Molecular genetic framework underlying pulmonary arterial hypertension

Laura Southgate[1,2], Rajiv D. Machado[1], Stefan Gräf[3,4,5] and Nicholas W. Morrell[3,5] *

[1]Molecular and Clinical Sciences Research Institute, St George's University of London, London, UK.

[2]Department of Medical and Molecular Genetics, King's College London, London, UK.

[3]Department of Medicine, University of Cambridge, Cambridge, UK.

[4]Department of Haematology, University of Cambridge, Cambridge, UK.

[5]NIHR BioResource, Cambridge, UK.

*e-mail: nwm23@cam.ac.uk

Abstract | Pulmonary arterial hypertension (PAH) is a rare, progressive disorder typified by occlusion of the pulmonary arterioles owing to endothelial dysfunction and uncontrolled proliferation of pulmonary artery smooth muscle cells and fibroblasts. Vascular occlusion can lead to increased pressure in the pulmonary arteries, often resulting in right ventricular failure with shortness of breath and syncope. Since the identification of *BMPR2,* which encodes a receptor in the transforming growth factor-β superfamily, the development of high-throughput sequencing approaches to identify novel causal genes has substantially advanced our understanding of the molecular genetics of PAH. In the past 6 years, additional pathways involved in PAH susceptibility have been described through the identification of deleterious genetic variants in potassium channels (*KCNK3* and *ABCC8*) and transcription factors (*TBX4* and *SOX17*), among others. Although familial PAH most often has an autosomal dominant pattern of inheritance, cases of incomplete penetrance and evidence of genetic heterogeneity support a model of PAH as a Mendelian disorder with complex disease features. In this Review, we outline the latest advances in the detection of rare and common genetic variants underlying PAH susceptibility and disease progression. These findings have clinical implications on lung vascular function and can help to identify mechanistic pathways amenable to pharmacological intervention.

[H1] Introduction

Pulmonary arterial hypertension (PAH) is a rare vascular disorder with an estimated annual incidence of 1–2 cases per million individuals[1]. Until 2019, PAH was defined clinically as mean pulmonary arterial pressure (mPAP) >25 mmHg at rest with normal left atrial pressure, but this definition has since been revised to mPAP >20 mmHg, a normal left atrial pressure and a pulmonary vascular resistance ≥3 Wood units[2]. Consistent features among patients with PAH include exertional weakness or syncope, dyspnoea and right ventricular hypertrophy in the absence of any underlying cardiac or pulmonary disease[3]. PAH was first described in 1891 in a case report of a patient with right ventricular failure in which the autopsy revealed no obvious reason for the observed pathology, referred to as "sclerosis of the pulmonary arteries"[4]. The term 'primary pulmonary hypertension' was subsequently employed in 1951 to set apart the elevated pulmonary arterial pressure observed in a cohort of 39 patients from other patients with a discernible cause of right ventricular failure[5]. The current terminology, adopted and adapted since 2004, takes into account the genetic contribution to disease and includes various categories of PAH, including idiopathic or sporadic cases (IPAH), heritable cases (HPAH; describing patients with a family history or identified germline mutation) and associated forms (APAH; attributable to specific risk factors, such as anorexigens, liver disease, congenital heart disease and connective tissue diseases)[6,7].

PAH typically manifests in the third to fourth decade of life, but can also be diagnosed in children, in whom prognosis is more severe[8]. Clinical features of disease in HPAH and IPAH are largely indistinguishable clinically, although patients with HPAH have an earlier age of onset with worse haemodynamic features upon initial diagnosis[9]. In the absence of contemporary ameliorative treatment, PAH is frequently fatal[10]. Histopathological examination of PAH indicates the presence of occluded pulmonary arterioles resulting from the proliferation of pulmonary artery endothelial cells (PAECs), pulmonary artery smooth muscle cells (PASMCs) and fibroblasts, which leads to right heart hypertrophy and eventual cardiac failure[11].

Familial segregation has been reported in approximately 6–10% of all patients with PAH, wherein the disease segregates as a monogenic, autosomal dominant trait with incomplete penetrance, ranging from 14% in males to 42% in females[12]. PAH is observed much more frequently in females, with a ratio of occurrence in women to men ranging between 2:1 and 4:1, which reflects the distinctions across different patient populations and PAH subtypes[13]. Both intrafamilial and interfamilial variability of phenotypic expression and differential age

of onset have also been also observed. Taken together, these clinical findings indicate that the pathophysiology of PAH is complex, wherein the development and progression of this condition can be influenced by both environmental and genetic modifying factors. In this Review, we discuss the genetic aetiology of PAH and the factors that might influence screening and treatment strategies in the future.

[H1] TGFβ signalling in PAH

In 1997, the linkage interval for PAH was established on the chromosomal locus 2q33–34, leading to an international positional cloning approach to determine the causal gene[14]. Subsequent systematic analyses of the genes contained within the minimal linkage interval in two independent, multi-generational familial cohorts resulted in the identification of heterozygous germline mutations in the gene encoding bone morphogenetic protein receptor type 2 (BMPR2), a member of the transforming growth factor-β (TGFβ) superfamily of signalling molecules[15,16]. Subsequent interrogation of *BMPR2* in patients with IPAH provided independent validation of a pathogenic variation in a subset of patients[17,18]. The mature BMPR2 polypeptide comprises 1,038 amino acids encompassing an extracellular signal peptide, a ligand-binding domain, a single-pass transmembrane domain, an intracellular catalytic kinase domain and an atypically long cytoplasmic tail[19]. Canonical bone morphogenetic protein (BMP) signal transduction is initiated at the plasma membrane upon ligand binding with the formation of a heteromeric complex of BMPR2 and a cognate type I receptor, namely activin receptor-like 1 (ACVRL1, also known as ALK1), activin receptor type 1 (ACVR1, also known as ALK2), BMPR1A (also known as ALK3) or BMPR1B (also known as ALK6). BMP signalling is cell-specific; for example, signalling in endothelial cells is triggered by the BMP9 ligand and relayed via a BMPR2–ACVRL1–endoglin receptor complex[20,21]. Next, the constitutively active BMPR2 transphosphorylates its type I receptor partner, which, in turn, activates a BMP-specific series of cytoplasmic intermediaries, namely mothers against decapentaplegic homolog 1 (SMAD1), SMAD5 and SMAD8. In conjunction with the nuclear chaperone SMAD4, this macromolecular complex translocates to the nucleus where it combines with a series of co-activators and co-repressors to regulate the transcription of a limited set of target genes[22].

[H2] BMPR2 mutations in PAH

Since the identification of *BMPR2*, numerous screening studies have been undertaken in patient cohorts of various sizes and diverse ethnicities, which have together established that mutations in *BMPR2* account for 53–86% of patients with a family history of PAH and 14–35% of patients with IPAH.[23] Indeed, *BMPR2* mutations of likely pathogenicity have been detected in >800 patients with PAH, of which 486 involve distinct, nonrecurrent variants. These data indicate that the spectrum of genetic variants is similar in both idiopathic and familial forms of PAH, and that these mutations fall into the major categories. Specifically, 25% of PAH-specific variations in *BMPR2* are missense mutations that result in amino acid substitutions (Table 1). By contrast, the vast majority of variants are predicted to lead to premature protein truncation owing to nonsense mutations (27%), frameshift mutations arising from small nucleotide insertions or deletions (23%), gene rearrangements (14%) or splice-site mutations (10%). Given that truncating mutations spare the terminal exon 13, mutant transcripts are most probably degraded via the nonsense-mediated decay pathway. Recurrent mutations have been observed in approximately 40% of patients with PAH-specific *BMPR2* variations, most likely as a result of small insertions and deletions at sites of low genomic complexity, such as mononucleotide runs, or as a consequence of spontaneous cytosine deamination[23]. The highly recurrent mutation c.1471C>T (p.R491W), which affects a critical residue within the catalytic domain, is most likely to be a result of the latter mechanism. Indeed, a study involving the genetically isolated Finnish population indicated that detected *BMPR2* mutations in PAH arose on distinct haplotypes, thereby precluding the presence of a founder mutation[24].

Whereas missense mutations are widely distributed across *BMPR2* exons, the majority are localized within key functional domains, specifically the ligand-binding domain encoded by exons 2–3 and the highly conserved catalytic kinase region functionally delineated by exons 6–9 and 11[23]. Of note, the majority of studies to identify genetic variants in *BMPR2* have historically been confined to the exonic space. The limited analyses of non-coding regions have identified substitutions within the *BMPR2* 5′ UTR, including the recurrent variant c.−669G>A, which has been predicted to abrogate specificity for an SP3 transcription factor binding site[25,26]. However, the high frequency at which this variant is present in control populations indicates that further investigation is required to establish its pathogenicity. Together, this extended dataset consolidates *BMPR2* haploinsufficiency as the primary molecular mechanism underlying hereditary PAH[18].

[H3] PAH-associated congenital heart disease. Congenital heart disease is an established complication of PAH. In 2004, an analysis of *BMPR2* variation was performed on a combined panel of 40 adults and 66 children with APAH encompassing a range of congenital heart conditions including patent ductus arteriosus, and atrial and ventricular septal defects[27]. Direct screening of protein-coding exons and flanking intronic sequences led to the detection of six missense mutations, representing 6% of the total cohort. Three defects within the BMPR2 extracellular domain (c.125A>G, p.Q42R; c.304A>G, p.T102A; and c.319T>C, p.S107P) were observed in adults with complete type C atrioventricular canal defects. Of the patients with mutation-positive, childhood-onset APAH, two presented with an atrial septal defect and patent ductus arteriosus, whereas one had an aortopulmonary window and a ventricular septal defect[27]. In contrast to the adult-onset form of the disease, BMPR2 mutations in patients with childhood-onset APAH were not clustered within a specific exon or functional domain of the protein (c.140G>A, p.G47N; c.556A>G, p.M186V; and c.1509A>C, p.E503D). Despite the limitations of small sample size and potential bias of ascertainment for PAH, this study is important in that these mutations are unique to PAH-associated congenital heart disease and are potentially under-represented because the authors did not screen for gene rearrangements or promoter variants. These genetic data are reinforced by a study in a mouse model that revealed the need for BMP signalling in the developing embryonic heart[28].

[H2] Rare TFGβ family alleles in PAH

Hereditary haemorrhagic telangiectasia (HHT) is an inherited vascular condition clinically characterized by arteriovenous malformations and subcutaneous telangiectasia. HHT and PAH infrequently co-present in families, suggesting a common molecular aetiology. HHT has autosomal dominant transmission and is caused by defects in the type I receptor ACVRL1 and the type III accessory receptor endoglin, which are members of the TGFβ signalling family, and thus represent good candidate genes for PAH. Analyses of patients with both PAH and HHT confirmed the existence of a pathogenic variant in *ACVRL1* and, at lower frequency, in *ENG*[29]. Since then, *ACVRL1* mutations have been reported in patients with evidence of PAH without concomitant HHT, typically in early-onset disease[30,31]. The vast majority of *ACVRL1* variants are missense mutations, of which approximately 89% are contained within the vital catalytic domain, indicating a high likelihood of pathogenicity[23].

Subsequent interrogation of the BMP-specific pathway genes *SMAD1*, *SMAD4, SMAD5* and *SMAD9* in European and Asian populations led to the identification of novel independent variants that were absent from ethnically matched healthy individuals[32,33]. Functional analyses of these variants by in vitro luciferase SMAD-responsive element reporter assays confirmed that mutations in *SMAD9* (encoding SMAD8) markedly impaired SMAD transcriptional activity, further validated by downregulation of the BMP target gene *ID1*[33]. By contrast, defects in *SMAD1* and *SMAD4* caused less severe suppression of transcriptional activation and have consequently been defined as variants of unknown significance, albeit with the caveat that SMAD-independent pathways were not explored[33]. Taken together with the discovery of a novel homozygous c.76C>T (p.Q26*) nonsense mutation in *GDF2* (which encodes the ligand BMP9) in a Hispanic boy aged 5 years[34], these genetic findings underline the critical role of the wider TGFβ signalling network in the precipitation of PAH.

[H1] Expanding the genetic architecture of PAH

Whereas classic gene identification studies have proven invaluable in identifying causal factors underlying the development of many diseases, technological advances have revolutionised the way in which the genetics and genomics of PAH have been analysed. Next-generation sequencing technologies, in particular whole-exome sequencing and whole-genome sequencing, have been powerful tools for the detection of rare, highly penetrant gene variants in PAH, and have thus established PAH as a genetically heterogeneous condition through analysis of families negative for mutations in known TGFβ pathway genes. Employing a direct approach of variant filtering based on a model of autosomal dominant inheritance has led to the isolation of rare familial genetic variants in the genes encoding caveolin 1 (*CAV1*) and the potassium channel subfamily K member 3 (*KCNK3*), both of which have been validated as causative genes by replication in patients with IPAH[35,36].

[H2] Caveolin 1

Implicated causal factors for PAH fall within cellular pathways pertinent to PAH pathology, particularly *CAV1*. Caveolin 1 is critical in forming the flask-shaped invaginations required for the regulation of BMP-driven, SMAD-dependent signalling in lung mesenchymal cells and endothelial cells[37,38]. Variants in *CAV1* were initially implicated in PAH pathogenesis by exome sequencing of four patients with PAH within a three-generational family who were

negative for the established variants in the TGFβ family[35]. Further examination of remaining affected members confirmed segregation of the likely truncating mutation in exon 3 (c.474delA, p.L159Sfs*22). Replication of *CAV1* variants was achieved through the identification of an adjacent de novo mutation in a patient with IPAH (c.473delC, p.P158Hfs*23)[35]. A separate study has further confirmed a third *CAV1* frameshift mutation (c.471delC; p.D157fs) in an adult patient with PAH[39]. Although reported as a novel mutation, on the basis of the standardized mutation nomenclature from the Human Genome Variation Society[40], this variant is actually identical to the previously described c.473delC mutation[35]. Assuming that these two patients are unrelated, the three mutations identified thus far are harboured in the terminal exon of *CAV1* and the gene product would, as such, not be predicted to undergo transcript degradation via the nonsense-mediated decay pathway. Indeed, functional analysis of a PAH-specific frameshift mutation in *CAV1* has indicated demonstrable impairment of caveolae assembly owing to retention of the mutant protein in the endoplasmic reticulum, with concomitant sequestration of the wild-type protein form[41]. Moreover, the c.474delA *CAV1* mutation leads to hyperphosphorylation of SMAD1, SMAD5 and SMAD8, consequently resulting in a reduction of the anti-proliferative function of caveolin 1, thereby supporting SMAD gain-of-function as the underlying molecular mechanism of disease in patients with this *CAV1* variant[42]. Although uncommon in human PAH, a causal role for *CAV1* is consolidated by the observation that caveolin 1-deficient mice develop clinical features of pulmonary vascular disease analogous to PAH[43-45].

[H2] Potassium channel dysregulation

Membrane potential is essential for contraction of PASMCs, and the dysregulation of potassium channels has been implicated in the pathogenesis of PAH. Transcript levels of *KCNA5*, which encodes the voltage-gated potassium channel subfamily A member 5, are reduced in PASMCs from patients with PAH compared with controls[46], suggesting that the delayed-rectifier voltage-gated potassium current in PASMCs from patients with PAH is likely to be attenuated[47]. Of note, potassium channel expression is regulated by BMP signalling, providing a putative association with established molecular defects in BMPR2[48]. Defects in *KCNA5* have also been identified in patients with PAH[49,50]. However, in the absence of any independent replication studies or comprehensive examination of functional consequences, this gene remains an unvalidated causal factor. In 2013, Ma and colleagues reported in three independent families with PAH the segregation of *KCNK3* mutations

leading to amino acid substitutions[36]. Moreover, this finding was supported by the discovery of additional deleterious missense variants in *KCNK3* in three patients with IPAH by the direct sequencing of 230 individuals. Electrophysiological analyses to examine the functional effect of these mutations showed a variable loss of function of the mutant channels that probably cause resting membrane potential depolarization[36,51]. Of note, the observed reduction in current across the potassium channel was reversed after treatment with a phospholipase inhibitor. The subsequent identification of a homozygous mutation underlying an aggressive form of HPAH has served to further consolidate a causal role for *KCNK3* in this disorder[52]. Taken together, the detection of germline mutations in *KCNK3* with deleterious functional effects reinforces a pathological role for potassium channel dysregulation in both familial and idiopathic forms of PAH.

[H2] Autosomal recessive PAH subtypes

Under the present clinical classification of pulmonary hypertension, pulmonary veno-occlusive disease (PVOD) and pulmonary capillary haemangiomatosis (PCH) fall within PAH subgroup 1.6, namely PAH with overt features of venous or capillary involvement[2]. Typically, PVOD and PCH have histological characteristics that are indicative of pulmonary venous and capillary proliferation with a diffuse presentation[53,54]. Patients with PVOD or PCH respond poorly to available vasodilatating agents, and these drugs might provoke life-threatening pulmonary oedema[54,55]. Occasionally, pulmonary arterial involvement is also observed[53,56]. In contrast to PAH, both PVOD and PCH are inherited in an autosomal recessive manner. To determine the genetic basis of PVOD, exome sequencing was performed to identify homozygous and/or compound heterozygous variations[57]. Mutations in *EIF2AK4*, which encodes eIF2α kinase GCN2 (GCN2), were identified in all of the analysed kindred ($n = 5$) and in 25% of those with sporadic disease ($n = 20$). Concurrently, Best and colleagues independently detected biallelic *EIF2AK4* mutations in heritable and sporadic PCH[58]. GCN2 is essential for the activation of the integrated stress response pathway, which is triggered by environmental stress via the phosphorylation of the α subunit of eukaryotic translation initiation factor 2[59]. The clinical and histological overlap between PAH, PVOD and PCH prompted the assessment of potential pathogenic *EIF2AK4* defects in patients diagnosed with PAH. Analysis of a HPAH and IPAH patient cohort ($n = 66$) who did not have mutations in any of the major causative genes for PAH led to the identification of a novel homozygous *EIF2AK4* mutation (c.257+4A>C) in two sisters with severe disease,

indicating that variations in this gene constitute a rare molecular trigger for PAH[60]. Further validation for a pathogenic role of *EIF2AK4* mutations in patients clinically diagnosed with PAH was provided by a study that identified nine patients (out of 864) with biallelic *EIF2AK4* mutations[61]. Notably, these individuals had clinical characteristics distinct from patients with PAH who did not have *EIF2AK4* mutations, namely a younger age at PAH diagnosis (29 (interquartile range 23–38) years versus 51 (interquartile range 37–65) years) and reduced survival[61]. Taken together, these data emphasise the importance of an accurate genetic diagnosis for the efficient management and treatment of patients with PAH.

[H2] Sample size and gene variant detection

The locus heterogeneity present in PAH has reduced the statistical power of genetic studies to detect causal genes, which is further confounded by small or clinically heterogeneous patient cohorts and the limited effect size of underlying variants. To advance our understanding of the increasingly complex aetiology of Mendelian disease, genetic studies are now contingent on the development of large, multicentre consortia such as the National Institute for Health Research BioResource for Rare Diseases. This resource is focussed on novel gene variant identification in conjunction with prospective phenotypic analyses through the generation of a unique biobank of patients and families with rare diseases, such as PAH[62-64]. To evaluate existing loci and expand our understanding of the genetics of PAH, a large cohort of 1,038 patients with PAH from the UK and Western Europe was recruited for whole-genome sequencing studies[65]. The cohort predominantly comprised patients with IPAH (87%; $n = 908$) of European descent (89%; $n = 806$). Initial screening of the established PAH gene variants confirmed *BMPR2* mutation as the major risk factor, with a mutation frequency of 12% in patients with IPAH and 76% in patients with HPAH[65]. The distribution of variation across the other PAH risk factors indicated that variants in *ACVRL1*, *ENG*, *SMAD9*, *KCNK3* and *TBX4* had a modest contribution to the disease, with variants in each gene accounting for ~1% of cases. Of note, no pathogenic coding variation was identified in *CAV1*, *SMAD1* or *SMAD4*, further reiterating the need for larger PAH sample sets in such studies[65].

Genome-wide gene burden tests, employed to identify novel gene variants through a comparison of variant frequencies between patients with PAH and control individuals, detected statistically significant enrichment of rare deleterious variants in *ATP13A3*, *AQP1* and *SOX17*[65]. In keeping with previous studies[23], the heritability ascribed to each novel gene variant was relatively small, accounting for 0.5–1.0% of the overall cohort. The majority of

reported mutations were missense and nonsense variants, with *ATP13A3* additionally harbouring three frameshift and two splice region defects, probably reflecting loss of protein function. *AQP1* and *SOX17* mutation carriers had a significantly younger age at diagnosis, putatively suggesting early-onset PAH; however, this observation requires validation in larger patient cohorts. The identified variants in *AQP1* and *SOX17* were further demonstrated to segregate with the phenotype in a limited set of HPAH pedigrees[65]. Importantly, this study also provided further evidence of a prominent role for *GDF2* as a key PAH risk factor, with the identification of one frameshift mutation and seven missense mutations, all of which are predicted with high confidence scores to be deleterious. Subsequently, the causal role of *GDF2* mutations was confirmed in a prospective study that employed a gene panel[66], and a whole-exome sequencing study in Chinese patients with PAH, in which *GDF2* mutations were detected in 6.7% of 331 patients with IPAH[67]. Moreover, preliminary evidence suggests a role for two putative loss-of-function mutations of BMP10 in disease susceptibility, further implicating the endothelial cell as the primary site of PAH initiation[66].

[H2] Functional effect of novel gene defects

Probable cation-transporting ATPase 13A3 (ATP13A3), a member of the P5B subfamily of ATPases involved in ion channel transport, is highly expressed in major vascular cell types and represents a strong candidate for involvement in PAH pathogenesis[65]. Many of the ATP13A3 mutations identified in PAH cluster within the catalytic phosphorylation domain, indicating a likely important effect on protein function. However, the cellular function and transported substrate of ATP13A3 remain poorly understood. PAH is considered to be a neoplastic growth disorder and, of note, ATP13A3 promotes cell proliferation in multiple cancers[68,69]. Indeed, preliminary analyses have demonstrated reduced proliferation and increased apoptosis of endothelial cells following *ATP13A3* knockdown with small interfering RNA[65], consistent with reports that PAH is initiated by endothelial dysfunction-mediated survival of apoptotic-resistant cells, which subsequently results in the uncontrolled growth of pre-proliferative cells at key sites of disease[70,71].

Expression of *AQP1*, which encodes the plasma membrane water channel aquaporin 1, is integral to the maintenance of vascular tone[72]. Animal studies indicate that abrogation of aquaporin 1 in PASMCs results in an attenuation of clinical features in hypoxia-induced PAH, providing the potential for targeted development of therapeutic options[73]. *Aqp1*-null mice have impaired endothelial cell migration and angiogenesis[74]. By contrast, elevated

aquaporin 1 levels promote PASMC proliferation and migration through upregulation of the expression of the β-catenin targets Myc proto-oncogene protein and cyclin D1[75].

The majority of variants detected in *SOX17* are also predicted to affect β-catenin function through either loss of the β-catenin binding region in SOX17 or abrogation of OCT4 (also known as POU domain, class 5, transcription factor 1)-mediated β-catenin interaction[76,77]. The endothelial transcription factor SOX17 depends on β-catenin binding to activate the transcription of target genes. Given that missense variations within the conserved high mobility group box of SOX17 are anticipated to affect DNA binding of the transcription factor, all identified variants are likely to result in loss-of-function defects.[65] A whole-exome sequencing study involving 256 patients with PAH strongly implicated *SOX17* mutations as a major risk factor in PAH-associated congenital heart disease, and also provided independent validation for their role in IPAH[78]. Likely pathogenic variants were detected in 0.7% of patients with IPAH and 3.2% of patients with PAH-associated congenital heart disease, with the vast majority clustered within the high mobility group box. A similar study in a cohort of Japanese patients with HPAH or IPAH identified three additional *SOX17* variants[79]. Of interest, gene-based, case-control analyses also revealed an over-representation of rare deleterious variations in a number of putative SOX17 target genes, indicating a key role for this pathway in PAH aetiology[78]. The SOX protein family have critical functions during human development and SOX17 is fundamental in angiogenic processes, including arteriovenous differentiation and development of the lung microvasculature[80-82]. In conjunction with Notch signalling, SOX17 is involved in the regulation of tip-cell and stalk-cell specification of endothelial cells, however, the exact mechanisms by which *SOX17* mutations can precipitate PAH remain to be determined[81,83,84].

GDF2 encodes BMP9, a major ligand of the BMPR2–ACVRL1 receptor complex and a potent inhibitor of endothelial cell migration and growth[85]. BMP9 is processed by the serine endoprotease furin to its fully activated form, which is secreted as a mature prodomain-bound protein[86]. Mapping of identified amino acid substitutions onto the crystal structure of the BMP9 homodimer revealed a grouping of PAH-associated variants on the interface of the prodomain and growth factor domain[65,86,87]. Consistent with the hypothesis that PAH-specific variants reduce ligand stability, analysis of BMP9 levels in HEK293T cells demonstrated significantly reduced secretion of mature ligands in the supernatants of cells transfected with mutant BMP9 constructs compared with cells transfected with wild-type BMP9[65]. In support

of these findings, plasma levels of BMP9 were also reduced in patients with *GDF2* mutations[67].

The importance of these newly identified gene variants in the pathology of PAH is yet to be fully defined. These data provide a compelling framework of PAH as a genetically heterogeneous disorder comprising low frequency mutations of high impact in seemingly disparate genes. Taken together, these findings provide important insights into the distinct mutation spectrum across different clinical categories for each causal gene, thereby facilitating focused diagnostic protocols (Table 2).

[H1] Genetics of childhood-onset PAH

The clinical features of childhood-onset and adult PAH have conspicuous differences. Specifically, poor survival and resistance to conventional therapies in paediatric patients have long indicated a distinct genetic aetiology to the disease in adult patients[39]. Candidate-gene analyses in limited sample sets have established *BMPR2* variants as the major risk factor in familial and idiopathic early-onset PAH, with *ACVRL1* variants also having a prominent role in conferring susceptibility[31,32,88–90]. In addition, rare PAH-associated alleles have been identified in *ENG*, *KCNK3*, *SMAD9* and *BMPR1B*[30,39,89,91]. Of note, the mutant allele frequency in these genes is comparable between adult and paediatric patients.

To obviate the restrictions of previous investigations and more comprehensively assess the mutation spectrum in childhood disease, whole-exome sequencing was performed in a cohort of 155 patients with childhood-onset and 257 patients with adult-onset PAH.[39] The analysis revealed a substantial enrichment of likely disease-causing alleles in *TBX4* (encoding the T-box transcription factor TBX4) among the childhood-onset IPAH sample set (10 of 130 patients) compared with the adult-onset IPAH cohort (none of 178 patients; *P* <0.0001). *TBX4* mutations have previously been identified as a prominent risk factor in small patella syndrome, a rare developmental bone dysplasia characterized by small or absent kneecaps[92], and heterozygous loss-of-function TBX4 mutations have been described in childhood-onset PAH, with or without associated small patella syndrome[90,93,94]. These data, generated from the largest study of its kind, provide compelling evidence that, after *BMPR2* mutations, variants in *TBX4* confer the highest degree of genetic risk in childhood-onset PAH (9.2% versus 7.7%)[39]. TBX4 has an important role in the regulation of embryogenesis, in particular, the formation of the hindlimb[95]. Additionally, the gene is expressed in the lung and trachea

mesenchyme, and in the atria of the heart[96]. *TBX4* mutations associated with early-onset PAH predominantly cluster in the critical T-box domain and typically result in damaging amino acid substitutions or premature truncation of the mutant transcript. Moreover, PAH has a significantly earlier age of onset in *TBX4* carriers (7.9±9.0 (0.6–33.0) years) compared with individuals with *BMPR2* heterozygous mutations (28.2±15.4 (1.5–72.0) years; *P* <0.0001), suggestive of a specific genotype–phenotype correlation[39].

A follow-up report in a cohort of patients with childhood-onset (*n* = 99) and adult-onset (*n* = 134) PAH who were negative for *BMPR2* or *ACVRL1* mutations, identified a de novo missense variant in *ABCC8* (encoding the ATP binding cassette subfamily C member 8 (ABCC8); also known as sulfonylurea receptor 1) in a patient with childhood-onset IPAH[97]. The mutation (c.2873G>A; p.R958H) was predicted to be deleterious, occurring at a highly conserved site within a critical domain of the encoded protein, which is a regulatory subunit of the ATP-sensitive potassium channel[97]. Screening of *ABCC8* mutations in an extended cohort identified likely deleterious, heterozygous, missense mutations distributed across the length of the gene in paediatric patients with IPAH, familial PAH or congenital heart disease-associated PAH. Expression of the identified *ABCC8* variants in a fibroblast-like cell line decreased the function of the ATP-sensitive potassium channel, further indicating that a subset of PAH might be mechanistically described as a potassium channelopathy[97]. Of interest, treatment of these cells with the pharmacological agent diazoxide, a selective ABCC8 activator, resulted in the recovery of channel function to normal levels[97].

In addition to being an important determinant of early-onset PAH, *TBX4* and *ABCC8* mutations are also key to the pathogenesis of the adult form of the disease[39,97]. Indeed, in an extensive study of adult-onset IPAH, heterozygous *TBX4* defects were observed in 14 patients (accounting for 1.3% of the total cohort), constituting the joint second most substantive risk factor for PAH development after *BMPR2* abnormalities[65]. These data serve to underscore the central importance of extensive study populations to investigate the composition of deleterious genetic variants in rare diseases with marked genetic heterogeneity.

[H1] Common variation in PAH

The first genome-wide association study (GWAS) in PAH was conducted in 2013 by Germain and colleagues in a French patient cohort comprising unrelated patients with IPAH

or familial PAH without identifiable *BMPR2* or *ACVRL1* mutations[98]. Using a discovery dataset of ~400,000 single nucleotide polymorphisms (SNPs) genotyped across 340 patients and 1,068 healthy individuals, an association was identified with two SNPs located 52kb downstream of *CBLN2* on chromosome 18q22.3. This signal was independently replicated in a US-based validation cohort of 284 patients with IPAH or HPAH and 456 controls[98]. Across the total sample set, the risk allele was associated with a two-fold increased risk of disease with an odds ratio of 1.97 (95% CI 1.59–2.45; $P = 7.47 \times 10^{-10}$). *CBLN2* encodes the precursor of cerebellin 2, which is predominantly expressed in the brain. Analysis of *CBLN2* expression in the whole lung revealed significantly higher mRNA levels in explants of patients with PAH and in cultured pulmonary arterial cells; *CBLN2* was predominantly expressed in PAECs, and poorly expressed in PASMCs. Treatment of PASMCs with cerebellin 2 led to a significant reduction in cell proliferation in a dose-dependent manner, supporting a potential role for this protein in the physiopathology of PAH[98].

A GWAS has also been conducted in a Japanese cohort of 44 patients with IPAH or HPAH and 2,993 control individuals[99]. Replication analysis in 21 patients and 991 control individuals led to the validation of two novel susceptibility SNPs located midway between *PDE1A* (encoding calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase 1A (PDE1A)) and *DNAJC10* (encoding DnaJ homolog subfamily C member 10) on chromosome 2q32.1[99]. Odds ratios obtained from the validation study were 5.18 (95% CI 1.86–14.42) and 5.16 (95% CI 1.86–14.37) for rs71427857 and rs13023449, respectively. Based on immunohistochemical studies demonstrating increased immunoreactivity for PDE1A in distal pulmonary arteries from patients with IPAH, the authors suggest that PDE1A could be a novel therapeutic target for PAH[99]. However, owing to the small sample size of this GWAS, these findings will require further validation in a larger cohort with different ethnic groups.

In 2019, a comprehensive GWAS involving >2,000 patients of European ancestry with PAH was conducted[100]. The study design, which involved two independent discovery datasets, followed by a meta-analysis to cross-validate loci attaining genome-wide significance, was unique in that it utilised whole genome sequencing data to obtain genotypes for one cohort. The cohort comprised patients from the UK, USA and Western Europe, including patients from the previously published French-led GWAS[98]. A combined association comparison of 2,085 patients with PAH against 9,659 population-matched control individuals demonstrated two regions of genome-wide significance across the *HLA–DPA1/DPB1* locus ($P = 7.65 \times 10^{-20}$)

and upstream of *SOX17*[100]. Of note, the previously reported associations between *CBNL2* and *PDE1A–DNAJC10* were not replicated in this study, indicating these might represent population-specific signals. An investigation of putative correlation between the *HLA–DPA* and *DPB1* risk variant and patient survival indicated a significantly improved prognosis with two copies of the risk allele. The *SOX17* risk variants were located within an enhancer region, and functional evaluation of PAECs supported a role of these variants in the regulation of gene expression, with PAH risk alleles demonstrating reduced enhancer activity compared with non-risk alleles[100]. Taken together with the report of rare, deleterious *SOX17* variants underlying PAH[65], this finding suggests that both common and rare variants might predispose to PAH.

[H1] Clinical implications

The assessment of the genetic architecture of PAH has facilitated the directed mutation analysis of established causal genes, which provides molecular confirmation of the diagnosis. Moreover, numerous studies have identified exonic regions that are likely to harbour recurrent mutations, defects clustering within conserved domains and areas of functional importance, thereby expediting the accurate assessment of deleterious variants in contrast to variants of unknown significance. Importantly, predictions of pathogenicity must be evaluated in the context of the guidelines provided by the American College of Medical Genetics and Genomics for interpretation of sequence variants[101]. Within families, the precise determination of causal mutations provides an opportunity for informed genetic counselling, carrier testing and preimplantation genetic diagnosis in at-risk individuals[102,103].

Aligned with these opportunities, the integration of genetic clinics into routine patient care within specialist PAH centres will be critical to support the development of mutation screening programmes and, as a progressive goal, the identification of early biomarkers of disease. For patients with a documented family history of PAH, clinical genetic testing is of exceptional importance to facilitate early monitoring and treatment intervention of relatives harbouring known gene defects. However, given the established role of heritable risk factors in PAH, we would also advocate genetic screening of all patients with IPAH to support risk calculations for family members and successive generations, consistent with international guidelines[104]. In the case of *BMPR2*, genetic data indicate that genotype–phenotype correlations differ between mutation carriers and noncarriers, and these correlations influence features of disease, such as age of onset and severity, that have the potential to direct specific

treatment strategies. Further, these molecular data permit differential diagnosis between subtypes of pulmonary hypertension, for example PAH and PVOD, thus informing efficacious patient management[53]. Ultimately, the identification of specific genetic defects in a given patient might provide options for personalised medicine.

Notably, the elucidation of disease genes has provided key insights into the pathways central to the maintenance of the pulmonary vascular bed, and enabled the identification of molecular targets for the development of novel treatment strategies (Figure 1). Proof-of-concept studies for restoration of BMPR2 function as a therapeutic approach in PAH holds promise, as has been demonstrated in patient-derived cells and in animal models[105-109]. These studies have been reviewed in detail previously[110]. To date, trials in patients based on genetically isolated targets have been focused on use of the calcineurin inhibitor tacrolimus which, among other effects, activates BMP signalling to reverse vascular occlusion in rat models, and seems to be well tolerated in patients with PAH[111]. Although preliminary in vitro findings support the potential restoration of potassium channel activity after pharmacological intervention in loss-of-function *KCNK3* and *ABCC8* cell models, further preclinical testing is required to evaluate the feasibility of manipulating these targets for patient benefit[36,97,112]. The endothelial cell-specific ligand BMP9, a known susceptibility factor for PAH, has been found to reverse cellular characteristics of disease in three established experimental models, including a murine *Bmpr2*-knock-in version of the recurrent human mutation p.R899*[113]. Critically, the fundamental understanding of the BMP9, ACVRL1 and BMP receptor signalling nexus has been integral to the progression of research from laboratory testing to clinical application, with a protein-engineered variant of BMP9 now in preclinical development. Based on this proof-of-principle notion, the novel genetic findings and wealth of genomic information that are now accessible lend promise to the progressive development of precision medicine treatments targeted to multiple pathways[114].

[H1] Future perspectives

To date, 16 genetic risk factors for PAH have been identified. For the majority of genes in which familial inheritance has been assessed, transmission occurs in an autosomal dominant manner with incomplete penetrance. This method of transmission is particularly relevant in the case of *BMPR2* variants, which suggests that although germline mutations represent a significant genetic risk, additional environmental and genetic factors might be required to precipitate manifest disease. Indeed, anorexigens confer an established environmental risk to

PAH development[2]. The identification of PAH-associated variants through GWAS analysis has demonstrated the likely existence of genetic modifiers[98-100]. Therefore, the genetic background of a mutation carrier in concert with environmental insults might explain whether the condition is potentiated. Taken together, these observations indicate that protein function needs to fall below a certain level for PAH to occur.

Missing heritability in PAH remains a vexed question. Although large datasets have been analysed for exonic and flanking variation with rigour, inadequate systematic examination of copy number variations across implicated susceptibility loci remains. Indeed, large deletions and duplications are estimated to comprise approximately 14% of the mutation spectrum in *BMPR2* alone (Table 1). Further, the functional genomics of non-coding variations remains to be robustly analysed to best interrogate the contribution of impaired regulatory elements to disease precipitation, for example, via a combination of transcriptomics and chromatin-conformation capture technologies. For the genetic architecture of a rare, heterogeneous disorder to be fully annotated, a global initiative is required. Specifically, international patient ascertainment, data collection and amalgamation onto a joint platform is central to identification of novel rare disease alleles. The formation of the International Consortium for Genetic Studies in PAH aims to provide the necessary statistical power to advance discovery of the complete genetic architecture of PAH towards smaller effect sizes (http://www.pahicon.com).

Critically, the functional analysis of detected variants is a requisite for the reliable determination of pathogenicity (Table 1). To investigate the effect of gene defects in the poorly characterized genes, an important first step would be the assessment of mRNA levels compared with wild-type controls to establish the stability and abundance of mutant transcripts. Similarly, the analysis of protein stability by Western blotting and determination of subcellular localization by immunofluorescence is desirable. Where truncating or splice-site mutations are suspected, experiments aimed at examining nonsense-mediated decay might be of value. Further to that, in vitro assays will need to be specifically tailored to the known or predicted protein function on a gene-specific basis. Specific examples include electrophysiological studies to examine channel function[36,97] and ligand-binding assays[115] or luciferase assays to evaluate signalling activity.[32,33] Employing a relevant cell type, for example PAECs and patient-derived cells, is an important factor in the accurate interpretation of results in the context of disease. Long-term analyses might include the use of model systems to demonstrate in vivo effects of identified variants.

In addition to genomics-focused strategies to understanding the pathogenesis of PAH, additional insights can be leveraged through a more holistic approach by combining multi-layered omics studies with comprehensive clinical phenotyping. In 2014, the Redefining Pulmonary Hypertension through Pulmonary Vascular Disease Phenomics (PVDOMICS) initiative was launched in the USA, with the aim of supporting clinical research with a molecular-based classification of pulmonary vascular disease. Over the next few years, the PVDOMICS consortium plan to integrate existing genomic data with patient transcriptomic, proteomic, metabolomic, coagulomic and cell biomic profiles, which can be influenced by environmental and genetic factors[116,117]. In combination with improvements in tissue functioning and imaging, these studies will help to precisely categorize features of PAH and pulmonary vascular disease, providing a foundation to facilitate the development of precision medicine and personalised therapeutic interventions.

[H1] Conclusion

The genetics of PAH has evolved rapidly since the introduction of massively parallel sequencing technologies that facilitate the evaluation of genomic variation across multiple genes concurrently. Active collaboration between molecular biologists and dedicated clinicians has facilitated the identification of novel causal genes and pathways, highlighting important genotype–phenotype correlations that will inform diagnosis and improve clinical management. Dysregulation of BMP signalling remains the predominant risk factor for the development of PAH and the seminal identification of *BMPR2* haploinsufficiency as the primary molecular mechanism of disease is now evolving into the development of targeted therapeutic options. Continued gene discovery alongside mutation screening of newly identified genes will likely support the transition of molecular testing into PAH clinics. Nonetheless, there remains a substantial gap with regard to providing a definitive molecular genetic diagnosis to all patients. The continued efforts of global consortia are, therefore, central to fully resolving the molecular mechanisms of this devastating disease.

References

1. Peacock, A. J., Murphy, N. F., McMurray, J. J. V., Caballero, L. & Stewart, S. An epidemiological study of pulmonary arterial hypertension. *Eur. Respir. J.* **30**, 104–109 (2007).

2. Simonneau, G. *et al*. Haemodynamic definitions and updated clinical classification of pulmonary hypertension. *Eur. Respir. J.* **53**, 1801913 (2019).

3. Gaine, S. P. & Rubin, L. J. Primary pulmonary hypertension. *Lancet* **352**, 719–725 (1998).

4. Romberg, E. Ueber sklerose der lungen arterie. *Dtsch. Arch. Klin. Med.* **48**, 197–206 (1891).

5. Dresdale, D. T., Schultz, M. & Michtom, R. J. Primary pulmonary hypertension. I. Clinical and hemodynamic study. *Am. J. Med.* **11**, 686–705 (1951).

6. Simonneau, G. *et al.* Clinical classification of pulmonary hypertension. *J. Am. Coll. Cardiol.* **43**, S5–12 (2004).

7. Simonneau, G. *et al.* Updated clinical classification of pulmonary hypertension. *J. Am. Coll. Cardiol.* **62**, D34–41 (2013).

8. Ivy, D. Pulmonary hypertension in children. *Cardiol. Clin*. **34**, 451–472 (2016).

9. Humbert, M. *et al.* Pulmonary arterial hypertension in France: results from a national registry. *Am. J. Respir. Crit. Care Med.* **173**, 1023–1030 (2006).

10. Badesch, D. B. *et al.* Pulmonary arterial hypertension: baseline characteristics from the REVEAL Registry. *Chest* **137**, 376–387 (2010).

11. Tuder, R. M. *et al.* Relevant issues in the pathology and pathobiology of pulmonary hypertension. *J. Am. Coll. Cardiol.* **62**, D4–12 (2013).

12. Larkin, E. K. *et al.* Longitudinal analysis casts doubt on the presence of genetic anticipation in heritable pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* **186**, 892–896 (2012).

13. Batton, K. A. *et al.* Sex differences in pulmonary arterial hypertension: role of infection and autoimmunity in the pathogenesis of disease. *Biol. Sex Differ.* **9**, 15 (2018).

14. Machado, R. D. *et al.* A physical and transcript map based upon refinement of the

critical interval for PPH1, a gene for familial primary pulmonary hypertension. The International PPH Consortium. *Genomics* **68**, 220–228 (2000).

15. International PPH Consortium *et al.* Heterozygous germline mutations in BMPR2, encoding a TGF-beta receptor, cause familial primary pulmonary hypertension. *Nat. Genet.* **26**, 81–84 (2000).

16. Deng, Z. *et al.* Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. *Am. J. Hum. Genet.* **67**, 737–744 (2000).

17. Thomson, J. R. *et al.* Sporadic primary pulmonary hypertension is associated with germline mutations of the gene encoding BMPR-II, a receptor member of the TGF-beta family. *J. Med. Genet.* **37**, 741–745 (2000).

18. Machado, R. D. *et al.* BMPR2 haploinsufficiency as the inherited molecular mechanism for primary pulmonary hypertension. *Am. J. Hum. Genet.* **68**, 92–102 (2001).

19. Liu, F., Ventura, F., Doody, J. & Massagué, J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. *Mol. Cell. Biol.* **15**, 3479–3486 (1995).

20. David, L., Feige, J.-J. & Bailly, S. Emerging role of bone morphogenetic proteins in angiogenesis. *Cytokine Growth Factor Rev.* **20**, 203–212 (2009).

21. Rigueur, D. *et al.* The type I BMP receptor ACVR1/ALK2 is required for chondrogenesis during development. *J. Bone Miner. Res.* **30**, 733–741 (2015).

22. Shi, Y. & Massagué, J. Mechanisms of TGF-β signaling from cell membrane to the nucleus. *Cell* **113**, 685–700 (2003).

23. Machado, R. D. *et al.* Pulmonary arterial hypertension: a current perspective on established and emerging molecular genetic defects. *Hum. Mutat.* **36**, 1113–1127 (2015).

24. Sankelo, M. *et al.* BMPR2 mutations have short lifetime expectancy in primary pulmonary hypertension. *Hum. Mutat.* **26**, 119–124 (2005).

25. Aldred, M. A., Machado, R. D., James, V., Morrell, N. W. & Trembath, R. C. Characterization of the BMPR2 5′-untranslated region and a novel mutation in pulmonary hypertension. *Am. J. Respir. Crit. Care Med.* **176**, 819–824 (2007).

26. Wang, H. *et al.* Novel promoter and exon mutations of the BMPR2 gene in Chinese patients with pulmonary arterial hypertension. *Eur. J. Hum. Genet.* **17**, 1063–1069 (2009).

27. Roberts, K. E. BMPR2 mutations in pulmonary arterial hypertension with congenital heart disease. *Eur. Respir. J.* **24**, 371–374 (2004).

28. Beppu, H. *et al.* BMP type II receptor regulates positioning of outflow tract and remodeling of atrioventricular cushion during cardiogenesis. *Dev. Biol.* **331**, 167–175 (2009).

29. Harrison, R. E. *et al.* Molecular and functional analysis identifies ALK-1 as the predominant cause of pulmonary hypertension related to hereditary haemorrhagic telangiectasia. *J. Med. Genet.* **40**, 865–871 (2003).

30. Harrison, R. E. *et al.* Transforming growth factor-beta receptor mutations and pulmonary arterial hypertension in childhood. *Circulation* **111**, 435–441 (2005).

31. Fujiwara, M. *et al.* Implications of mutations of activin receptor-like kinase 1 gene (ALK1) in addition to bone morphogenetic protein receptor II gene (BMPR2) in children with pulmonary arterial hypertension. *Circ. J.* **72**, 127–133 (2008).

32. Shintani, M., Yagi, H., Nakayama, T., Saji, T. & Matsuoka, R. A new nonsense mutation of SMAD8 associated with pulmonary arterial hypertension. *J. Med. Genet.* **46**, 331–337 (2009).

33. Nasim, M. T. *et al.* Molecular genetic characterization of SMAD signaling molecules in pulmonary arterial hypertension. *Hum. Mutat.* **32**, 1385–1389 (2011).

34. Wang, G. *et al.* Novel homozygous BMP9 nonsense mutation causes pulmonary arterial hypertension: a case report. *BMC Pulm. Med.* **16**, 17 (2016).

35. Austin, E. D. *et al.* Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. *Circ. Cardiovasc. Genet.* **5**, 336–343 (2012).

36. Ma, L. *et al.* A novel channelopathy in pulmonary arterial hypertension. *N. Engl. J. Med.* **369**, 351–361 (2013).

37. Nohe, A., Keating, E., Underhill, T. M., Knaus, P. & Petersen, N. O. Dynamics and interaction of caveolin-1 isoforms with BMP-receptors. *J. Cell Sci.* **118**, 643–650

(2005).

38. Saldanha, S. *et al.* Caveolae regulate Smad signaling as verified by novel imaging and system biology approaches. *J. Cell. Physiol.* **228**, 1060–1069 (2013).

39. Zhu, N. *et al.* Exome sequencing in children with pulmonary arterial hypertension demonstrates differences compared with adults. *Circ. Genom. Precis. Med.* **11**, e001887 (2018).

40. den Dunnen, J. T. *et al*. HGVS recommendations for the description of sequence variants: 2016 update. *Hum. Mutat.* **37**, 564–569 (2016).

41. Copeland, C. A. *et al.* A disease-associated frameshift mutation in caveolin-1 disrupts caveolae formation and function through introduction of a de novo ER retention signal. *Mol. Biol. Cell* **28**, 3095–3111 (2017).

42. Marsboom, G. *et al.* Aberrant caveolin-1-mediated Smad signaling and proliferation identified by analysis of adenine 474 deletion mutation (c.474delA) in patient fibroblasts: a new perspective on the mechanism of pulmonary hypertension. *Mol. Biol. Cell* **28**, 1177–1185 (2017).

43. Drab, M. *et al.* Loss of caveolae, vascular dysfunction, and pulmonary defects in caveolin-1 gene-disrupted mice. *Science* **293**, 2449–2452 (2001).

44. Zhao, Y.-Y. *et al.* Defects in caveolin-1 cause dilated cardiomyopathy and pulmonary hypertension in knockout mice. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11375–11380 (2002).

45. Maniatis, N. A. *et al.* Increased pulmonary vascular resistance and defective pulmonary artery filling in caveolin-1-/- mice. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **294**, L865–873 (2008).

46. Yuan, X.-J., Wang, J., Juhaszova, M., Gaine, S. P. & Rubin, L. J. Attenuated K+ channel gene transcription in primary pulmonary hypertension. *Lancet* **351**, 726–727 (1998).

47. Yuan, J. X. *et al.* Dysfunctional voltage-gated K+ channels in pulmonary artery smooth muscle cells of patients with primary pulmonary hypertension. *Circulation* **98**, 1400–1406 (1998).

48. Young, K. A., Ivester, C., West, J., Carr, M. & Rodman, D. M. BMP signaling controls

PASMC KV channel expression in vitro and in vivo. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **290**, L841–848 (2006).

49. Remillard, C. V. *et al.* Function of Kv1.5 channels and genetic variations of KCNA5 in patients with idiopathic pulmonary arterial hypertension. *Am. J. Physiol. Cell Physiol.* **292**, C1837–1853 (2007).

50. Wang, G. *et al.* Early onset severe pulmonary arterial hypertension with 'two-hit' digenic mutations in both BMPR2 and KCNA5 genes. *Int. J. Cardiol.* **177**, e167–169 (2014).

51. Bohnen, M. S. *et al.* The impact of heterozygous KCNK3 mutations associated with pulmonary arterial hypertension on channel function and pharmacological recovery. *J. Am. Heart Assoc.* **6**, e006465 (2017).

52. Navas Tejedor, P. *et al.* An homozygous mutation in KCNK3 is associated with an aggressive form of hereditary pulmonary arterial hypertension. *Clin. Genet.* **91**, 453–457 (2017).

53. Montani, D. *et al.* Pulmonary veno-occlusive disease. *Eur. Respir. J.* **47**, 1518–1534 (2016).

54. O'Keefe, M. C. & Post, M. D. Pulmonary capillary hemangiomatosis: a rare cause of pulmonary hypertension. *Arch. Pathol. Lab. Med.* **139**, 274–277 (2015).

55. Montani, D. *et al.* Clinical phenotypes and outcomes of heritable and sporadic pulmonary veno-occlusive disease: a population-based study. *Lancet Respir. Med.* **5**, 125–134 (2017).

56. Dorfmüller, P. *et al.* Fibrous remodeling of the pulmonary venous system in pulmonary arterial hypertension associated with connective tissue diseases. *Hum. Pathol.* **38**, 893–902 (2007).

57. Eyries, M. *et al.* EIF2AK4 mutations cause pulmonary veno-occlusive disease, a recessive form of pulmonary hypertension. *Nat. Genet.* **46**, 65–69 (2014).

58. Best, D. H. *et al.* EIF2AK4 mutations in pulmonary capillary hemangiomatosis. *Chest* **145**, 231–236 (2014).

59. Dever, T. E. *et al.* Phosphorylation of initiation factor 2 alpha by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. *Cell* **68**, 585–596

(1992).

60. Best, D. H. *et al.* EIF2AK4 mutations in patients diagnosed with pulmonary arterial hypertension. *Chest* **151**, 821–828 (2017).

61. Hadinnapola, C. *et al.* Phenotypic characterization of mutation carriers in a large cohort of patients diagnosed clinically with pulmonary arterial hypertension. *Circulation* **136**, 2022–2033 (2017).

62. Meyer, E. *et al.* Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. *Nat. Genet.* **49**, 223–237 (2017).

63. Carss, K. J. *et al.* Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. *Am. J. Hum. Genet.* **100**, 75–90 (2017).

64. Westbury, S. K. *et al.* Expanded repertoire of variants responsible for platelet dysfunction and severe bleeding. *Blood* **130**, 1026–1030 (2017).

65. Gräf, S. *et al.* Identification of rare sequence variation underlying heritable pulmonary arterial hypertension. *Nat. Commun.* **9**, 1416 (2018).

66. Eyries, M. *et al.* Widening the landscape of heritable pulmonary hypertension mutations in pediatric and adult cases. *Eur. Respir. J.* **53**, 1801371 (2019).

67. Wang, X.-J. *et al.* Germline *BMP9* mutation causes idiopathic pulmonary arterial hypertension. *Eur. Respir. J.* **53**, 1801609 (2019).

68. Madan, M. *et al.* ATP13A3 and caveolin-1 as potential biomarkers for difluoromethylornithine-based therapies in pancreatic cancers. *Am. J. Cancer Res.* **6**, 1231–1252 (2016).

69. Archer, S. L. Pyruvate kinase and Warburg metabolism in pulmonary arterial hypertension: uncoupled glycolysis and the cancer-like phenotype of pulmonary arterial hypertension. *Circulation* **136**, 2486–2490 (2017).

70. Taraseviciene-Stewart, L. *et al.* Inhibition of the VEGF receptor 2 combined with chronic hypoxia causes cell death-dependent pulmonary endothelial cell proliferation and severe pulmonary hypertension. *FASEB J.* **15**, 427–438 (2001).

71. Teichert-Kuliszewska, K. *et al.* Bone morphogenetic protein receptor-2 signaling promotes pulmonary arterial endothelial cell survival: implications for loss-of-function

mutations in the pathogenesis of pulmonary hypertension. *Circ. Res.* **98**, 209–217 (2006).

72. Sui, H., Han, B. G., Lee, J. K., Walian, P. & Jap, B. K. Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**, 872–878 (2001).

73. Schuoler, C. *et al.* Aquaporin 1 controls the functional phenotype of pulmonary smooth muscle cells in hypoxia-induced pulmonary hypertension. *Basic Res. Cardiol.* **112**, 30 (2017).

74. Saadoun, S., Papadopoulos, M. C., Hara-Chikuma, M. & Verkman, A. S. Impairment of angiogenesis and cell migration by targeted aquaporin-1 gene disruption. *Nature* **434**, 786–792 (2005).

75. Yun, X., Jiang, H., Lai, N., Wang, J. & Shimoda, L. A. Aquaporin 1-mediated changes in pulmonary arterial smooth muscle cell migration and proliferation involve β-catenin. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **313**, L889–898 (2017).

76. Remenyi, A. Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* **17**, 2048–2059 (2003).

77. Sinner, D., Rankin, S., Lee, M. & Zorn, A. M. Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes. *Development* **131**, 3069–3080 (2004).

78. Zhu, N. *et al.* Rare variants in SOX17 are associated with pulmonary arterial hypertension with congenital heart disease. *Genome Med.* **10**, 56 (2018).

79. Hiraide, T. *et al.* SOX17 mutations in Japanese patients with pulmonary arterial hypertension. *Am. J. Respir. Crit. Care Med.* **198**, 1231–1233 (2018).

80. Matsui, T. *et al.* Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. *J. Cell Sci.* **119**, 3513–3526 (2006).

81. Corada, M. *et al.* Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nat. Commun.* **4**, 2609 (2013).

82. Lange, A. W. *et al.* Sox17 is required for normal pulmonary vascular morphogenesis. *Dev. Biol.* **387**, 109–120 (2014).

83. Lee, S.-H. *et al.* Notch pathway targets proangiogenic regulator Sox17 to restrict angiogenesis. *Circ. Res.* **115**, 215–226 (2014).

84. Goveia, J. *et al.* Endothelial cell differentiation by SOX17: promoting the tip cell or stalking its neighbor instead? *Circ. Res.* **115**, 205–207 (2014).

85. David, L., Mallet, C., Mazerbourg, S., Feige, J.-J. & Bailly, S. Identification of BMP9 and BMP10 as functional activators of the orphan activin receptor-like kinase 1 (ALK1) in endothelial cells. *Blood* **109**, 1953–1961 (2007).

86. Bidart, M. *et al.* BMP9 is produced by hepatocytes and circulates mainly in an active mature form complexed to its prodomain. *Cell. Mol. Life Sci.* **69**, 313–324 (2012).

87. Mi, L.-Z. *et al.* Structure of bone morphogenetic protein 9 procomplex. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3710–3715 (2015).

88. Chida, A. *et al.* Outcomes of childhood pulmonary arterial hypertension in BMPR2 and ALK1 mutation carriers. *Am. J. Cardiol.* **110**, 586–593 (2012).

89. Pfarr, N. *et al.* Hemodynamic and genetic analysis in children with idiopathic, heritable, and congenital heart disease associated pulmonary arterial hypertension. *Respir. Res.* **14**, 3 (2013).

90. Levy, M. *et al.* Genetic analyses in a cohort of children with pulmonary hypertension. *Eur. Respir. J.* **48**, 1118–1126 (2016).

91. Chida, A. *et al.* Missense mutations of the BMPR1B (ALK6) gene in childhood idiopathic pulmonary arterial hypertension. *Circ. J.* **76**, 1501–1508 (2012).

92. Bongers, E. M. H. F. *et al.* Mutations in the human TBX4 gene cause small patella syndrome. *Am. J. Hum. Genet.* **74**, 1239–1248 (2004).

93. Kerstjens-Frederikse, W. S. *et al.* TBX4 mutations (small patella syndrome) are associated with childhood-onset pulmonary arterial hypertension. *J. Med. Genet.* **50**, 500–506 (2013).

94. Vanlerberghe, C. *et al.* Small patella syndrome: New clinical and molecular insights into a consistent phenotype. *Clin. Genet.* **92**, 676–678 (2017).

95. Glaser, A. *et al.* Tbx4 interacts with the short stature homeobox gene Shox2 in limb development. *Dev. Dyn.* **243**, 629–639 (2014).

96. Arora, R., Metzger, R. J. & Papaioannou, V. E. Multiple roles and interactions of Tbx4 and Tbx5 in development of the respiratory system. *PLoS Genet.* **8**, e1002866 (2012).

97. Bohnen, M. S. *et al.* Loss-of-function ABCC8 mutations in pulmonary arterial

hypertension. *Circ. Genom. Precis. Med.* **11**, e002087 (2018).

98. Germain, M. *et al.* Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. *Nat. Genet.* **45**, 518–521 (2013).

99. Kimura, M. *et al.* A genome-wide association analysis identifies PDE1A|DNAJC10 locus on chromosome 2 associated with idiopathic pulmonary arterial hypertension in a Japanese population. *Oncotarget* **8**, 74917–74926 (2017).

100. Rhodes, C. J. *et al.* Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir. Med.* **7**, 227–223 (2019).

101. Richards, S. *et al.* 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 (2015).

102. Chung, W. K., Austin, E. D., Best, D. H., Brown, L. M. & Elliott, C. G. When to offer genetic testing for pulmonary arterial hypertension. *Can. J. Cardiol.* **31**, 544–547 (2015).

103. Girerd, B. *et al.* Genetic counselling in a national referral centre for pulmonary hypertension. *Eur. Respir. J.* **47**, 541–552 (2015).

104. Morrell, N. W. *et al.* Genetics and genomics of pulmonary arterial hypertension. *Eur. Respir. J.* **53**, 1801899 (2019).

105. Dunmore, B. J. *et al.* The lysosomal inhibitor, chloroquine, increases cell surface BMPR-II levels and restores BMP9 signalling in endothelial cells harbouring BMPR-II mutations. *Hum. Mol. Genet.* **22**, 3667–3679 (2013).

106. Drake, K. M., Dunmore, B. J., McNelly, L. N., Morrell, N. W. & Aldred, M. A. Correction of nonsense BMPR2 and SMAD9 mutations by ataluren in pulmonary arterial hypertension. *Am. J. Respir. Cell Mol. Biol.* **49**, 403–409 (2013).

107. Long, L. *et al.* Chloroquine prevents progression of experimental pulmonary hypertension via inhibition of autophagy and lysosomal bone morphogenetic protein type II receptor degradation. *Circ. Res.* **112**, 1159–1170 (2013).

108. Spiekerkoetter, E. *et al.* FK506 activates BMPR2, rescues endothelial dysfunction, and

reverses pulmonary hypertension. *J. Clin. Invest.* **123**, 3600–3613 (2013).

109. Hurst, L. A. *et al.* TNFα drives pulmonary arterial hypertension by suppressing the BMP type-II receptor and altering NOTCH signalling. *Nat. Commun.* **8**, 14079 (2017).

110. Morrell, N. W. *et al.* Targeting BMP signalling in cardiovascular disease and anaemia. *Nat. Rev. Cardiol.* **13**, 106–120 (2016).

111. Spiekerkoetter, E. *et al.* Randomised placebo-controlled safety and tolerability trial of FK506 (tacrolimus) for pulmonary arterial hypertension. *Eur. Respir. J.* **50**, 1602449 (2017).

112. Sitbon, O. *et al.* Clinical trial design and new therapies for pulmonary arterial hypertension. *Eur. Respir. J.* **53**, 1801908 (2019).

113. Long, L. *et al.* Selective enhancement of endothelial BMPR-II with BMP9 reverses pulmonary arterial hypertension. *Nat. Med.* **21**, 777–785 (2015).

114. Ormiston, M. L., Upton, P. D., Li, W. & Morrell, N. W. The promise of recombinant BMP ligands and other approaches targeting BMPR-II in the treatment of pulmonary arterial hypertension. *Glob. Cardiol. Sci. Pract.* **2015**, 47 (2015).

115. Upton, P.D., Long, L., Trembath, R.C., & Morrell, N.W. Functional characterization of bone morphogenetic protein binding sites and Smad1/5 activation in human vascular cells. *Mol. Pharmacol.* **73**, 539–552 (2008).

116. Hemnes, A. R. *et al.* PVDOMICS: a multi-center study to improve understanding of pulmonary vascular disease through phenomics. *Circ. Res.* **121**, 1136–1139 (2017).

117. Newman, J. H. *et al.* Enhancing insights into pulmonary vascular disease through a precision medicine approach. A joint NHLBI-Cardiovascular Medical Research and Education Fund workshop report. *Am. J. Respir. Crit. Care Med.* **195**, 1661–1670 (2017).

Author contributions

All the authors researched data for the article, discussed its content, wrote the manuscript, and reviewed and edited it before submission.

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Cardiology thanks E. Austin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Key points

- Heterozygous germline mutations in *BMPR2* represent the central susceptibility factor in the precipitation and progression of pulmonary arterial hypertension (PAH).
- Causal rare disease alleles have been identified in both bone morphogenetic protein (BMP) signalling and non-BMP pathways, confirming locus heterogeneity in PAH.
- Next-generation sequencing has been instrumental in expanding the genetic architecture of PAH by broadening the mutation spectrum in known genes and identifying novel genetic risk alleles.
- Childhood-onset PAH is associated with greater morbidity and mortality than adult-onset disease and has a distinctive genetic signature.
- PAH is a Mendelian disorder with complex disease traits, indicating a role for modifying common variation in disease development.
- Elucidating the genetic architecture of PAH provides unprecedented potential for the development of novel, precision medicine options in disease management.

Figure 1. **Major pathways and gene variants in heritable pulmonary arterial hypertension**. Several gene variants associated with heritable pulmonary arterial hypertension (HPAH) converge within the bone morphogenetic protein 9 (BMP9) and BMP10 signalling pathways, including *BMPR2*, *ACVRL1* (also known as *ALK1*), *ENG*, *SMAD4* and *SMAD9*. Caveolin 1 (encoded by *CAV1*) can also participate in BMP signalling. *KCNK3* encodes the $K_{2P}3.1$ potassium channel and *ABCC8* encodes a subunit of the ATP-sensitive potassium (K_{ATP}) channel. The *ATP13A3* gene product is an ATPase that resides on the endosomal compartment. *SOX17* and *TBX4* encode transcription factors critical for organogenesis. Biallelic mutations in *EIF2AK4*, which encodes an important kinase in the integrated stress response, are implicated in the pathogenesis of pulmonary veno-occlusive disease (PVOD) and pulmonary capillary haemangiomatosis (PCH). ATP13A3, probable cation-transporting ATPase 13A3; BMPR2, bone morphogenetic protein receptor type 2; SMAD, mothers against decapentaplegic homologue; BRE, BMP-responsive element; GCN2, eIF2α kinase GCN2.

Table 1. **Genetic and molecular features of the 16 genes underlying PAH**

Gene	Allelic series[a]	Mutation number[b]	Mutation categories (%)						Mode of inheritance	Independent validation	Functional assessment	Molecular mechanism	Refs
			Missense	Nonsense	Frameshift	Splice site	CNV	Other					
BMPR2	486	806	25.3	27.4	22.6	10	13.9	0.7	AD	Y	Y	Haploinsufficiency	[15, 16, 18, 23]
EIF2AK4	58	101	26.7	26.7	30.7	14.9	–	1.0	AR/AD	Y	Y	Loss-of-function	[57, 58, 61]
ACVRL1	54	76	85.5	7.9	5.3	–	–	1.3	AD	Y	Y	Haploinsufficiency	[29–31]
TBX4	44	48	31.3	10.4	37.5	4.2	12.5	4.2	AD	Y	N	NK	[39, 93]
GDF2	30	34	70.6	17.6	11.8	–	–	–	AD/AR	Y	Y	Haploinsufficiency	[34, 65, 67]
SOX17	24	26	57.7	7.7	30.8	3.9	–	–	AD	Y	N	NK	[65, 78]
ENG	16	17	47.1	5.9	29.4	5.9	–	11.8	AD	Y	N	NK	[29, 30]
KCNK3	15	16	93.8	–	–	–	–	6.3	AD	Y	Y	Haploinsufficiency	[36, 51]
ABCC8	12	12	91.7	–	–	8.3	–	–	AD	Y	Y	Haploinsufficiency	[97]
ATP13A3	11	11	36.4	18.2	27.3	9.1	–	9.1	AD	N	N	NK	[65]
SMAD9	10	11	54.5	36.4	–	–	9.1	–	AD	Y	Y	Haploinsufficiency	[32, 33, 106]
AQP1	5	9	88.9	11.1	–	–	–	–	AD	N	N	NK	[65]
CAV1	3	4	25	–	75	–	–	–	AD	Y	Y	Gain-of-function; dominant negative	[35, 41, 42]
BMP10[c]	2	2	50	50	–	–	–	–	AD	N	N	NK	[66]
SMAD4[c]	2	2	50	–	–	50	–	–	AD	N[c]	Y	NK	[33]
SMAD1[c]	1	1	100	–	–	–	–	–	AD	N[c]	Y	NK	[33]

[a]Allelic series refers to the number of distinct mutations, excluding recurrent variations. [b]Mutation number refers to the total number of independent alleles reported, excluding related individuals. For biallelic mutations in *EIF2AK4* and *GDF2*, each allele was counted as independent (that is, $n = 2$), except for five patients with documented consanguinity. [c]Variants of unknown significance. AD, autosomal dominant; AR, autosomal recessive; CNV, copy number variation; N, no; NK, not known; PAH, pulmonary arterial hypertension; Y, yes.

Table 2. **Clinical and demographic features associated with reported PAH mutations**

Gene	Number of reported mutation-positive cases by PAH category[a]						Female (%)	Mean age of onset and range (years)	Associated clinical features
	HPAH	IPAH	APAH	PCH / PVOD	Unknown	Total			
BMPR2	284	415	22	3	82	806	67.9	33 (1– 72)	CHD, IPF, PVOD
EIF2AK4	8	6	0	38	14	66	48.1	32 (11–72)	PCH, PVOD
ACVRL1	8	21	42	0	5	76	78.9	27 (1–63)	HHT
TBX4	4	30	14	0	0	48	58.5	25 (<1 to 72)	SPS
GDF2	1	32	0	0	0	33	71	33 (3–54)	–
SOX17	0	13	13	0	0	26	55.2	19 (<1 to 51)	CHD
ENG	0	8	9	0	0	17	62.5	30 (<1 to 59)	HHT, CHD
KCNK3	5	11	0	0	0	16	58.8	23 (<1 to 44)	–
ABCC8	2	8	2	0	0	12	58.3	29 (<1 to 79)	CHD
ATP13A3	0	11	0	0	0	11	90.9	48 (42–58)	–
SMAD9	1	8	2	0	0	11	80	25 (<1 to 58)	CHD
AQP1	3	6	0	0	0	9	44.4	32 (25–46)	–
CAV1	2	1	1	0	0	4	75	28 (1 – 67)	CHD
BMP10	0	1	1	0	0	2	100	20 (11– 28)	CHD
SMAD4	0	2	0	0	0	2	50	35	–
SMAD1	0	1	0	0	0	1	100	47	–

[a]Clinical category data were generated from a total of 1,140 unrelated cases in which pathogenic variants have been previously reported in the literature. Sex and age of onset were compiled using available data from previously published datasets. APAH, associated pulmonary arterial hypertension (including those listed under associated clinical features); CHD, congenital heart disease; HHT, hereditary haemorrhagic telangiectasia; HPAH, hereditary pulmonary arterial hypertension; IPAH, idiopathic pulmonary arterial hypertension; IPF, idiopathic pulmonary fibrosis; PCH, pulmonary capillary haemangiomatosis; PVOD, pulmonary veno-occlusive disease; SPS, small patella syndrome.