Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 18 June 2026
Revised 27 July 2026
Accepted 06 August 2026

Natural Resources for Human Health 2026, Vol. 6 (6): 651–656

KEYWORDS:
Bioinformatics, congenital heart diseases, genome-wide association study, hub genes, pulmonary arterial hypertension.

Integrative Genomic Profiling Reveals Novel Diagnostic Signatures in Congenital Heart Disease with Pulmonary Arterial Hypertension

Gina Noor Djalilah[1,5], Bagus Setyoeboedi[2*], Meity Ardiana[3], Ali Rohman[4], Mahrus A Rahman[2], I Ketut Alit Utamayasa[2], Taufiq Hidayat[2], Clevia Revi[5]

[1]Doctoral Program of Medical Science, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia
[2]Department of Paediatrics, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia
[3]Department of Cardiology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia
[4]Department of Chemistry, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
[5]Faculty of Medicine, Universitas of Muhammadiyah Surabaya, Surabaya, Indonesia
*Corresponding Author: bagus.setyoboedi1@gmail.com

ABSTRACT: High pressure in the lung arteries that keeps getting worse is called pulmonary arterial hypertension (PAH). Children with congenital heart disease (CHD) are commonly affected. PAH-CHD accounts for approximately 28% of all PAH cases, significantly impacting patients' quality of life. Exact mechanism still unclear. Studies point to cellular and molecular factors, genetic changes among them. This study sought genetic variants tied to PAH-CHD occurrence, using a genetic-driven biomarker approach. A bioinformatic-based method was employed, integrating multiple genomic databases. Data were drawn from several resources, including the Genome-Wide Association Studies (GWAS) catalog, GTEx portal, HaploReg version 4.1, and Ensembl databases. The aim was to explore genes that heighten vulnerability to PAH-CHD. Using integrated bioinformatics analysis, several key genes were identified as potentially involved in the pathogenesis of PAH-CHD. Notably, HLA-DPA1, BMPR2, TBX4, and SOX17 emerged as candidate genes that may contribute to disease susceptibility. These genes have been previously implicated in vascular and cardiac developmental pathways, supporting their possible roles in PAH-CHD. This study demonstrates the utility of a bioinformatics-driven approach to uncover genetic variants potentially involved in PAH-CHD. The identified genes may serve as valuable biomarkers, offering insights into disease mechanisms and potential targets for further investigation and therapeutic development.

1. INTRODUCTION

High pressure in the lung arteries causes pulmonary arterial hypertension (Hansmann et al., 2019; Zhu et al., 2018). Congenital heart disease ties to this condition, which brings high death rates in adults and children alike (W. Zhao et al., 2025). Affecting approximately 0.0001%–0.0002% of the population, PAH is a rare vascular disease (M. Zhao et al., 2023).

The leading contributor to PAH is congenital heart disease (CHD), a cardiovascular structural defect that is present at birth (Varela et al., 2018). Based on database data, 28% of cases involve PAH associated with PAH-CHD (Awad et al., 2022; Jone et al., 2023). PAH develops when structural heart defects drive up pulmonary blood flow, PAD, and cardiac workload. Then PAH

feeds the cycle, pushing more blood to the heart and increasing its workload again. Life quality worsens as a result (Jone et al., 2023; Karma & Yurianda, 2024; Montani et al., 2022).

Pulmonary vascular remodeling may come from cellular and molecular factors, per some sources, though the exact mechanism remains unknown (Awad et al., 2022; Widiastari et al., 2022). First described in 2000, a pathogenic BMPR2 variant is the leading cause of familial or hereditary PAH (Lane et al., 2000; Tatius et al., 2021). Since then, pathogenic variants in several other genes have been identified (Best et al., 2014; Morrell et al., 2019; Taha & Southgate, 2022). In determining genes as the cause of PAH-CHD, it is expected to help investigate genetic variants that are at risk of developing PAH-CHD (Hayvaci Canbeyli et al., 2025). Therefore, several genomic databases were combined in a systematic way to find biomarkers behind PAH-CHD.

Pairing the Genome-Wide Association Study (GWAS) Catalog with the GTEx Portal, this study sought genetic markers for specific PAH-CHD. Overall, merging these data may uncover PAH-CHD biomarkers and offer new understanding of the cellular mechanism underneath.

2. MATERIALS AND METHODS

The workflow outlined in Figure 1 guided the screening of variants linked to PAH-CHD, with each step of the methodology summarized therein. To consolidate PAH-CHD variants, a range of bioinformatic tools was consulted, namely the GWAS catalog, GTEx portal, HaploReg version 4.1, and Ensembl databases.

Single Nucleotide Polymorphisms (SNPs) associated with PAH-CHD first underwent screening via the GWAS catalog database [https://www.ebi.ac.uk/gwas/] (accessed on September 27, 2025), with "Pulmonary Arterial Hypertension" serving as the search term. To ensure that only unique SNPs were retained, every duplicate variant was removed from the dataset. The expression of these variants was then investigated using a bioinformatic approach.

Next, the rigorously filtered SNPs were validated against the GTEx portal database [https://www.gtexportal.org/home/] (accessed on September 27, 2025). This stage was designed to assess how these genetic variants relate to gene expression profiles across a range of body tissues. Once the GTEx portal search was complete, the genes encoded by the SNPs were retrieved so that gene expression profiles could be identified. Regions harboring nucleotides correlated with changes in gene expression, known as expression quantitative trait loci (eQTL), were applied. As a result, the variants may drive shifts in gene expression within the tissues concerned.

The epidemiological distribution of the genetic basis of allele frequency across populations was then mapped with the aid of HaploReg version 4.1 and Ensembl genomic databases. Genetic databases, including HaploReg version 4.1 and Ensembl Genome, served as the source of allele frequencies associated with PAH-CHD. Confirming allele frequencies among populations spanning multiple continents (Africa, America, Asia, and Europe) is the reason for determining each variant's allele frequency.

3. RESULTS AND DISCUSSION

3.1 Results

From the GWAS catalog database, SNPs associated with pulmonary arterial hypertension were collected. Of these, 38 SNPs were determined to have passed the inclusion criteria. Tissue expression of five SNPs linked to PAH-CHD was then evaluated using the GTEx portal database. This procedure is designed so that a close association between the five SNPs and gene expression in lung and heart tissues can be verified.

The tissue expression of PAH-CHD was proposed to be affected by several SNPs, namely rs285683 or rs9277336, rs13266183, rs10103692, and rs3744439.

The GTEx portal database was employed to assess how related genes are expressed across human tissues. HLA-DPA1, BMPR2, TBX4, and SOX17, four gene symbols associated with PAH-CHD, were detected through the database and were shown to be expressed in the heart. For this reason, gene and variant expression in heart tissues was made the focus of the present study.

Natural Resources for Human Health
eISSN: 2583-1194

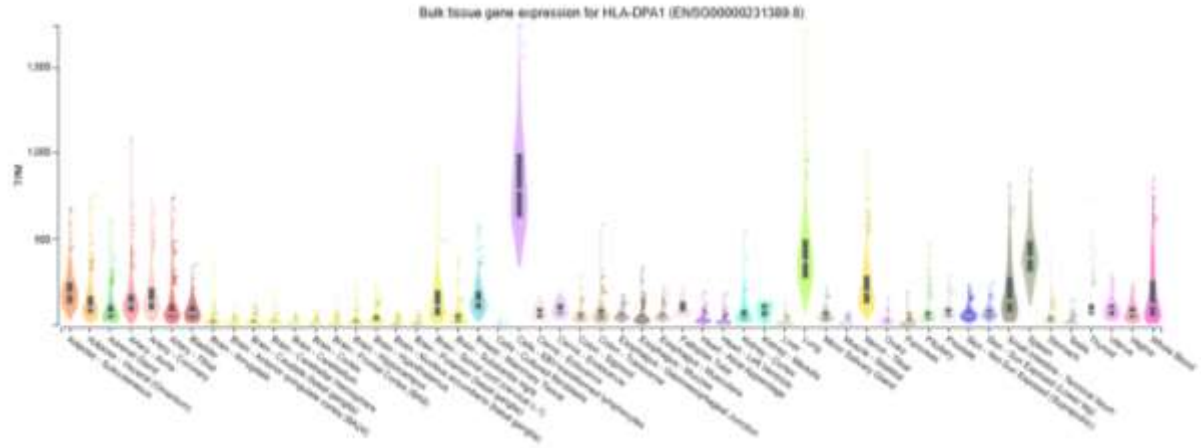


Figure 1. Gene expressions for HLA-DPA1

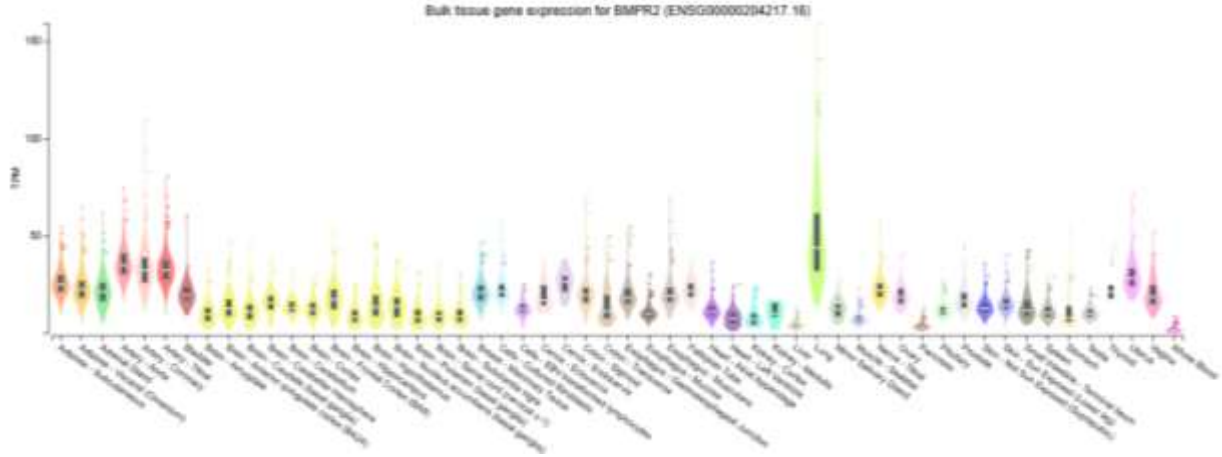


Figure 2. Gene expressions for BMPR2

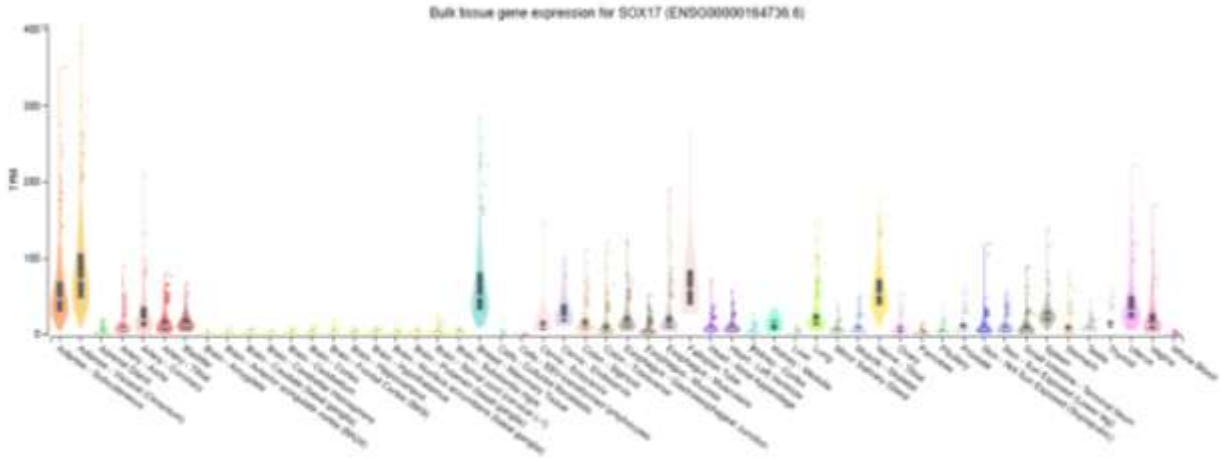


Figure 3. Gene expressions for SOX17

Natural Resources for Human Health
eISSN: 2583-1194

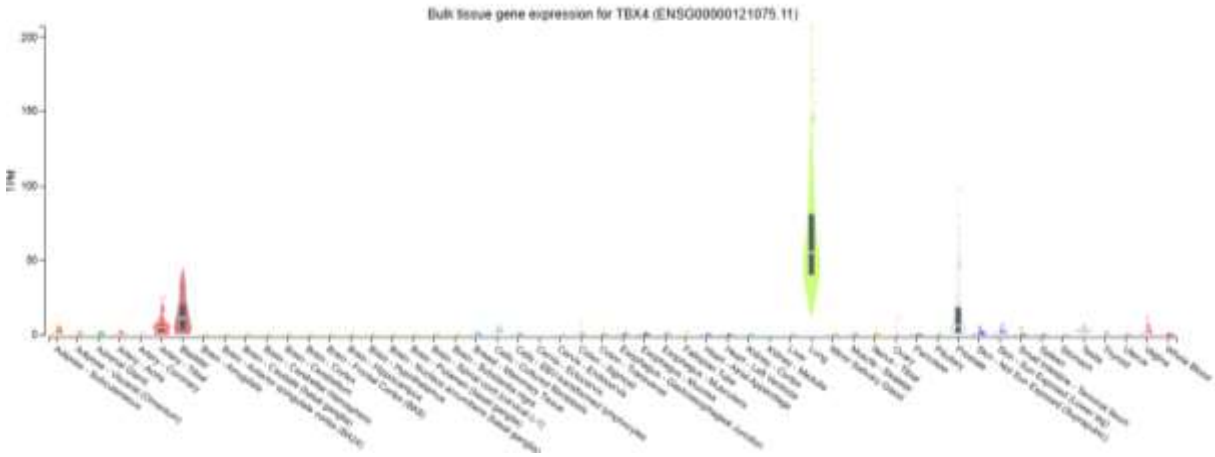


Figure 4. Gene expressions for TBX4

Minor alleles of HLA-HLADPA-1 rs285683 or rs9277336, BMPR2, TBX4 rs3744439, and SOX17 rs13266183 rs10103692 were identified as respectively associated with PAH-CHD using the publicly available GTEx portal database, as shown in Table 1.

Table 1. Variants associated with pulmonary arterial hypertension in congenital heart disease

SNP ID (Example)	GeneCode (Ensembl)	Gene Symbol	Tissue (Dominant Expression)
rs285683 or rs9277336	ENSG00000231389	**HLA-DPA1**	Lung, immune system, leukocyte., spleen
rs3744439	ENSG00000182634	**TBX4**	Lungs, smooth muscle, embryonic tissue
rs13266183 rs10103692	ENSG00000164484	**SOX17**	Vascular endothelium, liver, kidney, lung
rs10510747	ENSG00000204217	**BMPR2**	Lungs, pulmonary arteries, heart, connective tissue

3.2 Discussion

The potential role of genetic variants in the pathogenesis of pulmonary arterial hypertension associated with congenital heart disease (PAH-CHD) is brought into focus by the findings of this study. The research results revealed several variants within HLA-DPA1, TBX4, SOX17, and BMPR2 that could contribute to disease susceptibility and progression. Consistent with previous studies, BMPR2 remains one of the most frequently implicated genes in hereditary and idiopathic PAH, with mutations leading to impaired signaling in the TGF-β/BMP pathway and vascular remodeling (Best et al., 2014; Lane et al., 2000; Tatius et al., 2021; Welch et al., 2021). Similarly, TBX4 variants have been increasingly reported in pediatric PAH and are associated with abnormal lung development and vascular changes, underscoring their relevance in PAH-CHD (Al Hafiz et al., 2024; Austin & Elliott, 2020; Hernandez-Gonzalez et al., 2020; Taha & Southgate, 2022).

Interestingly, research analysis identified variants in SOX17, a transcription factor critical for vascular development and endothelial homeostasis. Rare deleterious variants in SOX17 have previously been reported in PAH-CHD, supporting its role in congenital heart defects and pulmonary vascular remodeling (Rhodes et al., 2019; Zhu et al., 2018). The present study further emphasizes SOX17 as a potential biomarker that bridges developmental and vascular pathways (Adu-Amankwaah et al., 2025; Zhu et al., 2018). Meanwhile, variants in HLA-DPA1, a gene involved in immune regulation, may indicate an immunogenetic component in PAH pathogenesis. Given the emerging evidence of immune dysregulation and inflammation in PAH, the association of HLA variants could represent an additional pathogenic mechanism linking congenital heart defects with pulmonary vascular disease (Prapa et al., 2022; Zhu et al., 2018).

The expression profiles retrieved from the GTEx database revealed tissue-specific patterns that strengthen the functional plausibility of these genes. For instance, BMPR2 and SOX17 were predominantly expressed in pulmonary vascular endothelium

and cardiac tissues, while TBX4 was expressed in the lung mesenchyme and embryonic tissues, consistent with their developmental roles. HLA-DPA1 expression was mainly localized in immune-related tissues such as leukocytes and spleen, reinforcing the concept that immune responses may influence PAH-CHD progression (Prapa et al., 2022).

Allele frequency analysis further demonstrated population-specific distributions of these variants, which may partially explain ethnic differences in PAH prevalence and outcomes (Rhodes et al., 2019). For example, HLA alleles display notable variability across continents, suggesting that genetic background could influence individual susceptibility to PAH-CHD. These findings emphasize the importance of incorporating population genetics into biomarker discovery and risk stratification.

Taken together, research results suggest that PAH-CHD is a genetically heterogeneous condition involving pathways of vascular remodeling, cardiopulmonary development, and immune regulation (Best et al., 2014; Hayvaci Canbeyli et al., 2025; Taha & Southgate, 2022). The integration of multi-database bioinformatics approaches provides a comprehensive overview of potential biomarkers and highlights candidate genes for future functional studies. Nonetheless, the current findings should be interpreted cautiously, as bioinformatic analyses are hypothesis-generating and require experimental validation in patient cohorts and cellular models.

4. CONCLUSION

In conclusion, the identification of variants in HLA-DPA1, TBX4, SOX17, and BMPR2 provides new insights into the molecular mechanisms of PAH-CHD. These genes may serve as promising biomarkers for early detection, risk prediction, and personalized therapeutic strategies. Future studies integrating genomic, transcriptomic, and clinical data are warranted to translate these genetic discoveries into clinical practice.

REFERENCES

[1] Adu-Amankwaah, J., You, Q., Liu, X., Jiang, J., Yang, D., Liu, K., Yuan, J., Wang, Y., Hu, Q., & Tan, R. (2025). Pulmonary Hypertension: Molecular Mechanisms and Clinical Studies. *MedComm*, *6*(3), e70134. https://doi.org/10.1002/mco2.70134

[2] Al Hafiz, A. H., Supriyadi, F. L., & Raymond, B. (2024). Gene Expressions in the Incidence of Non-Syndromic Cleft Lip and Palate. *PHARMACOLOGY, MEDICAL REPORTS, ORTHOPEDIC, AND ILLNESS DETAILS*, *3*(3), 122–138. https://doi.org/10.55047/comorbid.v3i3.1656

[3] Austin, E. D., & Elliott, C. G. (2020). TBX4 syndrome: a systemic disease highlighted by pulmonary arterial hypertension in its most severe form. In *European Respiratory Journal* (Vol. 55, Issue 5). European Respiratory Society. https://doi.org/10.1183/13993003.00585-2020

[4] Awad, A., Elnemr, S., Hodeib, H., & El Amrousy, D. (2022). Platelet activation markers in children with pulmonary arterial hypertension associated with congenital heart disease. *Pediatric Cardiology*, *43*(6), 1264–1270. https://doi.org/10.1007/s00246-022-02847-7

[5] Best, D. H., Austin, E. D., Chung, W. K., & Elliott, C. G. (2014). Genetics of pulmonary hypertension. *Current Opinion in Cardiology*, *29*(6), 520–527. https://doi.org/10.1097/hco.0000000000000105

[6] Hansmann, G., Koestenberger, M., Alastalo, T.-P., Apitz, C., Austin, E. D., Bonnet, D., Budts, W., D'Alto, M., Gatzoulis, M. A., Hasan, B. S., Kozlik-Feldmann, R., Kumar, R. K., Lammers, A. E., Latus, H., Michel-Behnke, I., Miera, O., Morrell, N. W., Pieles, G., Quandt, D., … Zartner, P. (2019). 2019 updated consensus statement on the diagnosis and treatment of pediatric pulmonary hypertension: The European Pediatric Pulmonary Vascular Disease Network (EPPVDN), endorsed by AEPC, ESPR and ISHLT. *The Journal of Heart and Lung Transplantation*, *38*(9), 879–901. https://doi.org/10.1016/j.healun.2019.06.022

[7] Hayvaci Canbeyli, F., Secgen, K., Ezgu, F. S., Tacoy, G., Unlu, S., Arabacı, H. O., Pektas, A., Inci, A., Kaya, E. B., & Sinan, U. Y. (2025). The Role of Genetics in Congenital Heart Disease-Associated Pulmonary Arterial Hypertension. *Pediatric Cardiology*, *47*, 1–9. https://doi.org/10.1007/s00246-025-03847-z

[8] Hernandez-Gonzalez, I., Tenorio, J., Palomino-Doza, J., Martinez Meñaca, A., Morales Ruiz, R., Lago-Docampo, M., Valverde Gomez, M., Gomez Roman, J., Enguita Valls, A. B., & Perez-Olivares, C. (2020). Clinical heterogeneity of pulmonary arterial hypertension associated with variants in TBX4. *PLoS One*, *15*(4), e0232216. https://doi.org/10.1371/journal.pone.0232216

[9] Jone, P.-N., Ivy, D. D., Hauck, A., Karamlou, T., Truong, U., Coleman, R. D., Sandoval, J. P., del Cerro Marin, M. J., Eghtesady, P., & Tillman, K. (2023). Pulmonary hypertension in congenital heart disease: a scientific statement from the American Heart Association. *Circulation: Heart Failure*, *16*(7), e00080. https://doi.org/10.1161/HHF.0000000000000080

[10] Karma, T., & Yurianda, Z. (2024). The Relationship between Physical Activity and the Incidence of Hypertension in the Elderly in Lamme Village, Jaya District, Aceh Jaya Regency. *PHARMACOLOGY, MEDICAL REPORTS, ORTHOPEDIC, AND ILLNESS DETAILS*, *3*(1), 33–40. https://doi.org/10.55047/comorbid.v3i1.1123

[11] Lane, K. B., Machado, R. D., Pauciulo, M. W., Thomson, J. R., Phillips, J. A., Loyd, J. E., Nichols, W. C., & Trembath,

R. C. (2000). Heterozygous germline mutations in BMPR2, encoding a TGF-β receptor, cause familial primary pulmonary hypertension. *Nature Genetics*, *26*(1), 81–84. https://doi.org/10.1038/79226

[12] Montani, D., Lechartier, B., Girerd, B., Eyries, M., Ghigna, M.-R., Savale, L., Jaïs, X., Seferian, A., Jevnikar, M., Boucly, A., Riou, M., Traclet, J., Chaouat, A., Levy, M., Le Pavec, J., Fadel, E., Perros, F., Soubrier, F., Remy-Jardin, M., … Humbert, M. (2022). An emerging phenotype of pulmonary arterial hypertension patients carrying SOX17 variants. *European Respiratory Journal*, *60*(6), 2200656. https://doi.org/10.1183/13993003.00656-2022

[13] Morrell, N. W., Aldred, M. A., Chung, W. K., Elliott, C. G., Nichols, W. C., Soubrier, F., Trembath, R. C., & Loyd, J. E. (2019). Genetics and genomics of pulmonary arterial hypertension. *European Respiratory Journal*, *53*(1). https://doi.org/10.1183/13993003.01899-2018

[14] Prapa, M., Lago-Docampo, M., Swietlik, E. M., Montani, D., Eyries, M., Humbert, M., Welch, C. L., Chung, W. K., Berger, R. M. F., Bogaard, H. J., Danhaive, O., Escribano-Subías, P., Gall, H., Girerd, B., Hernandez-Gonzalez, I., Holden, S., Hunt, D., Jansen, S. M. A., Kerstjens-Frederikse, W., … Valverde, D. (2022). First Genotype–Phenotype Study in TBX4 Syndrome: Gain-of-Function Mutations Causative for Lung Disease. *American Journal of Respiratory and Critical Care Medicine*, *206*(12), 1522–1533. https://doi.org/10.1164/rccm.202203-0485oc

[15] Rhodes, C. J., Batai, K., Bleda, M., Haimel, M., Southgate, L., Germain, M., Pauciulo, M. W., Hadinnapola, C., Aman, J., Girerd, B., Arora, A., Knight, J., Hanscombe, K. B., Karnes, J. H., Kaakinen, M., Gall, H., Ulrich, A., Harbaum, L., Cebola, I., … Wilkins, M. R. (2019). Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *The Lancet. Respiratory Medicine*, *7*(3), 227–238. https://doi.org/10.1016/S2213-2600(18)30409-0

[16] Taha, F., & Southgate, L. (2022). Molecular genetics of pulmonary hypertension in children. *Current Opinion in Genetics & Development*, *75*, 101936. https://doi.org/10.1016/j.gde.2022.101936

[17] Tatius, B., Wasityastuti, W., Astarini, F. D., & Nugrahaningsih, D. A. A. (2021). Significance of BMPR2 mutations in pulmonary arterial hypertension. *Respiratory Investigation*, *59*(4), 397–407. https://doi.org/10.1016/j.resinv.2021.03.011

[18] Varela, D. L., Teleb, M., & El-Mallah, W. (2018). Advanced therapies for the management of adults with pulmonary arterial hypertension due to congenital heart disease: a systematic review. *Open Heart*, *5*(1). https://doi.org/10.1136/openhrt-2017-000744

[19] Welch, C. L., Austin, E. D., & Chung, W. K. (2021). Genes that drive the pathobiology of pediatric pulmonary arterial hypertension. *Pediatric Pulmonology*, *56*(3), 614–620. https://doi.org/10.1002/ppul.24637

[20] Widiastari, E. F., Graha, W. A., & Pardede, M. (2022). Lung Bullae Due to Septic Pulmonary Embolism in a 4-Year-Old Child: A Case Report. *Pharmacology, Medical Reports, Orthopedic, and Illness Details (COMORBID)*, *1*(3). https://doi.org/10.55047/comorbid.v1i3.355

[21] Zhao, M., Liu, J., Xin, M., Yang, K., Huang, H., Zhang, W., Zhang, J., & He, S. (2023). Pulmonary arterial hypertension associated with congenital heart disease: An omics study. *Frontiers in Cardiovascular Medicine*, *10*, 1037357. https://doi.org/10.3389/fcvm.2023.1037357

[22] Zhao, W., Chen, Y., Yu, Z., Wang, Z., He, R., Cai, S., Chen, Z., & Cheng, L. (2025). Risk associated circulating biomarkers S100A3 identified in congenital heart disease-associated pulmonary arterial hypertension. *Intractable & Rare Diseases Research*, *14*(1), 36–45. https://doi.org/10.5582/irdr.2024.01064

[23] Zhu, N., Welch, C. L., Wang, J., Allen, P. M., Gonzaga-Jauregui, C., Ma, L., King, A. K., Krishnan, U., Rosenzweig, E. B., Ivy, D. D., Austin, E. D., Hamid, R., Pauciulo, M. W., Lutz, K. A., Nichols, W. C., Reid, J. G., Overton, J. D., Baras, A., Dewey, F. E., … Chung, W. K. (2018). Rare variants in SOX17 are associated with pulmonary arterial hypertension with congenital heart disease. *Genome Medicine*, *10*(1), 1–11. https://doi.org/10.1186/s13073-018-0566-x