The pathophysiological role of novel pulmonary arterial hypertension gene *SOX17*

Yukyee Wu[1], John Wharton[1], Rachel Walters[1], Eleni Vasilaki[1], Jurjan Aman[1,2], Lan Zhao[1], Martin R. Wilkins[1] and Christopher J. Rhodes[1]

[1]National Heart and Lung Institute, Imperial College London, London, UK. [2]VUmc, Amsterdam University Medical Center, Amsterdam, The Netherlands.

Corresponding author: Christopher J. Rhodes (crhodes@imperial.ac.uk)

Shareable abstract (@ERSpublications)
***SOX17*, a risk gene in PAH, manifests *in vivo* phenotypes and interacts with key signalling pathways and transcriptional targets in the pathobiology of PAH. Restoration of *SOX17* gene expression and signalling may represent a new therapeutic strategy in PAH.** https://bit.ly/37ldkIL

Cite this article as: Wu Y, Wharton J, Walters R, *et al*. The pathophysiological role of novel pulmonary arterial hypertension gene *SOX17*. *Eur Respir J* 2021; 58: 2004172 [DOI: 10.1183/13993003.04172-2020].

Received: 12 Nov 2020
Accepted: 8 Feb 2021

Abstract

Pulmonary arterial hypertension (PAH) is a progressive disease predominantly targeting pre-capillary blood vessels. Adverse structural remodelling and increased pulmonary vascular resistance result in cardiac hypertrophy and ultimately failure of the right ventricle. Recent whole-genome and whole-exome sequencing studies have identified *SOX17* as a novel risk gene in PAH, with a dominant mode of inheritance and incomplete penetrance. Rare deleterious variants in the gene and more common variants in upstream enhancer sites have both been associated with the disease, and a deficiency of *SOX17* expression may predispose to PAH. This review aims to consolidate the evidence linking genetic variants in *SOX17* to PAH, and explores the numerous targets and effects of the transcription factor, focusing on the pulmonary vasculature and the pathobiology of PAH.

Introduction

Pulmonary arterial hypertension (PAH) is a rare condition characterised by structural remodelling of pulmonary arterioles and arteries, endothelial dysfunction, proliferation of mesenchymal cells in the vessel wall, and inflammation [1]. About a third of patients with PAH are classified as idiopathic PAH (IPAH), heritable PAH (HPAH) or drug-induced PAH and the annual mortality, even in the most experienced treatment centres and with a typical age at diagnosis in the mid-40s, averages 10% [2]. Despite common features in the vascular pathology, there is evidence of heterogeneity among these patients, as seen in the individual response to specific drugs and in the emerging underlying genetic architecture [3].

A major breakthrough in our understanding of the pathobiology of PAH was the discovery two decades ago of heterozygous germline mutations in the gene encoding bone morphogenetic protein (BMP) receptor type 2 (*BMPR2*), a member of the transforming growth factor (TGF)-β superfamily [4, 5]. Mutations in *BMPR2* are recognised as the most common cause of PAH, occurring in approximately 70% of HPAH and 10–20% of IPAH cases, and their discovery has suggested potential therapies designed to enhance BMP signalling in PAH [6]. In addition to *BMPR2*, at least 16 other genes have now been implicated in the genetic architecture of PAH [3, 7]. These include mutations in TGF-β-related genes, such as those encoding activin receptor type II-like 1 (*ACVRL1*), endoglin (*ENG*), BMP receptor type 1A and type 1B (*BMPR1A* and *BMPR1B*), as well as variants in caveolin-1 (*CAV1*), eukaryotic initiation translation factor 2 α kinase 4 (*EIF2AK4*), potassium subfamily K member 3 (*KCNK3*), SMAD family members 4 and 9 (*SMAD4* and *SMAD9*), and T-box 4 (*TBX4*) [3, 7]. Deleterious variants in other previously unidentified genes were recently found in a large-scale European study in PAH, encoding the transcription factor SRY-box 17 (*SOX17*), ATPase 13A3 (*ATP13A3*), aquaporin 1 (*AQP1*) and growth differentiation factor 2/BMP9 (*GDF2*) [8]. The evidence level for the causal role of mutations in a particular gene needs to be high for it to be used for clinical screening, and this has not yet been met for *SMAD4*, *SMAD1*, *KLF2*

(Kruppel-like factor 2), *BMPR1B* and *KCNA5* (potassium subfamily A member 5) [7]. Most recently, evidence has been presented for *BMP10* [9, 10] and *KDR* (kinase insert domain receptor) [11, 12] as novel PAH genes. With the exception of *EIF2AK4*, which is associated with pulmonary veno-occlusive disease in a recessive manner, mutations in all of these PAH-associated genes are inherited in an autosomal dominant manner and exhibit reduced penetrance, suggesting that additional genetic, epigenetic and/or environmental factors are involved in modifying the development and progression of the disease [7].

SOX family of transcription factors

The SRY-box (SOX) family of transcription factors was first described for its role in sex determination, and its members are now known to be critical regulators of development and regeneration. The family is characterised by a highly conserved high mobility group (HMG)-box DNA-binding domain and includes 20 members, almost all showing approximately 50% sequence homology to the HMG-box domain of the *SRY* (sex determining region Y) gene [13, 14]. The SOX family is divided into subgroups (SoxA to SoxH) according to sequence homology and functional characteristics. SOX17 belongs to the SoxF subgroup, which also includes SOX7 and SOX18 [15]. Members of this group have been shown to be essential regulators of several developmental processes, with SOX17 most prominent in the formation of the definitive endoderm and gastrointestinal tract [16, 17], cardiovascular development [18], and pulmonary vascular morphogenesis [19]. SOX17 also plays a key role in priming haemogenic potential in endothelial cells and regulating haematopoietic development from human embryonic and induced pluripotent stem cells [20, 21].

The human *SOX17* gene was sequenced in 2002, localised to chromosome 8q12–q13 (now more specifically 8q11.23). It comprises two exons, with transcripts demonstrated in adult human heart, lung, spleen, testis, ovary and placenta, in fetal lung and kidney, and in the gastrointestinal tract [22]. The deduced 414-amino-acid protein shares 41% and 43.5% identity with SOX7 and SOX18, respectively, and contains an HMG-box domain located close to the N-terminus and β-catenin-binding domain in the C-terminal region. *SOX17* is also an important downstream transcriptional target of the Wnt/β-catenin pathway [13]. The HMG-box domain forms an L-shaped region composed of three helices and includes a signature amino acid sequence (RPMNAFMVW) that mediates binding to consensus motifs in the minor groove of the DNA helix, causing DNA bending and the formation of nucleoprotein complexes [13]. SOX proteins generally also require the binding of partner transcription factors to elicit transcriptional activation or repression; for the SoxF subgroup this includes specific factors that belong to the POU/OCT and PAX protein families [23, 24]. While in development *SOX17* is more widely expressed in healthy adult tissues, including the lung [19], it displays an endothelial-specific profile.

SOX17 in disease

Variants of *SOX17* were first described in patients with congenital abnormalities of the kidney and urinary tract. Over-accumulation of the mutated protein (SOX17-p.Y259N) is associated with vesicoureteral reflux and attributed to its ability to inhibit the canonical Wnt/β-catenin signalling pathway [25]. However, the missense variant (c.775T>A, p.Tyr259Asn), which caused excessive accumulation of the SOX17 protein *in vitro*, was subsequently found in an individual without the associated congenital phenotype, leading to the suggestion that this and other *SOX17* missense variants outside the HMG-box domain may represent "at-risk" rather than causative alleles [26]. Some of these correspond to polymorphisms in the promoter (rs55815819:C>T), in the 5′-untranslated region (UTR) (rs62516525:C>T) and in the 3′-UTR (rs117273864: T>G). Others were identified in the coding sequence: c.479C>T (p.Ala160Val), c.775T>A (p.Tyr259Asn), c.942G>A (p.Gln314Gln) and c.971_972dupGCACCA (p.His325insGlnHis). Susceptibility to intracranial aneurysm (figure 1) has also been linked to single nucleotide polymorphisms on chromosome 8q and SOX17 signalling [27]. *SOX17* is the only gene in the 8q interval, located between two association peaks at 43 kb from rs10958409 and 64 kb from rs9298506. Endothelial-specific deletion of *SOX17* leads to impaired endothelial cell junctional assembly, cell–matrix adhesion, regeneration capacity, and paracrine secretion and aneurysm formation in hypertensive mice [28]. Numerous studies have implicated *SOX17* in cancer, with many indicating that it acts as a tumour suppressor. Multiple mechanisms are reported to downregulate *SOX17* expression in human tumours, including chromosome loss and dysregulation by microRNAs and long noncoding RNAs, but accumulating evidence suggests that DNA hypermethylation plays the dominant role [29]. Indeed, *SOX17* hypermethylation may be a useful molecular biomarker for the detection and diagnosis of various tumours, predicting response to therapy and prognosis [29]. In contrast, high *Sox17* expression drives tumour angiogenesis and its suppression in tumour endothelium may be beneficial [30].

Rare deleterious *SOX17* variants in PAH

In contrast to vesicoureteral reflux in which an elevated SOX17 level/accumulation has been observed, all PAH-associated variants predict loss of function, emphasising that tight control of SOX17 is required to

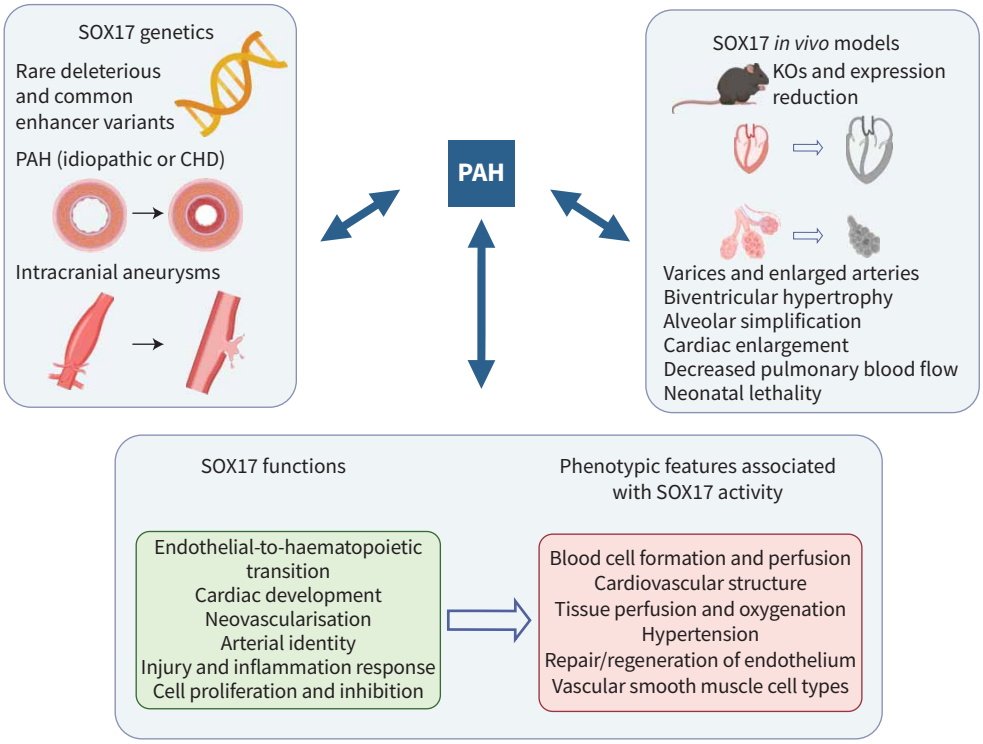

FIGURE 1 SOX17 has been linked to pulmonary arterial hypertension (PAH) through genetic studies and *in vivo* models have demonstrated its key role in pulmonary vascular development. CHD: congenital heart disease; KO: knockout. As an endothelial-specific transcription factor, SOX17 acts through multiple signalling pathways and has functions relevant to the pathophysiology of PAH, identifying it, or downstream signalling, as a clear target for therapeutic investigations.

avoid a variety of pathological phenotypes. They are also distinct from variants associated with other diseases. Recently, *SOX17* heterozygous variants were associated with PAH (figure 1) [8, 31]. As part of the UK National Institute for Health Research BioResource Rare Diseases project [32], a whole-genome sequence case–control analysis was conducted on 1038 patients with IPAH, HPAH and anorexigen-associated PAH, demonstrating statistically significant enrichment of rare deleterious variants in *SOX17*, *ATP13A3* and *AQP1* [8]. This was picked up by SKAT-O (optimal sequencing kernel association test) following protein-truncating variants impact analysis, with the exclusion of rare variants and deletions in previously reported PAH genes, such as *BMPR2*, to unmask novel causative genes. The heritability ascribed to each gene was relatively low, with *SOX17* variants representing approximately 0.9% of the overall cohort. Both *SOX17* and *AQP1* mutation carriers had a significantly younger age at diagnosis and the variants segregated with the phenotype in available pedigrees [8]. The *SOX17* variants included nonsense variants, predicted to lead to loss of β-catenin binding, and missense variants, predicted to disrupt interactions with β-catenin and the transcription factor OCT4 (figure 2). Missense variations in the conserved HMG-box domain are also likely to affect SOX17 DNA binding and lead to loss-of-function defects. In lung tissue sections, SOX17 was predominantly localised to the pulmonary endothelium, and *in vitro* was highly expressed in pulmonary endothelial and blood outgrowth endothelial cells compared with pulmonary smooth muscle cells [8].

Independent whole-exome sequencing studies have also implicated *SOX17* mutations as a risk factor in patients with PAH-associated congenital heart disease (PAH-CHD) [31]. Rare deleterious variants contributed to approximately 3.2% of the 256 PAH-CHD cases studied, compared with approximately 0.7% in a separate group of 413 IPAH/HPAH patients, and included missense variants in the HMG-box domain (figure 2). *SOX17* variants were additionally linked to a lower mean age of onset at 26 years compared with 41–46 years in pathogenic *KCNA5*, *ATP13A3* and *SMAD1* variant carriers. Rare predicted deleterious variants were observed in only five out of 7509 (0.07%) controls and without characterisation of specific variants it is unknown whether they are functional. The study also noted an enrichment of 163 rare deleterious variants among 149 putative targets of SOX17, with most expressed in the developing

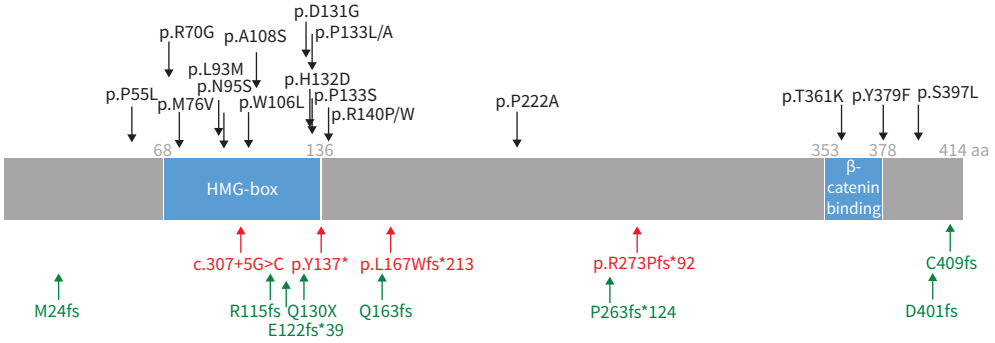

FIGURE 2 Schematic of the *SOX17* gene indicating variants identified by sequencing in pulmonary arterial hypertension patients [8, 31, 33, 34]. aa: amino acids; HMG: high mobility group. Black: missense mutations; red: protein-truncating variants; green: likely gene-disrupting variants.

murine heart and/or adult human pulmonary artery endothelial cells [31]. Another whole-exome sequencing study found four (three unrelated) *SOX17* variant carriers among 140 Japanese IPAH/HPAH patients, where one patient's healthy mother also carried the variant, emphasising the incomplete penetrance in PAH [33]. Similar analysis of 2572 cases in the US National Biological Sample and Data Repository for PAH reported rare deleterious *SOX17* variants in 10 patients, again with a significantly younger mean age of disease onset [34].

Common variants upstream of the *SOX17* gene in PAH

Common variants contribute to complex and multifactorial diseases, but the impact of common genetic variation on the risk of developing PAH and its natural history has not been widely explored. The largest genome-wide association study to date included two separate analyses and cross-validation of loci reaching genome-wide significance by meta-analysis, using data from four international case–control studies across 11 744 individuals with European ancestry and included 2085 patients with PAH [35]. One analysis used genotypes from 5895 whole-genome sequences and the other used genotyping array data from an additional 5849 individuals. A locus 100–200 kb upstream of *SOX17* (rs10103692; OR 1.80, 95% CI 1.55–2.08; $p=5.13\times10^{-15}$) in regions characterised by enhancer epigenomic markers in endothelial cells was found to associate with PAH. Conditional analysis showed that it comprised two independent signals located 100–103 kb (signal 1) and 106–200 kb (signal 2) upstream of *SOX17*, respectively. The lead single nucleotide polymorphisms for the two signals were rs13266183 (signal 1; OR 1.36, 95% CI 1.25–1.48; $p=1.69\times10^{-12}$) and rs10103692 (signal 2). The functional effect of the novel *SOX17* locus on PAH susceptibility was further confirmed by examining protein/DNA complexes and chromatin interaction using Hi-C conformation capture technology to determine DNA folding patterns. The first signal at locus rs13266183 was shown by CRISPR-mediated inhibition of the enhancer to exclusively regulate *SOX17*, reducing *SOX17* expression in human pulmonary artery endothelial cells. When the rs13266183 region was transfected into these cells, a 3–6-fold induction in luciferase reporter assays confirmed allele-specific enhancer activity.

Importantly, the *SOX17* genotype did not associate with survival. Impairment of SOX17 function might be more common in PAH than suggested by rare mutations in the gene itself and we anticipate that SOX17 has a role in susceptibility to the disease [35]. As the initiation of the pathophysiological processes in PAH remains an enigma, the discovery of *SOX17* may direct further research to elucidate the early beginnings of PAH.

Roles of SOX17 in cardiovascular and pulmonary vascular development and homeostasis

The three SoxF transcription factors have distinct as well as overlapping roles in cardiovascular development and postnatal neovascularisation (figure 1). SOX17 is necessary for the formation and maintenance of the endoderm [36], and comparative analyses of embryonic cardiovascular phenotypes in *Sox17*/*Sox18* double- and *Sox17* single-knockout mice indicated redundant roles for SOX17 and SOX18 in the formation of the anterior dorsal aorta and differentiation of the endocardial heart tube, with anomalies in anterior vascular formation corresponding with the sites of low *Sox7* expression [37]. In the early mouse embryo, *Sox17* expression is localised in the endoderm. It then increases in the dorsal aorta during vascular development and is subsequently preferentially expressed in arterial vascular endothelial and not venous cells [38]. SOX17 has a distinct role in developing and maintaining arterial endothelial cell

specificity, compared with the broader functions of SOX7 and SOX18 in the vasculature [31]. Whereas SOX17 and SOX18 are both required in vascular endothelial cells for postnatal angiogenesis in mice [39], low SOX17 and high SOX7/SOX18 levels are essential for systemic vein development and maintenance [40]. SOX18 is specifically required for lymphatic development [39].

SOX17 is required for the normal formation of the pulmonary vasculature and postnatal cardiovascular homeostasis (figure 1). The knockout of *Sox17* in developing pulmonary vascular endothelial cells (splanchnic mesenchyme-specific knockout using *Dermo1-Cre* mice) has potent effects on vascular development and causes perinatal defects. The abnormalities before birth include pulmonary vein varices, enlarged pulmonary arteries and decreased perfusion of the lung microvasculature, with alveolar simplification, biventricular cardiac hypertrophy and valvular regurgitation after birth [19].

The stability and permeability of the blood vessel wall in the lung relies on VE-cadherin [41] and endothelial cell-selective adhesion molecule (ESAM) [42], and the genes for both proteins contain consensus sequences for SOX17 in their promoters [42]. SOX17 has also been shown to be a key player in multiple signalling pathways of relevance to PAH, including TGF-β signalling, BMPs, Wnt/β-catenin, Notch and vascular endothelial growth factor (VEGF) signalling, which are themselves interlinked (figure 3) [43–45].

SOX17 in disease pathways related to PAH

SOX17 targets, such as nicotinamide phosphoribosyltransferase (NAMPT) and nestrin, have been implicated in PAH. Extracellular NAMPT is a pro-inflammatory cytokine that promotes endothelial-to-mesenchymal transition and pulmonary vascular remodelling [46]. SOX17 and SOX18 exhibit opposing effects on endothelial NAMPT expression; silencing of *SOX17* results in attenuation of VEGF-induced NAMPT promoter activity, whereas silencing of *SOX18* has the opposite effect [46]. Nestrin is a type IV intermediate filament that regulates sprouting and remodelling of the developing coronary vasculature [47]. Interestingly, nestrin has been found in proliferating vascular smooth muscle cells, notably during the development of experimental pulmonary hypertension, as well as in the media and neointima in human pulmonary vascular lesions [48]. Nestrin is also localised to the endothelium in pulmonary vascular lesions and implicated in endothelial proliferation and angiogenesis in PAH [49].

SOX17 and endothelin

The gene encoding endothelin converting enzyme 1 (*ECE1*) is a central component of the endothelin signalling pathway, being required for embryonic development and regulation of vascular tone. It is under the control of Forkhead and ETS-domain transcription factors, specifically FOXC2 and ETV2, that act synergistically *via* a FOX:ETS motif [50]. Unlike other FOX:ETS-dependent enhancers, ECE1 expression is restricted to the embryonic arterial endothelium and endocardium, and the enhancer region is bound and activated by *Sox17*, which functions together with *FoxC2* and *Etv2* to promote ECE1 expression [50]. Endothelin-1 (ET-1) is a potent vasoconstrictor peptide derived from the endothelium and demonstrates increased expression in the lungs of PAH patients, notably in endothelial cells of pulmonary arteries with medial thickening and intimal fibrosis [51]. There is a strong correlation with elevated plasma levels of ET-1 and PAH phenotypes, such as pulmonary vascular resistance, mean pulmonary arterial pressure, cardiac output, cardiac index and 6-min walk data [52]. Endothelin receptor antagonism is a validated therapeutic strategy for PAH. The effects of perturbation of SOX17 could well be mediated, in part, through the endothelin pathway.

SOX17 and TGF-β

SOX17 has various cell cycle mediating functions *via* TGF-β signalling. TGF-β is a potent inhibitor of proliferation in multiple epithelial cell types, including alveolar type II cells [53]. This anti-proliferative effect has been linked to SMAD-dependent transcriptional induction of cyclin-dependent kinase inhibitors. These include p15 and p21, and their induction leads to G_1 phase arrest. In the adult mouse lung, SOX17 decreased the expression of cell cycle inhibitors p15, p21 and p57 [54]. Co-immunoprecipitation assays showed that SOX17 interacts with SMAD3, thereby inhibiting TGF-β1-mediated transcription [55]. Inactivation of SMAD2/3 or SMAD4 causes downstream loss of expression of TGF-β target genes (including those encoding sphingosine-1-phosphate receptor 1 and N-cadherin) in mouse endothelial cells and results in vascular defects [20, 55, 56]. Rare variants of endoglin and β-glycan, both type III receptors for the TGF-β pathway, have been associated with intracranial aneurysm [57]. Loss of SOX17 expression has been linked to disruption of cell–cell adhesion *via* its modulation of VE-cadherin in precursor endothelial cells [20]. This mechanism could also be of relevance to aneurysm development. Within the context of endothelial-to-haematopoietic transition, SOX17 becomes a downstream effector of activated TGF-β signalling, where increased SOX17 expression impairs transition and blocks commitment to the

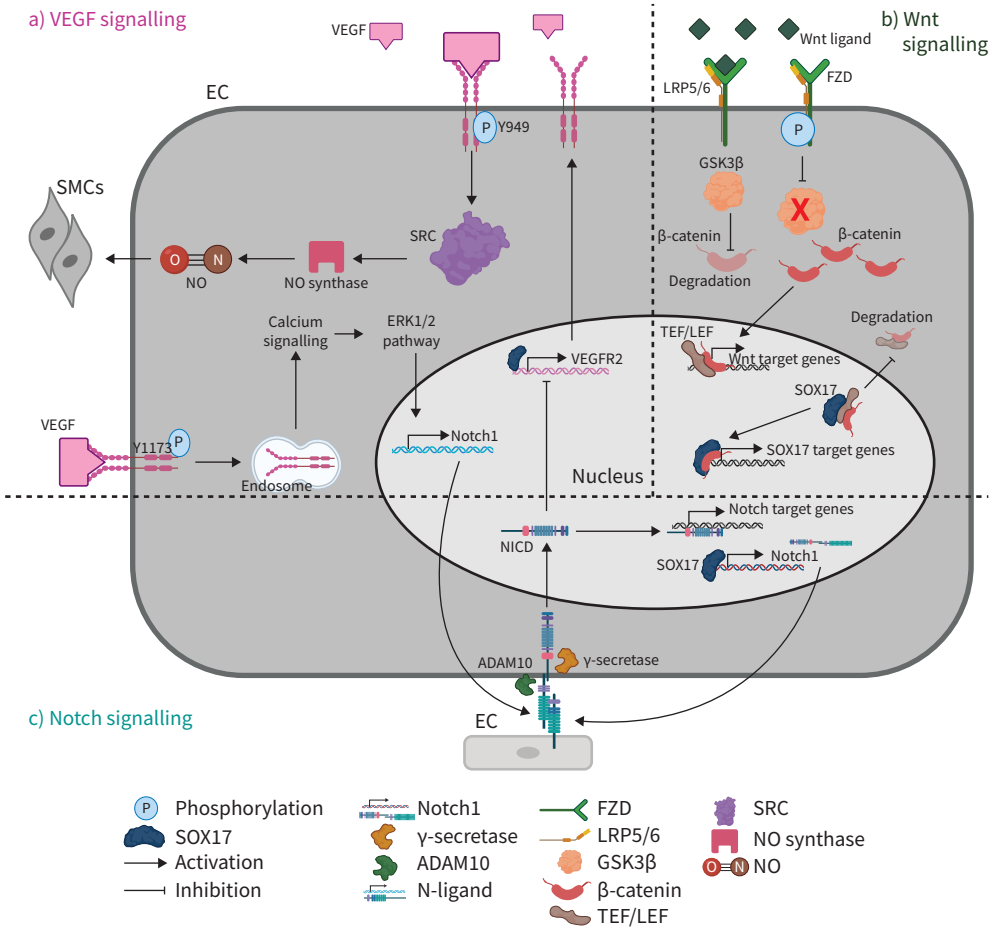

FIGURE 3 SOX17-related pathways in endothelial cells (ECs). VEGF: vascular endothelial growth factor; LRP: lipoprotein receptor-related protein; FZD: Frizzled; GSK3β: glycogen synthase kinase 3β; SMC: smooth muscle cell; NO: nitric oxide; ERK: extracellular signal-regulated kinase; TEF/LEF: T-cell factor/lymphoid enhancer factor; VEGFR2: VEGF receptor 2; NICD: Notch intracellular domain; ADAM10: a disintegrin and metalloproteinase domain containing protein 10; DLL: Delta-like protein. a) VEGF signalling. The phosphatidylinositol 3-kinase–AKT–mammalian target of rapamycin pathway is important in cell survival and regulation of barrier function, and involves activation of the SRC transcription factor through the phosphorylation of VEGFR2 at Y949. One of the downstream effects of this pathway is SMC vasodilation through activation of endothelial NO synthase [43]. VEGF signalling is regulated in ECs by the DLL4–Notch pathway. VEGF-induced angiogenesis is negatively regulated by Notch signalling, a pathway which SOX17 is potentially also negatively regulated by. b) Canonical Wnt signalling pathway. In the canonical Wnt signalling pathway, the FZD receptors bind to low-density LRP5/6 co-receptors and form a cell surface complex [44]. In the absence of a Wnt ligand, this complex does not repress GSK3β, which targets β-catenin for proteasome-mediated degradation through phosphorylation of serine residues. When the FZD/LRP5 receptor is bound, it represses the activity of GSK3β, and β-catenin stabilises and relocates to the nucleus. The SOX17 protein has a similar high mobility group binding domain structure to TCF/LEF transcription factors and is well known to interact with β-catenin. c) Notch signalling. The Notch receptor can be bound by Notch ligands Jagged1 and Jagged2 and DLL-1, -3 and -4 [45]. This causes the double cleavage of the Notch receptor extracellularly by ADAM10 and intracellularly by γ-secretase, which leads to the release of the NICD. The NICD then travels into the nucleus where it has a transcription factor role.

haematopoietic lineage [54]. As such, SOX17 is a key player in contributing to TGF-β signalling, and affects the proliferation of both pulmonary epithelial and haematopoietic cells.

SOX17 and BMPs

BMPs are highly conserved throughout evolution and account for 20 of the 33 known members of the TGF-β superfamily in humans [58]. BMPs are usually secreted as active dimeric complexes and some, like

BMP9, are bound to a prodomain. As well as *BMPR2*, rare, inactivating mutations in *GDF2* (encoding BMP9) have recently been associated with PAH development [8]. BMP signalling represses SOX17 expression, whereas TGF-β signalling activates SOX17 in endodermal development in the zebrafish [59]. In contrast, BMP2 and SOX17 form a positive feedback loop that triggers lineage commitment in human embryonic and induced pluripotent stem cells to become cardiac progenitors; BMP2 and OCT4 upregulate SOX17 expression [60]. SOX17 then promotes further expression and accumulation of BMP2 driving cardiogenesis. SOX17 and BMP2 demonstrated reciprocal induction in embryonic stem cells made resistant to cardiac differentiation by histone deacetylase 1 (*HDAC1*) knockdown, returning expression to levels comparable to wild-type cells [61]. BMPs and SOX17 also overlap in their signalling effector molecules with common effects on vascular development and homeostasis pathways, including VEGF, Notch, SMAD and Wnt signalling [62]. SOX17 and BMP9 share the same regulators and downstream effectors. BMP9 suppresses VEGF expression and VEGF-induced angiogenesis *via* ALK1 and BMPR2 signalling [58], which contrasts with promotion of VEGF signalling by SOX17. BMP9 treatment of endothelial cells resulted in restriction of apoptosis by inhibiting tumour necrosis factor-α-induced phosphorylation of c-Jun N-terminal kinase and prevented the loss of monolayer integrity driven by treatment with lipopolysaccharide [63]. Likewise, SOX17 is a key player in cell–cell adhesion and vascular endothelial remodelling. Whether or not one of the key features of loss of *SOX17* in heritable cases is an imbalance in TGF-β and BMP signalling, associated with dysfunctional pulmonary vascular homeostasis, is a topic worthy of further studies.

SOX17 and Wnt/β-catenin

SOX17 influences the regulation of cell proliferation and migration *via* Wnt/β-catenin signalling. SOX17 expression is linked to progression and prognosis of various forms of cancer, including colon cancer and endometrial cancer [64]. SOX17 is antagonistic to Wnt signalling in this context, displaying a tumour suppressor effect; the overexpression of SOX17 in SW480 colon carcinoma cells inhibited β-catenin/T-cell factor activity in a dose-dependent manner and reduced proliferation [65]. SOX17 inhibited the canonical Wnt pathway by forming a complex with TEF/LEF and β-catenin and targeting them for GSK3β-independent degradation (figure 3). A decrease in SOX17 levels has also been linked to endometrial cancer development and progression, accompanied by increased Wnt signalling [66]. SOX17 inhibits cell migration by inactivating the Wnt/β-catenin–epithelial-to-mesenchymal transition axis in endometrial cancer cells [67].

Endothelial-specific inactivation of SOX17 has profound effects on the vasculature. In the brain, it leads to increased permeability of the microcirculation. RNA expression analysis of brain endothelial cells in *Sox17* conditional knockout mice has identified members of the Wnt/β-catenin signalling pathway, such as NKD1 and LEF1, as downstream targets of SOX17 [41]. In this case, in contrast to studies of its role in cancer, SOX17 is a positive inducer of Wnt/β-catenin signalling, and it acts in synchrony with this pathway to induce and maintain integrity of the blood–brain barrier. In human endothelial cells, the canonical Wnt pathway controls the expression of target genes, such as those for c-Myc, cyclin D1, VEGF, survivin and axin 2, which are involved in cell maintenance, proliferation and survival [68–70], and interleukin-8, which is important in angiogenic regulation [71]. SOX17 was shown to play an essential role in cardiac muscle cell formation *via* interaction with the Wnt/β-catenin pathway, acting at a gastrulation-like stage and mediating mesoderm formation and patterning, which are prerequisites for cardiac myogenesis [72].

SOX17 and Notch signalling

The Notch pathway is an important downstream target of SOX17, binding directly to regulatory loci and promoting expression of Notch/Delta-like protein pathway genes (*Notch1*, *Notch4*, *Dll1* and *Dll4*) in adult mouse vascular endothelial cells [38]. Knockout of *SOX17* not only reduces the expression of arterial-specific genes but also induces the expression of venous-specific genes such as the gene encoding nuclear receptor subfamily 2 group F member 2 (*NR2F2*) [73]. In other situations, the Notch pathway can antagonise SOX17 signalling, supressing its endothelial expression at the post-transcriptional level, and thereby restrict angiogenesis [74]. SOX17 is involved in regulating migration of human umbilical vein endothelial cells *via* its interactions with Notch signalling (figure 3), destabilising endothelial junctions and rearranging cytoskeletal structure and upregulating expression of several genes preferentially expressed in tip cells for angiogenesis, including those for angiopoietin 2, platelet-derived growth factor-β and VEGF receptor 2 (VEGFR2) [74]. In this context, Notch and SOX17 are antagonistic, and SOX17 expression is suppressed in stalk cells where Notch signalling is relatively high. The Notch pathway regulates SOX17 expression mainly at the post-transcriptional level; activation of Notch in human umbilical vein endothelial cells by DLL4 markedly decreased SOX17 protein levels while mRNA levels were only moderately reduced [74]. In other instances, SOX17 can enhance Notch signalling. In the maintenance of

haematopoietic stem and progenitor cells in mouse intra-aortic haematopoietic clusters, SOX17 directly binds to the *Notch1* promoter and induces expression of Notch1 and downstream activation of the HES1 pathway [75].

SOX17 and VEGF

SOX17 involvement in the adult vasculature and PAH-related signalling is a relatively new field. Previous studies have focused on the role of SOX17 in development and cancer. VEGF stimulates SOX17 expression in cultured human endothelial cells [76] and upregulation of SOX17 expression at the protein level promotes VEGFR2 expression in angiogenic vessels [77]. SOX17 is also more directly involved in a positive feedback loop with VEGF (figure 3); VEGF upregulates SOX17 expression at the protein level in cultured endothelial cells *via* the mammalian target of rapamycin (mTOR) pathway while SOX17 promotes VEGFR2 expression in angiogenic vessels [77]. VEGF promotes DLL4 expression, which in turn activates Notch signalling and suppresses VEGFR2 expression. This is required to stabilise newly formed vessels [78, 79]. It underscores a link between loss of SOX17 and disrupted VEGF-regulated angiogenesis.

SOX17 and cyclins

SOX17 affects cell proliferation *via* cyclin signalling. A decrease in SOX17 levels with progression of endometrial cancer correlated with the increased expression of cyclin D1 across 30 tissue specimens [80]. In mature epithelial cell specification from progenitor cells and the cellular response to lung injury, SOX17 activates cell proliferation pathways involving cyclin E1 [81] and cyclin D1 [54]. Both are essential for the control of the cell cycle at the G_1/S transition. Cyclin proteins are downstream signalling molecules for Notch3, which regulates vascular smooth muscle cells switching between contractile and synthetic phenotypes and plays an important role in vascular remodelling, such as seen in PAH [82]. In chronic hypoxia-induced PAH mouse models, cyclin D1 expression is also elevated along with increased Notch3 expression in pulmonary vessels [83, 84]. In response to lung injury, endotoxaemia activates hypoxia-induced factor (HIF)-1α, which transcriptionally activates SOX17 to mediate cyclin E1-dependent endothelial cell regeneration [81]. This is especially relevant to PAH, where normoxic activation of HIF-1α is also observed. It may be more accurate to define the role of SOX17 in cell proliferation as a homeostatic role to maintain a healthy status quo, which remains vulnerable to environmental and cell status triggers.

Future directions

These pathways offer important clues as to how *SOX17* mutations and consequent loss of SOX17 expression and function might drive altered pulmonary endothelial function, but the variability in interactions between pathways in different cell and tissue types emphasises the importance of establishing their individual relevance and importance in pulmonary vascular cells. In particular, due to the importance of *BMPR2* in HPAH and the current development of therapeutics targeting BMP/TGF-β signalling in this condition, a detailed analysis of the role of SOX17–BMP–TGF-β functions in pulmonary endothelial cell function should be a priority. The availability of animal models of *Sox17* loss offers opportunities to test effects on PAH development in established systems including hypoxia. As with *BMPR2*, the presence of patients with *SOX17* rare pathogenic variants suggests therapeutic strategies aimed at restoring *SOX17* could be an early focus.

Conclusions

SOX17 expression is essential for normal cardiac and pulmonary vascular development, and animal models have demonstrated that disruption can produce phenotypes relevant to PAH. *SOX17* is specifically expressed in arterial endothelial cells in the lung and has emerged as a gene that delineates risk of PAH. Both rare and common genetic variation contributes to the disease. Indeed, SOX17 interacts with key signalling pathways and transcriptional targets that are known to be involved in the pathobiology of PAH. The restoration of *SOX17* expression and signalling may represent a promising new therapeutic strategy in PAH.

Author contributions: All authors contributed to the conception, approved the final version and take responsibility for the article. C.J. Rhodes and Y. Wu drafted the original manuscript, and all other authors provided critical revisions.

Conflict of interest: Y. Wu has nothing to disclose. J. Wharton has nothing to disclose. R. Walters has nothing to disclose. E. Vasilaki has nothing to disclose. J. Aman has nothing to disclose. L. Zhao has nothing to disclose. M.R. Wilkins has nothing to disclose. C.J. Rhodes reports personal fees from Actelion and United Therapeutics, outside the submitted work.

Support statement: This work was supported by the Academy of Medical Sciences (grant SBF004\1095) and the British Heart Foundation (grants FS-15-59-31839 and PG/19/17/34275). Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Humbert M, Guignabert C, Bonnet S, *et al.* Pathology and pathobiology of pulmonary hypertension: state of the art and research perspectives. *Eur Respir J* 2019; 53: 1801887.

2 Galiè N, Humbert M, Vachiery J-L, *et al.* 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS). *Eur Respir J* 2015; 46: 903.

3 Southgate L, Machado RD, Gräf S, *et al.* Molecular genetic framework underlying pulmonary arterial hypertension. *Nat Rev Cardiol* 2020; 17: 85–95.

4 Lane KB, Machado RD, Pauciulo MW, *et al.* Heterozygous germline mutations in *BMPR2*, encoding a TGF-β receptor, cause familial primary pulmonary hypertension. *Nat Genet* 2000; 26: 81–84.

5 Deng Z, Morse JH, Slager SL, *et al.* Familial primary pulmonary hypertension (PPH1) is caused by mutations in the bone morphogenetic protein receptor II gene. *Am J Hum Genet* 2000; 67: 737–744.

6 Morrell NW, Bloch DB, ten Dijke P, *et al.* Targeting BMP signalling in cardiovascular disease and anaemia. *Nat Rev Cardiol* 2016; 13: 106–120.

7 Morrell NW, Aldred MA, Chung WK, *et al.* Genetics and genomics of pulmonary arterial hypertension. *Eur Respir J* 2019; 53: 1801899.

8 Gräf S, Haimel M, Bleda M, *et al.* Identification of rare sequence variation underlying heritable pulmonary arterial hypertension. *Nat Commun* 2018; 9: 1416.

9 Eyries M, Montani D, Nadaud S, *et al.* Widening the landscape of heritable pulmonary hypertension mutations in paediatric and adult cases. *Eur Respir J* 2019; 53: 1801371.

10 Hodgson J, Swietlik EM, Salmon RM, *et al.* Characterisation of *GDF2* mutations and levels of BMP9 and BMP10 in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2019; 201: 575–585.

11 Eyries M, Montani D, Girerd B, *et al.* Familial pulmonary arterial hypertension by *KDR* heterozygous loss of function. *Eur Respir J* 2020; 55: 1902165.

12 Swietlik EM, Greene D, Zhu N, *et al.* Bayesian inference associates rare KDR variants with specific phenotypes in pulmonary arterial hypertension. *Circ Genom Precis Med* 2021; 14: e003155.

13 Kamachi Y, Kondoh H. Sox proteins: regulators of cell fate specification and differentiation. *Development* 2013; 140: 4129–4144.

14 Sarkar A, Hochedlinger K. The Sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* 2013; 12: 15–30.

15 Bowles J, Schepers G, Koopman P. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* 2000; 227: 239–255.

16 Séguin CA, Draper JS, Nagy A, *et al.* Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* 2008; 3: 182–195.

17 Pfister S, Jones VJ, Power M, *et al. Sox17*-dependent gene expression and early heart and gut development in *Sox17*-deficient mouse embryos. *Int J Dev Biol* 2011; 55: 45–58.

18 Francois M, Koopman P, Beltrame M. *SoxF* genes: key players in the development of the cardio-vascular system. *Int J Biochem Cell Biol* 2010; 42: 445–448.

19 Lange AW, Haitchi HM, LeCras TD, *et al.* Sox17 is required for normal pulmonary vascular morphogenesis. *Dev Biol* 2014; 387: 109–120.

20 Nakajima-Takagi Y, Osawa M, Oshima M, *et al.* Role of SOX17 in hematopoietic development from human embryonic stem cells. *Blood* 2013; 121: 447–458.

21 Clarke RL, Yzaguirre AD, Yashiro-Ohtani Y, *et al.* The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol* 2013; 15: 502–510.

22 Katoh M. WNT and FGF gene clusters. *Int J Oncol* 2002; 21: 1269–1273.

23 Hou L, Srivastava Y, Jauch R. Molecular basis for the genome engagement by Sox proteins. *Semin Cell Dev Biol* 2017; 63: 2–12.

24 Julian LM, McDonald ACH, Stanford WL. Direct reprogramming with SOX factors: masters of cell fate. *Curr Opin Genet Dev* 2017; 46: 24–36.

25 Gimelli S, Caridi G, Beri S, *et al.* Mutations in *SOX17* are associated with congenital anomalies of the kidney and the urinary tract. *Hum Mutat* 2010; 31: 1352–1359.

26 Combes P, Planche V, Eymard-Pierre E, *et al.* Relevance of *SOX17* variants for hypomyelinating leukodystrophies and congenital anomalies of the kidney and urinary tract (CAKUT). *Ann Hum Genet* 2012; 76: 261–267.

27 Bilguvar K, Yasuno K, Niemelä M, *et al.* Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat Genet* 2008; 40: 1472–1477.

28 Seungjoo L, Il-Kug K, Sung AJ, *et al.* Deficiency of endothelium-specific transcription factor Sox17 induces intracranial aneurysm. *Circulation* 2015; 131: 995–1005.
29 Higashijima Y, Kanki Y. Molecular mechanistic insights: the emerging role of SOXF transcription factors in tumorigenesis and development. *Semin Cancer Biol* 2019; 67: 39–48.
30 Yang H, Lee S, Lee S, *et al.* Sox17 promotes tumor angiogenesis and destabilises tumor vessels in mice. *J Clin Invest* 2013; 123: 418–431.
31 Zhu N, Welch CL, Wang J, *et al.* Rare variants in *SOX17* are associated with pulmonary arterial hypertension with congenital heart disease. *Genome Med* 2018; 10: 56.
32 Turro E, Astle WJ, Megy K, *et al.* Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* 2020; 583: 96–102.
33 Hiraide T, Kataoka M, Suzuki H, *et al. SOX17* mutations in Japanese patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2018; 198: 1231–1233.
34 Zhu N, Pauciulo MW, Welch CL, *et al.* Novel risk genes and mechanisms implicated by exome sequencing of 2572 individuals with pulmonary arterial hypertension. *Genome Med* 2019; 11: 69.
35 Rhodes CJ, Batai K, Bleda M, *et al.* Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir Med* 2019; 7: 227–238.
36 Kim I, Saunders TL, Morrison SJ. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell* 2007; 130: 470–483.
37 Sakamoto Y, Hara K, Kanai-Azuma M, *et al.* Redundant roles of *Sox17* and *Sox18* in early cardiovascular development of mouse embryos. *Biochem Biophys Res Commun* 2007; 360: 539–544.
38 Corada M, Orsenigo F, Morini MF, *et al.* Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nat Commun* 2013; 4: 2609.
39 Matsui T, Kanai-Azuma M, Hara K, *et al.* Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. *J Cell Sci* 2006; 119: 3513–3526.
40 Zhou Y, Williams J, Smallwood PM, *et al.* Sox7, Sox17, and Sox18 cooperatively regulate vascular development in the mouse retina. *PLoS One* 2015; 10: e0143650.
41 Corada M, Orsenigo F, Bhat GP, *et al.* Fine-tuning of Sox17 and canonical Wnt coordinates the permeability properties of the blood–brain barrier. *Circ Res* 2019; 124: 511–525.
42 Takahashi S, Nobuhisa I, Saito K, *et al.* Sox17-mediated expression of adherent molecules is required for the maintenance of undifferentiated hematopoietic cluster formation in midgestation mouse embryos. *Differentiation* 2020; 115: 53–61.
43 Jin Z-G, Wong C, Wu J, *et al.* Flow shear stress stimulates Gab1 tyrosine phosphorylation to mediate protein kinase B and endothelial nitric-oxide synthase activation in endothelial cells. *J Biol Chem* 2005; 280: 12305–12309.
44 He X, Semenov M, Tamai K, *et al.* LDL receptor-related proteins 5 and 6 in Wnt/β-catenin signaling: arrows point the way. *Development* 2004; 131: 1663–1677.
45 Fleming RJ. Structural conservation of Notch receptors and ligands. *Semin Cell Dev Biol* 1998; 9: 599–607.
46 Sun X, Sun BL, Babicheva A, *et al.* Direct eNAMPT involvement in pulmonary hypertension and vascular remodeling: transcriptional regulation by SOX and HIF2α. *Am J Respir Cell Mol Biol* 2020; 63: 92–103.
47 González-Hernández S, Gómez MJ, Sanchez-Cabo F, *et al.* Sox17 controls emergence and remodeling of nestin-expressing coronary vessels. *Circ Res* 2020; 127: e252–e270.
48 Saboor F, Reckmann AN, Tomczyk CUM, *et al.* Nestin-expressing vascular wall cells drive development of pulmonary hypertension. *Eur Respir J* 2016; 47: 876–888.
49 Bhagwani AR, Hultman S, Farkas D, *et al.* Endothelial cells are a source of nestin expression in pulmonary arterial hypertension. *PLoS One* 2019; 14: e0213890.
50 Robinson AS, Materna SC, Barnes RM, *et al.* An arterial-specific enhancer of the human *endothelin converting enzyme 1* (*ECE1*) gene is synergistically activated by Sox17, FoxC2, and Etv2. *Dev Biol* 2014; 395: 379–389.
51 Giaid A, Yanagisawa M, Langleben D, *et al.* Expression of endothelin-1 in the lungs of patients with pulmonary hypertension. *N Engl J Med* 1993; 328: 1732–1739.
52 Rubens C, Ewert R, Halank M, *et al.* Big endothelin-1 and endothelin-1 plasma levels are correlated with the severity of primary pulmonary hypertension. *Chest* 2001; 120: 1562–1569.
53 Machado RD, Aldred MA, James V, *et al.* Mutations of the TGF-β type II receptor BMPR2 in pulmonary arterial hypertension. *Hum Mutat* 2006; 27: 121–132.
54 Lange AW, Keiser AR, Wells JM, *et al.* Sox17 promotes cell cycle progression and inhibits TGF-β/Smad3 signaling to initiate progenitor cell behavior in the respiratory epithelium. *PLoS One* 2009; 4: e5711.
55 Feng X-H, Lin X, Derynck R. Smad2, Smad3 and Smad4 cooperate with Sp1 to induce $p15^{Ink4B}$ transcription in response to TGF-β. *EMBO J* 2000; 19: 5178–5193.
56 Itoh F, Itoh S, Adachi T, *et al.* Smad2/Smad3 in endothelium is indispensable for vascular stability *via* S1PR1 and N-cadherin expressions. *Blood* 2012; 119: 5320–5328.
57 Teresa S-S, Sumy M-J, Hariyadarshi P, *et al.* Sequencing of TGF-β pathway genes in familial cases of intracranial aneurysm. *Stroke* 2009; 40: 1604–1611.

58 Long L, Ormiston ML, Yang X, *et al.* Selective enhancement of endothelial BMPR-II with BMP9 reverses pulmonary arterial hypertension. *Nat Med* 2015; 21: 777–785.
59 Alexander J, Stanier DY. A molecular pathway leading to endoderm formation in zebrafish. *Curr Biol* 1999; 9: 1147–1157.
60 Stefanovic S, Abboud N, Désilets S, *et al.* Interplay of Oct4 with Sox2 and Sox17: a molecular switch from stem cell pluripotency to specifying a cardiac fate. *J Cell Biol* 2009; 86: 665–673.
61 Hoxha E, Lambers E, Wasserstrom JA, *et al.* Elucidation of a novel pathway through which HDAC1 controls cardiomyocyte differentiation through expression of SOX-17 and BMP2. *PLoS One* 2012; 7: e45046.
62 Scharpfenecker M, van Dinther M, Liu Z, *et al.* BMP-9 signals *via* ALK1 and inhibits bFGF-induced endothelial cell proliferation and VEGF-stimulated angiogenesis. *J Cell Sci* 2007; 120: 964–972.
63 Hoeper MM, McLaughlin VV, Dalaan AMA, *et al.* Treatment of pulmonary hypertension. *Lancet Respir Med* 2016; 4: 323–336.
64 Kormish JD, Sinner D, Zorn AM. Interactions between SOX factors and Wnt/β-catenin signaling in development and disease. *Dev Dynam* 2010; 239: 56–68.
65 Sinner D, Kordich JJ, Spence JR, *et al.* Sox17 and Sox4 differentially regulate β-catenin/T-cell factor activity and proliferation of colon carcinoma cells. *Mol Cell Biol* 2007; 27: 7802–7815.
66 Zhang Y, Bao W, Wang K, *et al.* SOX17 is a tumor suppressor in endometrial cancer. *Oncotarget* 2016; 7: 76036–76046.
67 Zhou W, Wang K, Wang J, *et al.* SOX17 inhibits tumor metastasis *via* Wnt signaling in endometrial cancer. *Onco Targets Ther* 2019; 12: 8275–8286.
68 Chen C, Chen H, Sun J, *et al.* Smad1 expression and function during mouse embryonic lung branching morphogenesis. *Am J Physiol Lung Cell Mol Physiol* 2005; 288: L1033–L1039.
69 Bain G, Müller T, Wang X, *et al.* Activated β-catenin induces osteoblast differentiation of C3H10T1/2 cells and participates in BMP2 mediated signal transduction. *Biochem Biophys Res Commun* 2003; 301: 84–91.
70 Rawadi G, Vayssière B, Dunn F, *et al.* BMP-2 controls alkaline phosphatase expression and osteoblast mineralisation by a Wnt autocrine loop. *J Bone Miner Res* 2003; 18: 1842–1853.
71 Masckauchán TNH, Shawber CJ, Funahashi Y, *et al.* Wnt/β-catenin signaling induces proliferation, survival and interleukin-8 in human endothelial cells. *Angiogenesis* 2005; 8: 43–51.
72 Liu Y, Asakura M, Inoue H, *et al.* Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells. *Proc Natl Acad Sci USA* 2007; 104: 3859–3864.
73 Swift MR, Pham VN, Castranova D, *et al.* SoxF factors and Notch regulate *nr2f2* gene expression during venous differentiation in zebrafish. *Dev Biol* 2014; 390: 116–125.
74 Lee SH, Lee S, Yang H, *et al.* Notch pathway targets proangiogenic regulator Sox17 to restrict angiogenesis. *Circ Res* 2014; 115: 215–226.
75 Saito K, Nobuhisa I, Harada K, *et al.* Maintenance of hematopoietic stem and progenitor cells in fetal intra-aortic hematopoietic clusters by the Sox17–Notch1–Hes1 axis. *Exp Cell Res* 2018; 365: 145–155.
76 Kutschera S, Weber H, Weick A, *et al.* Differential endothelial transcriptomics identifies semaphorin 3G as a vascular class 3 semaphorin. *Arterioscler Thromb Vasc Biol* 2011; 31: 151–159.
77 Kim K, Kim IK, Yang JM, *et al.* SoxF transcription factors are positive feedback regulators of VEGF signaling. *Circ Res* 2016; 119: 839–852.
78 Thurston G, Noguera-Troise I, Yancopoulos GD. The Delta paradox: DLL4 blockade leads to more tumour vessels but less tumour growth. *Nat Rev Cancer* 2007; 7: 327–331.
79 Suchting S, Freitas C, Le Noble F, *et al.* The Notch ligand Delta-like 4 negatively regulates endothelial tip cell formation and vessel branching. *Proc Natl Acad Sci USA* 2007; 104: 3225–3230.
80 Liao YM, Song Y, Li YK, *et al.* SOX17, β-catenin and CyclinD1 expression in the endometrioid adenocarcinoma and influence of 5-AZA on expression. *Cancer Gene Ther* 2020; 27: 256–263.
81 Liu M, Zhang L, Marsboom G, *et al.* Sox17 is required for endothelial regeneration following inflammation-induced vascular injury. *Nat Commun* 2019; 10: 2126.
82 Morris HE, Neves KB, Montezano AC, *et al.* Notch3 signalling and vascular remodelling in pulmonary arterial hypertension. *Clin Sci* 2019; 133: 2481–2498.
83 Wang W, Liu J, Ma A, *et al.* mTORC1 is involved in hypoxia-induced pulmonary hypertension through the activation of Notch3. *J Cell Physiol* 2014; 229: 2117–2125.
84 Ivanovska J, Shah S, Wong MJ, *et al.* mTOR-Notch3 signaling mediates pulmonary hypertension in hypoxia-exposed neonatal rats independent of changes in autophagy. *Pediatr Pulmonol* 2017; 52: 1443–1454.