**Final Report**

Title: Transcriptomic Signatures Associated with Pulmonary Arterial Hypertension and Right Ventricular Maladaptation

Authors: Hongyang Pi, MD MS1,2; Lan Shui, MS2, Runjia Zou, MS2; Lu Xia, PhD1,2; Sina A. Gharib, MD1; Peter J. Leary, MD PhD1,3

1University of Washington, Department of Medicine; 2University of Washington, Department of Biostatistics; 3University of Washington, Department of Epidemiology

**Abstract**

RATIONALE

Pulmonary arterial hypertension (PAH) is a complex disease characterized by progressive right ventricular (RV) failure leading to significant morbidity and mortality. Investigating proteomic features and pathways associated with RV dilation, mortality, and measures of disease severity can provide insight into molecular mechanisms, identify sub-phenotypes, and suggest potential therapeutic targets.

METHODS

The Seattle Right Ventricle Translational Science Study (Servetus) is a single institution cohort enrolled between 2014 and 2016 and followed for three years. RNA extracted from isolated cells was assessed for purity and then hybridized to a HumanRef-8 BeadChip (Illumina) that was inclusive of 26,703 unique genes. Partial least squares discriminant analysis was applied to assess the distinguishability of relevant outcomes including 3-year mortality and RV dilation. Differentially abundant proteins were then identified with adjustments for age, sex, body mass index and PAH etiology. Pathway enrichment analysis was performed by Gene Enrichment Set Analysis to identify significantly dysregulated processes.

RESULTS

A total of 55 participants with World Health Organization Group 1 PAH were included. PLS-DA showed clear separations between survivors and non-survivors and between participants with dilated versus non-dilated RVs. A total of 202 transcripts were differentially expressed in mortality and 29 transcripts were differentially expressed in RV dilation. Pathways involving the tricarboxylic acid cycle were enriched for both outcomes.

CONCLUSIONS

Distinct plasma transcriptomic profiles are associated with mortality, and RV dilation. Oxidative stress and immune system should be further investigated as a potential target to improve prognosis for PAH patients.

**Background**

PAH is a rare but life-threatening disease characterized by progressive elevation of pulmonary arterial pressures1. Despite advances in pulmonary vasodilators over the past two decades, the 5-year survival remains low, at 57% in the national PAH registry2-5. Progressive elevation of the pulmonary arterial pressures increase afterload on the Right Ventricle (RV). The RV adapts to the increased afterload by increasing its wall thickness and contractility. When these compensatory mechanisms are insufficient, RV dysfunction occurs.6 Right heart failure tends to be the key driver for morbidity and mortality in PAH7— it accounts for more than 50% of hospitalizations and 14% of in-hospital mortality for patients with PAH8. Conventional heart failure management focuses mostly on the left heart9, while targeted therapy to promote or maintain RV adaptation over failure is lacking.

A major barrier to identifying effective, targeted therapeutics in the right heart failure is substantial clinical and pathophysiological heterogeneity between patients. Indeed, the timing of RV dilation and dysfunction vary from patient to patient, even among individuals with an otherwise similar severity of pulmonary vascular disease and right heart afterload. Differences in RV adaptation translate into significant differences in clinical outcomes10,11. Using a prospective cohort, we recently found that RV dilation, a marker of RV failure, can further stratify PAH patients who are deemed high risk by well-established REVEAL 2.0 score12,13. Defining subsets of patients with key distinguishing features, often called sub-phenotypes, has led to the identification of important new therapeutics in cancer, chronic obstructive lung disease (COPD) and asthma14-16. Identifying reliably defined and well validated RV sub-phenotypes indicative of RV failure or RV adaptation could provide biological insight and expand therapeutic targets specific to right heart failure.

A valuable approach to characterize sub-phenotypes is via transcriptomics-based analyses, which incorporate systems biology to identify relationships at molecular, cellular, and clinical levels. Despite its rising popularity in precision medicine, transcriptomics has been applied to PAH sparingly, with most studies focusing on molecular profiles for patients with PAH compared to healthy controls17,18. Key questions about molecular and phenotypic differences in RV failure and adaptation remain unanswered. The transcriptomic analysis may be helpful in complementing metabolomics and proteomics in identifying nuanced sub-phenotype like RV adaptation in a complex disorder like PAH.

**Methods**

The Seattle Right Ventricle Translational Science (Servetus) Study is a prospective observational cohort of participants with PAH at the University of Washington Medical Center. Participants with incident and prevalent PAH were enrolled from 2014 to 2016 and followed for at least 3 years. They were between 18 and 80 years of age and able to provide informed consent. All participants contributed to a blood-based biorepository. The inclusion criteria for Servetus study were: English-speaking; adults (age 18-80 years of age); able to provide informed consent; hemodynamics met World Health Organization (WHO) Group 1 Pulmonary Arterial Hypertension defined by right heart catheterization (RHC) within five years demonstrating a mean pulmonary arterial pressure of ≥ 25 mmHg, occlusion pressure of ≤ 15 mmHg, and pulmonary vascular resistance of ≥ 3 wood units). We excluded patients with pulmonary hypertension classified in other WHO Groups, those who had multi-physiology pulmonary hypertension including more than one WHO Group, and those who did not have pulmonary hypertension.

Demographic data including age, sex at birth, height and weight were recorded. PAH subtype was determined through chart review with adjudication by investigators. Etiology was classified as idiopathic or into one of several associated PAH groups including connective tissue disease (CTD), familial, or toxin-induced PAH. Six-minute walk distance (6MWD), transthoracic echocardiogram (TTE) and RHC results, and New York Heart Association (NYHA) Functional Class were collected. Baseline TTE was included if obtained within 3 months of enrollment. Two independent researchers provided research reads for all TTEs with available images. Baseline RHC was included if it occurred within one year of enrollment. NYHA Functional Class was documented with a standardized decision aid. N-terminal-pro hormone brain natriuretic peptide (NT-proBNP) was measured using banked samples at the time of enrollment for all participants.

We investigated two core outcomes of interest, including 3-year mortality and RV dilation. For visualization, variables were categorized using cut-offs from standard risk tools with a few exceptions. Clear definitions for RV failure are lacking; however, in patients with PAH, RV dilation is strongly suggestive of maladaptation and vulnerability to failure.8 RV basal diameter ≥ 5 cm, measured on apical 4-chamber view on TTE, was empirically chosen as surrogate for severe RV dilation given the lack of standardized cut-offs for this metric and our previous work in this cohort suggesting an association with outcomes at this threshold.9

Transcriptomics data was pre-processed and normalized in the R statistical environment (R Core Team 2020). Our overall approach to better understand the association between measured outcomes and the transcriptome included two core components (Figure 1). As a first step, we evaluated global gene variation to understand whether the transcriptome had any association with our outcomes of interest. Subsequently, we focused on the associations of individual RNA sequences (and assessed the comparability of the Servetus cohort with previously reported analyses of individual sequences) and pathways (involving a set of genes within each pathway) with outcomes.

To evaluate global metabolic variation associated with metrics of disease severity, multivariate analyses including Partial Least Squares Discriminant Analysis (PLS­-DA) (MetaboAnalyst 5.0), were performed.11To examine associations between outcomes and RNA sequences, we performed differential analyses (for mortality and RV dilation) adjusted for age, sex, body mass index (BMI) and PAH etiology. We fitted a negative binomial model for each, with transcript counts as our outcome, the coefficient of condition (survival vs death and RV dilation or not) as our parameter of interest, and adjusted for age, sex, BMI, and PAH etiology. To control for false discovery rate (FDR) in multiple comparisons while capturing important biological information, Benjamini-Hochberg’s method with adjusted p-value threshold < 0.05 was used to identify significant sequences. To determine associations between pathways and outcomes, dysregulated genetic pathways with annotations curated by a R package “biomart”) were identified using the Gene Set Enrichment Analysis (GSEA) with adjustment for age, sex, BMI, and PAH etiology.12 Benjamini-Hochberg’s method with adjusted p-value threshold < 0.05 was used to identify significant pathways.

**Results**

**Cohort characteristics**

A total of 55 participants with PAH were included in the study. Table 1 summarizes the clinical characteristics including demographics, PAH etiology, treatment, and the outcomes of interest. Eight participants (15%) died by the end of the 3-year follow-up. Ten (18%) participants had RV dilation by RV basal diameter no less than 5cm.

**Table 1: Summary of the Servetus cohort.**

|  |  |
| --- | --- |
| **Demographics** | **Total N = 55** |
| Age, years | 52 ± 14 |
| Female, n (%) | 46 (84%) |
| Race (white), n (%) | 40 (73%) |
| Body mass index (kilograms/meter2) | 29 ± 7 |
| **Etiology** | |
| Congenital, n (%) | 6 (11%) |
| Connective tissue disease, n (%) | 12 (22%) |
| Idiopathic, n (%) | 27 (49%) |
| Toxin, n (%) | 9 (16%) |
| Other, n (%) | 1 (2%) |
| **Outcomes of Interest** | |
| 3-year mortality, n (%) | 8 (15%) |
| 6-minute walk test, meters | 380 ± 104 |
| NT-proBNP (x1000) | 2.2 ± 2.8 |
| NYHA functional class III/IV, n (%) | 14 (25%) |
| RV basal diameter, cm | 4.5 ± 0.8 |
| Tricuspid annular plane systolic excursion (TAPSE), mm | 21 ± 6 |
| High-risk REVEAL 2.0 score ≥ 9, n (%) | 27 (49%) |

Missing data: Mortality, NT-proBNP and REVEAL 2.0 score were available for all participants, 6-minute walk test was available for 48 participants, functional class and TAPSE were available for 40 participants, RV basal diameter were available for 39 participants. NT-proBNP: N-terminal probrain natriuretic peptide; NYHA: New York Heart Association; RV: right ventricular; and TAPSE: tricuspid annular plane systolic excursion.

**Global transcriptomic signatures associated with clinical outcomes**

We initially investigated whether global metabolomic profiles (based on variability across all 26703 RNA sequences) distinguished the outcomes of interest. After filtering out the pseudogenes, we left with 22837 RNA sequences. Using PLS-DA, we found distinct clustering of participants based on 3-year mortality and RV dilation (Figure 2).

**Differentially abundant RNA sequences associated with metrics of disease severity**

During 36 months of follow-up, eight (15%) participants died. 202 RNA sequences were significantly associated with mortality (adjusted p < 0.05), 180 of which were associated with increased probability of death and 22 were associated with decreased probability of death (Figure 3-A). Ten (26%) participants had RV dilation by RV basal diameter no less than 5cm. 29 RNA sequences were significantly associated with RV dilation (adjusted P<0.05), 28 of which were associated with increased probability of RV dilation and one is associated with decreased probability of RV dilation (Figure 3-B).

Plot the most significant transcripts with log2 transformed counts, 40 transcripts for mortality and 39 transcripts for RV dilation (Figure 4). We performed a Z-score transformation before plotting a heatmap for each. We can see that the normalized counts of these significant transcripts are differentially associated with the conditions. Also, we noticed that the top significant transcripts are all up-regulated in people who died and in people who have RV dilation.

**Transcriptomic pathways dysregulated in participants with severe PAH**

In Figure 4 we provide the significant pathways that were enriched in mortality and RV dilation. Processes involving tricarboxylic acid (TCA) cycle and immune system were significantly associated with both outcomes of interest (adjusted p-value < 0.05). Multiple gene expression related pathways were enriched in participants who died but were not associated with RV dilation (Figure 4).

Mortality was associated with 288 pathways (adjusted p-value < 0.01) using gene set enrichment analysis (GSEA). Pathways within the tricarboxylic acid (TCA) cycle, signal transduction, immune system, and oxidative stress were among the most significant, followed by interferon response and cell cycle (Figure 4A).

RV dilation was associated with 104 significant pathways (adjusted p-value < 0.05). Multiple pathways within the TCA cycle, cell transduction and oxidative stress were upregulated, while the immune system and RNA metabolism were downregulated in RV dilation (Figure 4B).

**Discussion**

In this deeply phenotyped prospective cohort, we performed large-scale transcriptomic data analysis on 55 participants with PAH and found distinct transcripts and pathways associated with two metrics of disease severity. Specifically, key pathways in the TCA cycle were associated with both queried metrics of severity. This intuitive but novel finding suggests shared transcriptomic underpinning across a range of metrics of disease severity.

The TCA cycle, also known as the Krebs or citric acid cycle, is the main source of energy for cells and an important part of aerobic respiration. In our cohort, multiple TCA related transcripts and pathways including electron transportation and oxidative phosphorylation were upregulated in both mortality and RV dilation, indicating a possible overload of TCA metabolism in severe disease.

Similarly, the immune system is also highly involved in both mortality and RV dilation. Dysfunctional immunity has been well documented in connective tissue disease related to PAH, but our findings suggest that maybe there is overall immune dysregulation across different etiologies of PAH.

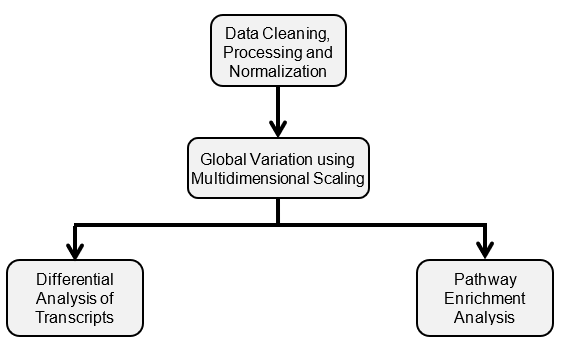
There are several limitations to this study. First, the sample size is relatively small in the context of high-dimensional data. As such, we used an adjusted p-value < 0.05 to limit the possibility of Type I error while also avoiding exclusion of biologically relevant pathways. While many of our associations remained significant at an adjusted p-value < 0.01 as well, we recognize that our use of a higher threshold also increases the possibility of a Type 1 error. Second, the small sample size prevented us from meaningfully comparing metabolic profiles across etiologic subtypes of PAH. Despite losing the nuances of PAH etiology, we believe our approach highlights gene expression alterations that represent common pathophysiological changes for patients as PAH severity progresses and RV maladaptation occurs. Third, there may be residual confounding in our RV dilation model given that TTE was available for 39 of 55 participants. Nevertheless, we were still able to identify a strong and distinct transcriptomic profile in RV dilation.

**Conclusion**

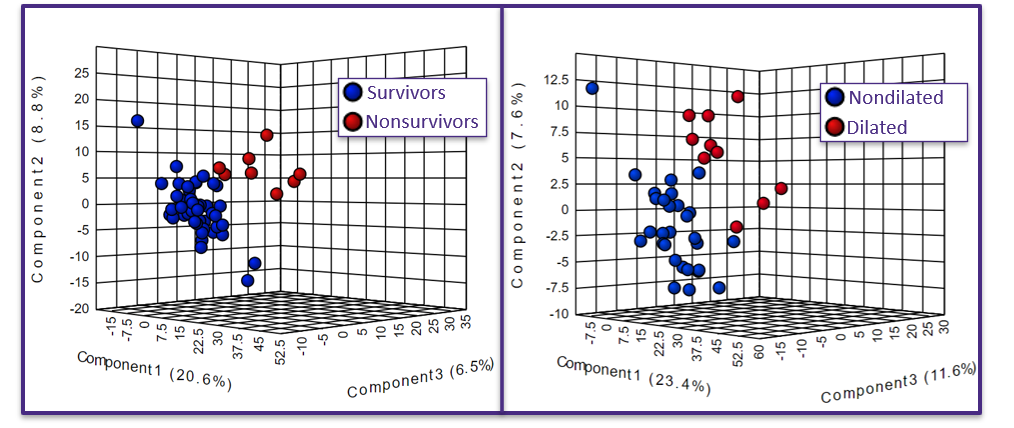
In conclusion, our study provides a detailed analysis to date of the transcriptomic signatures in all-cause mortality and RV dilation, both of which are clinically significant outcomes in PAH. Pathways in TCA cycle and immune system were most consistently associated with PAH severity. Our results highlight the potential of whole genome RNA sequencing in unraveling the disease mechanisms in PAH and in identifying prognostic and therapeutic targets in PAH outcomes.

**Figures**

**Figure 1: overview of analytical methods**

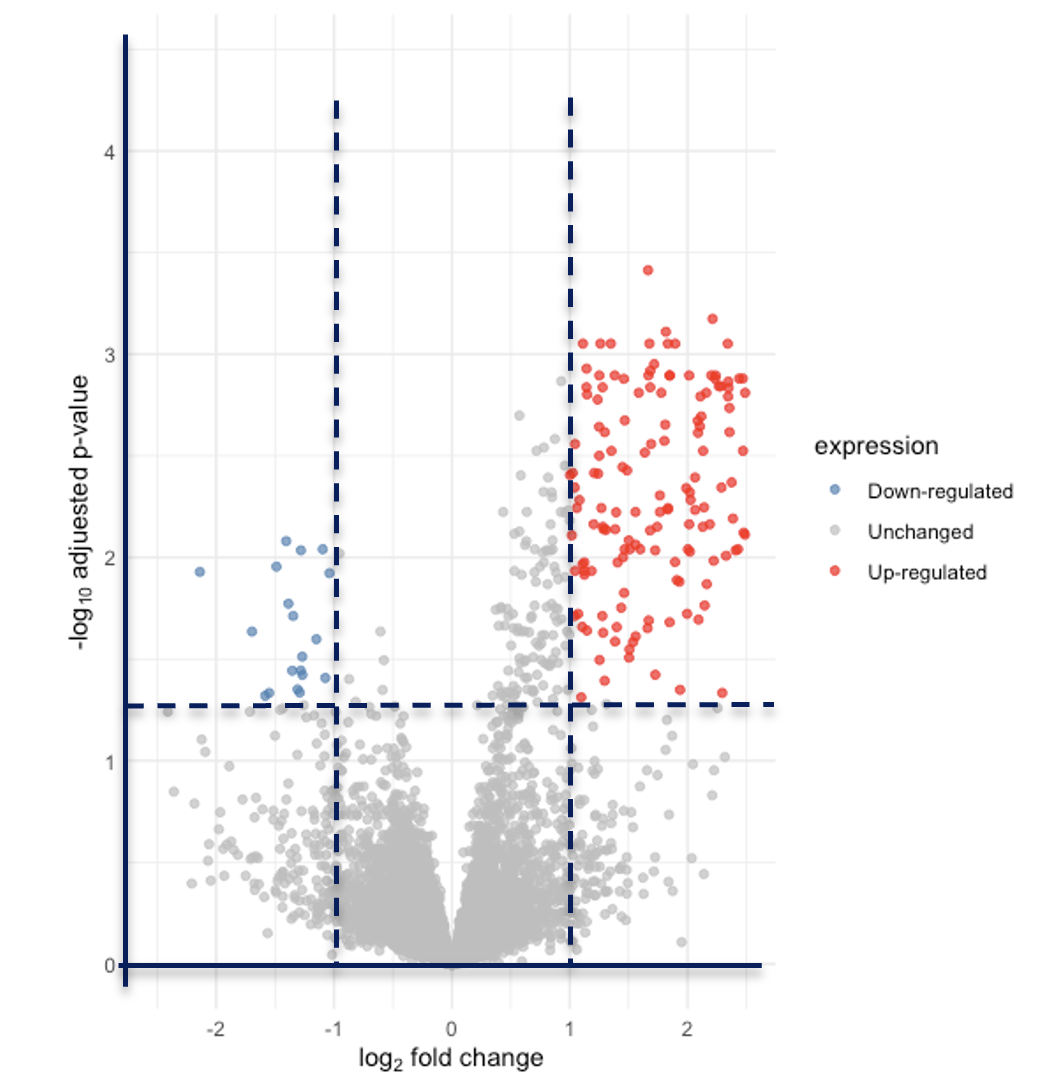


**Figure 2: PLS-DA plots for both metrics of PAH disease severity**

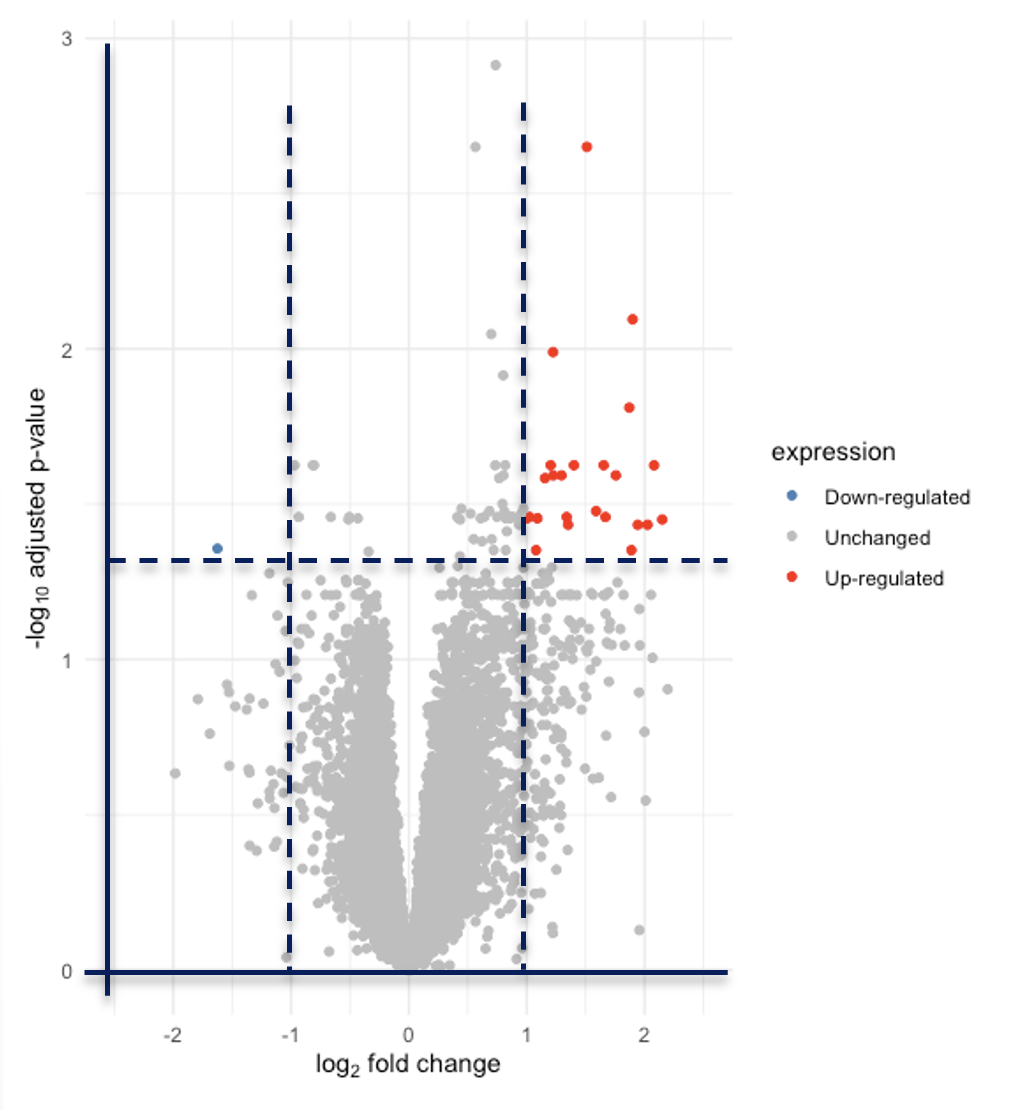


**Figure 3: Volcano plots for differentially abundant transcripts in mortality and RV dilation.**

1. **3-year Mortality**



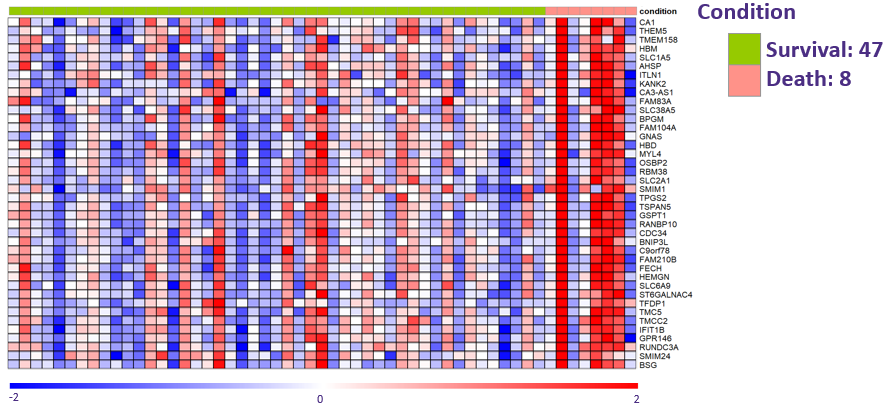
1. **RV Dilation**



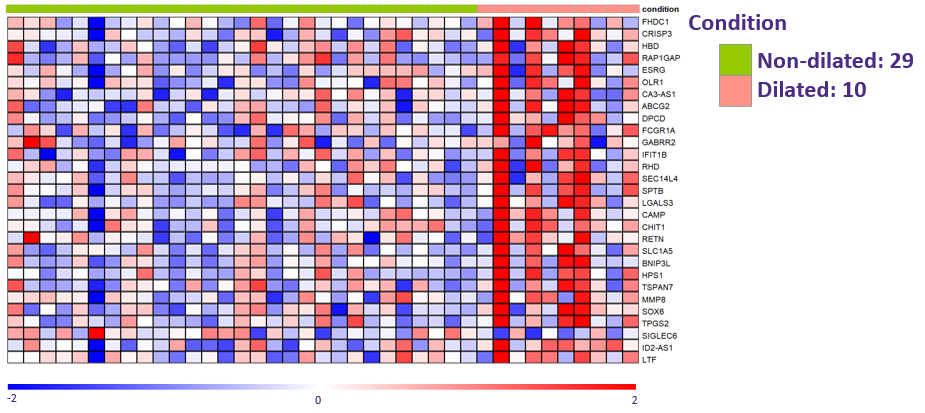
‘Down-regulated’ RNA sequences are those with adjusted p-value less than 0.05 and fold change less than -2; ‘Up-regulated’ RNA sequences are those with adjusted p-value less than 0.05 and fold change larger than 2.

**Figure 4: Heatmaps for top differentially expressed transcripts in mortality and RV dilation**

1. **3-year Mortality**

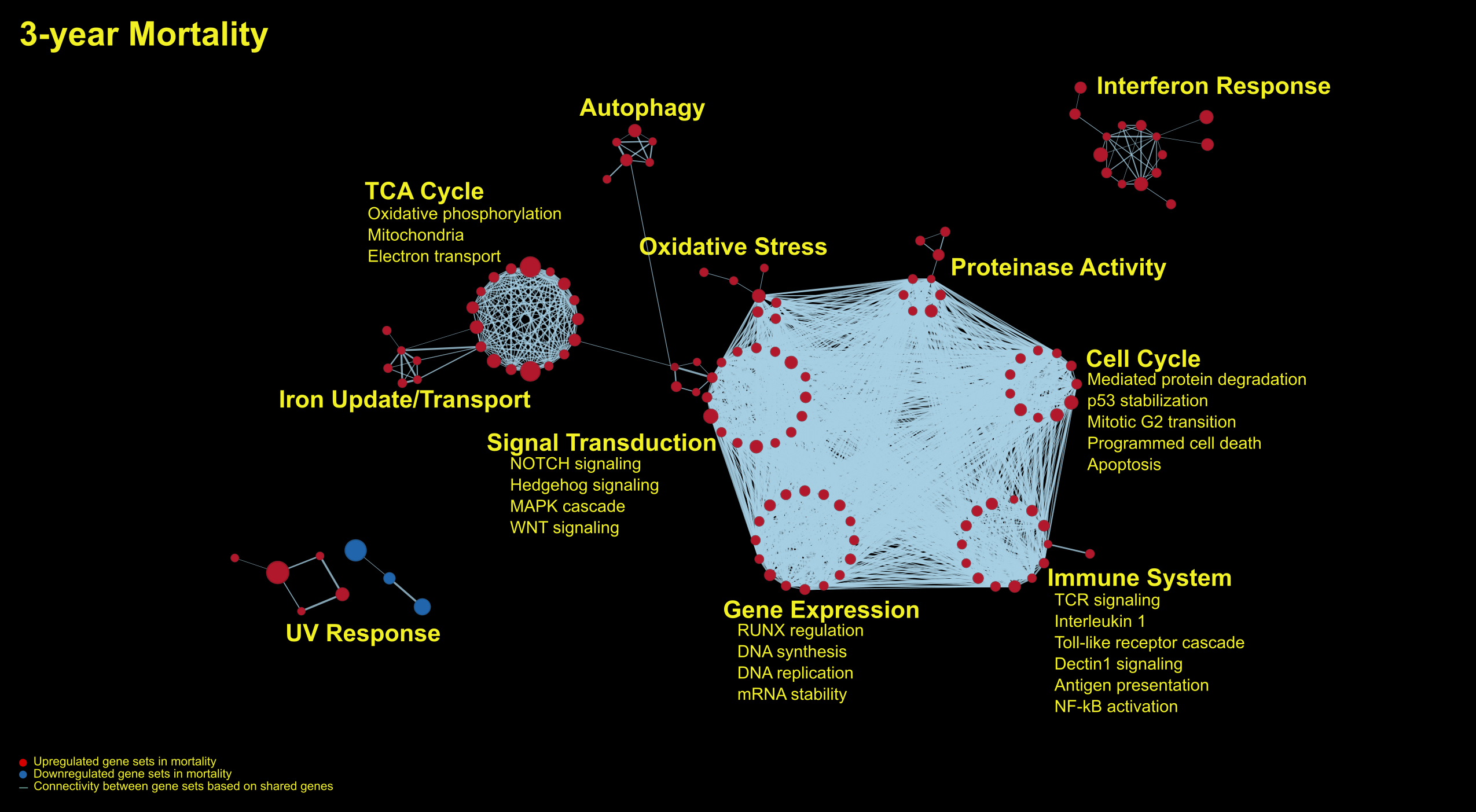


1. **RV Dilation**

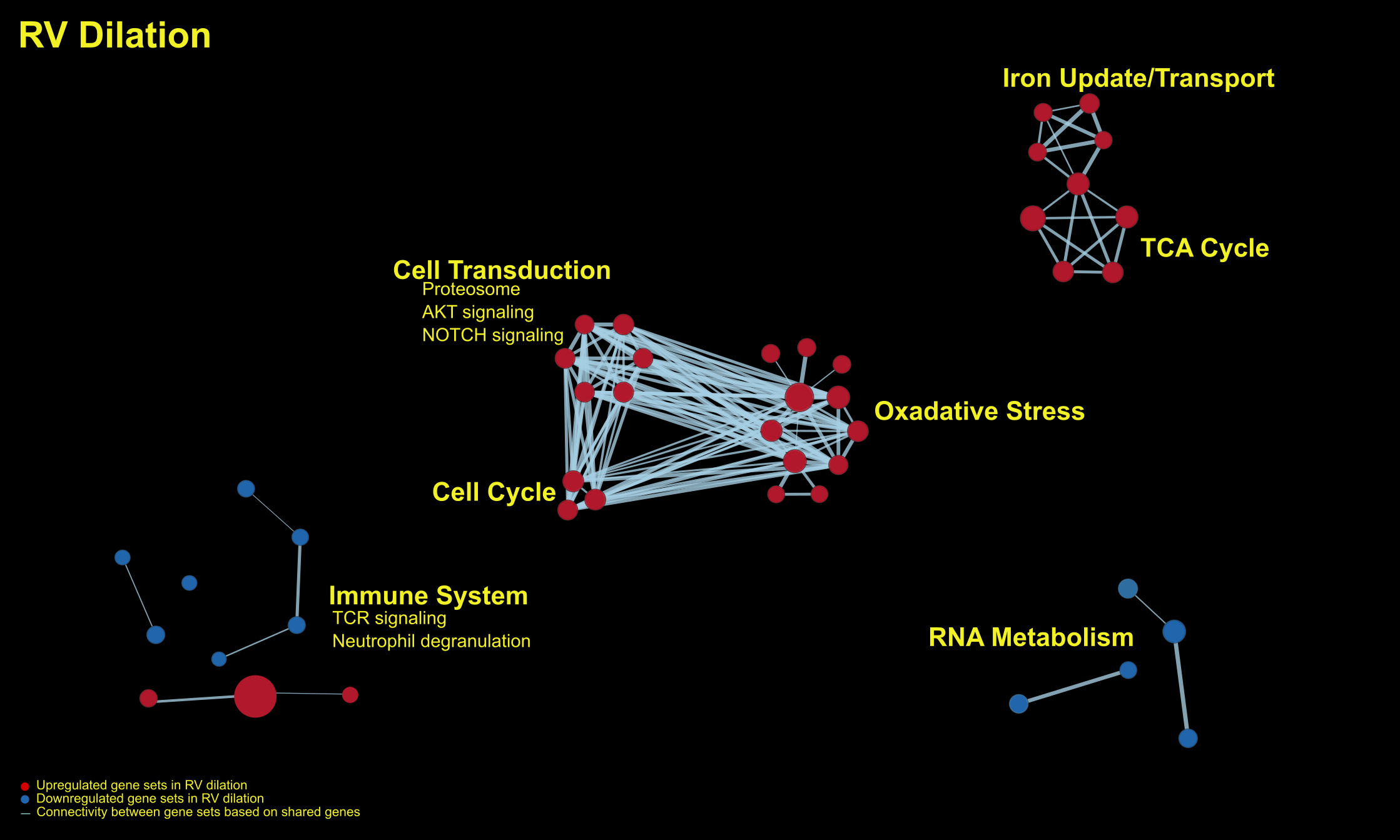


**Figure 5: Pathway enrichment analyses for mortality and RV dilation**

1. **3-year mortality**



1. **RV dilation**



**References:**

1. Galiè N, Humbert M, Vachiery J, Gibbs S, Lang I, Torbicki A, Simonneau G, Peacock A, Vonk Noordegraaf A, Beghetti M, et al. 2015 ESC/ERS Guidelines for the Diagnosis and Treatment of Pulmonary Hypertension. *Eur Respir J*. 2015 Oct;46(4):903-75. PMID: 26318161.
2. Badesch DB, Raskob GE, and Elliott CG. Pulmonary arterial hypertension: baseline characteristics from the REVEAL Registry. *Chest.*2010;137(2):376-87. PMID: 19837821.
3. D’Alonzo GE, Barst RJ, and Ayres SM. Survival in patients with primary pulmonary hypertension: results from a national prospective registry. *Ann Intern Med.* 1991;115(5):343-9. PMID: 1863023.
4. Maron BA and Galiè N. Pulmonary arterial hypertension diagnosis, treatment, and clinical management in the contemporary era. *JAMA Cardiol.*2016;1(9):1056-65. PMID: 27851839.
5. Pauwaa S, Machado RF, and Desai AA. Survival in pulmonary arterial hypertension: a brief review of registry data. *Pulm Circ.*2011;1(3):430-1. PMID: 22140634.
6. Vonk-Noordegraaf A, Haddad F, Chin KM, Forfia PR, Kawut SM, Lumens J, Naeije R, Newman J, Oudiz RJ, Provencher S, et al. Right heart adaptation to pulmonary arterial hypertension: physiology and pathobiology. *J Am Coll Cardiol*. 2013 Dec 24;62(25 Suppl):D22-33. PMID: 24355638.
7. D’Alonzo GE, Barts RJ, Ayres SM, Bergofsky EH, Brundage BH, Detre KM, Fishman AP, Goldring RM, Groves BM, Kernis JT. Survival in Patients with Primary Pulmonary Hypertension. *Ann Intern Med*. 1991 Sep 1;115(5):343-9. PMID: 1863023.
8. Campo A, Mathai SC, Le Pavec J, Zaiman AL, Hummers LK, Boyce D, Housten T, Lechtzin N, Chami H, Girgis RE, et al. Outcomes of hospitalisation for right heart failure in pulmonary arterial hypertension. *Eur Respir J*. 2011 Aug;38(2):359-67. PMID: 21310884.
9. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, Deswal A, Drazner MH, Dunlay SM, Evers LR, et al. 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2022;145:e895–e1032.
10. Vonk-Noordegraaf A, Westerhof BE, and Westerhof N. The Relationship Between the Right Ventricle and Its Load in Pulmonary Hypertension. [*J Am Coll Cardiol*](https://www.jacc.org/journal/jacc). 2017 Jan;69(2):236-43. PMID: 28081831.
11. Rich S, Pogoriler J, Husain AN, Toth PT, Gomberg-Maitland M, Archer SL. Long-Term Effects of Epoprostenol on the Pulmonary Vasculature in Idiopathic Pulmonary Arterial Hypertension. *Chest*. 2010 Nov;138(5):1234-9. PMID: 21051399.
12. Pi H, Carlson SD, Barros LM, Hogl L, Kirkpatrick JN, Nolley S, Ralph DD, Rayner SG, Leary PJ. Risk prediction with REVEAL 2.0 and right ventricular dilation in a single institution pulmonary arterial hypertension cohort. *Journal of American Heart Association*. 2022 July;0:e025521. PMID: 35861845.
13. Benza RL, Gomberg-Maitland M, Elliott CG, Farber HW, Foreman AJ, Frost AE, McGoon MD, Pasta DJ, Selej M, Burger CD, et al. Predicting Survival in Patients With Pulmonary Arterial Hypertension: The REVEAL Risk Score Calculator 2.0 and Comparison With ESC/ERS-Based Risk Assessment Strategies. *Chest.* 2019;156:323-37. PMID: 30772387.
14. Wu AC, Kiley JP, Noel PJ, Amur S, Burchard EG, Clancy JP, Galanter J, Inada M, Jones TK, Kropski JA, et al. Current Status and Future Opportunities in Lung Precision Medicine Research with a Focus on Biomarkers. An American Thoracic Society/National Heart, Lung, and Blood Institute Research Statement. *Am J Respir Crit Care Med*. 2018 Dec 15;198(12):e116–e136. PMID: 30640517.
15. Moore WC, Fitzpatrick AM, Li X, Hastie AT, Li H, Meyers DA, Bleecker ER. Clinical Heterogeneity in the Severe Asthma Research Program. *Ann Am Thorac Soc*. 2013 Dec;10(Suppl):S118–S124. PMID: 24313761.
16. Han MK, Agusti A, Calverley PM, Celli BR, Criner G, Curtis JL, Fabbri LM, Goldin JG, Jones PW, Macnee W, et al. Chronic obstructive pulmonary disease phenotypes: the future of COPD. *Am J Respir Crit Care Med*. 2010 Sep 1;182(5):598–604. PMID: 20522794.
17. Rhodes CJ, Wharton J, and Wilkins MR. Metabolomic Insights in Pulmonary Arterial Hypertension. Advances in Pulmonary Hypertension. 2018;17(3):103-9.
18. Rhodes CJ, Ghataorhe P, Wharton J, Rue-Albrecht KC, Hadinnapola C, Watson G, Bleda M, Haimel M, Coghlan G, Corris PA, et al. Plasma Metabolomics Implicates Modified Transfer RNAs and Altered Bioenergetics in the Outcomes of Pulmonary Arterial Hypertension. *Circulation.* 2017 Jan 31;135(5):460-75. PMID: 27881557.
19. Tian T, Fan J, and Elf SE. Metabolon: a novel cellular structure that regulates specific metabolic pathways. *Cancer Commun (Lond)*. 2021 Jun;41(6):439-41. PMID: 33939322.
20. Pang Z, Chong J, Zhou G, Morais D, Chang L, Barrette M, Gauthier C, Jacques PE, Li S, and Xia J. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucl Acids Res*. 2021 Jul;49(W1):W388-396. PMID: 34019663.
21. Liao Y, Wang J, Jaehnig E, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019 Jul 2;47(W1):W199-W205. PMID: 31114916.
22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102:15545-50. PMID: 16199517.
23. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021 Jan 8;49(D1):D605-12. PMID: 33237311.
24. Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, Chen RO, Brownstein BH, Cobb JP, Tschoeke SK, et al. A network-based analysis of systemic inflammation in humans. *Nature*. 2005;437:1032-7. PMID: 16136080.
25. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*. 2011;27:431-2. PMID: 21149340.
26. Ruiz-Perez, D., Guan, H., Madhivanan, P. et al. So you think you can PLS-DA?. BMC Bioinformatics 21 (Suppl 1), 2 (2020). <https://doi.org/10.1186/s12859-019-3310-7>
27. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001 Nov 1;125(1-2):279-84. doi: 10.1016/s0166-4328(01)00297-2. PMID: 11682119.