This article has been accepted for publication in JMEDGENET following peer review.
The definitive copyedited, typeset version is available online at
10.1136/jmg-2025-111027

Expanding the phenotypic spectrum of *MECOM*-associated syndrome: rare variants are associated with syndromic pulmonary arterial hypertension

Carrie L Welch,[1,2] Meriel McEntagart,[3] Shahin Moledina,[4] Cara Morgan,[4] Emilia Swietlik,[5,6] Chao Hou,[7] Lu Qiao,[7] Emily Callejo,[1] Savanna Craib,[8] Damian Smedley, [9] Emilia K. Bijlsma,[10] Patrice Bouvagnet,[11,12] Nahir Cortes-Santiago,[13,14] Tamir Dagan,[15] Jacqueline Eason,[16] Frances Flinter,[17] Aakash Joshi,[18] Jeremie Mortreux,[11] Fadel E Ruiz,[19,20] Deborah Shears,[18] Celia Azevedo Soares,[21] Nidhy P Varghese,[19,20] Wendy K Chung[1,2**]

1. Department of Pediatrics, Boston Children's Hospital, Boston, USA (CLW, EC, WKC).

2. Department of Pediatrics, Harvard Medical School, Boston, USA (CLW, WKC).

3. St George's University Hospitals NHS Foundation Trust, London, UK (MM).

4. Great Ormond Street Hospital, London, UK (SM, CM).

5. Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK (ES).

6. University of Warmia and Mazury, Olsztyn, Poland

7. Department of Systems Biology, Columbia University Irving Medical Center, New York, USA (CH, LQ).

8. Department of Pediatrics, Columbia University Irving Medical Center, New York, USA (SC).

9. William Harvey Research Institute, Queen Mary University of London, London, UK (DS)

10. Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands (EKB).

11. Laboratoire génétique moléculaire, Eurofins Biomnis, Lyon, France (PB, JM).

12. Hôpital MFME, CHU Martinique, Fort de France, France (PB).

13. Department of Pathology, Texas Children's Hospital, Texas, USA (NCS).

14. Department of Pathology & Immunology, Baylor College of Medicine, Texas, USA (NCS).

15. Schneider Children's Medical Center, Petah Tikva, Israel (TD).

16. Nottingham University Hospitals NHS Trust, Nottingham, UK (JE).

17. Department of Clinical Genetics, Guy's & St Thomas' NHS Foundation Trust, London, UK (FF).

18. Oxford University Hospital NHS Foundation Trust, Oxford, UK (AJ, DS).

19. Department of Pediatrics, Division of Pulmonology, Baylor College of Medicine, Houston, USA (FER, NPV).

20. Pulmonary Medicine and Pulmonary Hypertension Center, Texas Children's Hospital , Houston, USA (FER, NPV).

21. Medical Genetics Department, Centro de Genética Médica Dr. Jacinto Magalhães, Unidade Local de Saúde Santo António, Porto (CAS).

**Correspondence: Dr. Wendy K Chung, wendy.chung@childrens.harvard.edu

ABSTRACT

Background – *MECOM* encodes a developmental and hematopoietic transcription factor associated with a rare early-onset syndrome including bone marrow failure, skeletal and other congenital anomalies. Heterozygous *de novo* variants are the primary cause. We previously identified *MECOM* as a candidate gene for pediatric pulmonary arterial hypertension (PAH) using trio exome sequencing.

Methods –To test the role of *MECOM* in pediatric PAH, and further define the clinical phenotype of *MECOM*-associated syndrome, we queried GeneMatcher and screened rare disease databases for individuals with predicted deleterious *MECOM* variants. We analyzed the clinical spectrum of patients, performed protein modeling of genetic variants, and assessed cardiopulmonary expression.

Results – We identified fifteen individuals with *MECOM* variants, including eleven unrelated probands and eight *de novo* variants. Eleven individuals had severe or mild thrombocytopenia, nine had skeletal issues, eight had cardiac anomalies, six had PAH, and ten had additional conditions. Three were diagnosed *in utero* and died in the neonatal period. All missense variants map to the zinc finger 6 or zinc finger 8/9 region, a known hotspot for *MECOM*-associated syndrome. Protein modeling predicted that both regions are DNA-binding, and that the variants may interfere with binding to a *VEGFR2*/*KDR* enhancer. Data from LungMAP showed that *MECOM* is primarily expressed in pulmonary arterial endothelial cells.

Conclusion – Rare *MECOM* variants are associated with early-onset syndromic PAH. PAH monitoring should be considered for all individuals with rare *MECOM* variants. We

speculate that the pathogenetic mechanism for PAH and cardiac defects may be impaired *VEGFR2*/*KDR* signaling.

- **What is already known on this topic** – *MECOM*-associated syndrome (MIM: 616738) is a rare, early-onset syndrome associated with bone marrow failure, skeletal abnormalities, cardiac anomalies, and other congenital anomalies. Approximately 30% of pediatric pulmonary arterial hypertension (PAH) is syndromic in nature, but few of the causative genes have been identified.
- **What this study adds** – Rare variants in *MECOM* are associated with syndromic PAH with comorbidities overlapping the phenotypic spectrum of *MECOM*-associated syndrome. Missense variants in zinc finger 6 and the zinc finger 8/9 hotspot are predicted to interfere with binding to a *VEGFR2*/*KDR* enhancer.
- **How this study might affect research, practice or policy** – Genetic diagnoses for children with syndromic forms of PAH and other developmental diseases can provide critical information for clinical monitoring and management.

Introduction

MECOM-associated syndrome (MIM: 616738, also called RUSAT2, radioulnar synostosis with amegakaryocytic thrombocytopenia 2) is a rare, early-onset syndrome associated with bone marrow failure (BMF), skeletal anomalies, and other congenital anomalies. *MECOM*, the myelodysplastic syndrome and ecotropic virus integration site-1 complex locus, encodes a widely expressed transcription factor containing ten zinc finger (ZF) domains. Somatic inversions and translocation involving chromosome 3 (including *MECOM*) frequently cause *MECOM* activation leading to acute myeloid leukemia[1] whereas germline heterozygous loss of function variants and deletions cause reduced expression leading to BMF.[2] *De novo* variants are the primary cause of *MECOM*-associated syndrome, including both protein-truncating and missense variants. Most of the missense variants cluster within the ZF 8/9 region (amino acids 919-974).[3] The phenotypic spectrum of individuals with *MECOM*-associated syndrome has expanded to include congenital heart disease, genitourinary anomalies, developmental delay and hearing impairment.[3] Furthermore, individuals can exhibit severe early-onset BMF requiring hematopoietic stem cell transplantation within the first few years of life, mild or late-onset thrombocytopenia, or no hematopoietic deficiency.[3] Similarly, radioulnar synostosis is a common feature but not seen in all individuals, and the presence of other congenital anomalies or developmental delay is highly variable.[3]

Pulmonary arterial hypertension (PAH [MIM: 178600]) is a rare pulmonary vascular disease with high morbidity and mortality. The diagnosis of pediatric PAH is made by hemodynamic assessment and is defined mean by pulmonary artery pressure >20 mmHg,

indexed pulmonary vascular resistance >3 Wood units · m^2 ($WU \cdot m^2$), and pulmonary artery wedge pressure ≤15 mmHg.[4] PAH is caused by both genetic and environmental factors, with twelve established definitive genes harboring causal variants for PAH.[5] While PAH can manifest across the lifespan, child-onset PAH is particularly challenging because it is often syndromic involving other congenital anomalies and/or neurodevelopmental features.[6] As such, the underlying genes are likely important in early developmental pathways. Variants in developmental transcription factors *SOX17*[7-9] and *TBX4*[10] are known to cause PAH associated with congenital heart disease and developmental skeletal disorders, respectively, but the genetic causes of syndromic PAH are largely unknown.

In a previous analysis of 124 PAH child-parent trios, we identified *MECOM* and seven other genes as candidate PAH genes based on high heart/lung expression and identification of one affected child with a rare *de novo* variant for each gene.[11] To identify additional genetic evidence for a role of these genes in PAH, we queried GeneMatcher and received two potential matches for *MECOM*.

Methods

Identification of individuals with rare *MECOM* variants

We queried GeneMatcher (genematcher.org)[12] to identify individuals with rare variants in *MECOM*. We then screened genomic registries including DECIPHER (deciphergenomics.org)[13], the 100,000 Genomes Project (https://www.genomicsengland.co.uk/initiatives/100000-genomes-project)[14], the UK NIHR BioResource – Rare Diseases Study (UK NIHR BioResource) (https://www.nihr.ac.uk)[15], and

our own studies of rare neurodevelopmental disorders[16] and structural anomalies[17; 18]. Additional individuals were identified via these queries. Inclusion criteria included individuals with rare *de novo* or inherited variants (gnomAD v2.1.1/v4.1 allele frequency less than 0.01% and CADD score >20), with or without PAH. Patient expert clinicians and/or geneticists were identified and invited to collaborate on the study. During preparation of the manuscript, we realized that Individual 6 (c.2850-1G>A) had been reported previously,[19] but not in the context of PAH. Individual 8 had also been reported.[2] Genetic variants were identified by trio (n = 8) or duo (n = 1) exome sequencing, gene panel (n = 1, 483 BMF genes; n = 1, 530 skeletal dysplasia genes), or cascade testing after a variant was identified in an affected family member. Family 2 also had copy number variant screening. Genetic, demographic, family history, and clinical data for individuals with variants that met our criteria were collected from the patients' clinicians and/or geneticists. The study was approved by the institutional ethics committees of the participating centers and written informed consent was obtained from the families involved, in accordance with the Declaration of Helsinki.

Clinical classification of *MECOM* variants

Liftover of variant coordinates to human genome assembly GRCh38 was performed when needed. Transcript coordinates are according to MANE select NM_004991.4/ ENST00000651503.2 and variant nomenclature was checked using variant validator (variantvalidator.org). Variant curation was performed according to the American College of Medical Genetics/American Association of Medical Pathology (ACMG/AMP) 2015 guidelines,[20] and including updated recommendations of the ClinGen Sequence Variant

Interpretation Working Group (clinicalgenome.org). AlphaMissense[21] was used as the *in silico* pathogenicity prediction tool. A Bayesian points-based system for evidence code combinations[22] was applied to determine the clinical classifications. Briefly, a classification of pathogenic was defined by >=10 points, likely pathogenic 6-9 points, variant of uncertain significance -1 to 5 points, likely benign -2 to -6 points, and benign <= -7 points.

Protein modeling of novel variants

MECOM NP_004982 was used for all analyses. Variant mapping onto the two-dimensional protein structure was performed using the trackViewer package in R/Bioconductor,[23] with conserved domain coordinates based on UniProt Q03112.3 (https://www.uniprot.org/)[24]. The three-dimensional structure of MECOM was predicted using the AlphaFold3 web server (https://alphafoldserver.com).[25] Structure visualization and electrostatic surface analysis were conducted using UCSF ChimeraX version 1.8 (http://www.rbvi.ucsf.edu/chimerax).[26; 27] To enhance clarity in visualization, disordered and coil regions (residues 1-78, 428-919, and 1003-1239) were excluded from the final representations.

DNA binding analysis was performed using BindUP (http://bindup.technion.ac.il/),[28] a web server for non-homology-based prediction of DNA- and RNA-binding proteins. This analysis identified a large positively charged patch around ZF6/8/9, suggesting DNA-binding activity.

To model the MECOM-DNA complex, we used a previously characterized MECOM binding site[29] in the *VEGFR2* regulatory region (5'-CTGGTGACTCACAA-3', including 3-

nucleotide extensions on both ends). The initial structural model was generated using AlphaFold3, incorporating three components: full-length MECOM protein sequence, double-stranded DNA, and ten Zn^{2+} ions corresponding to the zinc finger domains. Analysis of this initial model confirmed that ZF6/8/9 interact with DNA. To refine these predictions, additional models were generated: ZF6 region (residues 374-399) with the same DNA sequence and one Zn^{2+} ion; ZF8/9 region with the same DNA sequence and two Zn^{2+} ions.

For visualizing variants affecting Zn^{2+} coordination, zinc finger domains were first predicted using AlphFold3. Variants were introduced using pdb4amber (http://ambermd.org)[30] followed by energy minimization with OpenMM (http://openmm.org)[31] to optimize the mutated structures.

Gene expression data

Quantitative gene expression data were retrieved from LungMAP Phase 2 LGEA reference data (lungmap.net).[32] Two data sets were queried: normal human lung (median age of donors, 41 years; interquartile range, 29-61 years; similar numbers of females and males) and mean percentage of gene-expressing cells/total number cells sampled. Differential gene set expression was queried using the LGEA gene list query of LungGENS data. Aorta/heart expression data were retrieved from *Tabula Muris* (https://tabula-muris.sf.czbiohub.org/).[33]

Results

Novel rare *MECOM* variants extend the phenotypic spectrum of *MECOM*-associated syndrome to include PAH

We identified fifteen individuals (5 females, 10 males) with rare, predicted deleterious *MECOM* variants, including eleven unrelated probands (Table 1 and Supplementary Table 1). The eleven variants included four likely gene disrupting (LGD: 2 nonsense, 1 frameshift, and 1 canonical splice) and seven missense variants. There were eight confirmed *de novo*, one presumed *de novo* (the father was not available for testing), and two inherited variants. All variants were absent from the population database, gnomAD v4.1. All missense variants were predicted deleterious by *in silico* tools (CADD and AlphaMissense[21]), and all were pathogenic/likely pathogenic by ACMG/AMP criteria (Supplementary Table 2).

Demographic and clinical characteristics for all individuals are summarized in Table 1 and Supplementary Table 1. Detailed clinical synopses for all individuals are provided in Supplementary Table 3 and pedigrees for the two families are in Supplementary Figure 1. Most individuals were of European ancestry, and there were one each of French Caribbean, Ashkenazi Jewish, Hispanic, or unknown ancestry. Three individuals were diagnosed *in utero* and died in the neonatal period. The remaining individuals were diagnosed within 1 week of birth (n = 3), at 1-5 years of age (n = 6), as adults (n = 2, fathers of affected children), or unknown (n = 1). At least eleven individuals had thrombocytopenia/pancytopenia, nine had skeletal issues (7 with radioulnar synostosis, 4 club foot, 3 hand anomalies, 1 patellar hypoplasia, and 1 patellar contracture), seven had cardiac anomalies (4 with patent ductus arteriosus, 4 ventricular septal defect, 2 truncus arteriosus, 1 each of aortic root dilation/atrial septal defect/atrioventricular valve

dysplasia/mitral valve prolapse), six had PAH, and ten had additional medical conditions. Three children had bone marrow transplantation (including 2 that still developed PAH), and one had a heart and lung transplant. Three newborns died at birth (bilateral renal agenesis leading to anhydramnios) or 1-3 weeks of life (n = 1, intracerebral hemorrhage; n = 1, following complex surgery for a type 4 truncus arteriosus that could have been worsened by undiagnosed PAH or other comorbidities). One child died at 5-10 years of age due to complications of progressive right ventricular failure and chronic respiratory failure. Both fathers had subclinical hematologic findings that were not diagnosed until after genetic diagnoses, and one had adult-onset unilateral hearing loss.

Six individuals were diagnosed with PAH by right heart catheterization (RHC) (Table 2). Mean pulmonary arterial pressure ranged from 35-62 mm Hg, pulmonary vascular resistance indexed to body surface area 7-30 $WU \cdot m^2$, and wedge pressures were in the normal range. Three infants died before hemodynamics could be assessed. Two babies were born with persistent pulmonary hypertension of the newborn and then diagnosed with severe PAH at the age of 1-5 years (Table 2). An additional four children were diagnosed at 1-5 years with PAH. All but one PAH case was associated with congenital cardiac anomalies. Vasodilator therapy regimens ranged from mono- to triple therapy. Since no one met criteria for vasoreactivity on catheterization, none were candidates for treatment with calcium channel blockers. The children were too young to perform diffusion capacity measurements of the lungs. The World Health Organization PAH functional classifications ranged from I-III. Five children had follow-up visits at 6-15 years of age, three with RHC. Despite combination therapy, mean pulmonary arterial pressures and

pulmonary vascular resistance remained high (50-60 mm Hg and 11.8-14.1 $WU \cdot m^2$, respectively). At follow-up, three children were able to have diffusion capacity measurements which were in the normal range. One child had a heart-lung transplant at 11-15 years of age and was alive at 21-25 years of age. One child, who had undergone successful bone marrow transplantation in infancy, died at 6-10 years of age from progressive right ventricular failure and acute-on-chronic respiratory failure (see below). One child was placed in hospice care at 11-15 years of age due to worsening clinical status; she was not considered a candidate for lung transplant and no further surgical or catheterization interventions were available. The three remaining children were alive and well (1-25 years of age).

Pathologic findings of the lung and heart from individual 6, diagnosed with PAH associated with congenital heart disease, are shown in Figures 1 and 2. Chest computed tomography showed diffuse multifocal pneumonia with effusion, and stable yet significant prominence of the cardiac silhouette and central pulmonary vasculature (Figure 1). Post-mortem pathologic examination of the lungs showed severe vasculopathic changes affecting the entire spectrum of pulmonary vessels, including pulmonary arteries, pulmonary microvasculature, and venules and veins (Figure 2A-D). The vessels showed variable occlusion, often severe and predominantly by intimal expansion. The intima had an unusual edematous to myxoid appearance in many of the smaller vessels, with variable but frequently brisk associated inflammation (Figure 2A). This inflammatory component is unusual in the context of classic PAH and in this patient it was unclear if it was related to his history of bone marrow transplantation or if it was a primary component of his

underlying *MECOM*-associated lung disease. Within medium to larger pulmonary artery branches and veins, classic fibrous to fibromyxoid intimal expansion was more frequently seen (Figure 2C-D). Plexiform lesions were infrequent but present and they occasionally showed associated fibrin thrombi (Figure 2E). A single medium-sized pulmonary artery branch showed features of necrotizing vasculitis without granulomas or infectious organisms noted with special stains (image not shown). Given the localized nature of this finding, and the absence of pulmonary hemorrhage, the etiology and significance remained uncertain. Finally, most alveolar septa were thin, delicate and appropriately developed with few relatively limited foci of markedly dilated and tortuous alveolar capillaries imparting a pulmonary capillary hemangiomatosis-like picture. However, no proliferation of capillaries was present on closer examination (Figure 2F). The dilated and tortuous alveolar capillaries were mostly seen near larger vessels with severe occlusive lesions.

2D protein mapping reveals *MECOM* missense genotype-phenotype relationships

MECOM is a complex locus with multiple alternative transcripts. The full-length transcript encodes a protein with conserved histone methyltransferase, DNA-binding, and protein-interaction domains. Protein locations of the rare variants identified in this study, and from previous reports of *MECOM*-associated syndrome ([2; 19; 34-42 3]), are provided in Figure 3.

The four LGD variants identified in this study were expected to result in nonsense-mediated decay, leading to haploinsufficiency. All four variants were associated with thrombocytopenia or pancytopenia, consistent with previous findings.[3] Monozygotic twin

fetuses diagnosed with p.(Gly219*) had a severe phenotype with unilateral or bilateral renal agenesis and other complications (Table 1); only one survived to birth but died on day one and therefore were incompletely clinically characterized. The father was found to have thrombocytopenia (130 x 10E9 platelets/microliter) and no renal findings. A healthy sibling had the same variant with a patent ductus arteriosus. The other three individuals with LGD variants ((p.(Lys835Thrfs*3), p.(Leu843*), and c.2850-1G>A)) had congenital cardiac anomalies (ventricular septal defect, n = 2; patent ductus arteriosus, n = 2; aortic root dilation, n =1) and PAH; one also had radioulnar synostosis and one had autism spectrum disorder.

MECOM contains ten ZF domains, and the missense variants reported herein map to the ZF6 (amino acids 376-398) or the ZF8/9 region (amino acids 919-974) (Figure 2A). Both individuals identified with ZF6 missense variants had the same congenital heart anomaly, truncus arteriosus, and these are the only reported individuals with MECOM-associated syndrome to have a variant in ZF6. Neither individual had hematological findings. The individual with p.(Tyr385 Cys) also had developmental issues including attention deficit disorder and intellectual disability. The individual with p.(Arg393His) also had a left club foot and died at 0-3 weeks of age from post-surgical cardiac complications, potentially worsened by PAH. We identified six individuals with missense variants in the hot spot ZF8/9 region (amino acid range 938-958). These individuals all had complex phenotypes including hematologic issues and radioulnar synostosis or hand/foot anomalies (n = 6), PAH (n = 4), and cardiac anomalies (n = 3).

We identified six individuals with PAH-associated variants, all *de novo* (3 LGD and 3 missense) (Table 1, Figure 3). The three missense variants all map to the ZF8/9 region. At least five had complex phenotypes including cardiac (n = 4), hematologic (n = 3), skeletal (n = 3), other pulmonary (n = 2), and neurodevelopmental (n = 2) issues.

Updated clinical classification of the previously published variants was performed (Supplementary Table 4) and only pathogenic/likely pathogenic variants were included in Figure 3. Four variants of uncertain significance and two benign/likely benign variants, all located outside of the zinc finger domains, were excluded. There were a total of 37 pathogenic/likely pathogenic variants identified in 66 unrelated individuals. Twenty-four individuals (36%) had LGD variants and forty-two (64%) had predicted deleterious missense variants. Across all studies, six individuals with LGD variants (6/24, 25%) were reported to have only hematological findings, eleven (46%) hematological plus skeletal findings, seven (29%) hematological plus congenital heart disease, six (25%) hematological plus PAH, and seven (29%) hematological plus at least two of these other anomalies.

Across all studies, within the ZF8/9 region (Figure 3B), twenty-five individuals with deleterious missense variants (25/40, 62.5%) were reported to have both hematologic and skeletal anomalies, with five individuals also having congenital heart disease and/or PAH, and only three were reported to have isolated pancytopenia. In contrast, eight out of nine individuals with variants located towards the C'-terminal end of ZF9, between amino acids 965-971, had skeletal defects alone.

3D protein modeling predicts that ZF6 and the ZF8/9 region are DNA-binding

We used protein structural modeling to predict functional consequences of the *MECOM* variants reported herein (Figure 4). Both ZF6 and the ZF8/9 region are predicted to be DNA-binding (Figure 4C). ZF8 variants, p.(His939Arg) and p.(His939Tyr), and ZF9 variants, p.(Cys951Tyr) and p.(Cys954Arg), were predicted to disrupt the ZF structure by disrupting native zinc coordination (Figure 4D). Vascular endothelial cells are a key cell type in cardiac and lung development, and we hypothesized that transcriptional dysregulation of one or more MECOM target genes expressed in endothelial cells could provide a pathogenetic mechanism for truncus arteriosus and PAH development. Based on a recent report[29] that MECOM is a critical regulator of endothelial lineage specification via transcriptional activation of *VEGFR2* (vascular endothelial growth factor receptor 2) and the *VEGF* (vascular endothelial growth factor) signaling pathway, we used AlphaFold3 to model the complex structure of MECOM ZF6 or ZF8/9 bound to the experimentally determined DNA-binding motif of the *VEGFR2*/KDR promoter. The model predicted direct DNA groove interactions for both ZF6 and ZF8/ZF9 within the binding motif (Figure 4E).

Endothelial-specific *MECOM* expression in lung

Single cell RNA-seq data from LungMAP[32] (February 5, 2025) showed that *MECOM* is expressed in a high percentage of PAH-associated pulmonary arterial endothelial cells but not pulmonary arterial smooth muscle cells or pericytes, and with relative expression level similar to that of other definitive PAH genes (Figure 4F). Further analysis indicated that *MECOM*, *VEGF2R*/*KDR*, and a subset of other definitive PAH genes (*BMPR2*, *ACVRL1*,

CAV1, *ENG*, *SMAD9*, *SOX17*) comprise a gene set that is selectively highly expressed in arterial endothelial cells compared to other type lung cells (i.e. epithelial, immune, mesenchymal, mixed control, and biopsy) from neonates (p = 1.55e-9), infants (p = 3.06e-9), and children (p = 9.16e-11). In adults, *MECOM* is also expressed in airway epithelial progenitor cells, goblet cells, and other mucus-secreting cells of the lung (data not shown). Data from *Tabula Muris*[33] (February 5, 2025) showed that in murine adult heart/aorta, *MECOM* was expressed in ~25% of endothelial and endocardial cells, and fibroblasts, but not cardiac muscle. In the same dataset, *KDR* was expressed in 84% of endothelial cells, 48% of endocardial cells, but only 5% of cardiac muscle cells and fibroblasts (Supplementary Figure 2).

Discussion

We report the genotypes and phenotypes of nine unrelated individuals and two families with rare *MECOM* variants and expand the phenotypic spectrum of *MECOM*-associated syndrome to include early onset PAH. All children were diagnosed before 5 years of age; two fathers with less severe phenotypes were diagnosed as adults. Among fifteen total individuals, common features were thrombocytopenia or pancytopenia (n = 11), skeletal anomalies (n = 9), cardiac anomalies (n = 8), and development of PAH (n = 6). Three individuals were diagnosed *in utero* and died in the neonatal period before the development of PAH. We identified four LGD and seven missense variants, all classified as pathogenic/likely pathogenic. The missense variants map to ZF6 or the ZF8/9 region, all predicted to affect regions that bind DNA.

The six variants associated with PAH included three LGD (1 frameshift, 1 nonsense, and 1 canonical splicing) and 3 missense variants. All PAH variants were *de novo*, surpassing the ClinGen threshold of >=2 *de novo* variants to establish a "true" PAH candidate gene.[11] Five of the individuals with PAH also had congenital heart anomalies. The three missense variants mapped to the ZF8/9 region and structural modeling predicted that ZF8 and ZF9 directly interact with a DNA binding motif in the *VEGFR2*/*KDR* promoter. It is plausible that the variants interfere with transactivation of *VEGFR2*/*KDR*, either by abolishing zinc binding or another mechanism. *KDR* is a definitive PAH gene,[5] and VEGF signaling and endothelial cell lineage specification are critical for normal lung and heart development. We hypothesize that the pathogenetic mechanism for *MECOM* variants is haploinsufficiency (LGD variants)/loss of function (missense variants), similar to *KDR* variants[11; 43].

Another novel finding of this study was the identification of two missense variants in ZF6 associated with the same congenital heart anomaly, truncus arteriosus, and skeletal or developmental issues but not thrombocytopenia or pancytopenia. A mouse line with a hypomorphic *Mecom* variant, produced by a deletion of exon 3 leading to a new cryptic start site, resulted in 100% (6/6 mice) with congenital heart defects including 50% (3/6 mice) with truncus arteriosus[44]. Pathogenic/likely pathogenic missense variants in ZF6 have not been reported previously. Protein modeling indicated that ZF6 is DNA binding and that p.(Tyr385Cys) and p.(Arg393His) can directly interact with the same DNA binding motif in the *VEGFR2*/*KDR* promoter as variants in the ZF8/9 region. Truncus arteriosus is a rare congenital heart anomaly that occurs when the single large blood vessel fails to divide into

the pulmonary artery and the aorta during embryonic development. This process involves differentiation of neural crest cells into endothelial cells. *MECOM* is expressed in neural crest-derived cells in the outflow tract by E9.5-10.5[45], and it is interesting to speculate that *MECOM* may be involved in the differentiation of neural crest cells to endothelial cells.

Family 1/twin B suffered an intracerebral hemorrhage and died shortly after birth. In a prenatal screen of 192 fetuses with intracranial hemorrhage, Coste et al[46] identified a causative *MECOM* frameshift variant (c.1268del; p[S423Lfs*6]) in a family with recurrent fatal intracranial hemorrhage in two fetus siblings. Fetal brain examination identified hemorrhagic lesions in the first sibling and intraparenchymal and subarachnoid lesions with intraparenchymal petechiae in the second sibling. The first sibling also had pulmonary atresia with ventricular septal defect, while the second one had only eleven pairs of ribs and radioulnar synostosis. These cases suggest that intracranial hemorrhage is a rare manifestation of *MECOM*-associated syndrome, likely due to fetal thrombocytopenia. Altogether, the expanding phenotypic spectrum of *MECOM*-associated syndrome indicates that this is a developmental vascular condition.

The development of PAH in very young children with other medical conditions, including rare syndromes, can be associated with worse prognosis.[4] We recommend that clinical assessment of all individuals with thrombocytopenia (or pancytopenia) and/or radioulnar synostosis and rare deleterious *MECOM* variants should include non-invasive screening for PAH. Further, genetic testing in children with PAH should consider rare *MECOM* variants. PAH patients identified to have rare deleterious *MECOM* variants should be assessed for the full spectrum of *MECOM*-associated syndrome phenotypes.

Our comparative analysis of genotype-phenotype relationships among all known individuals with pathogenic/likely pathogenic *MECOM* variants (Figure 3), confirmed thrombocytopenia/pancytopenia in all individuals with LGD variants. Most of these individuals have at least one other *MECOM*-associated syndrome phenotype. Newly identified individuals with missense variants in ZF6 share the phenotype of truncus arteriosus. Most individuals with missense variants in the proximal ZF8/9 region (residues 938-954), had both hematologic and skeletal phenotypes, and nearly half had more than two phenotypes, whereas all but one individual with variants in residues 965-971 had a skeletal phenotype alone (Figure 3B). This suggests that residues 965-971 may confer higher DNA-binding specificity than residues 938-954.

Phenotypic heterogeneity is well-documented for *MECOM*-associated syndrome, both among unrelated and related individuals with same variants. The two families with autosomal dominant paternal inheritance included in our study demonstrate intrafamilial phenotypic heterogeneity. In Family 1, the identical twin fetuses had severe phenotypes, with pancytopenia and intracerebral hemorrhage in at least one twin at birth (the other died *in utero*) and renal agenesis was present in both. The father was asymptomatic into adulthood and found to have asymptomatic, isolated moderate thrombocytopenia only after diagnosis of the twins; genetic testing of his parents confirmed that his variant occurred *de novo*. A healthy sibling was found to have patent ductus arteriosus, only after diagnosis of the twins. In Family 2, the son presented in the clinic under five years of age with skeletal dysplasia (radioulnar synostosis, brachydactyly and fifth finger clinodactyly). His symptomatic father (radioulnar synostosis) then underwent genetic testing which led

to the identification of a rare *MECOM* variant of unknown mode of inheritance. The father also had adult-onset hearing loss with bilateral dysmorphic semicircular ear canals and an aortic aneurysm. Following genetic diagnosis, the son was found to have intermittent platelet aggregates and the father recurrent thrombocytopenia. These data support clinical screening of family members identified with causal *MECOM* variants for multi-organ subclinical phenotypes.

Phenotypic heterogeneity in *MECOM*-associated syndrome is likely due to genetic modifiers. Niihori et al[41] suggested that variable expressivity of thrombocytopenia in one family was loss of heterozygosity in the hematopoietic system with reversion of the mutant allele to wild-type. Whether genetic rescue explains variable expressivity of hematopoietic output in other families has not been reported. Inherited genetic modifiers likely explain, at least in part, the observed multi-organ phenotypic heterogeneity. Genome-wide association studies have identified a common regulatory variant in the *MECOM* promoter associated with distal limb growth regulation[47], intronic variants associated with myeloproliferative neoplasms[48] and lung function[49], and variants in a differentially methylated promoter region in placenta associated with cognitive function of offspring.[50] The relevance of any of these, or other common *MECOM* variants, in *MECOM*-associated syndrome has not been explored.

We acknowledge several limitations of the study.There is ascertainment bias based on utilization of GeneMatcher, rare disease registries, and who accesses genetic testing. However, our query to GeneMatcher was phenotype agnostic, not restricted to presence of PAH. The registries were not enriched for PAH and the variants that we identified were

absent or ultra-rare from the gnomAD population database. Two of the registries queried (100,000 Genomes Project and the UK NIHR BioResource – Rare Diseases Study) are European ancestry enriched but GeneMatcher and other patients identified were of French Caribbean, Ashkenazi Jewish, and Hispanic ancestry. PAH is a rare disease and we estimate that *de novo* variants in many developmental genes underlie early-onset PAH, each explaining a small number of cases.[11] Statistical enrichment in case-control studies of pediatric PAH is not feasible with the cohort sizes currently available. Due to the low background rate of *de novo* variants,[51] the statistical evidence for a candidate risk gene is effectively equivalent to multiplicity (or identifying ≥2 unrelated cases in the same gene). The identification of six individuals with PAH and novel or ultra-rare deleterious *MECOM* variants exceeds this threshold.

Conclusions

Based on our detailed assessment of individuals with rare pathogenic/likely pathogenic *MECOM* variants, the phenotypic spectrum of *MECOM*-associated syndrome is extended from bone marrow failure and skeletal abnormalities to include PAH, especially associated with congenital heart disease, and intracranial hemorrhage. Prompt recognition of *MECOM*-associated clinical features is important for diagnosis and management. Identification of the underlying genetic cause can inform clinical management, improve accuracy of prognosis, and facilitate cascade testing of family members and reproductive planning. Ongoing identification of *MECOM*-associated syndrome in PAH is critical to enhancing understanding of this emerging phenotype.

Data Availability

All data relevant to the study are included in the article or uploaded as supplementary information. The *MECOM* genetic variants and variant curation evidence summaries reported herein have been deposited in ClinVar under accession numbers SCV006082477 - SCV006082487. Some of the data are also available in public, open access repositories.

Funding

None.

Contributorship statement

Conceptualization, study design, and project administration: CLW, WKC; writing –original draft preparation: CLW, MM, SM, CM, ES, CH, LQ, EC, SC, EKB, PB, NC-S, TD, JE, FER, DS, CAS, NPV, WKC; writing – review and editing: CLW, MM, WKC; clinical diagnosis or clinical data gathering: MM, ES, SM, CM, EKB, PB, NC-S, TD, JE, AJ, JM, FER, DS, CAS, NPV; genetic data acquisition or analysis: CLW, MM, CH, LQ, DS, PB, FF, JM, CAS, NPV, WKC. All authors have approved the final version of the manuscript.
WKC is the guarantor.

Conflict of Interest

The authors declare no conflicts of interest.

Acknowledgments

This study makes use of data generated by the DECIPHER community. A full list of centers who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. DECIPHER is hosted by EMBL-EBI and funding for the DECIPHER project was provided by the Wellcome Trust [grant number WT223718/Z/21/Z].

Molecular graphics and analyses performed with ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Disease, and HL153009. All research at Great Ormond Street Hospital NHS Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the NIHR Great Ormond Street Hospital Biomedical Research Centre. We thank NIHR BioResource volunteers for their participation, and gratefully acknowledge NIHR BioResource centres, NHS Trusts and staff for their contribution. We thank the National Institute for Health and Care Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. Part of this research was made possible through access to data in the National Genomic Research Library, which is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The National Genomic Research Library holds data provided by patients and collected by the NHS as part of their care and data collected as part of their participation in research. The National Genomic Research Library is funded by the National Institute for Health Research and NHS England.
The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure.

Ethics Approval

The institutional review board of Boston Children's Hospital approved this study (IRB-P00046556).

References

1. Glass, C., Wilson, M., Gonzalez, R., Zhang, Y., and Perkins, A.S. (2014). The role of EVI1 in myeloid malignancies. Blood Cells Mol Dis 53, 67-76.
2. Germeshausen, M., Ancliff, P., Estrada, J., Metzler, M., Ponstingl, E., Rutschle, H., Schwabe, D., Scott, R.H., Unal, S., Wawer, A., et al. (2018). MECOM-associated syndrome: a heterogeneous inherited bone marrow failure syndrome with amegakaryocytic thrombocytopenia. Blood Adv 2, 586-596.
3. Voit, R.A., and Sankaran, V.G. (2023). MECOM Deficiency: from Bone Marrow Failure to Impaired B-Cell Development. J Clin Immunol 43, 1052-1066.
4. Ivy, D., Rosenzweig, E.B., Abman, S.H., Beghetti, M., Bonnet, D., Douwes, J.M., Manes, A., and Berger, R.M.F. (2024). Embracing the challenges of neonatal and paediatric pulmonary hypertension. Eur Respir J 64.
5. Welch, C.L., Aldred, M.A., Balachandar, S., Dooijes, D., Eichstaedt, C.A., Graf, S., Houweling, A.C., Machado, R.D., Pandya, D., Prapa, M., et al. (2023). Defining the clinical validity of genes reported to cause pulmonary arterial hypertension. Genet Med 25, 100925.
6. Abman, S.H., Mullen, M.P., Sleeper, L.A., Austin, E.D., Rosenzweig, E.B., Kinsella, J.P., Ivy, D., Hopper, R.K., Raj, J.U., Fineman, J., et al. (2022). Characterisation of paediatric pulmonary hypertensive vascular disease from the PPHNet Registry. Eur Respir J 59.
7. 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., et al. (2018). Rare variants in SOX17 are associated with pulmonary arterial hypertension with congenital heart disease. Genome Med 10, 56.
8. Zhao, L., Jiang, W.F., Yang, C.X., Qiao, Q., Xu, Y.J., Shi, H.Y., Qiu, X.B., Wu, S.H., and Yang, Y.Q. (2021). SOX17 loss-of-function variation underlying familial congenital heart disease. Eur J Med Genet 64, 104211.
9. Mullen, M.P., Ivy, D.D., Varghese, N.P., Winant, A.J., Cortes-Santiago, N., Vargas, S.O., Porres, D., Maschietto, N., Critser, P.J., Hirsch, R., et al. (2024). SOX17 - Associated Pulmonary Hypertension in Children: A Distinct Developmental and Clinical Syndrome. J Pediatr, 114422.
10. Karolak, J.A., Welch, C.L., Mosimann, C., Bzdega, K., West, J.D., Montani, D., Eyries, M., Mullen, M.P., Abman, S.H., Prapa, M., et al. (2023). Molecular Function and Contribution of TBX4 in Development and Disease. Am J Respir Crit Care Med 207, 855-864.
11. Zhu, N., Swietlik, E.M., Welch, C.L., Pauciulo, M.W., Hagen, J.J., Zhou, X., Guo, Y., Karten, J., Pandya, D., Tilly, T., et al. (2021). Rare variant analysis of 4241 pulmonary

arterial hypertension cases from an international consortium implicates FBLN2, PDGFD, and rare de novo variants in PAH. Genome Med 13, 80.
12. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). Data from: GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat 36, 928-930. December 5, 2020. https://doi: 10.1002/humu.22844
13. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). Data from: DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet 84, 524-533. March 1, 2024. https://doi.10.1016/j.ajhg.2009.03.010
14. Turnbull, C., Scott, R.H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F.B., Halai, D., Baple, E., Craig, C., Hamblin, A., et al. (2018). The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. BMJ 361, k1687. March 31, 2022. https://doi: 10.1136/bmj.k1687
15. Graf, S., Haimel, M., Bleda, M., Hadinnapola, C., Southgate, L., Li, W., Hodgson, J., Liu, B., Salmon, R.M., Southwood, M., et al. (2018). Identification of rare sequence variation underlying heritable pulmonary arterial hypertension. Nat Commun 9, 1416.
16. Zhou, X., Feliciano, P., Shu, C., Wang, T., Astrovskaya, I., Hall, J.B., Obiajulu, J.U., Wright, J.R., Murali, S.C., Xu, S.X., et al. (2022). Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes. Nat Genet 54, 1305-1319.
17. Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., et al. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat Genet 49, 1593-1601.
18. Qiao, L., Welch, C.L., Hernan, R., Wynn, J., Krishnan, U.S., Zalieckas, J.M., Buchmiller, T., Khlevner, J., De, A., Farkouh-Karoleski, C., et al. (2024). Common variants increase risk for congenital diaphragmatic hernia within the context of de novo variants. Am J Hum Genet.
19. Lozano Chinga, M.M., Bertuch, A.A., Afify, Z., Dollerschell, K., Hsu, J.I., John, T.D., Rao, E.S., Rowe, R.G., Sankaran, V.G., Shimamura, A., et al. (2023). Expanded phenotypic and hematologic abnormalities beyond bone marrow failure in MECOM-associated syndromes. Am J Med Genet A 191, 1826-1835.
20. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17, 405-424.
21. Cheng, J., Novati, G., Pan, J., Bycroft, C., Zemgulyte, A., Applebaum, T., Pritzel, A., Wong, L.H., Zielinski, M., Sargeant, T., et al. (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science 381, eadg7492.

22. Tavtigian, S.V., Harrison, S.M., Boucher, K.M., and Biesecker, L.G. (2020). Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. Hum Mutat 41, 1734-1737.
23. Ou, J., and Zhu, L.J. (2019). trackViewer: a Bioconductor package for interactive and integrative visualization of multi-omics data. Nat Methods 16, 453-454.
24. UniProt, C. (2023). Data from: UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res 51, D523-D531. October 15, 2024. https://doi: 10.1093/nar/gkae1010
25. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Data from: Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493-500. October 15, 2024. https://doi: 10.1038/s41586-024-07487-w
26. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27, 14-25.
27. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci 30, 70-82.
28. Paz, I., Kligun, E., Bengad, B., and Mandel-Gutfreund, Y. (2016). BindUP: a web server for non-homology-based prediction of DNA and RNA binding proteins. Nucleic Acids Res 44, W568-574.
29. Lv, J., Meng, S., Gu, Q., Zheng, R., Gao, X., Kim, J.D., Chen, M., Xia, B., Zuo, Y., Zhu, S., et al. (2023). Epigenetic landscape reveals MECOM as an endothelial lineage regulator. Nat Commun 14, 2390.
30. Case, D.A., Aktulga, H.M., Belfon, K., Cerutti, D.S., Cisneros, G.A., Cruzeiro, V.W.D., Forouzesh, N., Giese, T.J., Gotz, A.W., Gohlke, H., et al. (2023). AmberTools. J Chem Inf Model 63, 6183-6191.
31. Eastman, P., Galvelis, R., Pelaez, R.P., Abreu, C.R.A., Farr, S.E., Gallicchio, E., Gorenko, A., Henry, M.M., Hu, F., Huang, J., et al. (2024). OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials. J Phys Chem B 128, 109-116.
32. Guo, M., Morley, M.P., Jiang, C., Wu, Y., Li, G., Du, Y., Zhao, S., Wagner, A., Cakar, A.C., Kouril, M., et al. (2023). Data from: Guided construction of single cell reference for human and mouse lung (LungMap). Nat Commun 14, 4566. January 18, 2024. https://doi: 10.1038/s41467-023-40173-5
33. Tabula Muris Consortium; Overall coordination; Logistical coordination; Organ collection and processing; Library preparation; sequencing; Computational data analysis; Cell type annotation; Writing group; Principal investigators. (2018). Data from: Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature 562, 367-372. February 27, 2025. https://doi: 10.1038/s41586-018-0590-4
34. Bluteau, O., Sebert, M., Leblanc, T., Peffault de Latour, R., Quentin, S., Lainey, E., Hernandez, L., Dalle, J.H., Sicre de Fontbrune, F., Lengline, E., et al. (2018). A landscape of germ line mutations in a cohort of inherited bone marrow failure patients. Blood 131, 717-732.

35. Lord, S.V., Jimenez, J.E., Kroeger, Z.A., Patrick, C.S., Sanchez-Pena, I., Ziga, E., Bademci, G., and Tekin, M. (2018). A MECOM variant in an African American child with radioulnar synostosis and thrombocytopenia. Clin Dysmorphol 27, 9-11.
36. Osumi, T., Tsujimoto, S.I., Nakabayashi, K., Taniguchi, M., Shirai, R., Yoshida, M., Uchiyama, T., Nagasawa, J., Goyama, S., Yoshioka, T., et al. (2018). Somatic MECOM mosaicism in a patient with congenital bone marrow failure without a radial abnormality. Pediatr Blood Cancer 65, e26959.
37. Ripperger, T., Hofmann, W., Koch, J.C., Shirneshan, K., Haase, D., Wulf, G., Issing, P.R., Karnebogen, M., Schmidt, G., Auber, B., et al. (2018). MDS1 and EVI1 complex locus (MECOM): a novel candidate gene for hereditary hematological malignancies. Haematologica 103, e55-e58.
38. Walne, A., Tummala, H., Ellison, A., Cardoso, S., Sidhu, J., Sciuccati, G., Vulliamy, T., and Dokal, I. (2018). Expanding the phenotypic and genetic spectrum of radioulnar synostosis associated hematological disease. Haematologica 103, e284-e287.
39. Loganathan, A., Munirathnam, D., and Ravikumar, T. (2019). A novel mutation in the MECOM gene causing radioulnar synostosis with amegakaryocytic thrombocytopenia (RUSAT-2) in an infant. Pediatr Blood Cancer 66, e27574.
40. Al-Abboh, H., Zahra, A., and Adekile, A. (2022). A novel MECOM variant associated with congenital amegakaryocytic thrombocytopenia and radioulnar synostosis. Pediatr Blood Cancer 69, e29761.
41. Niihori, T., Tanoshima, R., Sasahara, Y., Sato, A., Irie, M., Saito-Nanjo, Y., Funayama, R., Shirota, M., Abe, T., Okuyama, Y., et al. (2022). Phenotypic heterogeneity in individuals with MECOM variants in 2 families. Blood Adv 6, 5257-5261.
42. Shen, F., Yang, Y., Zheng, Y., Li, P., Luo, Z., Fu, Y., Zhu, G., Mei, H., Chen, S., and Zhu, Y. (2022). MECOM-related disorder: Radioulnar synostosis without hematological aberration due to unique variants. Genet Med 24, 1139-1147.
43. Swietlik, E.M., Greene, D., Zhu, N., Megy, K., Cogliano, M., Rajaram, S., Pandya, D., Tilly, T., Lutz, K.A., Welch, C.C.L., et al. (2020). Bayesian Inference Associates Rare KDR Variants with Specific Phenotypes in Pulmonary Arterial Hypertension. Circ Genom Precis Med 14, e003155.
44. Bard-Chapeau, E.A., Szumska, D., Jacob, B., Chua, B.Q., Chatterjee, G.C., Zhang, Y., Ward, J.M., Urun, F., Kinameri, E., Vincent, S.D., et al. (2014). Mice carrying a hypomorphic Evi1 allele are embryonic viable but exhibit severe congenital heart defects. PLoS One 9, e89397.
45. Hoyt, P.R., Bartholomew, C., Davis, A.J., Yutzey, K., Gamer, L.W., Potter, S.S., Ihle, J.N., and Mucenski, M.L. (1997). The Evi1 proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development. Mech Dev 65, 55-70.
46. Coste, T., Vincent-Delorme, C., Stichelbout, M., Devisme, L., Gelot, A., Deryabin, I., Pelluard, F., Aloui, C., Leutenegger, A.L., Jouannic, J.M., et al. (2022). COL4A1/COL4A2 and inherited platelet disorder gene variants in fetuses showing intracranial hemorrhage. Prenat Diagn 42, 601-610.
47. Li, J., Glover, J.D., Zhang, H., Peng, M., Tan, J., Mallick, C.B., Hou, D., Yang, Y., Wu, S., Liu, Y., et al. (2022). Limb development genes underlie variation in human fingerprint patterns. Cell 185, 95-112 e118.

48. Tapper, W., Jones, A.V., Kralovics, R., Harutyunyan, A.S., Zoi, K., Leung, W., Godfrey, A.L., Guglielmelli, P., Callaway, A., Ward, D., et al. (2015). Genetic variation at MECOM, TERT, JAK2 and HBS1L-MYB predisposes to myeloproliferative neoplasms. Nat Commun 6, 6691.
49. Higbee, D.H., Lirio, A., Hamilton, F., Granell, R., Wyss, A.B., London, S.J., Bartz, T.M., Gharib, S.A., Cho, M.H., Wan, E., et al. (2024). Genome-wide association study of preserved ratio impaired spirometry (PRISm). Eur Respir J 63.
50. Diez-Ahijado, L., Cilleros-Portet, A., Fernandez-Jimenez, N., Fernandez, M.F., Guxens, M., Julvez, J., Llop, S., Lopez-Espinosa, M.J., Subiza-Perez, M., Lozano, M., et al. (2024). Evaluating the association between placenta DNA methylation and cognitive functions in the offspring. Transl Psychiatry 14, 383.
51. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnstrom, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nat Genet 46, 944-950.

Table 1. Demographic, clinical, and genetic overview of the *MECOM* cohort (n = 15 all, n = 11 unrelated). Full details are provided in the supplemental data.

Case origin	GeneMatcher, genetic consult, 100,000 Genomes Project, UK NIHR, DECIPHER
Sex	9 male, 6 female
Genetic ancestry	7/11 European, 1 French-Caribbean, 1 Ashkenazi Jewish, 1 Hispanic, 1 unknown
$Age_{diagnosis}$	*in utero* (n=3), 0-3 wks (n=3), 1-5 yrs (n=6), 35-40 yrs (n=1), 41-45 yrs (n=1)
Hematologic anomaly	11/15 (73%)
Skeletal anomaly	9/15 (60%)
Cardiac defect	8/15 (53.3%)
Pulmonary phenotype	6/15 (40%)
Other medical condition(s)	9/15 (60%)
Bone marrow transplantation	3/15 (20%)
Heart/lung transplantation	1/15 (0.6%)
Vital status	alive (n=11), hospice (n=1 child), deceased (n=4)
Genetic variant type	missense (n=7), nonsense (n=2), frameshift (n=1), splicing (n=1)
Mode of inheritance	*de novo* (n=9), autosomal dominant (n=2)
Allele frequency	all variants absent from gnomAD v4.1
Clinical classification (ACMG/AMP)	all pathogenic/likely pathogenic

Table 2. Hemodynamic assessment of patients with rare *MECOM* variants associated with pulmonary arterial hypertension.

Patient ID	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7	Individual 9
PAH subtype	PAH-CHD	PAH-CHD	PAH-CHD	PAH-CHD	PAH-CHD	IPAH
Sex	F	F	M	M	F	M
Genetic ancestry	Ashkenazi Jewish	EUR	EUR	EUR	Hispanic	EUR
Age $_{diagnosis}$	1-5 yrs	1-5 yrs	1-5 yrs	0-3 wks, cath at 1-5 yrs	0-3 wks, cath at 6-10 yrs	1-5 yrs
WHO functional class	II	III	II	III	IV	I
MPAP $_{diagnosis}$ (mmHg)	50	57	62	43	35	62
PVRi $_{diagnosis}$ (Woods units $\cdot$ m^2)	14.4	21.4	20	10	7	30
$D_{LCO\ diagnosis}$	too young to test	too young to test	too young to test	too young to test	unable to perform	too young to test
Treatment	initial: sildenafil and bosentan; last follow-up: tadalafil and bosentan	initial: sildenafil; last follow-up: sildenafil and bosentan	initial: sildenafil; last follow-up: sildenafil and bosentan	tadalafil, macitentin, IV continuous treprostinil infusion, nocturnal oxygen	tadalafil, mecitentan, continuous SQ remodulin	sildenafil, bosentan, epoprostenol iv
Response to vasodilator?	mild	no	not tested	no	no	not tested
Age $_{follow-up}$	n/a	6-10 yrs	11-15 yrs	6-10 yrs	11-15 yrs	6-10 yrs
MPAP $_{follow-up}$ (mmHg)	n/a	not tested	not tested	56	60	50
PVRi $_{follow-up}$ (Woods units x m^2)	n/a	not tested	not tested	14.1	14	11.8
$D_{LCO\ follow-up}$ ml/min/kPa	n/a	3.25 (90% predicted)	6.23 (74% predicted)	too young to test	unable to perform	6.12 (83% predicted)
Transplant age	n/a	n/a	heart/lung, 15 yrs	n/a	n/a	n/a
Vital status	alive	alive, 11-15 yrs	alive, 21-25 yrs	deceased, 6-10 yrs	alive, 11-15 yrs	alive, 21-25 yrs

Abbreviations: cath, (right heart) catheterization; CHD, congenital heart disease; D_{LCO}, diffusion capacity of the lung for carbon monoxide; EUR, European; IPAH, idiopathic pulmonary arterial hypertension; mos, months; MPAP, mean pulmonary artery pressure; PVRi, indexed pulmonary vascular resistance; SQ, subcutaneous; WHO, World Health Organization; wk, week(s); yrs, years.

Figure legends

Figure 1. Chest computed tomography from individual 6 with PAH. Transaxial computed tomography of the chest showing diffuse heterogeneity of the lung parenchyma with bilateral diffuse ground glass opacities, enlarged main pulmonary arteries and cardiac enlargement.

Figure 2. Histopathologic findings in MECOM-associated pulmonary hypertension. A) Near complete occlusion of small parenchymal vasculature with edematous and myxoid intimal expansion with a brisk associated inflammatory infiltrate (Movat pentachrome 200x and H&E, inset, 200x). **B-C)** Variable occlusion of pulmonary artery profiles with fibrous to fibromyxoid intimal expansion and variable associated inflammatory infiltrate. The accompanying airways were patent with mild lymphocytic bronchiolitis (asterisks) (B, H&E 100x and C, H&E 200x). **D)** Venous involvement with near occlusion of venules and veins (Movat pentachrome 200x). Inset highlights the location within the interlobular septa (Movat pentachrome 100x). **E)** Plexiform-like lesion (H&E 200x). **F)** Rare foci of marked capillary congestion imparted a pulmonary capillary hemangiomatosis-like appearance on low power (H&E 100x).

Figure 3. Locations of rare *MECOM* syndrome variants identified in this study (outlined in bold) and reported in the literature. A) Full 2D protein map of predicted pathogenic/likely pathogenic *MECOM* variants with the amino acid positions and conserved protein domains on the x-axis and number of unrelated probands on the y-axis. Variant type is indicated by different shapes and phenotypes by colors. Likely gene disrupting variants and exon/multi-exon deletions are plotted above and missense variants

below the schematic. **B)** Magnification of MECOM aa920 – aa973, containing zinc finger (ZF) 8/9 region, to show variant and phenotype details. 2D map of predicted pathogenic/likely pathogenic *MECOM* variants with the amino acid positions on the x-axis and number of unrelated probands on the y-axis. Variant type is indicated by different shapes and phenotypes by colors. Likely gene disrupting variants are plotted above and missense variants below the schematic.

Figure 4. 3D protein modeling and cell-type specific lung expression of *MECOM*. A) AlphaFold3-predicted MECOM structure (NP_004982). For clarity, disordered and coil regions (residues 1-78, 428-919, and 1003-1239) were removed. The structure was visualized using ChimeraX and colored according to pLDDT scores using the ChimeraX AlphaFold palette: well-modeled backbone regions (pLDDT ≥ 70) are colored yellow to blue, while regions with lower confidence (pLDDT < 70) are yellow to orange/red. **B)** Localization of pathogenic/likely pathogenic *MECOM* variants identified in this study and reported in the literature. Red indicates residues 385, 393, 938, 944, 951, 954, 958, 939, 946, 947, 948, 949, and 950; pink indicates residues 965, 966, 969, 971. Zinc finger (ZF) domains 6, 8, and 9 are shown in green. **C)** Electrostatic surface representation of MECOM generated by ChimeraX. Negatively charged regions are shown in red and positively charged regions in blue, with positively charged regions corresponding to predicted DNA-binding sites. **D)** Superimposition of wild-type (grey) and variant (green) ZF8/ZF9 structures, illustrating structural differences due to the variants**.** Blue spheres represent zinc ions. **E)** AlphaFold3-predicted complex structures of MECOM ZF6 (residues 374-399), ZF8 (residues 919-942), and ZF9 (residues 949-974) bound to the DNA-binding motif of the

VEGFR2/KDR promoter as reported in Lv et al. Nature Comm 2023. Variant residues are colored as in (B). Two views of ZF9 highlight pathogenic residues 938-950 (red, a) and 965-971 (pink, b). **F)** *MECOM* expression compared to 21 PAH genes with definitive, moderate, or limited evidence for the gene-disease relationship. Single cell LungMAP data from normal human lung (265K cells from 72 donors). PAECs, pulmonary arterial endothelial cells; PASMCs, pulmonary arterial smooth muscle cells.