Circulation

ORIGINAL RESEARCH ARTICLE

SOX17 Enhancer Variants Disrupt Transcription Factor Binding And Enhancer Inactivity Drives Pulmonary Hypertension

Rachel Walters, PhD; Eleni Vasilaki, MSc; Jurjan Aman, PhD; Chien-Nien Chen, PhD; Yukyee Wu, PhD; Olin D. Liang, PhD; Ali Ashek, PhD; Olivier Dubois, PhD; Lin Zhao, PhD; Farah Sabrin, PhD; Inês Cebola, PhD; Jorge Ferrer, PhD; Nicholas W. Morrell, MD; James R. Klinger, MD; Martin R. Wilkins, MD; Lan Zhao, PhD*; Christopher J. Rhodes, PhD*

BACKGROUND: Pulmonary arterial hypertension (PAH) is a rare disease characterized by remodeling of the pulmonary arteries, increased vascular resistance, and right-sided heart failure. Genome-wide association studies of idiopathic/heritable PAH established novel genetic risk variants, including conserved enhancers upstream of transcription factor (TF) *SOX17* containing 2 independent signals. SOX17 is an important TF in embryonic development and in the homeostasis of pulmonary artery endothelial cells (hPAEC) in the adult. Rare pathogenic mutations in *SOX17* cause heritable PAH. We hypothesized that PAH risk alleles in an enhancer region impair TF-binding upstream of *SOX17*, which in turn reduces *SOX17* expression and contributes to disturbed endothelial cell function and PAH development.

METHODS: CRISPR manipulation and siRNA were used to modulate *SOX17* expression. Electromobility shift assays were used to confirm in silico–predicted TF differential binding to the *SOX17* variants. Functional assays in hPAECs were used to establish the biological consequences of *SOX17* loss. In silico analysis with the connectivity map was used to predict compounds that rescue disturbed *SOX17* signaling. Mice with deletion of the *SOX17* signal 1 enhancer region (*SOX17*-4593/enhKO) were phenotyped in response to chronic hypoxia and SU5416/hypoxia.

RESULTS: CRISPR inhibition of *SOX17*-signal 2 and deletion of *SOX17*signal 1 specifically decreased *SOX17* expression. Electromobility shift assays demonstrated differential binding of hPAEC nuclear proteins to the risk and nonrisk alleles from both *SOX17* signals. Candidate TFs HOXA5 and ROR-α were identified through in silico analysis and antibody electromobility shift assays. Analysis of the hPAEC transcriptomes revealed alteration of PAH-relevant pathways on *SOX17* silencing, including extracellular matrix regulation. *SOX17* silencing in hPAECs resulted in increased apoptosis, proliferation, and disturbance of barrier function. With the use of the connectivity map, compounds were identified that reversed the SOX17-dysfunction transcriptomic signatures in hPAECs. SOX17 enhancer knockout in mice reduced lung SOX17 expression, resulting in more severe pulmonary vascular leak and hypoxia or SU5416/hypoxia-induced pulmonary hypertension.

CONCLUSIONS: Common PAH risk variants upstream of the *SOX17* promoter reduce endothelial *SOX17* expression, at least in part, through differential binding of HOXA5 and ROR-α. Reduced SOX17 expression results in disturbed hPAEC function and PAH. Existing drug compounds can reverse the disturbed SOX17 pulmonary endothelial transcriptomic signature.

Key Words: hypertension, pulmonary ■ SOX17 protein, human

Pulmonary arterial hypertension (PAH) is a rare but lethal disease. With no intervention, the mean survival is 2.8 years,[1] and with modern therapeutic intervention, the rate of mortality in the first year is ≈15%.[2] Increased pulmonary vascular resistance in PAH is driven by vasoconstriction, inflammation, and proliferative

Correspondence to: Christopher J. Rhodes, PhD, or Lan Zhao, PhD, Imperial College London, Hammersmith Hospital, 72 Du Cane Road, London, W12 ONN, United Kingdom. Email c.rhodes07@imperial.ac.uk or l.zhao@imperial.ac.uk

*L. Zhao and C.J. Rhodes contributed equally.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/circulationaha.122.061940.

For Sources of Funding and Disclosures, see page XXX

Circulation is available at www.ahajournals.org/journal/circ

Clinical Perspective

What Is New?

- *SOX17* enhancer common variants associated with pulmonary arterial hypertension development alter binding of transcription factors and thereby levels of SOX17.
- Loss of *SOX17* enhancer region drives worsening of pulmonary hypertension in animal models.
- Gene and protein signatures driven by SOX17 dysfunction are identified, and potential rescue therapeutic candidates are proposed.

What Are the Clinical Implications?

- Higher confidence in the causal variants of common genetic risk of pulmonary arterial hypertension is provided.
- This research describes the biological pathways that are likely affected in patients with pulmonary arterial hypertension carrying risk genotypes of SOX17, thereby improving our understanding of the pathogenesis of this deadly condition.
- Compounds targeting the pathways to restore physiological SOX17 signaling in patients with pulmonary arterial hypertension are a priority for further investigation.

Nonstandard Abbreviations and Acronyms

BMP9	bone morphogenetic protein-9
COL18A1	collagen type xviii alpha-1
Cmap	connectivity map
CRISPR-I	CRISPR inhibition
EC	endothelial cell
ECE1	endothelin-converting enzyme
ECFC	endothelial colony-forming cells
ECM	extracellular matrix
EMSA	electromobility shift assay
hPAEC	homeostasis of pulmonary artery endothelial cell
IL5	interleukin-5
PAH	pulmonary arterial hypertension
PECAM1	platelet and endothelial cell adhesion molecule-1
PH	pulmonary hypertension
qPCR	quantitative polymerase chain reaction
RNAseq	RNA sequencing
SPARC	secreted protein acidic and cysteine-rich
STAB1	stabilin-1
TF	transcription factor

remodeling of the intima and media of precapillary arteries.[3,4] The endothelium of healthy pulmonary arteries forms a semipermeable barrier, which dynamically adapts to external stimuli such as shear stress or hypoxia. Injury or dysfunction of the endothelium is thought to be an early but poorly understood trigger in PAH development. Although genetic factors enhance susceptibility (eg, bone morphogenetic protein receptor 2 [*BMPR2*] variants), environmental factors such as hypoxia, change in shear stress, inflammation, drugs, or toxins can directly injure the endothelial barrier, leading to apoptosis, loss of barrier integrity, and vascular remodeling of the pulmonary artery wall.[5]

Rare pathogenic variants in several genes, most commonly *BMPR2,* are associated with PAH,[6] but ≈75% of idiopathic cases cannot be explained by these variants. A recent large genome-wide association study using data from 11 744 European individuals (2085 patients) identified 2 independent PAH risk variant–containing signals (*SOX17*-signal 1 and *SOX17*-signal 2) in a region located 106 to 200 kb upstream of the *SOX17* gene promoter. The risk alleles are common in the populations tested and enriched in PAH, with 59% of patients homozygous for the risk allele of both signals compared with 46% of control subjects.[7] In addition, whole-genome sequencing studies identified rare deleterious variants in the *SOX17* gene associated with the development of severe PAH. Therefore, *SOX17* has the potential to provide powerful insight into PAH risk through rare and common variants.

The *SOX17* gene encodes the transcription factor (TF) SOX17, which is a member of the SoxF protein subfamily. SoxF proteins are important regulators of cell fate and differentiation[8] and have key roles in cardiovascular development.[9] SOX17 is essential for developmental angiogenesis and arterial differentiation in the embryo. In the adult, SOX17 plays a role in maintaining arterial identity and tumor angiogenesis.[10,11] Endothelial cell (EC) deletion of SOX17 in mouse models leads to embryonic lethality due to underdeveloped arteries and a complete lack of large arteries.[10] Conditional deletion of *SOX17* in splanchnic mesenchyme derivatives leads to severe vascular abnormalities, including reduced branching of pulmonary arteries and dilated cardiomyopathy.[12] Thus far, the role of SOX17 in the human pulmonary arterial endothelium remains unclear. In addition, it remains unclear how upstream common variants increase the risk of PAH. Variation in *SOX17*-signal 1 has been shown to affect *SOX17* expression, but the cellular and in vivo functions of this element are still poorly understood[7]; clear function has not been defined for *SOX17*-signal 2 region to date. We hypothesized that the PAH variants upstream of *SOX17* drive allele-specific TF binding at the 2 signals, which affects *SOX17* expression, homeostasis of pulmonary artery EC (hPAEC) function, and PAH development.

METHODS

The Supplemental Material gives the full methods and materials. We have deposited the RNA sequencing (RNAseq) data to GEO with accession no. GSE214742. All other data are available on reasonable request.

Patient ECs

Individuals with a diagnosis of idiopathic PAH (n=11) diagnosed according to international guidelines[13] and healthy control subjects who did not self-report cardiovascular or respiratory conditions (n=5) were recruited between August 23, 2017, and September 18, 2019, from the National Pulmonary Hypertension service and staff at Hammersmith Hospital for derivation of endothelial colony-forming cells (ECFCs) from blood. Samples were obtained with written informed consent and local research ethics committee approval. ECFCs were isolated and cultured as previously described[14] and extracted for RNA after growth in 2% FBS-supplemented EGM-2 media plus indicated treatments or vehicle for 24 hours.

CRISPR Manipulation

To inhibit *SOX17*-signal 2 and -signal 1, CRISPR inhibition (CRISPR-I) was used as previously described.[7] To delete *SOX17*-signal 1, CRISPR deletion was used. Two guides (Table S1) against *SOX17*-signal 1 were cloned into a pSpCas9(BB)-T2A-HygR vector (No. 118153) using a one-step dual CRISPR/Cas9 guide RNA cloning protocol as previously described.[15]

Electromobility Shift Assays and Supershift

To investigate TFs with binding that may be affected by the *SOX17* variants, the TF databases CIS-BP, PROMO, and ConSite were used. To investigate the differential binding of TFs to the risk and nonrisk alleles present at the *SOX17* variants, an electromobility shift assay (EMSA) was performed with the LightShift (chemiluminescent) EMSA kit according to the manufacturer's instructions. To investigate which TF binds to the *SOX17* variants, supershift assays with anti–HOXa5/ROR-α antibodies were performed. Five hPAEC donors were also used for chromatin immunoprecipitation–quantitative polymerase chain reaction (qPCR) assays with the same antibodies.

siRNA Experiments

To assess the role of *SOX17* in hPAECs, siRNA experiments were performed with silencer select siRNA (ThermoFisher) targeting *SOX17* (No. s34626). Comparable control siRNA was also used and included 2 scrambled siRNA (Nos. ASO2FOQH and 4390847) and a control targeting GAPDH (No. ASO2FLIC).

RNAseq Analysis

To assess whole transcriptomic effects of knockdown of *SOX17* by siRNA and CRISPR-I *SOX17*-signal 2 and -signal 1, RNAseq was carried out by the Imperial College BRC Genomics Facility, and analysis of the data set was performed in RStudio. To assess gene ontology changes resulting from siRNA-*SOX17*, CRISPR-I *SOX17*-signal 2, or CRISPR-I *SOX17*-signal 1, overrepresentation analysis was performed with the WEB-based Gene SeTAnaLysis Toolkit (Webgestalt).

qPCR Analysis

To investigate the change in expression of target genes, reverse transcription–polymerase chain reaction–qPCR were performed with actin-β (*ACTβ*) used as a reference gene ($2^{-\Delta Ct}$).

Western Blotting

To assess the levels of SOX17 protein after siRNA transduction, total protein was extracted from cells with radioimmunoprecipitation assay buffer (10×, Sigma) supplemented with protease and phosphatase inhibitor cocktail (ThermoFisher) and immunoblotted with anti-SOX17 (ab224637, Abcam, 1:500).

Proteomic Analysis

SomaLogic SomaScan measurements were available from a recent proteomics study by Rhodes et al.[16] Patient characteristics are shown in Table S2. Peripheral venous blood was collected during patients' routine clinical appointments.

The single nucleotide polymorphism genotypes for *SOX17* signal 1 (rs13266183) were obtained from a whole-genome sequencing study from the UK National Institute for Health Research BioResource.[7] Linear regression models were conducted with *SOX17*-signal 1 genotypes as the independent variables, protein concentrations as the dependent variables, and age and sex included as covariates. The *P* values from the linear regression were corrected for multiple comparisons with the false discovery rate method. All analyses were completed in R with RStudio version 1.4.1106, and the volcano plots were designed with the EnhancedVolcano package.

Cellular Function Assays

To assess the effect of siRNA-*SOX17* on hPAEC function, assays investigating proliferation, apoptosis, cell viability, adhesion, and barrier function were used and are discussed in full in the Supplemental Material.

In Silico Analysis Using the Connectivity Map

To analyze the differential expression patterns that occur when *SOX17* expression is manipulated, RNAseq of *SOX17*-signal 1 CRISPR-I hPAECs was performed as previously stated. To discover compounds that could be repurposed for the treatment of SOX17 dysfunction, the connectivity map (Cmap) was used.[17] Differential gene expression lists (Table S3) were used to create queries for the 3 conditions: *SOX17*-promoter activation, *SOX17*-promoter repression, and *SOX17*-signal 1 repression. Candidate compounds with a tau score of >90 or <−90 were selected. The compounds selected were sirolimus, Aminopurvalanol-a, and YK-4279. The signature of each compound was compared with the condition signature to select genes for further analysis by qPCR.

Animal Models of Pulmonary Hypertension in *SOX17* Enhancer Knockout

SOX17 enhancer knockout mice generated by CRISPR-cas9 technology at the MRC Imperial College London were

phenotyped blinded both in London and at Brown University (Providence, RI) to normoxia, hypoxia, or combined vascular endothelial growth factor receptor 2 blockade SUGEN/SU5416 and hypoxia using standard measures of pulmonary hypertension (PH; Supplemental Material) and in accordance with institutional guidelines.

RESULTS

PAH Common Variant Signals rs10958403 and rs765727 Identify *SOX17* Enhancers

The function of the noncoding sequence containing rs765727 (*SOX17*-signal 2) is unknown. To test whether it is an enhancer that targets *SOX17*, we first targeted the region using CRISPR-I in hPAECs. Guide RNAs targeting either rs765727 or the *SOX17* promoter led to a significant decrease in *SOX17* expression (enhancer guide A, 0.73±0.055; guide B, 0.81±0.025 of negative controls; both $P<0.05$; Figure 1A) but did not affect nearby gene *TMEM68* (Figure S1A). One guide also decreased the expression of another nearby gene, *MRPL15* (guide B, 0.83±0.013; Figure S1A). Deletion of *SOX17*-signal 1 (rs10958403) with CRISPR deletion guides resulted in a significant decrease in *SOX17* but had no effect on nearby genes *TMEM68* or *MRLP15* (Figure 1B and Figure S1B).

In EMSA, nuclear protein from hPAEC bound to probes representing the nonrisk alleles of both *SOX17* signals, inducing a shift, but exhibited loss of binding to the risk alleles (Figure 1C and Figure S1E). Competition EMSAs of both loci showed removal of this shift with the addition of unlabeled competitive probes for the nonrisk allele but not the risk allele, confirming the specificity of the protein binding to the nonrisk sequences (Figure 1C and Figure S1E).

In silico analyses using the TF CIS-BP, PROMO, and ConSite databases predicted multiple TFs more likely to bind the nonrisk versus risk sequence at both *SOX17* signals (Figure 1D and Figure S1D). These TFs were subsequently prioritized by the level of gene expression in hPAECs using RNAseq and the predicted binding score (if available). TFs with no detectable or low expression in hPAECs were not investigated further. E47 was not selected despite having the highest expression in hPAECs because a splice variant (E12) was predicted to also bind to the risk allele (C). NKX2-5 was found to be undetectable in hPAECs in alternative public databases. Of the hPAEC-expressed candidates tested (HOXa5, ROR-α, Lin54, ZFX, RAR; Figure S1) only HOXa5 and ROR-α unlabeled probes were found to compete for nuclear protein binding to the nonrisk sequences. For rs10958403, a TF competition EMSA with a probe containing the HOXa5 consensus binding sequence prevented the shift seen with the nonrisk allele probe. Incubation with an antibody for HOXa5 produced a supershift pattern consistent with a probe-protein-antibody complex (Figure 1C). For rs765727, a TF competition EMSA with a probe containing the ROR-α consensus sequence prevented the shift seen with the nonrisk allele probe. Incubation with an antibody for ROR-α also showed removal of the shift pattern (Figure S1E). To validate the EMSA findings, chromatin immunoprecipitation was performed with an antibody against HOXa5 in hPAEC nuclear lysates and qPCR for the area containing rs10958403. The area containing single nucleotide polymorphism rs10958403 was amplified only in donors containing a nonrisk A allele, confirming that HOXa5 binds this allele only at rs10958403 (Figure 1E). For rs765727/ROR-α, none of the donors tested expressed a nonrisk allele, preventing a similar comparison.

Taken together, these experiments indicated that the PAH signals identify enhancers active in hPAECs that target *SOX17* and contain variants likely to drive differential binding of TFs, including HOXa5 and ROR-α.

PAH-Associated Stimuli Regulate Endothelial *SOX17* Expression in Cells of Patients With PAH

To determine whether *SOX17* is regulated by factors implicated in the development of PAH, we tested *SOX17* expression in ECFCs derived from healthy control subjects or patients with PAH after stimulation with the hypoxia mimic dimethyloxalylglycine, the inflammatory stimulus lipopolysaccharide, and the BMPR2 ligand BMP9 (bone morphogenetic protein-9). ECFCs from patients with idiopathic PAH have reduced barrier function compared with those from control subjects and are more susceptible to lipopolysaccharide-induced permeability.[18] Lipopolysaccharide treatment not only is a well-established means of stimulating permeability in pulmonary EC cultures but also is linked to selective hypoxia-inducible factor 1α stabilization (dimethyloxalylglycine treatment in vitro) and the induction of SOX17 expression in human pulmonary microvascular ECs.[19] Expression of SOX17 has a protective effect and is required for the restoration of barrier function. ECFCs derived from patients with PAH with pathogenic BMPR2 variants are more susceptible to lipopolysaccharide-induced permeability compared with control ECFCs, and the effect is blocked by cotreatment with BMP9.[20] Stimulation with lipopolysaccharide, dimethyloxalylglycine, or BMP9 significantly increased *SOX17* expression in both ECFCs from both control subjects (n=5) and patients with PAH (n=11; Figure 2).

SOX17 Regulates Pathological Downstream Molecular Pathways and Functions in hPAECs

The risk alleles are associated with reduced enhancer activity and therefore reduced *SOX17* expression. To

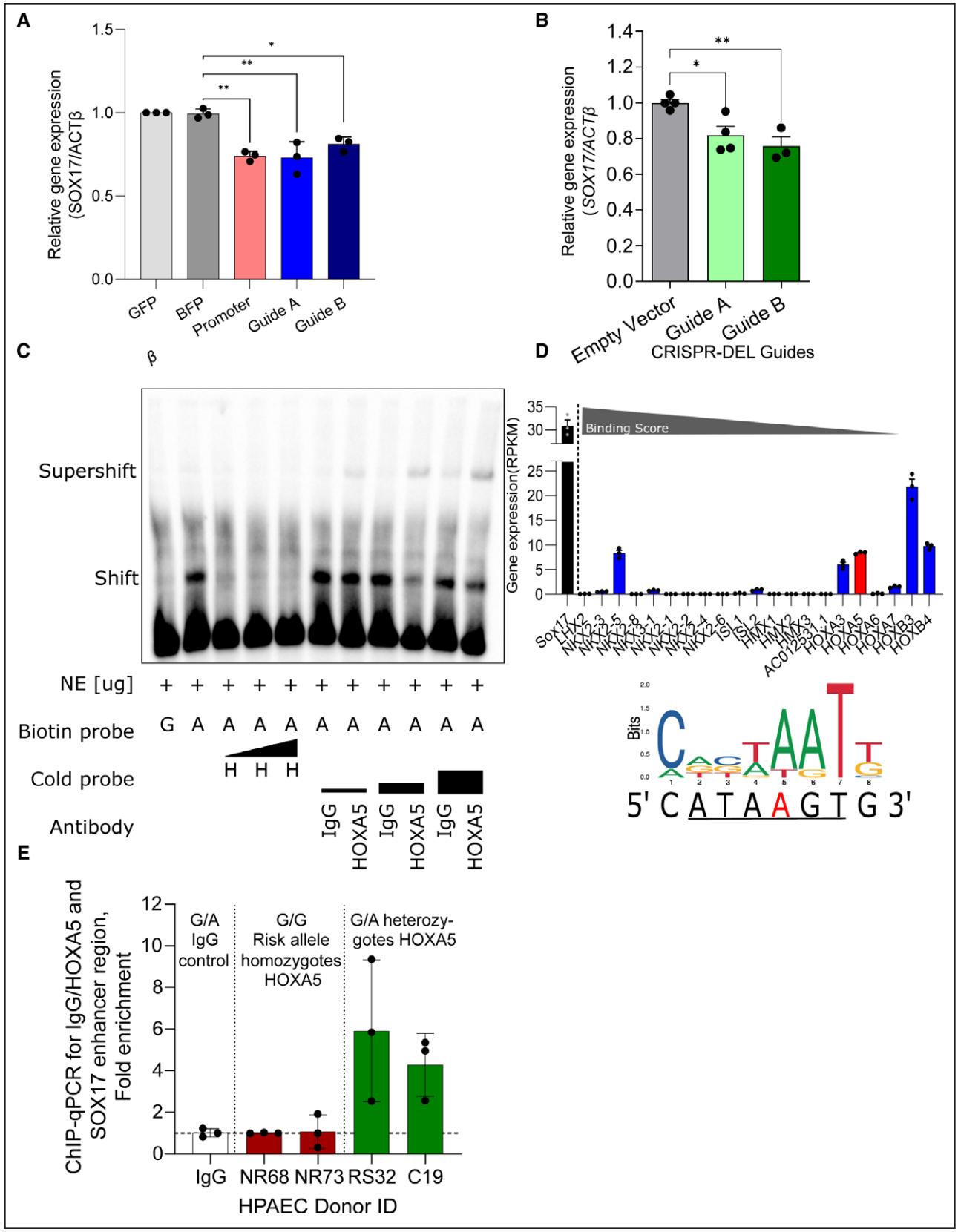

Figure 1. Defining upstream regulators of SOX17 and effects of common variation at the *SOX17* locus in PAH.
A, Knockdown of *SOX17* through CRISPR inhibition (CRISPRI) of *SOX17*-signal 2. Relative gene expression of *SOX17* compared with *ACTβ* in homeostasis of pulmonary artery endothelial cells (hPAECs). Ordinary 1-way ANOVA of conditions compared with BFP (blue fluorescent protein) condition with the Dunnett multiple-comparisons test. n=3 experiments performed in triplicate. **B**, Knockdown of *SOX17* through CRISPR deletion (CRISPR-DEL) of *SOX17*-signal 1 region. Relative gene expression of *SOX17* compared with *ACTβ* in hPAECs. (*Continued*)

Figure 1 Continued. CRISPR deletion guides A/B target *SOX17*-signal 1. Ordinary 1-way ANOVA of conditions compared with BFP condition with the Dunnett multiple-comparisons test. n=3. **C**, Electromobility shift assay (EMSA) showing binding of hPAEC nuclear proteins to 21-bp DNA probes containing the sequence at the rs1098403 region. Shift and supershift are highlighted. Biotin-probe, biotin-labeled probe (G/A, alleles of *SOX17* enhancer variant included in probe sequence); Cold-Probe, unlabeled probe. Black triangles show increasing molecular excess from **left** to **right**. **D**, RNA-sequencing expression in hPAECs (RPKM) of potential transcription factors (TFs) of interest for rs1098403. *HOXa5* is shown in red. *SOX17* is shown in black for reference. All other TFs are shown in blue and were found through CIS-BP. Binding score (taken from CIS-BP) refers to the predicted likelihood of the TF binding to the given sequence and decreases from **left** to **right**. Underlined sequence refers to the potential binding location of *HOXa5*, and the A in red is the site of rs1098403. Consensus binding sequence of HOXA5 was taken from JASPAR (jaspar.genereg.net) and is shown with the sequence of the region surrounding rs1098403 and rs765727 were taken from https://genome.ucsc.edu/. **P*<0.05. ***P*<0.01. n=3. **E**, Chromatin immunoprecipitation (ChIP–quantitative polymerase chain reaction (qPCR) for HOXA5 at rs10958493. Results of a qPCR (triplet measurements per donor, n=3) performed on precipitated fraction of ChIP with immunoglobulin G (IgG) or HOXa5 antibody in 4 hPAEC donors. The ChIP with IgG control was performed in hPAEC donor C19. GFP indicates green fluorescent protein; H, HOXa5 competitive probe; and ID, identification.

determine the downstream effects of *SOX17* depletion, we performed RNAseq analysis of hPAEC after modulation of *SOX17* by siRNA-mediated silencing or by CRISPR-I of *SOX17*-signal 1 and -signal 2 (Figure 3A and Figures S2 and S3). We identified 1717 genes that are differentially expressed after siRNA-*SOX17* knockdown (absolute log2-fold change >0.25 or <−0.25; false discovery rate q<0.05; Figure 3A). Gene ontology shows significant enrichment of the pathways involving cell adhesion and extracellular matrix (ECM) organization (Figure 3B). Four hundred fifty-one genes were significantly downregulated or upregulated after CRISPR-I of *SOX17*-signal 1; 786 genes were significantly downregulated or upregulated after CRISPR-I of *SOX17*-signal 2 (absolute log2-fold change >0.25 or <−0.25; *P*<0.05; Figure S3). There was a significant overlap of 81 genes differentially expressed in both si*SOX17* and *SOX17*-signal 1 CRISPR-I (*P*=0.0356; Supplemental Material). Consistent with the siRNA analysis, gene ontology analysis of these differentially expressed genes shows enrichment for pathways linked to the ECM organization and cell adhesion (Figure S3). qPCR was used to validate the effect of *SOX17*-siRNA on affected genes enriched in ECM and adhesion pathways. ECM and adhesion genes *ADAMTS12*, *MMP17*, and *LAMB3* were significantly increased compared with a negative control siRNA, whereas *CDH5* was decreased by *SOX17* knockdown (Figure 3C). These data demonstrate that *SOX17* loss in hPAEC drives gene expression changes in pathways relevant to PAH pathology.

Plasma Proteomic Differences in Patients With Differing *SOX17* Variant Genotypes

To further examine the potential effect of common variation in the enhancer area on downstream targets of SOX17, we analyzed the plasma proteome of 431 patients with PAH with a known genotype of the *SOX17*-signal 1 and *SOX17*-signal 2 using linear regression analysis. We identified 198 and 161 proteins for which plasma levels were significantly affected by the genotype in *SOX17*-signal 1 and *SOX17*-signal 2 single nucleotide polymorphisms, respectively (β estimate >|0.25| and *P*<0.05; Figure 4A and Figure S4). In line with our data obtained from genetic interference of *SOX17* in hPAEC, gene ontology analysis of the plasma proteomics identified enrichment in proteins involved in regulation of adhesion and ECM for both signals (Figure 4B and Figure S4B). In addition, enrichment in proteins involved in regulation of proliferation, migration, and apoptosis was found for *SOX17*-signal 2 (Figure S4). As similar processes and functions emerged from the

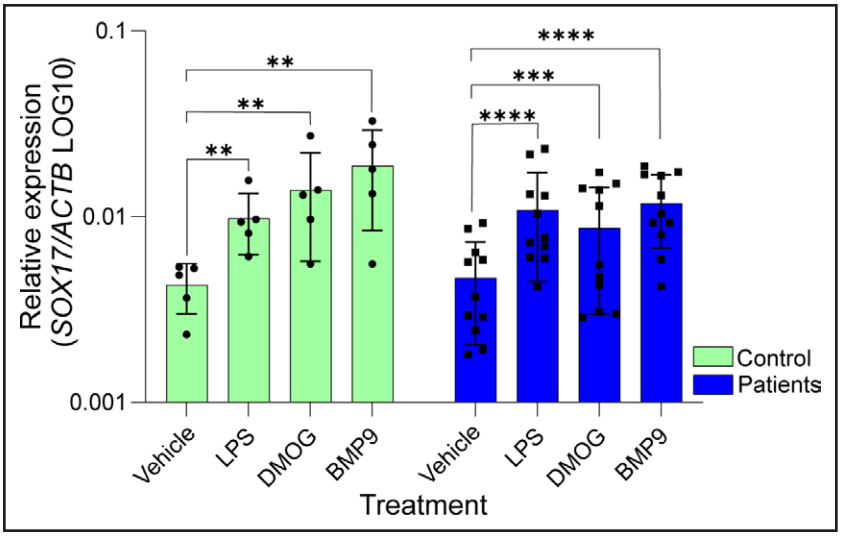

Figure 2. Defining the effect of upstream regulators and PAH-relevant stimuli on *SOX17*. Relative gene expression of *SOX17* compared with *ACTβ* in control and patient endothelial colony-forming cells after exposure to known pulmonary arterial hypertension (PAH) stimuli. Individual data points represent individuals, n=5 control subjects, n=11 patients. Vehicle/treatments in 2% FBS endothelial cell media. BMP9 indicates bone morphogenetic protein-9 (10 ng/mL); DMOG, dimethyloxalylglycine (100 µmol/L); and LPS, lipopolysaccharide (2 µg/mL). Ordinary 1-way ANOVA within groups compared with baseline condition with the Dunnett multiple comparisons test. ***P*<0.01. ****P*<0.005. *****P*<0.001.

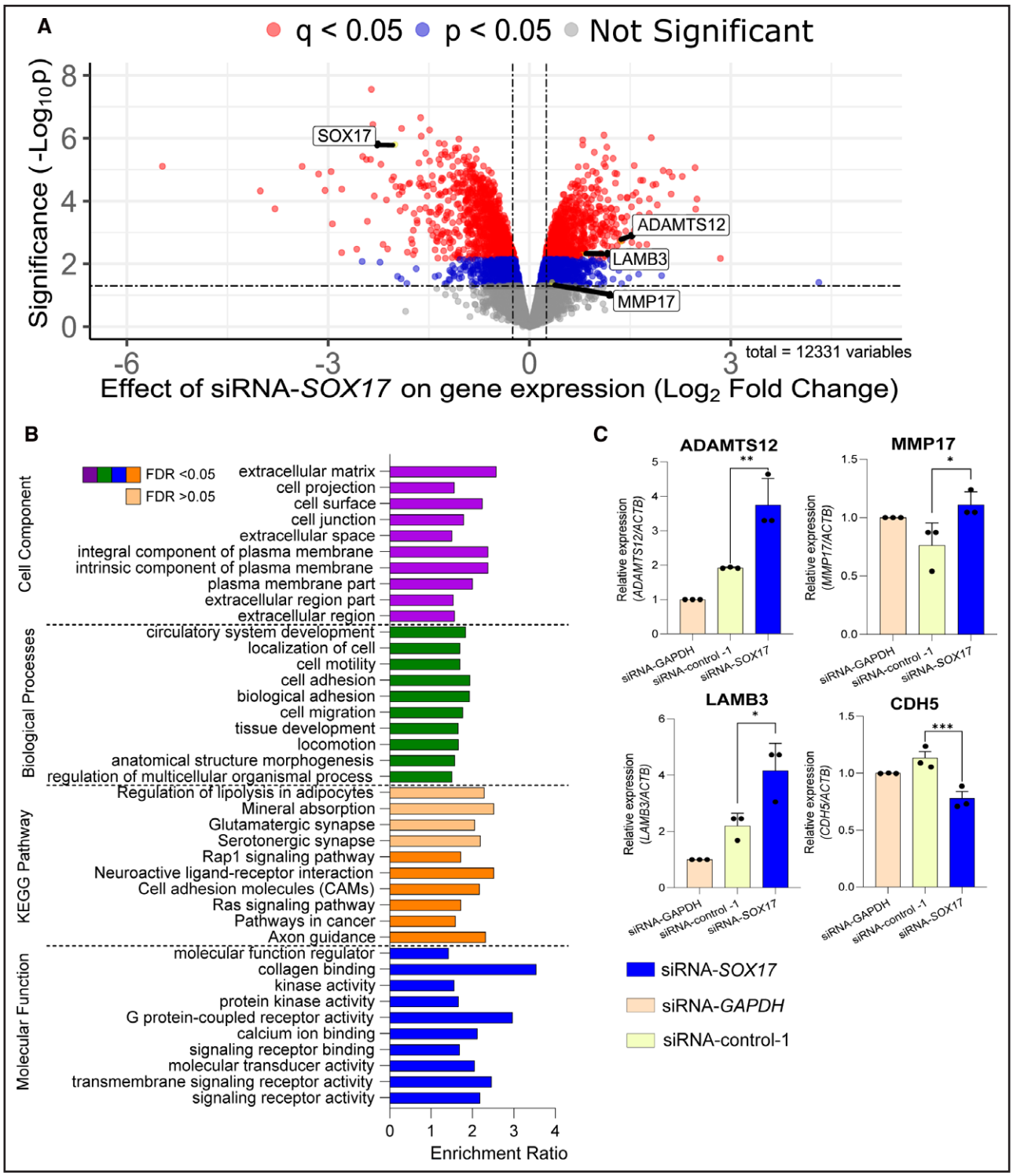

Figure 3. Analysis of Sox17 manipulation in pulmonary vascular cells.
A, Differentially expressed genes after *SOX17*-siRNA in homeostasis of pulmonary artery endothelial cells (hPAECs). Volcano plot of the $\log_2$ fold change (FC) between siRNA-*SOX17* and siRNA-negative control and the negative $\log_{10}$ *P* value. Differentially expressed genes shown in blue met the cutoff points: $P<0.05$ and $\log_2$ fold change <−0.25 or >0.25. Genes shown in red also met the cutoff of $q<0.05$. Genes of interest are highlighted in black boxes. n=4 with 12 196 variables. **B**, Overrepresentation analysis for enriched pathways and functions after *SOX17*-siRNA. Gene ontology analysis of cell component (purple), biological process (green), KEGG pathway (orange), and molecular function (blue) enrichment after *SOX17*-siRNA in hPAECs. Darker colors indicate $f<0.05$; lighter colors, $f>0.05$. Enrichment ratios were obtained from WebGestalt.
C, Relative gene expression of gene ontology target genes by quantitative polymerase chain reaction. Change in target gene expression is normalized to *ACTβ*, and all siRNA conditions are relative to the GAPDH-targeting siRNA control. Target genes are *ADAMTS12*, *MMP17*, *LAMB3*, and *CDH5*. All statistical tests shown are paired, 1-way Student *t* test. **P*<0.05. n=3. All in hPAECs after *SOX17*-siRNA treatment for 48 hours.

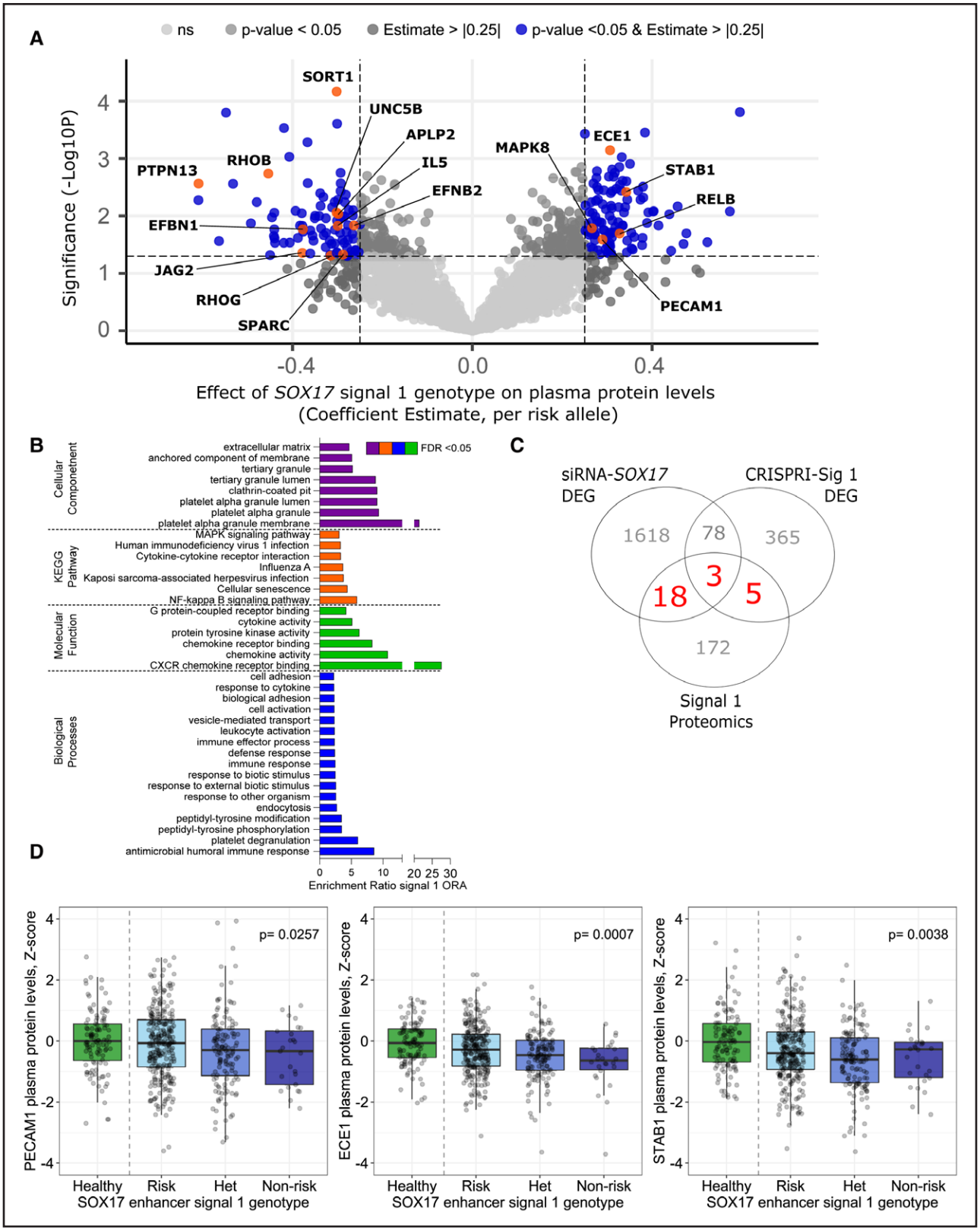

Figure 4. Effect of *SOX17* enhancer variant genotype on patient proteomics.

A, Linear regression for the effect of *SOX17*-signal 1 on the levels of serum proteins in the patient sample. Volcano plot of the coefficient estimate and the negative $\log_{10}$ *P* value. Corrected for age and sex. Protein shown in blue met threshold b estimate >|0.25| and *P*<0.05. Proteins of interest are labeled and shown in orange. n=431. **B**, Gene ontology analysis for the proteins significantly affected by signal 1 genotype. Gene ontology analysis of cell component (purple), biological process (green), KEGG pathway (orange), and molecular function (blue) enrichment. Proteins met the threshold b estimate >|0.25| and *P*<0.05. Enrichment ratios were obtained from WebGestalt. (*Continued*)

Figure 4 Continued. C, Comparisons of transcriptomic and proteomic analysis. Venn diagram showing overlapping differentially expressed genes (DEGs) and proteins from transcriptomic and proteomic analysis from CRISPR inhibition (CRISPRI) of *SOX17*-signal 1, siRNA-*SOX17*, and *SOX17*-signal 1 enhancer variant genotype on patient proteomics. Numbers in red show genes and proteins that are in common between all analyses. **D**, *Z*-scored proteins in healthy control subjects vs patients with different genotypes in proteins of interest. Proteins included are PECAM1 platelet and endothelial cell adhesion molecule-1), ECE1 (endothelin-converting enzyme-1), and STAB1 (stabilin-1). Risk, homozygous for risk allele, n=271. Homozygous for nonrisk allele (nonrisk), n=26; heterozygotes (Het), n=134; control, n=108.

transcriptome gene ontology analyses, the significantly affected proteins and genes from all conditions (CRISPRI, siRNA, and proteomics) were compared in detail (Figure 4C and Figure S4C). Six proteins were affected by patient genotypes at both *SOX17* signals (IL5 [interleukin-5], PTPN13, STAB1 [stabilin-1], SUGT1, GAPDH, and ADGRG5). There were 26 genes in common between *SOX17*-signal 1 proteomics and CRISPR-I or siRNA analyses (Figure 4A and Supplemental Material) and 23 genes in common between *SOX17*-signal 2 proteomics and CRISPR-I or siRNA analysis (Figure S4). These included SPARC (secreted protein acidic and cysteine-rich), PECAM1 (platelet and endothelial cell adhesion molecule-1), ECE1 (endothelin-converting enzyme-1), COL18A1(collagen type xviii alpha-1), IL5, and STAB1, which have been previously associated with PAH[21–24] (Figure 4D and Figure S4). These analyses suggest that differences in *SOX17* enhancer activity associated with PAH risk alleles lead to changes in the plasma proteome with pathologically relevant functions.

Functional Impact of Loss of SOX17 in Cultured hPAECs

To determine the functional impact of loss of *SOX17*, we exposed hPAECs to relevant stimuli after siRNA-mediated *SOX17* knockdown. SOX17 knockdown increased hPAEC apoptosis (caspase-3/7 activity) in response to either tumor necrosis factor-α or lipopolysaccharide compared with siRNA controls ($P<0.001$; Figure 5A), with hPAEC viability being either unchanged or decreased with siRNA-*SOX17* compared with siRNA controls (Figure 5E). Knockdown of *SOX17* led to an increase in hPAEC monolayer permeability, as measured by a transwell assay (both $P<0.05$; Figure 5B) and by electric impedance assays ($P<0.001$; Figure 5B bottom). Adhesion to collagen IV was significantly decreased in *SOX17*-depleted hPAECs compared with controls (Figure 5D), whereas SOX17 knockdown increased vascular endothelial growth factor–induced hPAEC proliferation as determined by MTT assays ($P<0.001$; Figure 5C). These results suggest that *SOX17* loss in hPAECs fundamentally changes their function, mirroring changes observed in patient pulmonary artery ECs.

Cmap Prediction of Rescue Compounds for Drug Repurposing

To predict whether available drug compounds could be repurposed to reverse the gene changes associated with *SOX17* dysfunction, we interrogated the Cmap database. The Cmap contains transcriptomic signatures of the effects of thousands of compounds on multiple cell lines, allowing comparison of user-generated signatures (Figure 6A). To test effects most relevant to the common and rare *SOX17* variants associated with PAH, we generated *SOX17* hPAEC signatures comprising differentially expressed genes after CRISPR-I of *SOX17*-signal 1 or the *SOX17* promoter or CRISPR activation of the *SOX17* promoter. In silico analysis of these signatures showed the compounds sirolimus, aminopurvalanol-a, and YK4279 to match our *SOX17* signature in hPAECs (Figure 6B). Sirolimus and aminopurvalanol-a are predicted to reverse (negatively connected to) *SOX17* promoter repression (τ score, −96.94 and −95.17, respectively). Aminopurvalanol-a is also negatively connected to *SOX17*-signal 1 repression (−99.65). YK4279 is predicted to mimic *SOX17* promoter activation (positive connection, 93.88). Comparisons of the signature of each compound from the Cmap and our RNAseq signature resulted in the gene lists shown in Table S3. For each compound, we selected a panel of genes that showed consistent directional changes across multiple cell lines in the Cmap and tested their expression by qPCR in hPAECs after compound exposure. For sirolimus, all tested genes showed the predicted expression change (Figure 6C). For aminopurvalanol-a, only half of the genes tested changed in the direction predicted by the Cmap signature (Figure S5). For YK4279, all the gene expression changes tested were as predicted (Figure S5). To determine whether the changes seen could be directly through *SOX17* effects, we measured *SOX17* levels in the treated hPAECs and found that both YK-4279 and sirolimus significantly increased *SOX17* relative expression compared with vehicle (Figure 6D). These data confirmed that Cmap-predicted compounds can successfully reverse some genetic changes associated with *SOX17* dysfunction in hPAECs.

Animal Knockout of *SOX17* Enhancer Worsens PH

We have successfully generated mice lacking the 747bp enhancer region containing *SOX17* genome-wide association study signal 1 using CRISPR/Cas9-mediated deletion on C57BL/6 background (SOX17-enhKO; Figure 7A and Figure S6). The SOX17-enhKO and wild-type (WT) mice kept in normoxia did not show significant differences in right ventricular systolic pressure and right ventricular hypertrophy index (right ventricle/left ventricle+septum; Figure S6B). We then exposed the mice to hypoxia

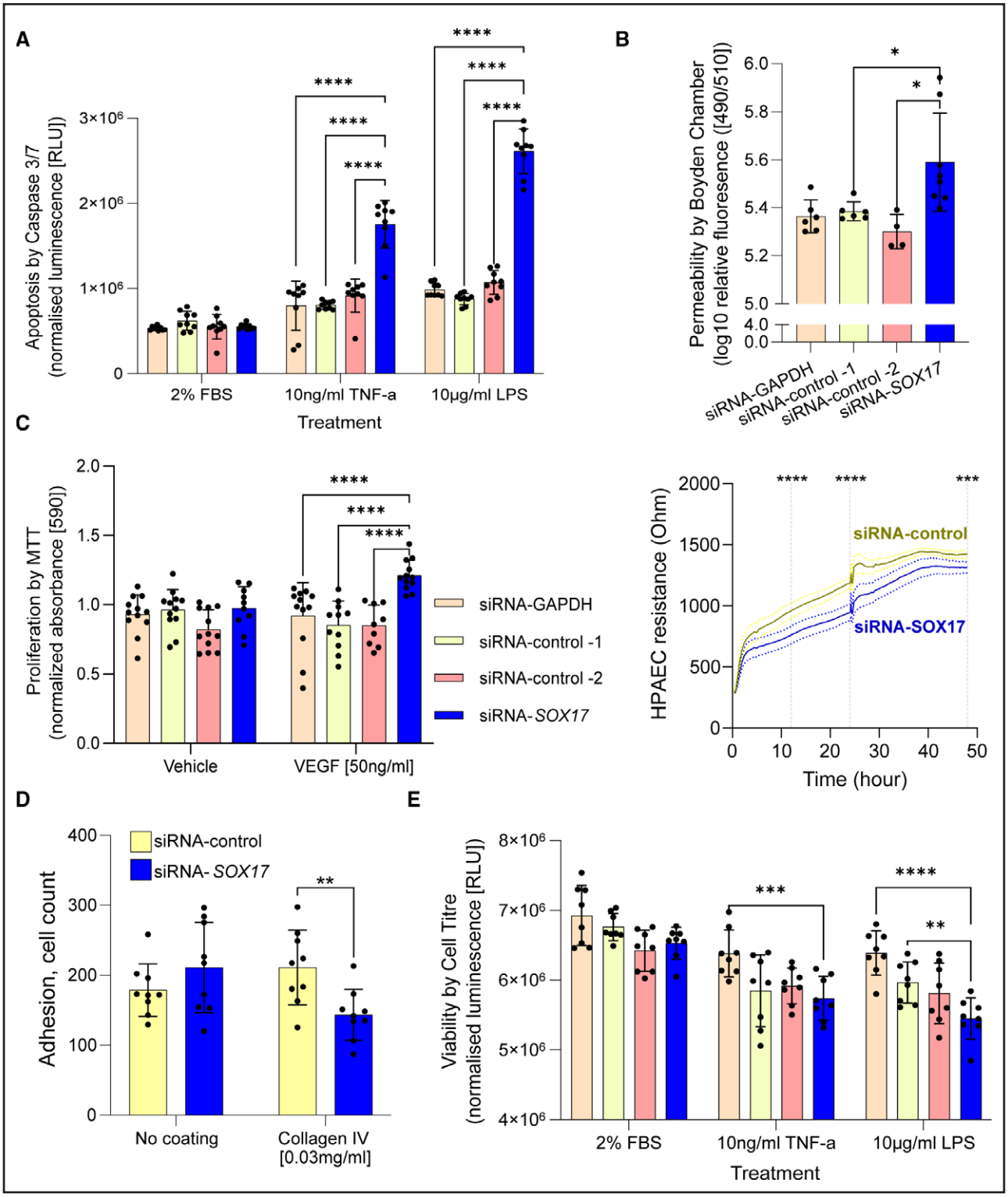

Figure 5. Functional analysis of Sox17 loss in homeostasis of pulmonary artery endothelial cells.
All after siRNA treatment for 48 hours with scrambled controls, targeting *SOX17* or unrelated gene *GAPDH*, with relevant stimuli as indicated. **A**, Caspase 3/7 apoptosis assay. FBS 2% was used as a proliferation control. Tumor necrosis factor-α (TNFα) and lipopolysaccharide (LPS) were used as proapoptotic inflammatory stimuli. **B**, Permeability barrier function assays. **Top,** Boyden chamber FITC dextran. **Bottom,** Electric cell substrate impedance sensing measurements at 0 to 48 hours. **C**, MTT proliferation assay. Vehicle, 0.1% BSA in PBS. Vascular endothelial growth factor (VEGF) was used as a proproliferative stimulus. **D**, Adhesion cell counting assay. Adhesion was compared between cells in wells with no coating and cells precoated with collagen IV. **E**, Cell titer viability assay. Performed under same conditions as caspase assay. Statistical tests used in **A**, **C**, and **D** are ordinary 2-way ANOVA with the Dunnett multiple-comparisons test. Statistical test used in **B** is an ordinary 1-way ANOVA with the Dunnett multiple-comparison test. Minimum n=3 experiments for all, culture replicates plotted. **P*<0.05. ***P*<0.01. ****P*<0.005. *****P*<0.001.

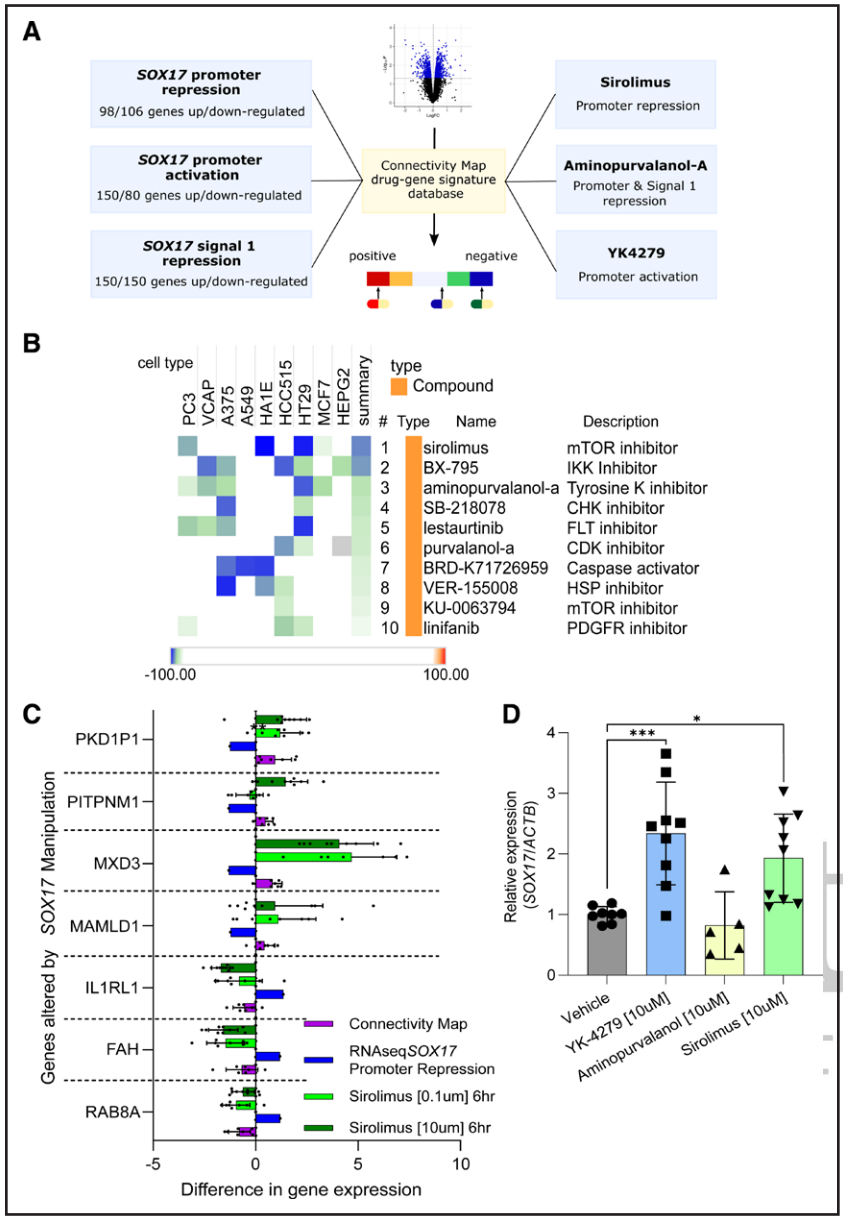

Figure 6. Repurposing of compounds to rescue loss of *SOX17* Function in pulmonary arterial hypertension. **A**, Summary diagram showing the prediction of compounds from omics signatures using connectivity map (Cmap) database signatures of drug or gene manipulation in reference cell lines. Omics signatures used to query the Cmap database are shown on the **left**. Predicted compounds are shown on the **right**. **B**, Cmap analysis results for *SOX17* promoter repression using CRISPR inhibition in human pulmonary artery endothelial cells (hPAECs). Cell types are shown. τ Scores within heat map indicate percentage of all possible compounds in Cmap, and cell lines tested the specific result is more connected than. Sirolimus overall is negatively connected more strongly than other compounds, with a summary τ of −96.94. **C**, Relative gene expression of target genes in the sirolimus perturbagen signature. Target genes were *PKD1P1*, *PITPNM1*, *MXD3*, *MAMLD1*, *IL1RL1*, *FAH*, and *RAB8A*. Cmap is perturbagen *z* score taken from the Cmap. RNA-sequencing (RNAseq) *SOX17* promoter repression through CRISPR inhibition (CRISPR-I; n=3); fold change is from RNAseq analysis of differentially expressed genes after CRISPR-I of the *SOX17* promoter (n=3). Sirolimus 0.1 μm/10 μm, n=3. the change in target gene expression is normalized to *ACTβ*, and all siRNA conditions are relative to the GAPDH-targeting siRNA control in hPAECs after sirolimus exposure at the stated concentrations. **D**, Relative expression of *SOX17* after Cmap compound exposure by quantitative polymerase chain reaction in hPAECs. Change in *SOX17* expression is normalized to *ACTβ*, and all compounds are relative to the vehicle control. Ordinary 1-way ANOVA of conditions compared with vehicle with the Dunnett multiple-comparisons test. **P*<0.05. ****P*<0.005. n=3.

(normobaric, 10% oxygen) for 1 and 3 weeks. SOX17 protein levels in the lungs of the SOX17-enhKO mice were reduced compared with WT mice when exposed to hypoxia (Figure 7B). After 1 week of hypoxia exposure, SOX17-enhKO mice demonstrated significantly increased lung vascular permeability compared with WT mice (Figure 7C). At 3 weeks, chronic hypoxia-induced PH severity was intensified in the SOX17-enhKO animals as demonstrated by increased peripheral pulmonary vessel muscularization (Figure 7D), right ventricular systolic pressure (mean±SD, 26.93±3.916 mm Hg versus 30.66±3.856 mm Hg; *P*=0.004), right ventricular hypertrophy, and right ventricle/body weight (Figure 7E and Figure S6C).

To confirm these findings with a different model of PH, the same SOX17-enhKO mice were studied under different levels of hypoxia and SUGEN (SU-5416). Using the standard protocol of 20 mg SUGEN and 3 weeks of hypoxia, we found no significant differences in the severity of PH that developed between SOX17-enhKO and WT mice. However, SOX17-enKO mice were more susceptible to developing SUGEN/hypoxia-PH as evidenced by the development of severe PH at lower levels of SUGEN-hypoxia (5 mg/kg and 12% O_2), which did not produce PH in WT littermates (right ventricular systolic pressure, 21.66±4.43 mm Hg versus 35.94±15.36 mm Hg; *P*=0.006; Figure 7F and Figure S7). Thus, in 2 independent laboratories using different PH models, it was shown that *SOX17* enhancer knockout increases susceptibility to and severity of PH.

DISCUSSION

Here we provide novel insight into how 2 independent common genetic variants upstream of the key endothelial TF *SOX17* can increase susceptibility to PH (Figure 8). In

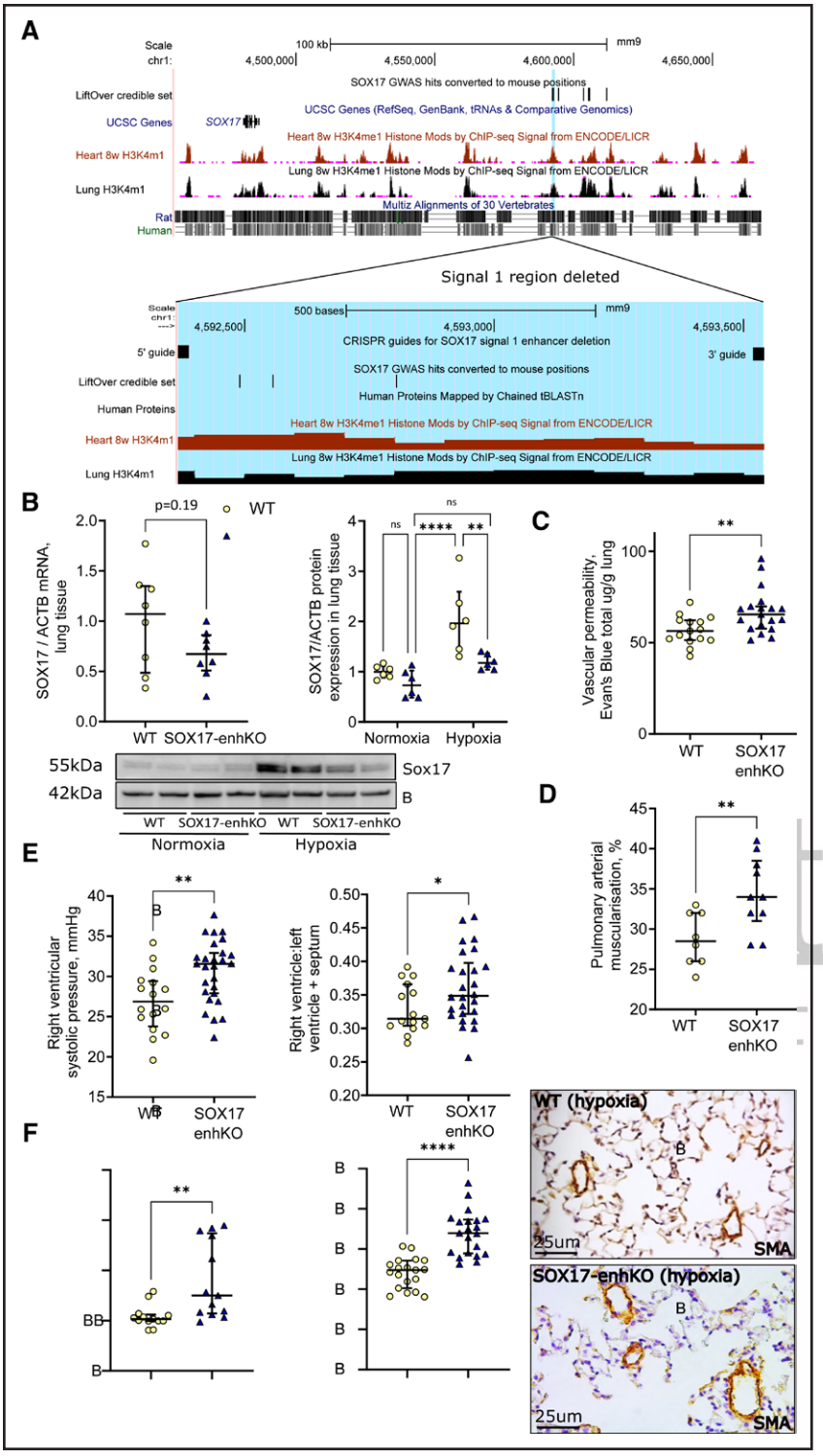

Figure 7. *SOX17*-signal 1 enhancer knockout mice develop more severe PH in hypoxia.

A, Map of pulmonary arterial hypertension (PAH) *SOX17* enhancers in mouse genome. Black bars labeled enhancer and genome-wide association study (GWAS) hits LiftOver and User track indicate conserved genomic regions from the human *SOX17* enhancer peaks. Lines indicate positions of variants associated with PAH in the human GWASs. Mouse epigenomic H3K4m1 data show that this area is also likely to be an active regulatory region in mice. Blue region highlights enhancer targeted for deletion. **B**, *SOX17* mRNA and protein expression. Taken from lung tissue after 3 weeks hypoxia and compared to β-actin (ACTB) housekeeping gene/protein. **P*<0.05, ***P*<0.01 vs wild-type (WT) mice (unpaired *t* tests). n=6 and n=8. **C**, Lung vascular permeability. Determined by Evan's blue dye in WT and *SOX17* enhancer knockout lung tissue after 1 week of hypoxia. **P*<0.05, ***P*<0.01 vs WT (unpaired *t* tests). n=15 and n=18. **D**, Pulmonary vascular muscularization. Determined from smooth muscle actin (SMA) and elastic Van Gieson staining. **P*<0.05, ***P*<0.01 vs WT (unpaired *t* tests). n=7 and n=10. **E**, Right ventricular systolic pressure (RVSP) and right ventricular hypertrophy (RVH) indices of pulmonary hypertension (PH) severity in WT and SOX17 enhancer knockout mice after chronic hypoxia (10% O_2 for 3 weeks). **P*<0.05, ***P*<0.01 vs WT (unpaired *t* tests). **F**, RVSP and RVH indices of PH severity in WT and SOX17 enhancer knockout mice after SUGEN 5 mg/kg and 12% O_2 for 3 weeks (SuHx). ChIP seq indicates chromatin immunoprecipitation sequencing. ***P*<0.01, *****P*<0.001 vs WT SuHx (unpaired *t* tests).

brief, variation at rs765727 and rs10958403 in putative *SOX17* enhancer signals 1 and 2 determines the binding of 2 TFs, ROR-α and HOXA5, respectively. Allele-specific reduced binding of either factor leads to reduced *SOX17* expression. SOX17 is crucial for maintaining EC homeostasis, and its loss drives abnormal proliferation, apoptosis, and adhesion and impairs endothelial barrier integrity. Our prediction that this would increase susceptibility to PH was confirmed in mice lacking *SOX17*-signal 1 enhancer signaling exposed to hypoxia with and without Sugen.

Defining the importance and biological function of genome-wide association study signals in complex diseases is a challenge. Many are located in noncoding regions of the genome, which complicates interpretation. Confirmed examples of variation in enhancer regions

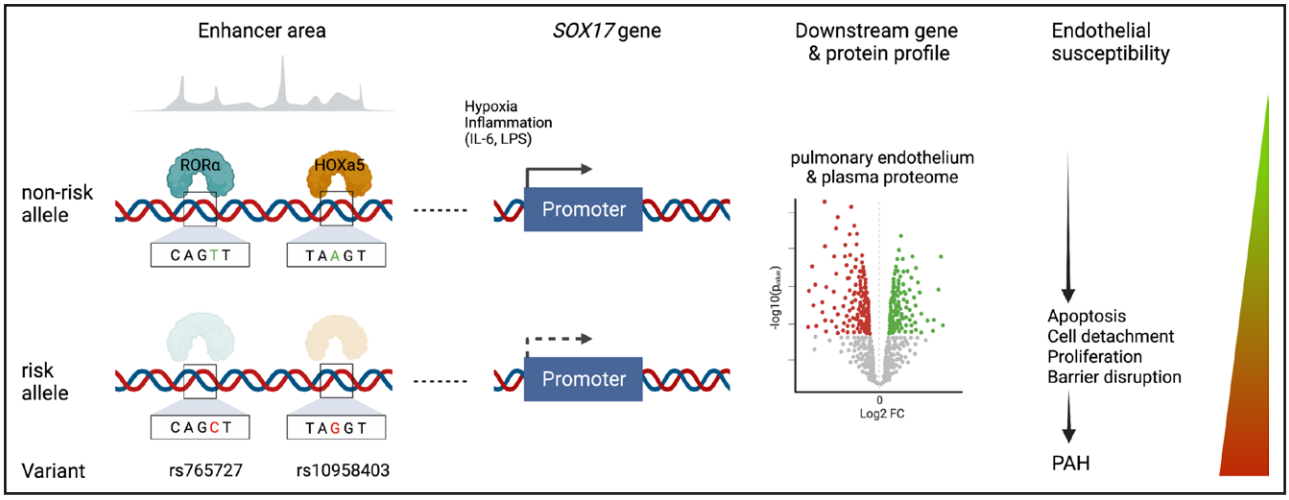

Figure 8. Summary figure.
Schematic depicting overall study findings from identification of ROR-α and HOXa5 as transcription factors binding pulmonary arterial hypertension (PAH)–associated variants in enhancers upstream of *SOX17*, through regulation of *SOX17* by PAH stimuli and downstream effects of *SOX17* on gene and protein expression profiles and endothelial cell behavior, culminating in worsened PAH in *SOX17* enhancer knockout mice.

causing disease are few, including BCL11 in sickle cell disease[25] and FTO in obesity.[26] *SOX17*-signal 1 and -signal 2 are located inside a topologically associated domain in which *SOX17* is the only gene, making it the most likely target for these 2 enhancers.[7] A significant and specific decrease in *SOX17* expression was observed after CRISPR-I of *SOX17*-signal 2 and CRISPR deletion of *SOX17*-signal 1, establishing that these genomic signals are associated with the regulation of *SOX17* expression. EMSA (and chromatin immunoprecipitation–qPCR for HOXa5) demonstrated that HOXa5 or ROR-α binds to the nonrisk alleles present at *SOX17*-signal 1 and signal 2, respectively. Thus, an individual homozygous for 1 or both risk alleles would be more resistant to HOXa5- or ROR-α–induced SOX17 expression than an individual hosting nonrisk alleles. Although it is not a requirement that these TFs are themselves associated with PAH, HOXa5 TF expression has been shown to differ in PAH lung tissue,[27] which may further exacerbate the effect of the differential binding affected by the disease-driving variant.

SOX17 has established roles in systemic artery ECs, but little is known about its role in the pulmonary vasculature. Establishing the downstream targets of *SOX17* in hPAECs is vital to understanding how it mediates the risk of developing PAH. Our transcriptomic analysis identified several specific candidates and enrichment of gene pathways, implicating dysregulation of endothelial functions and ECM associated with loss of *SOX17*. This was supported by our plasma proteomic analysis between patients with risk or nonrisk *SOX17* enhancer genotypes, in which we found an enrichment of adhesion- and ECM-associated proteins. The basement membrane is thicker in the lungs of patients with idiopathic PAH, and regulation of the ECM is established as an important factor in PAH.[28] We have demonstrated here that the ECM-linked genes *LAMB3*, *ADAMTS12*, and *MMP17* were affected at the mRNA level by *SOX17* knockdown in independent experiments. *ADAMTS12* and *MMP17* expression increased after a loss of *SOX17*, suggesting that they may be important contributors to the functional effects of *SOX17* loss in hPAECs. *ADAMTS12* is a disintegrin and matrix metalloproteinase gene with an important role in ECM composition.[29] Coupled with *MMP17*, it is part of a large family of matrix metalloproteinase genes that may have important roles in pathophysiological functions in hPAECs in PAH.[30] Although *MMP17* has not been directly linked to PAH, loss-of-function variants, found in familial studies, confer a greater risk of aortic aneurysm in mice through dysfunctional ECM filament deposition and an enlarged aortic lumen.[31]

Analysis of the effect of patient *SOX17* enhancer genotypes at *SOX17*-signal 1 and -signal 2 on plasma protein levels identified a large number of significantly affected proteins enriched in the regulation of adhesion and ECM, proliferation, migration, and apoptosis, all processes crucial for the development of PAH. There were proteins and genes that were significantly affected both by the *SOX17*-signal 1 and -signal 2 patient genotypes and by the loss of *SOX17* expression via siRNA or CRISPR-I in hPAEC, suggesting circulating proteins might reflect *SOX17* dysfunction in PAH. Interesting candidates among these are SPARC, which was found at increased levels in lung of patients with idiopathic PAH and is involved in the regulation of pulmonary artery smooth muscle cell proliferation[21]; ECE1, which is involved in the cleavage/transformation of endothelin 1 to its active form[22]; PECAM-1, which plays a role in the adaptation of endothelium to shear stress[23]; and COL18A1, which when cleaved produces endostatin, a

protein with serum levels that are correlated with disease severity and survival in PAH.[24]

The fact that rare pathogenic variants in *SOX17* also drive PAH development and are associated with more severe PAH and younger age[32] emphasizes the importance of this gene and related pathways as a therapeutic target. To pursue this, we explored Cmap for novel candidates that might rescue *SOX17* activity. Three candidate compounds emerged from our screen using hPAECs, all suitable for exploratory studies in humans. Sirolimus is a specific inhibitor of mammalian target of rapamycin and an allosteric inhibitor of mammalian target of rapamycin complex 1. Another immunosuppressor, tacrolimus, has entered clinical trials in PAH. Tacrolimus is a calcineurin inhibitor but was selected for study at low doses expected to modulate *BMPR2* signaling on the basis of a screening strategy and hence has a mechanism distinct from sirolimus. Tacrolimus was found to be well tolerated and to improve 6-minute walk distance and echocardiographic parameters of heart failure in patients with PAH in a phase IIa safety and tolerability study. Although the small number of patients studied did not result in statistical significance, the findings did support the study of tacrolimus in a phase IIb efficacy trial.[33] Several prior studies have shown the efficacy of sirolimus in reversing animal models of PH,[34] and an albumin-bound nanoparticle form of sirolimus is currently undergoing a phase I/II clinical trial and has so far shown no safety concerns and an early efficacy signal.[35] It would be of interest to gauge the importance of regulation of *SOX17* and its signaling in the efficacy of sirolimus, and the addition of biomarker measurements (eg, SPARC) in future trials would be valuable.

Aminopurvalanol-a (also known as purvalanol-a) is a CDK1/cyclin B inhibitor that arrests proliferating cells in the G2/M stage of the cell cycle and prevents proliferation. It has been investigated in human microvascular ECs as an antiangiogenic drug and was shown to inhibit proliferation, to increase apoptosis, and to prevent tube formation.[36] Some CDKs are upregulated after *SOX17* silencing in arterial endothelium.[37] In both monocrotaline and Sugen/hypoxia PAH rat models, the CDK inhibitor palbociclib reversed PAH pathology, including right-sided heart hypertrophy and pulmonary remodeling.[38] Both CDK inhibitors dinaciclib (inhibits CDK1, 2, 5, and 9) and palbociclib (inhibits CDK4 and 6) reduced proliferation in smooth muscle cells.[38]

A third compound from the Cmap screen, YK-4279, is an ETS family inhibitor that has undergone preclinical efficiency trials as an antilymphoma drug.[39] ETS family members are known oncogenes with an aberrant expression is found in many solid tumours.[40] They also have roles in vascular development and maintenance (Craig and Sumanas[41] provide a full review). YK-4279 specifically inhibits ERG transcription and ERG-mediated cell migration and proliferation in prostate cancer.[42] The ERG TF is essential for EC homeostasis and has recently been shown to bind to a superenhancer upstream of the *SOX17* gene in human umbilical vein ECs, suggesting a possible role for ERG in the regulation of *SOX17* in this cell type.[43]

SOX17 knockdown in mice results in embryonic lethality with heart defects and enlarged veins.[44] Endothelium-specific knockdown in either embryonic or adult mice causes defects in artery specification and in utero lethality.[10,45] SOX17 endothelial-(Cdh-CreER)-knockout exacerbated hypoxic PH, which was sustained despite a return to normoxia for 3 weeks. Consistent with our findings in si-SOX17–treated human pulmonary artery ECs, hyperproliferation of ECs was prominent in Sox17knockout/hypoxic mice by Ki67 staining.[46] These studies support a role for *SOX17* in arterial ECs in vivo. We can report that mice lacking *SOX17*-signal 1 are viable but show increased susceptibility to hypoxia- or combined SUGEN-hypoxia–associated PH. Consistent with our in vitro data and prior reports in other vascular beds,[47,48] the *SOX17* enhancer knockout mice exhibited elevated vascular permeability compared with WT animals. This illustrates the importance of fine-tuning *SOX17* levels to modulate endothelial barrier function under pathological conditions. Although no phenotype was apparent in the mice under normoxic conditions or more severe SUGEN-hypoxia, it is remarkable that deletion of an enhancer alone, rather than a complete or partial gene deletion (as required for *BMPR2*), was sufficient to lead to a worse PH phenotype in 2 independent laboratories, under hypoxia or lower levels of SUGEN-hypoxia than drove PH in WT animals. The concept that a second hit may be required to exhibit the PH phenotype is well understood, and in patients harboring a *SOX17* enhancer risk allele, this may include inflammation or drug toxicity as well as hypoxia.

It remains possible that other TFs also bind differentially to PAH-associated *SOX17* enhancer variants. In our EMSA experiments, use of an ROR-α antibody led to loss of signal rather than a clear supershift, which is most likely due to the antibody-preventing formation of the ROR-α probe complex. SOX17 is not included in the proteomics platform used in this study. The numbers of homozygotes for rarer variant alleles are small, so variability in these measurements is high. Validation of key findings in hPAECs with (naturally occurring or knocked-in) variants in *SOX17*-signal 1 or 2, in addition to the complete deletion of the enhancer, would have even more robustly supported functional conclusions. The effects would likely be more subtle than deletion or inhibition of the enhancer requiring larger numbers. Aminopurvalanol-a did not affect SOX17 expression but was identified as a drug that can alter the downstream transcriptomic signature produced by SOX17 loss.

CONCLUSIONS

We provide comprehensive insight into how common variation influences the binding of HOXa5 and ROR-α to enhancers upstream of *SOX17* and can reduce susceptibility to PH. Loss of *SOX17* leads to downstream alterations in ECM regulation and hPAEC function. *SOX17* is a priority for therapeutic rescue, and predicted compounds that restore endothelial gene expression offer candidates for future investigation.

ARTICLE INFORMATION

Received August 4, 2022; accepted March 15, 2023.

Affiliations

National Heart and Lung Institute, Hammersmith Hospital, Imperial College, London, United Kingdom (R.W., E.V., J.A., C.-N.C., Y.W., A.A., O.D., L.Z., F.S., M.R.W., L.Z., C.J.R.). Department of Pulmonary Medicine, Amsterdam University Medical Center, The Netherlands (J.A.). Division of Hematology/Oncology, Department of Medicine (O.D.L.) and Division of Pulmonary, Sleep and Critical Care Medicine, Department of Medicine (J.R.K.), Rhode Island Hospital and Warren Alpert Medical School of Brown University, Providence. Section of Genetics & Genomics, Department of Metabolism, Digestion & Reproduction, Hammersmith Hospital, Imperial College, London, United Kingdom (I.C., J.F.). Bioinformatics and Genomics Program, Centre for Genomic Regulation, Barcelona Institute of Science and Technology, Spain (J.F.). Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas, Barcelona, Spain (J.F.). Department of Medicine, University of Cambridge, United Kingdom (N.W.M.). NIHR BioResource for Translational Research, University of Cambridge, United Kingdom (N.W.M.). On Behalf of the British Heart Foundation/Medical Research Council UK PAH Cohort Consortium (N.W.M., M.R.W., C.J.R.).

Acknowledgments

The authors thank the National Institute for Health Research (NIHR) BioResource volunteers for their participation and gratefully acknowledge NIHR BioResource centers, NHS Trusts, and staff for their contribution. They thank the NIHR Imperial Clinical Research Facility, NIHR Sheffield Biomedical Research Centre, Cambridge NIHR Cardiorespiratory BRC, and NHS Blood and Transplant. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. The authors thank the members of the UK PAH Cohort Consortium listed in full in the Supplemental Material.

Sources of Funding

This work was supported by the National Institute for Health and Care Research BioResource, which supports the UK National Cohort of Idiopathic and Heritable PAH; the British Heart Foundation (BHF PG/19/17/34275, SP/12/12/29836), and the UK Medical Research Council (MR/K020919/1). This work was supported in part by a British Heart Foundation Centre for Research Excellence award to Dr Wilkins (RE/18/4/34215). Dr Rhodes is supported by a BHF Intermediate Basic Science Research fellowship (FS/15/59/31839) and Academy of Medical Sciences Springboard fellowship (SBF004\1095). Dr Morrell is a BHF Professor and NIHR Senior Investigator. Dr Aman is supported by an European Respiratory Society/European Molecular Biology Organisation Longterm Research Fellowship and by an Nederlandse Organisatie voor Wetenschappelijk Onderzoek VENI grant (09150161910155). Dr Ferrer is supported by Medical Research Council (MR/L02036X/1), Wellcome Trust Senior Investigator Award (WT101033), and European Research Council Advanced Grant (789055) awards. This work was also supported in part by grants from the National Institutes of Health P20 GM119943 (Drs Liang and Klinger), by the American Heart Association, Transformational Project Award 18TPA34110329 (Dr Liang), and by the Brown Physicians, Inc, Academic Assessment Research Award (Drs Klinger and Liang).

Disclosures

None.

Supplemental Material

Expanded Methods
Figures S1–S7
Tables S1–S8
References 49–54
UK National PAH Cohort Study Consortium contributors list

REFERENCES

1. Nakhleh MK, Haick H, Humbert M, Cohen-Kaminsky S. Volatolomics of breath as an emerging frontier in pulmonary arterial hypertension. *Eur Respir J*. 2017;49:1601897. doi: 10.1183/13993003.01897-2016
2. Thenappan T, Shah SJ, Rich S, Gomberg-Maitland M. A USA-based registry for pulmonary arterial hypertension: 1982-2006. *Eur Respir J*. 2007;30:1103–1110. doi: 10.1183/09031936.00042107
3. Schermuly RT, Ghofrani HA, Wilkins MR, Grimminger F. Mechanisms of disease: pulmonary arterial hypertension. *Nat Rev Cardiol*. 2011;8:443–455. doi: 10.1038/nrcardio.2011.87
4. Lan N, Massam B, Kulkarni S, Lang C. Pulmonary arterial hypertension: pathophysiology and treatment. *Diseases*. 2018;6:3838. doi: 10.3390/diseases6020038
5. Vaillancourt M, Ruffenach G, Meloche J, Bonnet S. Adaptation and remodelling of the pulmonary circulation in pulmonary hypertension. *Can J Cardiol*. 2015;31:407–415. doi:10.1016/j.cjca.2014.10.023
6. Southgate L, Machado RD, Gräf S, Morrell NW. Molecular genetic framework underlying pulmonary arterial hypertension. *Nat Rev Cardiol*. 2020;17:85–95. doi:10.1038/s41569-019-0242-x
7. Rhodes CJ, Batai K, Bleda M, Haimel M, Southgate L, Germain M, Pauciulo MW, Hadinnapola C, Aman J, Girerd B, et al; UK NIHR BioResource Rare Diseases Consortium. Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir Med*. 2019;7:227–238. doi: 10.1016/S2213-2600(18)30409-0
8. Kamachi Y, Kondoh H. Sox proteins: regulators of cell fate specification and differentiation. *Development*. 2013;140:4129–4144. doi: 10.1242/dev.091793
9. 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. doi: 10.1016/j.biocel.2009.08.017
10. Corada M, Orsenigo F, Morini MF, Pitulescu ME, Bhat G, Nyqvist D, Breviario F, Conti V, Briot A, Iruela-Arispe ML, et al. Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nat Commun*. 2013;4:2609. doi: 10.1038/ncomms3609
11. Zhou Y, Williams J, Smallwood PM, Nathans J. Sox7, Sox17, and Sox18 cooperatively regulate vascular development in the mouse retina*PLoS One*. 2015;10:e0143650. doi: 10.1371/journal.pone.0143650
12. Lange AW, Haitchi HM, LeCras TD, Sridharan A, Xu Y, Wert SE, James J, Udell N, Thurner PJ, Whitsett JA. Sox17 is required for normal pulmonary vascular morphogenesis. *Dev Biol*. 2014;387:109–120. doi: 10.1016/j.ydbio.2013.11.018
13. Galiè N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, Simonneau G, Peacock A, Vonk Noordegraaf A, Beghetti M, et al. 2015 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J*. 2015;46:903–975. doi: 10.1183/13993003.01032-2015
14. Wojciak-Stothard B, Abdul-Salam VB, Lao KH, Tsang H, Irwin DC, Lisk C, Loomis Z, Stenmark KR, Edwards JC, Yuspa SH, et al. Aberrant chloride intracellular channel 4 expression contributes to endothelial dysfunction in pulmonary arterial hypertension. *Circulation*. 2014;129:1770–1780. doi: 10.1161/circulationaha.113.006797
15. Cebola I. Deletion of regulatory elements with all-in-one CRISPR-Cas9 vectors. *Methods Mol Biol*. 2021;2351:321–334. doi: 10.1007/978-1-0716-1597-3_18/COVER
16. Rhodes CJ, Wharton J, Swietlik EM, Harbaum L, Girerd B, Coghlan JG, Lordan J, Church C, Pepke-Zaba J, Toshner M, et al. Using the plasma proteome for risk stratifying patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2022;205:1102–1111. doi: 10.1164/rccm.202105-1118OC
17. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell*. 2017;171:1437–1452.e17. doi: 10.1016/j.cell.2017.10.049
18. Tsang H, Leiper J, Lao KH, Dowsett L, Delahaye MW, Barnes G, Wharton J, Howard L, Iannone L, Lang NN, et al. Role of asymmetric methylarginine and connexin 43 in the regulation of pulmonary endothelial function. *Pulm Circ*. 2013;3:675–691. doi: 10.1086/674440
19. Liu M, Zhang L, Marsboom G, Jambusaria A, Xiong S, Toth PT, Benevolenskaya EV, Rehman J, Malik AB. Sox17 is required for endothelial regeneration following inflammation-induced vascular injury. *Nat Commun*. 2019;10:2126. doi: 10.1038/s41467-019-10134-y
20. Long L, Ormiston ML, Yang X, Southwood M, Gräf S, Machado RD, Mueller M, Kinzel B, Yung LM, Wilkinson JM, et al. Selective enhancement of endothelial BMPR-II with BMP9 reverses pulmonary arterial hypertension. *Nat Med*. 2015;21:777–785. doi: 10.1038/nm.3877

21. Veith C, Vartürk-Özcan I, Wujak M, Hadzic S, Wu CY, Knoepp F, Kraut S, Petrovic A, Gredic M, Pak O, et al. SPARC, a novel regulator of vascular cell function in pulmonary hypertension. *Circulation*. 2022;145:916–933. doi: 10.1161/CIRCULATIONAHA.121.057001
22. Galié N, Manes A, Branzi A. The endothelin system in pulmonary arterial hypertension. *Cardiovasc Res*. 2004;61:227–237. doi: 10.1016/j.cardiores.2003.11.026
23. Szulcek R, Happe CM, Rol N, Fontijn RD, Dickhoff C, Hartemink KJ, Grünberg K, Tu L, Timens W, Nossent GD, et al. Delayed microvascular shear adaptation in pulmonary arterial hypertension: Role of platelet endothelial cell adhesion molecule-1 cleavage. *Am J Respir Crit Care Med*. 2016;193:1410–1420. doi: 10.1164/rccm.201506-1231OC
24. Damico R, Kolb TM, Valera L, Wang L, Housten T, Tedford RJ, Kass DA, Rafaels N, Gao L, Barnes KC, et al. Serum endostatin is a genetically determined predictor of survival in pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2015;191:208–218. doi: 10.1164/rccm.201409-1742OC
25. Xu J, Peng C, Sankaran VG, Shao Z, Esrick EB, Chong BG, Ippolito GC, Fujiwara Y, Ebert BL, Tucker PW, et al. Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. *Science*. 2011;334:993. doi: 10.1126/SCIENCE.1211053
26. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*. 2014;507:371–375. doi: 10.1038/nature13138
27. Golpon HA, Geraci MW, Moore MD, Miller HL, Miller GJ, Tuder RM, Voelkel NF. HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. *Am J Pathol*. 2001;158:955–966. doi: 10.1016/s0002-9440(10)64042-4
28. Jandl K, Marsh LM, Hoffmann J, Mutgan AC, Baum O, Bloch W, Thekkekara-Puthenparampil H, Kolb D, Sinn K, Klepetko W, et al. Basement membrane remodeling controls endothelial function in idiopathic pulmonary arterial hypertension. *Am J Respir Cell Mol Biol*. 2020;63:104–117. doi: 10.1165/rcmb.2019-0303OC
29. Mohamedi Y, Fontanil T, Cal S, Cobo T, Obaya AJ. ADAMTS-12: functions and challenges for a complex metalloprotease. *Front Mol Biosci*. 2021;8:378. doi: 10.3389/FMOLB.2021.686763/BIBTEX
30. Chelladurai P, Seeger W, Pullamsetti SS. Matrix metalloproteinases and their inhibitors in pulmonary hypertension. *Eur Respir J*. 2012;40:766–782. doi: 10.1183/09031936.00209911
31. Alonso M, García-Redondo AB, Guo D, Camafeita E, Martínez F, Alfranca A, Méndez-Barbero N, Pollán Á, Sánchez-Camacho C, Denhardt DT, et al. Deficiency of MMP17/MT4-MMP proteolytic activity predisposes to aortic aneurysm in mice. *Circ Res*. 2015;117:e13–e26. doi: 10.1161/CIRCRESAHA.117.305108
32. Zhu N, Welch CL, Wang J, Allen PM, Gonzaga-Jauregui C, Ma L, King AK, Krishnan U, Rosenzweig EB, Ivy DD, et al. Rare variants in SOX17 are associated with pulmonary arterial hypertension with congenital heart disease. *Genome Med*. 2018;10:56. doi: 10.1186/s13073-018-0566-x
33. Spiekerkoetter E, Sung YK, Sudheendra D, Scott V, del Rosario P, Bill M, Haddad F, Long-Boyle J, Hedlin H, Zamanian RT. Randomised placebo-controlled safety and tolerability trial of FK506 (tacrolimus) for pulmonary arterial hypertension. *Eur Respir J*. 2017;50:1602449. doi: 10.1183/13993003.02449-2016
34. Paddenberg R, Stieger P, von Lilien AL, Faulhammer P, Goldenberg A, Tillmanns HH, Kummer W, Braun-Dullaeus RC. Rapamycin attenuates hypoxia-induced pulmonary vascular remodeling and right ventricular hypertrophy in mice. *Respir Res*. 2007;8:1–12. doi: 10.1186/1465-9921-8-15/FIGURES/6_553
35. Simon M, Gomberg-Maitland M, Oudiz RJ, Machado RF, Rischard FP, Elinoff JM, et al. ABI-009, nab-Sirolimus, an mTOR inhibitor with high lung accumulation in preclinical models is active in patients with severe pulmonary arterial hypertension. *American Thoracic Society International Conference Meetings Abstracts*. 2019;A4409. doi: 10.1164/AJRCCM-CONFERENCE.2019.199.1_MEETINGABSTRACTS.A4409
36. Zahler S, Liebl J, Fürst R, Vollmar AM. Anti-angiogenic potential of small molecular inhibitors of cyclin dependent kinases in vitro. *Angiogenesis*. 2010;13:239–249. doi: 10.1007/s10456-010-9181-1
37. Lee S, Kim IKI, Ahn JS, Woo DC, Kim ST, Song S, Koh GY, Kim H-S, Jeon BH, Kim I. Deficiency of endothelium-specific transcription factor Sox17 induces intracranial aneurysm. *Circulation*. 2015;131:995–1005. doi: 10.1161/CIRCULATIONAHA.114.012568
38. Weiss A, Neubauer MC, Yerabolu D, Kojonazarov B, Schlueter BC, Neubert L, Jonigk D, Baal N, Ruppert C, Dorfmuller P, et al. Targeting cyclin-dependent kinases for the treatment of pulmonary arterial hypertension. *Nat Commun*. 2019;10:1–17. doi: 10.1038/s41467-019-10135-x
39. Lamhamedi SE, Menegaz BA, Ramamoorthy V, Aiyer RA, Maywald RL, Buford AS, Doolittle DK, Culotta KS, O'Dorisio JE, Ludwig JA. An oral formulation of YK-4-279: preclinical efficacy and acquired resistance patterns in Ewing sarcoma. *Mol Cancer Ther*. 2015;14:1591–1604. doi: 10.1158/1535-7163.MCT-14-0334
40. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. *Nat Rev Cancer*. 2017;17:337–351. doi: 10.1038/nrc.2017.20
41. Craig MP, Sumanas S. ETS transcription factors in embryonic vascular development. *Angiogenesis*. 2016;19:275–285. doi: 10.1007/s10456-016-9511-z
42. Wei G, Srinivasan R, Cantemir-Stone CZ, Sharma SM, Santhanam R, Weinstein M, Muthusamy N, Man AK, Oshima RG, Leone G, et al. Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood*. 2009;114:1123–1130. doi: 10.1182/blood-2009-03-211391
43. Kalna V, Yang Y, Peghaire CR, Frudd K, Hannah R, Shah AV, Almagro LO, Boyle JJ, Göttgens B, Ferrer J, et al. The transcription factor ERG regulates super-enhancers associated with an endothelial-specific gene expression program. *Circ Res*. 2019;124:1337–1349. doi: 10.1161/CIRCRESAHA.118.313788
44. Azuma M, Kanai Y, Gad JM, Tajima Y, Taya C, Kurohmaru M, Sanai Y, Yonekawa H, Yazaki K, Tam PPL, et al. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development*. 2002;129:2367–2379. doi: 10.1242/dev.129.10.2367
45. Clarke RL, Yzaguirre AD, Yashiro-Ohtani Y, Bohnsack BL, Blanpain C, Pear WS, Speck NA, Keller G. The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol*. 2013;15:502–510. doi: 10.1038/ncb2724
46. Soon Park C, Hyun Kim S, Young Yang H, Kim JH, Theo Schermuly R, Seul Cho Y, Kang H, Park JH, Lee E, Park H, et al. *Sox17* deficiency promotes pulmonary arterial hypertension via HGF (hepatocyte grow factor)/c-Met signaling. *Circ Res*. 2022;131:792–806. doi: 10.1161/CIRCRESAHA.122.320845
47. Liu M, Zhang L, Marsboom G, Jambusaria A, Xiong S, Toth PT, Benevolenskaya EV, Rehman J, Malik AB. Sox17 is required for endothelial regeneration following inflammation-induced vascular injury. *Nat Commun*. 2019;10:2126. doi: 10.1038/s41467-019-10134-y
48. Corada M, Orsenigo F, Bhat GP, Conze LL, Breviario F, Cunha SI, Claesson-Welsh L, Beznoussenko GV, Mironov AA, Bacigaluppi M, et al. Fine-tuning of Sox17 and canonical Wnt coordinates the permeability properties of the blood-brain barrier. *Circ Res*. 2019;124:511–525. doi: 10.1161/circresaha.118.313316
49. Beucher A, Cebola I. One-step dual CRISPR/Cas9 guide RNA cloning protocol. *Research Square*. Preprint posted online June 26, 2019. doi:10.21203/RS.2.1831/V1
50. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14:417–419. doi: 10.1038/nmeth.4197
51. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*. 2016;4:1521. doi: 10.12688/f1000research.7563.2
52. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139–140. doi: 10.1093/bioinformatics/btp616
53. Blighe K, Sharmila R, Myles L. EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. 2018. Accessed xxxx. https://bioconductor.org/packages/devel/bioc/vignettes/EnhancedVolcano/inst/doc/EnhancedVolcano.html
54. Yates B, Braschi B, Gray KA, Seal RL, Tweedie S, Bruford EA. Genenames.org: the HGNC and VGNC resources in 2017. *Nucleic Acids Res*. 2017;45:D619–D625. doi: 10.1093/nar/gkw1033