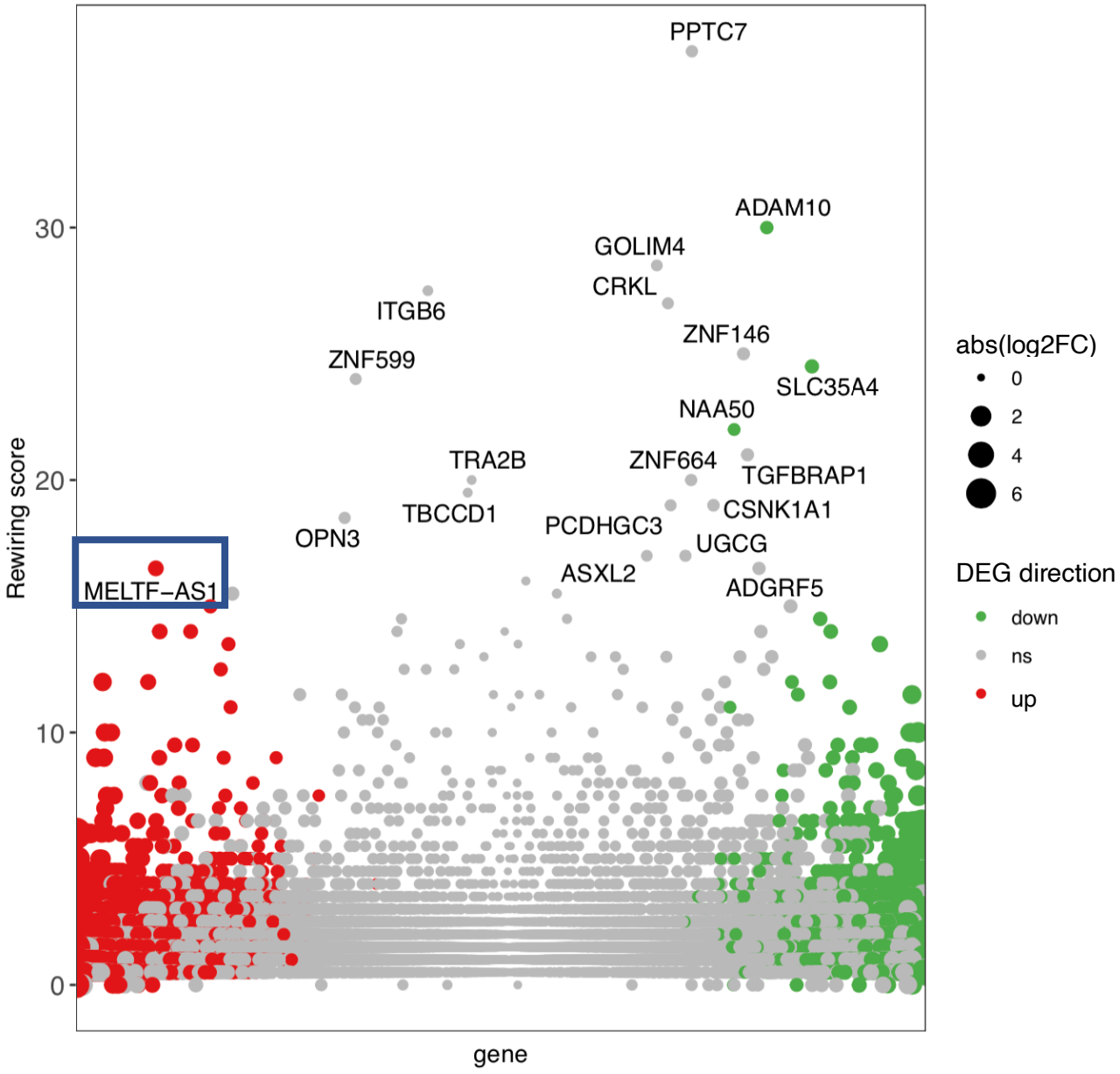


Figure 3.13. MELTF-AS1 was the most rewired node in PAH whose expression is also upregulated in PAH lungs. Dot plot showing DyNet rewiring score of genes between PAH and FD lung networks. Red and green indicate up- and downregulation in PAH, respectively. X-axis represents genes ordered by fold change. MELTF-AS1 highlighted by a blue box. ns, not significant; FC = fold change.

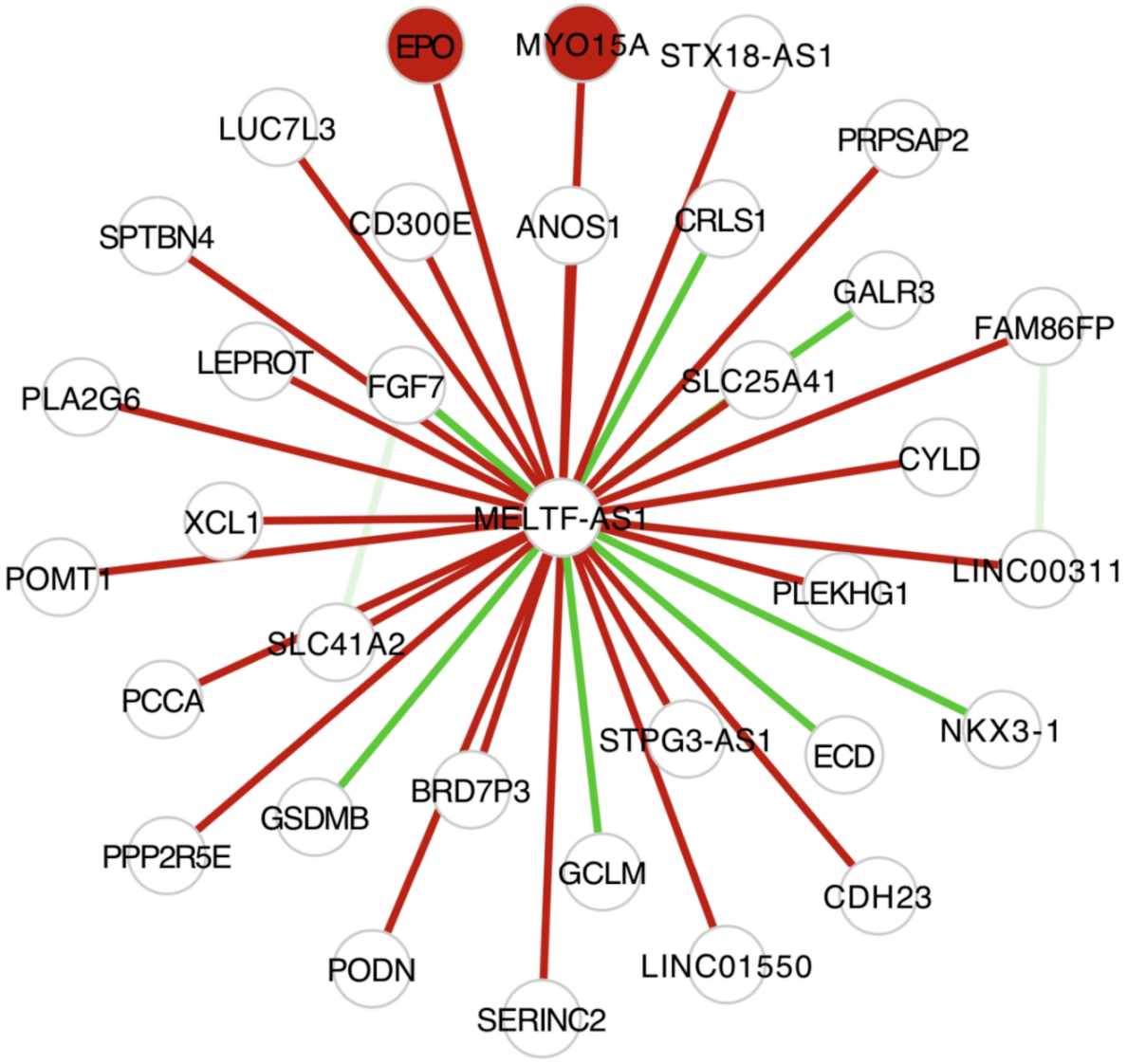


Figure 3.14. MELTF-AS1 subnetwork in which most regulatory connections to other genes are only present in the PAH network. White nodes were present in both PAH and FD networks. Red nodes were only present in the PAH network. Red and green edges were present only in the PAH or FD networks, respectively.

scRNAseq integration identifies cell types involved in tissue-level networks

Enrichment testing of WGCNA modules with cell type-specific PAH signatures from our previously generated rat scRNAseq PAH DEGs[40] revealed specificity in which cell types may be mediating the genes co-expressed in each WGCNA module (**Figure 3.15**). For example, the pink module was enriched for the monocrotaline (MCT) PAH smooth muscle cell (SMC) signature and Sugen-hypoxia (SuHx) PAH signatures of fibroblasts (Fb), SMC, and endothelial cells.

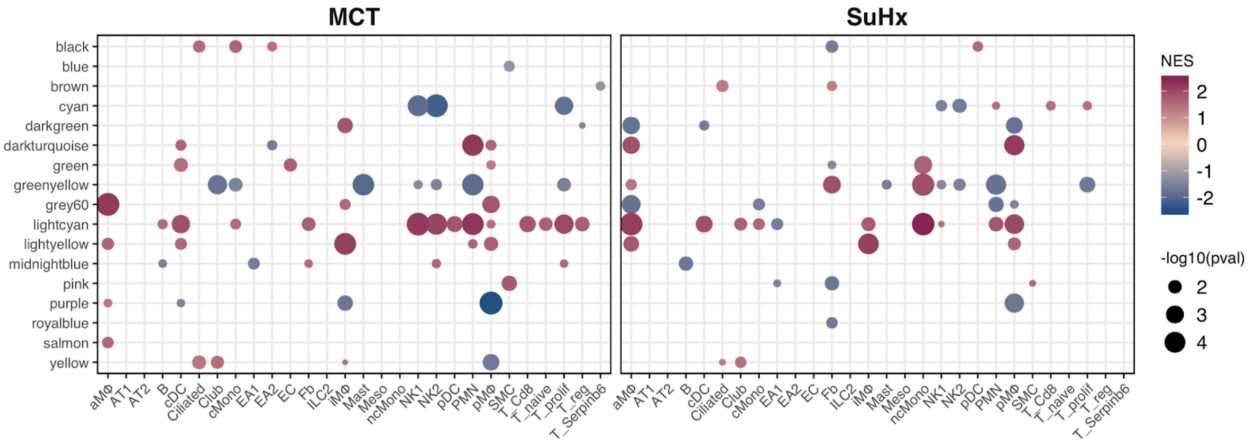


Figure 3.15. Rat lung scRNAseq integration infers cell types mediating human lung WGCNA modules. Dot plot showing upregulation (red) or downregulation (blue) of human orthologs of Rat lung scRNAseq DEGs in lung WGCNA modules as determined by GSEA. Dots represent FDR < 0.05 with larger size corresponding to stronger statistical significance. NES = normalized enrichment score; pval = p value; aMΦ = alveolar macrophages; AT1 = alveolar type 1 cells; AT2 = alveolar type 2 cells; cDC = conventional dendritic cells; cMono = classical monocytes; EA1 = endothelial arterial 1; EA2 = endothelial arterial 2; EC = endothelial capillary; Fb = fibroblast; ILC2 = innate lymphoid cell type 2; iMΦ = interstitial macrophages; Meso = mesothelial; ncMono = non-classical monocytes; NK1 = natural killer 1; NK2 = natural killer 2; pDC = plasmacytoid dendritic cell; PMN = polymorphonuclear neutrophil; pMΦ = proliferating macrophages; SMC = smooth muscle cell; T_prolif = proliferating T cells; T_reg = regulatory T cells.

As an alternative approach to infer which cell types may be coordinating the activity of lung co-expression modules, we first deconvoluted each bulk RNAseq sample using CIBERSORT[41] to estimate relative cell proportions. To ensure rigorous deconvolution estimates, we curated 7 distinct single-cell datasets totaling over 500,000 cells from 154 human lungs[42–48]

and integrated them using Seurat[49] to generate a high-confidence cell type reference map for the human lung (**Figure 3.16**). We identified similar cell types using this approach as compared to our rat scRNAseq integration. For example, using deconvolution, we found the pink module to be most strongly correlated with myofibroblasts (MyoFb), which express both SMC and fibroblast markers, and endothelial cells (**Figure 3.17**).

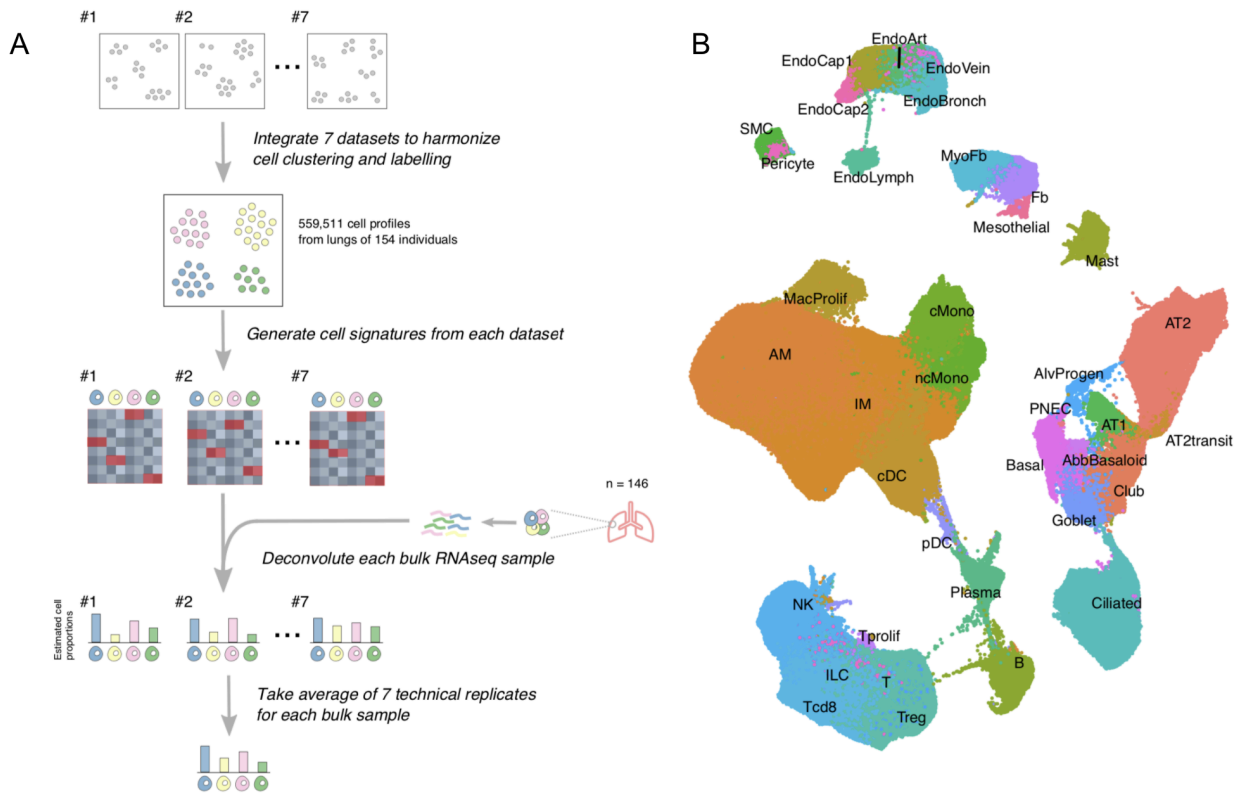


Figure 3.16. Generating a cell type reference map for the human lung. A) Schematic of construction of a high-confidence cell type reference map for the human lung for deconvolution. B) Uniform manifold approximation and projection (UMAP) plot showing integration of 7 distinct single-cell datasets totaling over 500,000 cells from 154 human lungs[42–48] . AM = alveolar macrophages; AT1 = alveolar type 1 cells; AT2 = alveolar type 2 cells; AT2 = alveolar type 2 transitional cells; cDC = conventional dendritic cells; cMono = classical monocytes; EndoArt = endothelial arterial; EndoVein = endothelial vein; EndoBronch = endothelial bronchial; EndoCap = endothelial capillary; Fb = fibroblast; ILC = innate lymphoid cells; IM = interstitial macrophages; MyoFb = myofibroblast; ncMono = non-classical monocytes; NK = natural killer; pDC = plasmacytoid dendritic cell; PMN = polymorphonuclear neutrophil; MacProlif = proliferating macrophages; SMC = smooth muscle cell; Treg = regulatory T cells; PNEC = pulmonary neuroendocrine cells.

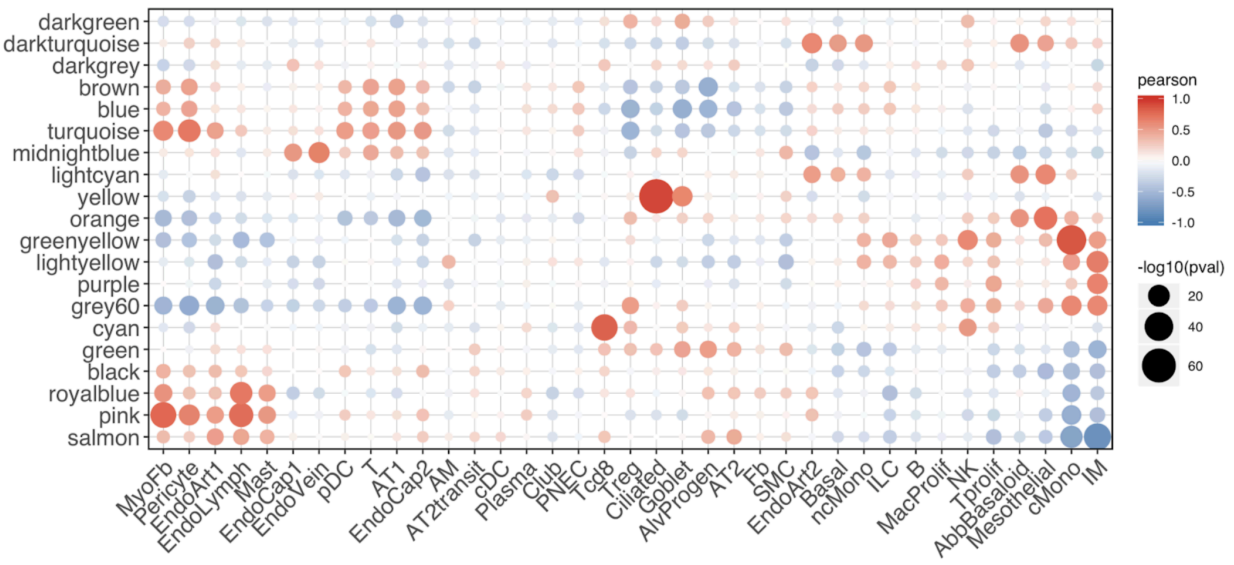


Figure 3.17. Deconvolution using human lung scRNAseq infers cell types mediating human lung WGCNA modules. Dots represent statistically significant Pearson correlation between estimated cell type proportions by deconvolution across samples and the first principal component of WGCNA modules across samples (p value < 0.05) where red indicates positive correlation, blue indicates negative correlation, and larger dot sizes indicate stronger statistical significance. AM = alveolar macrophages; AT1 = alveolar type 1 cells; AT2 = alveolar type 2 cells; AT2 = alveolar type 2 transitional cells; cDC = conventional dendritic cells; cMono = classical monocytes; EndoArt = endothelial arterial; EndoVein = endothelial vein; EndoCap = endothelial capillary; Fb = fibroblast; ILC = innate lymphoid cells; IM = interstitial macrophages; MyoFb = myofibroblast; ncMono = non-classical monocytes; NK = natural killer; pDC = plasmacytoid dendritic cell; PMN = polymorphonuclear neutrophil; MacProlif = proliferating macrophages; SMC = smooth muscle cell; Treg = regulatory T cells; PNEC = pulmonary neuroendocrine cells.

Discussion

In this chapter, we conducted a well-powered bulk RNAseq analysis of PAH and healthy control human lungs utilizing a variety of integrative multiomics and systems biology approaches to uncover the tissue-level alterations of genes, pathways, and cell types in PAH lungs. We uncovered thousands of genes to be dysregulated in PAH lungs and found that these genes are involved in pathways known to be reprogrammed in PAH such as mesenchymal transition. Grouping genes into modules based on how they co-express with each other, we found specific modules to be associated with PAH diagnosis, severity, and risk. Integrating these modules with a gene-gene regulatory network of the human lung, we found hub genes likely playing a central role in mediating the functions of these PAH modules. Finally, integration with scRNAseq revealed which cell types might also be involved in the PAH modules.

The pink module of 266 co-expressed genes was one of the 20 modules most strongly associated with PAH diagnosis, disease severity, and disease risk. This module was not only positively correlated with mPAP and PVR, both well-established hemodynamic parameters of PAH severity measured by right heart catheterization, and FVC/DLCO, a proxy for disease severity measured by pulmonary function testing, but the pink module was also the only module negatively correlated with cardiac output. Reduced cardiac output in PAH patients is due to RV failure seen in the most severe cases. These findings suggest that patients who had stronger expression of the pink module genes in their lungs had worse disease. Interestingly, the pink module was the only module associated with PAH that was also negatively correlated with epoprostenol, which is the most effective therapy currently available for PAH that has also been shown to improve survival. In other words, PAH patients whose lungs had stronger transcriptional activity of the pink module were less likely to be on epoprostenol treatment. This

raises the question of whether epoprostenol is somehow protective of the gene programs associated with the pink module. The most strongly upregulated pathway in the pink module genes was the Hallmark pathway of "epithelial mesenchymal transition" which encompasses endothelial mesenchymal transition (EndMT) and was also the top pathway enriched in the transcriptome-wide DEG analysis. EndMT is widely known to be a critical process by which pulmonary arteries undergo vascular remodeling in PAH[50] further supporting the importance of the pink module.

Given that the RNAseq was performed on explanted lungs of patients undergoing lung transplantation and were thus likely advanced stage PAH, we performed GWAS enrichment analysis to assess for a causal molecular link between the risk of PAH and advanced-stage PAH lungs. We found the pink module to be significantly enriched for PAH GWAS in multiple GWAS cohorts using two distinct GWAS enrichment methods. This suggests that the pink module genes are not merely just associated with severe disease but also likely causal in PAH pathogenesis.

Key driver analysis utilizing a comprehensive gene-gene regulatory network constructed from over 2000 human lungs identified *PDE7B,* a gene that encodes the phosphodiesterase 7B protein which hydrolyzes the second messenger cAMP, as a central hub gene of the pink module but has never before been implicated in PAH aside from a finding from one study that its expression was upregulated in the lungs of MCT rats[51]. However, *PDE7B* has been implicated in cancer, which is increasingly recognized to have similarity to PAH in the dysregulation of many biological processes such as cell survival and proliferation[52]. For example, one study implicated PDE7B as a drug target in chronic lymphocytic leukemia[53]. In PAH lungs, while *PDE7B* may be playing a central role in the pink module, understanding which cell types are involved is also

critically important. Using two distinct and novel approaches to scRNAseq integration, we found that endothelial cells, SMCs, and fibroblasts/myofibroblasts likely mediate the pink module genes, consistent with EndMT as the most upregulated pathway in this module. Further supporting this finding, a recent PAH human lung scRNAseq study found that PDE7B had increased expression in endothelial cells, fibroblasts, and pericyte/SMCs[54].

The strengths of this study include the large sample size and cutting edge multiomic methods utilized to uncover new biological insights into PAH pathogenesis. Although these lung samples derived from advanced stage PAH lungs, we were able to establish a casual molecular link to PAH pathogenesis using innovative GWAS enrichment methods. Another potential limitation is that the RNAseq data was at the bulk tissue level. However, we were able to infer cell types of relevant co-expression modules using two novel scRNAseq integration methods.

Overall, this integrative multiomics and systems biology study of human PAH lungs revealed the dysregulation of many genes, pathways, and cell types in PAH lungs at the tissue level. This study further identified and prioritized new candidate PAH genes such as PDE7B and the potential cell types it acts in, which warrant further investigation as novel therapeutic targets.

References

1. Stearman RS, Bui QM, Speyer G, Handen A, Cornelius AR, Graham BB, Kim S, Mickler EA, Tuder RM, Chan SY, Geraci MW. Systems Analysis of the Human Pulmonary Arterial Hypertension Lung Transcriptome. *Am J Respir Cell Mol Biol* 2019;60:637–649.

2. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* 2016;11:1650–1667.

3. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8:118–127.

4. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014;15:550.

5. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.

6. Weirauch MT. Gene Coexpression Networks for the Analysis of DNA Microarray Data. *Applied Statistics for Network Biology* John Wiley & Sons, Ltd; 2011. p. 215–250.doi:10.1002/9783527638079.ch11.

7. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;102:15545–15550.

8. Shu L, Zhao Y, Kurt Z, Byars SG, Tukiainen T, Kettunen J, Orozco LD, Pellegrini M, Lusis AJ, Ripatti S, Zhang B, Inouye M, Mäkinen V-P, Yang X. Mergeomics: multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genomics* 2016;17:874.

9. Leeuw CA de, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* 2015;11:e1004219.

10. Yoon S, Nguyen HCT, Yoo YJ, Kim J, Baik B, Kim S, Kim J, Kim S, Nam D. Efficient pathway enrichment and network analysis of GWAS summary data using GSA-SNP2. *Nucleic Acids Res* 2018;46:e60.

11. Rhodes CJ, Batai K, Bleda M, Haimel M, Southgate L, Germain M, Pauciulo MW, Hadinnapola C, Aman J, Girerd B, Arora A, Knight J, Hanscombe KB, Karnes JH, Kaakinen M, Gall H, Ulrich A, Harbaum L, Cebola I, Ferrer J, Lutz K, Swietlik EM, Ahmad F, Amouyel P, Archer SL, Argula R, Austin ED, Badesch D, Bakshi S, *et al.* Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir Med* 2019;7:227–238.

12. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, Johnson R, Segrè AV, Djebali S, Niarchou A, Consortium TGte, Wright FA, Lappalainen T, Calvo M, Getz G, Dermitzakis ET, Ardlie KG, Guigó R. The human transcriptome across tissues and individuals. *Science* 2015;348:660–665.

13. Hao K, Bossé Y, Nickle DC, Paré PD, Postma DS, Laviolette M, Sandford A, Hackett TL, Daley D, Hogg JC, Elliott WM, Couture C, Lamontagne M, Brandsma C-A, Berge M van den, Koppelman G, Reicin AS, Nicholson DW, Malkov V, Derry JM, Suver C, Tsou JA, Kulkarni A, Zhang C, Vessey R, Opiteck GJ, Curtis SP, Timens W, Sin DD. Lung eQTLs to Help Reveal the Molecular Underpinnings of Asthma. *PLOS Genetics* 2012;8:e1003029.

14. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, Zhang C, Lamb J, Edwards S, Sieberts SK, Leonardson A, Castellini LW, Wang S, Champy M-F, Zhang B, Emilsson V, Doss S, Ghazalpour A, Horvath S, Drake TA, Lusis AJ, Schadt EE. Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008;452:429–435.

15. Yang X, Deignan JL, Qi H, Zhu J, Qian S, Zhong J, Torosyan G, Majid S, Falkard B, Kleinhanz RR, Karlsson J, Castellani LW, Mumick S, Wang K, Xie T, Coon M, Zhang C, Estrada-Smith D, Farber CR, Wang SS, van Nas A, Ghazalpour A, Zhang B, Macneil DJ, Lamb JR, Dipple KM, Reitman ML, Mehrabian M, Lum PY, *et al.* Validation of candidate causal

genes for obesity that affect shared metabolic pathways and networks. *Nat Genet* 2009;41:415–423.

16. Yang X, Zhang B, Molony C, Chudin E, Hao K, Zhu J, Gaedigk A, Suver C, Zhong H, Leeder JS, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich RG, Slatter JG, Schadt EE, Kasarskis A, Lum PY. Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. *Genome Res* 2010;20:1020–1036.

17. Zhu J, Sova P, Xu Q, Dombek KM, Xu EY, Vu H, Tu Z, Brem RB, Bumgarner RE, Schadt EE. Stitching together multiple data dimensions reveals interacting metabolomic and transcriptomic networks that modulate cell regulation. *PLoS Biol* 2012;10:e1001301.

18. Zhu J, Wiener MC, Zhang C, Fridman A, Minch E, Lum PY, Sachs JR, Schadt EE. Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLoS Comput Biol* 2007;3:e69.

19. Zhu J, Zhang B, Smith EN, Drees B, Brem RB, Kruglyak L, Bumgarner RE, Schadt EE. Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat Genet* 2008;40:854–861.

20. Zhang B, Gaiteri C, Bodea L-G, Wang Z, McElwee J, Podtelezhnikov AA, Zhang C, Xie T, Tran L, Dobrin R, Fluder E, Clurman B, Melquist S, Narayanan M, Suver C, Shah H, Mahajan M, Gillis T, Mysore J, MacDonald ME, Lamb JR, Bennett DA, Molony C, Stone DJ, Gudnason V, Myers AJ, Schadt EE, Neumann H, Zhu J, *et al.* Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 2013;153:707–720.

21. Chan KHK, Huang Y-T, Meng Q, Wu C, Reiner A, Sobel EM, Tinker L, Lusis AJ, Yang X, Liu S. Shared molecular pathways and gene networks for cardiovascular disease and type 2 diabetes mellitus in women across diverse ethnicities. *Circ Cardiovasc Genet* 2014;7:911–919.

22. Chella Krishnan K, Kurt Z, Barrere-Cain R, Sabir S, Das A, Floyd R, Vergnes L, Zhao Y, Che N, Charugundla S, Qi H, Zhou Z, Meng Y, Pan C, Seldin MM, Norheim F, Hui S, Reue K, Lusis AJ, Yang X. Integration of Multi-omics Data from Mouse Diversity Panel Highlights Mitochondrial Dysfunction in Non-alcoholic Fatty Liver Disease. *Cell Systems* 2018;6:103-115.e7.

23. Huan T, Meng Q, Saleh MA, Norlander AE, Joehanes R, Zhu J, Chen BH, Zhang B, Johnson AD, Ying S, Courchesne P, Raghavachari N, Wang R, Liu P, International Consortium for Blood Pressure GWAS (ICBP), O'Donnell CJ, Vasan R, Munson PJ, Madhur MS, Harrison DG, Yang X, Levy D. Integrative network analysis reveals molecular mechanisms of blood pressure regulation. *Mol Syst Biol* 2015;11:799.

24. Huan T, Zhang B, Wang Z, Joehanes R, Zhu J, Johnson AD, Ying S, Munson PJ, Raghavachari N, Wang R, Liu P, Courchesne P, Hwang S-J, Assimes TL, McPherson R, Samani NJ, Schunkert H, Coronary ARteryDIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) Consortium, International Consortium for Blood Pressure GWAS (ICBP), Meng Q, Suver C, O'Donnell CJ, Derry J, Yang X, Levy D. A systems biology framework identifies molecular underpinnings of coronary heart disease. *Arterioscler Thromb Vasc Biol* 2013;33:1427–1434.

25. Mäkinen V-P, Civelek M, Meng Q, Zhang B, Zhu J, Levian C, Huan T, Segrè AV, Ghosh S, Vivar J, Nikpay M, Stewart AFR, Nelson CP, Willenborg C, Erdmann J, Blakenberg S, O'Donnell CJ, März W, Laaksonen R, Epstein SE, Kathiresan S, Shah SH, Hazen SL, Reilly MP, Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Consortium, Lusis AJ, Samani NJ, Schunkert H, Quertermous T, *et al.* Integrative genomics

reveals novel molecular pathways and gene networks for coronary artery disease. *PLoS Genet* 2014;10:e1004502.

26. Meng Q, Ying Z, Noble E, Zhao Y, Agrawal R, Mikhail A, Zhuang Y, Tyagi E, Zhang Q, Lee J-H, Morselli M, Orozco L, Guo W, Kilts TM, Zhu J, Zhang B, Pellegrini M, Xiao X, Young MF, Gomez-Pinilla F, Yang X. Systems Nutrigenomics Reveals Brain Gene Networks Linking Metabolic and Brain Disorders. *EBioMedicine* 2016;7:157–166.

27. Meng Q, Zhuang Y, Ying Z, Agrawal R, Yang X, Gomez-Pinilla F. Traumatic Brain Injury Induces Genome-Wide Transcriptomic, Methylomic, and Network Perturbations in Brain and Blood Predicting Neurological Disorders. *EBioMedicine* 2017;16:184–194.

28. Shu L, Chan KHK, Zhang G, Huan T, Kurt Z, Zhao Y, Codoni V, Trégouët D-A, Cardiogenics Consortium, Yang J, Wilson JG, Luo X, Levy D, Lusis AJ, Liu S, Yang X. Shared genetic regulatory networks for cardiovascular disease and type 2 diabetes in multiple populations of diverse ethnicities in the United States. *PLoS Genet* 2017;13:e1007040.

29. Shu L, Meng Q, Diamante G, Tsai B, Chen Y-W, Mikhail A, Luk H, Ritz B, Allard P, Yang X. Prenatal Bisphenol A Exposure in Mice Induces Multitissue Multiomics Disruptions Linking to Cardiometabolic Disorders. *Endocrinology* 2019;160:409–429.

30. Zhao Y, Chen J, Freudenberg JM, Meng Q, Rajpal DK, Yang X. Network-Based Identification and Prioritization of Key Regulators of Coronary Artery Disease Loci. *Arterioscler Thromb Vasc Biol* 2016;36:928–941.

31. Zhao Y, Jhamb D, Shu L, Arneson D, Rajpal DK, Yang X. Multi-omics integration reveals molecular networks and regulators of psoriasis. *BMC Syst Biol* 2019;13:8.

32. Arneson D, Zhang G, Ying Z, Zhuang Y, Byun HR, Ahn IS, Gomez-Pinilla F, Yang X. Single cell molecular alterations reveal target cells and pathways of concussive brain injury. *Nature Communications* 2018;9:3894.

33. Goenawan IH, Bryan K, Lynn DJ. DyNet: visualization and analysis of dynamic molecular interaction networks. *Bioinformatics* 2016;32:2713–2715.

34. Ching T, Huang S, Garmire LX. Power analysis and sample size estimation for RNA-Seq differential expression. *RNA* 2014;20:1684–1696.

35. Piñero J, Bravo À, Queralt-Rosinach N, Gutiérrez-Sacristán A, Deu-Pons J, Centeno E, García-García J, Sanz F, Furlong LI. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res* 2017;45:D833–D839.

36. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J, Wiegers TC, Mattingly CJ. The Comparative Toxicogenomics Database: update 2019. *Nucleic Acids Res* 2019;47:D948–D954.

37. Li C, Tan F, Pei Q, Zhou Z, Zhou Y, Zhang L, Wang D, Pei H. Non-coding RNA MFI2-AS1 promotes colorectal cancer cell proliferation, migration and invasion through miR-574-5p/MYCBP axis. *Cell Prolif* 2019;52:e12632.

38. Flippot R, Mouawad R, Spano J-P, Rouprêt M, Compérat E, Bitker M-O, Parra J, Vaessen C, Allanic F, Manach Q, Tannir NM, Khayat D, Su X, Malouf GG. Expression of long non-coding RNA MFI2-AS1 is a strong predictor of recurrence in sporadic localized clear-cell renal cell carcinoma. *Scientific Reports* 2017;7:8540.

39. Wei Y, Wang Z, Zong Y, Deng D, Chen P, Lu J. LncRNA MFI2-AS1 promotes HCC progression and metastasis by acting as a competing endogenous RNA of miR-134 to upregulate FOXM1 expression. *Biomedicine & Pharmacotherapy* 2020;125:109890.

40. Hong J, Arneson D, Umar S, Ruffenach G, Cunningham CM, Ahn IS, Diamante G, Bhetraratana M, Park JF, Said E, Huynh C, Le T, Medzikovic L, Humbert M, Soubrier F, Montani D, Girerd B, Trégouët D-A, Channick R, Saggar R, Eghbali M, Yang X. Single-cell Study of Two Rat Models of Pulmonary Arterial Hypertension Reveals Connections to Human Pathobiology and Drug Repositioning. *Am J Respir Crit Care Med* 2020;doi:10.1164/rccm.202006-2169OC.

41. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* 2015;12:453–457.

42. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, Fernandez R, Akbarpour M, Chen C-I, Ren Z, Verma R, Abdala-Valencia H, Nam K, Chi M, Han S, Gonzalez-Gonzalez FJ, Soberanes S, Watanabe S, Williams KJN, Flozak AS, Nicholson TT, Morgan VK, Winter DR, Hinchcliff M, Hrusch CL, Guzy RD, Bonham CA, Sperling AI, Bag R, *et al.* Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2018;doi:10.1164/rccm.201712-2410OC.

43. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, Chu SG, Raby BA, DeIuliis G, Januszyk M, Duan Q, Arnett HA, Siddiqui A, Washko GR, Homer R, Yan X, Rosas IO, Kaminski N. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Science Advances* 2020;6:eaba1983.

44. Madissoon E, Wilbrey-Clark A, Miragaia RJ, Saeb-Parsy K, Mahbubani KT, Georgakopoulos N, Harding P, Polanski K, Huang N, Nowicki-Osuch K, Fitzgerald RC, Loudon KW, Ferdinand JR, Clatworthy MR, Tsingene A, van Dongen S, Dabrowska M, Patel M, Stubbington MJT, Teichmann SA, Stegle O, Meyer KB. scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biology* 2019;21:1.

45. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, Chen H, Wang J, Tang H, Ge W, Zhou Y, Ye F, Jiang M, Wu J, Xiao Y, Jia X, Zhang T, Ma X, Zhang Q, Bai X, Lai S, Yu C, Zhu L, Lin R, Gao Y, Wang M, Wu Y, Zhang J, Zhan R, *et al.* Construction of a human cell landscape at single-cell level. *Nature* 2020;581:303–309.

46. Morse C, Tabib T, Sembrat J, Buschur K, Bittar HT, Valenzi E, Jiang Y, Kass DJ, Gibson K, Chen W, Mora A, Benos PV, Rojas M, Lafyatis R. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *European Respiratory Journal* 2019;1802441.doi:10.1183/13993003.02441-2018.

47. Raredon MSB, Adams TS, Suhail Y, Schupp JC, Poli S, Neumark N, Leiby KL, Greaney AM, Yuan Y, Horien C, Linderman G, Engler AJ, Boffa DJ, Kluger Y, Rosas IO, Levchenko A, Kaminski N, Niklason LE. Single-cell connectomic analysis of adult mammalian lungs. *Science Advances* 2019;5:eaaw3851.

48. Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, Peter L, Chung M-I, Taylor CJ, Jetter C, Raju L, Roberson J, Ding G, Wood L, Sucre JMS, Richmond BW, Serezani AP, McDonnell WJ, Mallal SB, Bacchetta MJ, Loyd JE, Shaver CM, Ware LB, Bremner R, Walia R, Blackwell TS, Banovich NE, Kropski JA. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Science Advances* 2020;6:eaba1972.

49. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. *Cell* 2019;177:1888-1902.e21.

50. Ranchoux B, Antigny F, Rucker-Martin C, Hautefort A, Péchoux C, Bogaard HJ, Dorfmüller P, Remy S, Lecerf F, Planté S, Chat S, Fadel E, Houssaini A, Anegon I, Adnot S, Simonneau G, Humbert M, Cohen-Kaminsky S, Perros F. Endothelial-to-mesenchymal transition in pulmonary hypertension. *Circulation* 2015;131:1006–1018.

51. Tian X, Vroom C, Ghofrani HA, Weissmann N, Bieniek E, Grimminger F, Seeger W, Schermuly RT, Pullamsetti SS. Phosphodiesterase 10A Upregulation Contributes to Pulmonary Vascular Remodeling. *PLoS One* 2011;6:.

52. Boucherat O, Vitry G, Trinh I, Paulin R, Provencher S, Bonnet S. The cancer theory of pulmonary arterial hypertension. *Pulmonary Circulation* 2017;7:285–299.

53. Zhang L, Murray F, Zahno A, Kanter JR, Chou D, Suda R, Fenlon M, Rassenti L, Cottam H, Kipps TJ, Insel PA. Cyclic nucleotide phosphodiesterase profiling reveals increased expression of phosphodiesterase 7B in chronic lymphocytic leukemia. *PNAS* 2008;105:19532–19537.

54. Saygin D, Tabib T, Bittar HET, Valenzi E, Sembrat J, Chan SY, Rojas M, Lafyatis R. Transcriptional profiling of lung cell populations in idiopathic pulmonary arterial hypertension. *Pulm Circ* 2020;10:.

CHAPTER 4

Conclusion

This dissertation investigated the transcriptional alterations in the lungs of PAH animal models and patients using integrative multiomics to identify and prioritize candidate genes, pathways, and cell types implicated in PAH pathobiology. We identified reprogramming of genes and pathways in various cell types in SuHx and MCT lungs. We also found that genes dysregulated in SuHx nonclassical monocytes were significantly enriched for PAH-associated genes and GWAS variants. We further identified candidate drugs predicted to reverse the dysregulated gene programs. This rat study revealed the distinct and shared reprogramming of genes and pathways in two commonly used PAH models for the first time at single-cell resolution and demonstrated their relevance to human PAH and utility for drug repositioning.

We then dissected the human PAH lung transcriptional landscape at the tissue level using an innovative network and systems biology methods on a well-powered RNAseq dataset of human PAH lungs. We discovered many DEGs and pathways in human PAH lungs at the tissue level, and through integration with clinical data and PAH GWAS, our network analysis revealed modules of co-expressed genes that are not only associated with PAH diagnosis and severity, but also risk of PAH implicating their causal role in PAH pathogenesis. Key driver analysis utilizing a comprehensive gene-gene regulatory network of the human lung identified and prioritized candidate genes. Furthermore, we integrated the tissue-level networks with scRNAseq to uncover the specific cell types mediating the tissue-level gene programs. Thus, our findings implicate novel genes, pathways, and cell types in PAH pathobiology.

While our rat single-cell study demonstrated connections to human pathobiology of PAH, direct comparison to our human PAH lung study is challenging given the different resolution (cell vs tissue) and different analytical methods (i.e. using cell type-specific rat DEGs vs human tissue-level WGCNA co-expression modules). The relative timing of lung sampling with respect

to the disease course could also be very different between the rat and human studies and the dysregulated pathways could be quite different earlier vs later in the disease course. For example, inflammatory pathways were prominently upregulated across a number of cell types in both PAH rat models. However, such pathways were downregulated in the overall lung tissue from our human dataset, which could reflect that these patients had advanced disease requiring lung transplantation. Obtaining human lung samples from patients earlier in their disease course would be difficult since lung biopsies are not routine. Thus, other approaches are needed to link molecular alterations from advanced disease to disease risk and pathogenesis, such as GWAS integration as implemented in this study. A comprehensive human lung single-cell study in PAH employing such integrative methods will facilitate a more direct comparison to and further validation of findings from our rat single-cell study, such as the importance of non-classical monocytes.

Future directions include single-nucleus RNAseq and spatial transcriptomics of human PAH lungs, as well as experimental validation of candidate genes. Single-nucleus RNAseq enables use of archived frozen tissues while retaining similar cellular resolution and gene coverage as scRNAseq. Spatial transcriptomics will localize the key genes and pathways within the lung cellular architecture which as of yet has never been performed in PAH. Furthermore, genetic perturbation in relevant cells *in vitro* will elucidate the functional role of candidate genes in PAH pathogenesis. Such state-of-the-art omics and experimental studies will unravel new causal insights into the specific dysregulation of genes and pathways within the diseased PAH lung.