Science Translational Medicine
AAAS

Supplementary Materials for

Allele-specific control of rodent and human lncRNA KMT2E-AS1 promotes hypoxic endothelial pathology in pulmonary hypertension

Yi-Yin Tai *et al.*

Corresponding author: Stephen Y. Chan, chansy@pitt.edu

Sci. Transl. Med. **16**, eadd2029 (2024)
DOI: 10.1126/scitranslmed.add2029

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

SUPPLEMENTAL MATERIALS AND METHODS

Cell culture and reagents

Primary human (Lonza, #CC-2530) or C57BL/6 mouse (Cell Biologics, #C57-6059) pulmonary arterial endothelial cells (PAECs) were grown in EBM-2 basal medium supplemented with EGM-2 MV BulletKit (Lonza). Human pulmonary artery adventitial fibroblasts (PAAFs, ScienCell, #3120) and smooth muscle cells (PASMCs, Lonza, #CC-2581) were grown in FBM-2 basal medium supplemented with FGM-2 BulletKit and SmBM basal medium supplemented with SmGM BulletKit (Lonza), respectively. Experiments were performed at passages 5 to 8. HEK293FT (Thermo Fisher Scientific, #R70007) cells were cultivated in DMEM containing 10% fetal bovine serum (FBS). Mouse embryonic fibroblasts (MEF, ATCC, #PTA-9386) were grown in DMEM containing 10% FBS, 20 mmol/L glutamine and 10 mmol/L sodium pyruvate. Recombinant human IL-1β (10 ng/mL) and IL-6 (100 ng/mL) used were from PeproTech. Actinomycin D (Sigma-Aldrich) was used at 1μg/mL for 4h. All human and mouse cell lines were male.

Exposure to hypoxia

As we have described (*32*), cultured cells were exposed to either normoxic conditions with 20% O_2 at 37 ˚C or to hypoxic conditions with 0.2% O_2 in a hypoxia incubator chamber (Stemcell Technologies, #27310) for 24 hours before harvesting for RNA or protein.

Oligonucleotides and transfection

Silencer or Silencer select siRNAs for *KMT2E-AS1* (n263822), KMT2E (132615), and scrambled control (12935300) were purchased from Thermo Fisher Scientific. siRNAs for HIF-1α (sc-35561), HIF-2α (sc-35316) and scrambled control (sc-37007) were purchased from Santa Cruz Biotechnology. PAECs were plated in collagen-coated plastic and transfected 16h later at 70-80% confluence using siRNA (40nM) or gapmers (5nM) and Lipofectamine 2000 reagent (Thermo Fisher Scientific), according to the manufacturer's instructions. 48 hours post-transfection, cells were lysed for RNA or protein. Antisense oligonucleotides gapmers were designed and provided by Sanofi

(sequences are listed in the **Table S9. Key Resource Table**).

RNA-Sequencing

RNA profiles of whole lung from mice with hypoxia+SU5416-induced PH vs. control were generated by deep sequencing, in triplicate, using Illumina HiSeq 2000. Briefly, 150ng of total RNA input was used by Illumina's TruSeq RNA Sample Preparation v2 kit to construct libraries according to the manufacturer protocol. The libraries were quantified using KAPA library quantification kit. Unstranded, paired-end sequencing was then done on Illumina HiSeq 2000 at 50 cycles/base-pair to generate 50bp paired-end reads. Around 14-19 million reads per sample were generated. Experimental data and associated designs were deposited in the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under series GSE61828 (*17*).

In a separate study, cultured HPAECs were transfected with siRNAs targeting either *KMT2E-AS1* or KMT2E (si-NC serves as negative control), and then exposed to hypoxic stress (0.2% oxygen) for 24 hours. Cells were harvested for total RNA extraction, and all total RNA samples received ribosomal depletion and directional (stranded) RNA library prep. mRNA and lncRNA profiles were generated by deep sequencing, using Illumina NextSeq 2000 P3 (2 x 101bp, 40M reads/sample). The raw fastq reads were mapped against the human genome build hg19 using STAR aligner (*48*), and the read counts for each gene were calculated using featureCounts. Gene expression was normalized by DESeq2 with variance-stabilizing transformation (VST) (*49*) and the low expression genes with total counts across all samples less than 10 were excluded. The differential gene expression analysis was performed using DESeq2 with the adjusted p-value cutoff of 0.05. The Benjamin-Hochberg method was used for multiple test correction. GSEA was conducted on differentially expressed genes with the enrichR R package (*50*) using the "GO_Biological_Process_2021", "Reactome_2022", and "KEGG_2021_Human" databases.

Chromatin immunoprecipitation sequencing (ChIP-Seq)

Chromatin Immunoprecipitation (ChIP) was performed as described below. Human pulmonary artery

endothelial cells (HPAECs) cultured in 15-cm-dishes under normoxia or hypoxia conditions were directly fixed with 1% PFA at room temperature for 10 min. Cross-linking was stopped by 125 mM glycine. Cells were harvested and lysed in Farnham Lysis Buffer (500 mM PIPES pH 8, 1M KCl, 0.5% NP-40, 1X protease inhibitor cocktail). Chromatin was then sonicated with a Bioruptor Pico Sonication Device (Diagenode) to obtain chromatin fragments of 200-500 bp. 10% of the fragmented chromatin was used as the INPUT control. Chromatin was precipitated with Protein G Dynabeads (Invitrogen, #10004D) and anti-H3K4me3 antibody (Active Motif, #39915). DNA was eluted from captured chromatin and extracted by ChIP DNA Clean & Concentrator kit (Zymo, #D5201). Library preparation and sequencing of purified DNA were performed by BGI (Shenzhen, China) using DNBseq^{TM} ChIP-seq (single end 50bp, 20M reads per sample). All samples were aligned to the hg19 build of the Human genome with STAR (*48*). Using the input sample as control, normoxia and hypoxia samples (N = 3 per group) were analyzed by MACS, as above, to find peaks (*51*). Differential peaks were identified using DiffBind (*52*). Peaks were visualized with the Integrative Genomics Viewer.

Plasmid construction and lentivirus production

Full-length *KMT2E-AS1* was PCR cloned (Aldevron) using primers AS1_Lenti_F AGATTCTAGAGGTACGCCCGTCC, AS1_Lenti_R CCTTTCTCCCCGCAAAGTATACTTAT, while full length *5031425E22Rik* was PCR cloned (Aldevron) using primers E22_Lenti_F ACTCCGTCTTTCCCACAACA, E22_Lenti_R TGGGCAGGAGTATGAGTTCC. *KMT2E-AS1* and *5031425E22Rik* were then sub-cloned into the pCDH-CMV-MCS-EF1-copGFP (System Biosciences) using XbaI/BstZ17I and NheI/BsrGI-HF restriction sites respectively. The constitutively active HIF-1α plasmid and HIF-2α plasmid (*20*) were kindly gifted from W. G. Kaelin (Harvard Medical School) and W. Kim (University of North Carolina, Chapel Hill). HIF-1α and HIF-2α genes were sub-cloned into the pCDH-CMV-MCS-EF1-copGFP (System Biosciences). All cloned plasmids were confirmed by DNA sequencing at the Genomics Research Core, University of Pittsburgh. Lentiviral vectors were transfected into HEK293FT cells with Lipofectamine 2000 (Thermo Fisher Scientific) and packing packaging plasmids (pPACK, System Biosciences), according to manufacturer's instructions. Stable

expression of lentiviral constructs as compared to GFP lentiviral vector control in PAECs was achieved by lentiviral transduction for 2-3 days; GFP signal was used for normalizing titers across lentiviruses. Infection efficacy was evaluated by GFP fluorescence signal under a fluorescence microscope (EVOS), accompanied by subsequent RNA and protein expression measurement.

Deletion analysis

A site-directed mutagenesis kit (New England Biolabs, E0554S) was used to introduce deletion of 1150-1750bp to full-length *KMT2E-AS1* and deletion of 950-1500bp to *5031425E22Rik* with mutagenic primers (Q5_AS1 Del_1150-1750_F: TCAAGGAGGATCGATTACTTTCTG, Q5_AS1 Del_1150-1750_R: TGGGTTCCATATCGGCCG, Q5_E22 Del_950-1500_F: ACTCCTTTCGGGGTGTGC and Q5_E22 Del_950-1500_R: CGCCCGACAGAAATCACC), following manufacturer's instructions (**Fig. S8**). PCR products were ligated using enzyme mix and reaction buffers and transformed into NEB 5-alpha competent cells (all provided by the mutagenesis kit, New England Biolabs). Deleted versions of *KMT2E-AS1* and *5031425E22Rik* were sub-cloned into the pCDH-CMV-MCS-EF1-copGFP (System Biosciences) using XbaI/BstZ17I and NheI/BsrGI-HF restriction sites, respectively. All cloned plasmids were confirmed by DNA sequencing.

KMT2E promoter luciferase reporter assay

To test whether the conserved sequence of the mouse lncRNA *E22* regulates the *cis* activity of the neighboring KMT2E promoter, we cloned this sequence and adjacent promoter (2500 bp) according to their natural genomic position upstream to a *NanoLuc* luciferase reporter gene. Specifically, we generated a full-length 2557 bp sequence and a truncated version carrying a large deletion (551 bp) of the conserved sequence of the lncRNA *E22* gene. The luciferase reporter gene in the pCDH lenti-viral backbone was obtained from pCDH-CMV-Nluc (Addgene #73038). This was digested by SnaBI/XbaI to remove 293 bp of the CMV promoter. We then cloned a small double-stranded fragment including a 87 bp truncated mouse KMTE2E promoter with SnaBI/XbaI cohesive ends at N- and C-terminus, respectively, to the same restriction enzyme sites in the above lentivirus vector. Such a truncated KMTE2E promoter fragment with cohesive ends was composed of two synthesized

oligonucleotides (5'-gtaGTCCCCTCCCCCCTGCCCCTGCGGAGCGGGGCGGGCGGGCCCCGTGCGCTTGCCGCCG CCGCCGCCGCCGCTGCCGCCGCCGCCGCCt and 5'-ctagaGGCGGCGGCGGCGGCAGCGGCGGCGGCGGCGGCGGCAAGCGCACGGGGCCCGCCC GCCCCGCTCCGCAGGGGCAGGGGGGAGGGGACtac) by boiling it in a water bath for 10 minutes, then placing it in a 70-degree water bath, and letting the temperature gradually fall to room temperature. Subsequently, we digested pCDH-mKMT2E with SnaBI at the N-terminus of the mKMT2E promoter to generate a blunt end of the lentivector backbone to which was ligated with the blunt end fragment of the full-length lncRNA *E22* 2557 bp sequence (wmE22lnRNA) or the 2006 bp truncated version (dmE22lnRNA). Above full-length and truncated lncRNA *E22* fragments were obtained (pCDH-wtmELincRNA-mKMT2E pro-Nluc and pCDH-wtmELincRNA-dmKMT2E pro-Nluc) through the digestion with XbaI at N-terminus and NotI C-terminus and then were filled by DNA polymerase. HEK293FT cells were plated in 96-well plates and transfected with 1μg of pCDH-wtmELincRNA-mKMT2E pro-Nluc vs. pCDH-wtmELincRNA-dmKMT2E pro-Nluc, using Lipofectamine 2000 (Thermo Fisher Scientific). The medium was replaced 8 hours after transfection. 48 hours after transfection, luminescence was measured using the Dual-Glo Luciferase assay system (Promega, #N1610), and normalized to protein concentration.

KMT2E SNV rs73184087 luciferase reporter assay

A luciferase gene reporter vector (pEZX-PG04.1) was purchased from GeneCopoeia (#HPRM49354-PG04) carrying a luciferase reporter gene driven by the KMT2E promoter and a constitutively expressed alkaline phosphatase gene driven by a CMV promoter. A 471bp segment of the human KMT2E intron 10 containing SNV rs73184087 was generated by PCR from human PAEC genomic DNA. The SNV rs73184087 A or G allele was generated with a forward primer 5'- cta gtc tag aCC TCC TGT GTT CAA GCA AGC AAT TCT CCT-3' and two separate reverse primers (A allele: 5'- CCG GCG CGC CGT TTA AAC CAA GAG AAA ACA TAA TGA ACA TTT TAA AAA CTT CAT TTA AAA AGA GTA TCA CAC AAA GTT TAA AAA TAT ATA GAA TAA GAG G-3', G allele: 5'- CCG GCG CGC CGT

TTA AAC CAA GAG AAA ACA TAA TGA ACA TTT TAA AAA CTT CAT TTA AAA AGA GTA TCA CAC AAA GTT TAA AAA TAT ACA GAA TAA GAG G-3'). The PCR products with A or G allele were purified and then cloned into pEZX-PG04.1 downstream of the luciferase reporter gene at sites (XbaI 2963 and PmeI 2990). The luciferase reporter construct DNA was transfected into HEK293FT cells by Lipofectamine 2000 (Thermo Fisher Scientific). The medium was replaced 8 hours after transfection. 48 hours after transfection, the activity of secreted Gaussian luciferase was measured using a Secrete-Pair Dual Luminescence Assay Kit (GeneCopoeia, #LF031). Luciferase activity was normalized to the constitutively expressing secreted alkaline phosphatase under CMV.

Messenger RNA extraction and quantitative RT-PCR

RNA extraction was performed using Qiagen RNeasy extraction kit (#74106), followed by cDNA synthesis. Quantitative RT-PCR (RT-qPCR) was performed on an Applied Biosystems Quantstudio 6 Flex Fast Real Time PCR device, as we previously described (*40*). Fold-change of messenger RNA was calculated using the formula ($2^{-\Delta\Delta Ct}$), normalized to actin expression. TaqMan primers and custom-made primers are listed in **Table S9. Key Resource Table**.

Proteasome inhibitor assays

Human PAECs were transfected with scramble control vs. si-*KMT2E-AS1*, followed by hypoxic exposure (0.2% O_2) vs. normoxia (20% O_2) for 24 hours. 46 hours post-transfection, cells were treated with proteosome inhibitor MG132 (5 μM, Sigma, #M7449) and DMSO vehicle control for 2 hours under hypoxia before harvesting cell lysate for immunoblot.

Immunoblotting and antibodies

Whole-cell lysates were prepared using RIPA buffer supplemented with protease inhibitor cocktail (Sigma-Aldrich). The nuclear extract was isolated using NE-PER nuclear and cytoplasmic extraction reagents (Thermo Fisher Scientific). Protein concentration was determined using the BCA Protein Assay Kit (BioRad). Protein lysates were resolved by SDS-PAGE gel electrophoresis and transferred onto a PVDF membrane (BioRad). Membranes were blocked with 5% non-fat milk or bovine serum

albumin (BSA) in PBS containing 0.1% Tween 20 (PBST), followed by primary antibody incubation overnight at 4°C and then secondary antibody incubation for 1 hour at room temperature. Primary antibodies for HIF-1α (NB100-134, 1:1000) and HIF-2α (NB100-122, 1:1000) were obtained from Novus biologicals. VEGF (ab183100, 1:1000), H3K4me3 (ab8580, 1:1000), H3K9me3 (ab8898, 1:1000), H3K27me3 (ab6002, 1:1000), PARP-1 (ab151794, 1:1000) were obtained from Abcam. Antibodies against KMT2E (sc-377182, 1:200) and β-actin (sc-47778; 1:1000) were obtained from Santa Cruz Biotechnology. Histone (9715s, 1:1000) was from Cell Signaling Technology. ELOC (12450-1-AP, 1:2000) was from Proteintech. Corresponding secondary antibodies (anti-rabbit, anti-mouse and anti-goat) coupled to HRP were used (Thermo Fisher Scientific). Immunoreactive bands were visualized using SuperSignal™ West Pico PLUS or Femto Chemiluminescent Substrate (Thermo Fisher Scientific) and detected with ChemiDoc MP Imager (Bio-Rad). Signals were quantified using NIH ImageJ software (https://imagej.nih.gov/ij/). All immunoblot analyses were performed at least three times.

RNA fluorescent in situ hybridization (FISH)

Locked nucleic acid-in situ hybridization (LNA-ISH) with biotin/streptavidin signal amplification was performed to detect lncRNA. All LNA probes were synthesized by Qiagen, including double digoxigenin (DIG)-labeled probes against *KMT2E-AS1* and *E22*, and DIG-labeled scramble probe (negative control), and used at the concentration of 20nM. Anti-DIG biotin conjugate (Novus biologicals, #BAM7520), and Streptavidin-Alexa594 (Thermo Fisher Scientific, #S11227) were used for signal amplification. Confocal fluorescence microscopy was performed using a Zeiss LSM 780 confocal microscope.

Proximity ligation assay

Proximity ligation assay was done by using Duolink PLA assay kit (Sigma Aldrich, DUO92102-1KT) with modifications. Briefly, human PAECs cultured and treated on 8-well Chamber Slide (Thermo Fisher Scientific, 154534PK) were washed and fixed with 4% PFA for 10 min, followed by permeabilization of the cells with PBS containing 0.25% Triton X-100 for 15 min. Cells were then

blocked with Duolink Blocking Solution for 1 h at 37°C, and incubated with KMT2E (Santa Cruz Biotechnology, #sc-377182, 0.2 mg/mL), H3K4me3 (Abcam, #ab8580, 1 mg/mL) and rabbit and mouse IgG antibodies (Thermo Fisher Scientific, 02-6102 and 02-6502, 2.5 mg/mL) diluted in Duolink Antibody Diluent overnight at 4°C. After washing steps, PLUS and MINUS PLA probes in Duolink Antibody Diluent were added and incubated at 37°C for 1 h, followed by ligation mixture incubation for 30 min at 37°C. Amplification solution was then applied for 1h 40 min at 37°C. Cells were mounted with Duolink In Situ Mounting Medium with DAPI for 15 min and imaged using Nikon A1 confocal microscope.

RNA-immunoprecipitation analyses

RNA-binding protein immunoprecipitation in human PAECs was performed by following the instructions of the Magna RIP kit (Millipore Sigma, #17-701). Briefly, PAECs were transfected with scramble control or si-*KMT2E-AS1*. 48 hours after transfection, cells were exposed to normoxia or hypoxia (24 h). Before harvest, cells were exposed to 1 μg/mL of actinomycin D or DMSO for 4 hours (Sigma-Aldrich). Cells were then lysed using RIP lysis buffer with protease inhibitor and RNase inhibitor. Cell lysates were immunoprecipitated by protein A/G magnetic beads with 5 μg of IgG or KMT2E antibodies. 10% of the immunoprecipitated proteins were used for immunoblotting. RNA was purified from the immunoprecipitated proteins, followed by qPCR quantification of *KMT2E-AS1*.

Chromatin immunoprecipitation (ChIP)-qPCR

Chromatin immunoprecipitation (ChIP) was performed using Pierce magnetic ChIP kit (Thermo Fisher Scientific, #26157) with modifications. Briefly, cells were fixed and cross-linked in 1% paraformaldehyde (PFA) for 10 minutes at room temperature. Cells were then harvested and lysed in membrane extraction buffer containing 1X protease/phosphatase inhibitors (Pierce magnetic ChIP kit) for 10 minutes on ice. After centrifugation, nuclei were resuspended in IP dilution buffer with protease inhibitor cocktail, followed by sonication (Bioruptor) in an ice bath at high power (7 cycles of 1 minute with 20 seconds on and 40 seconds off). After centrifugation at 20,000 g for 5 minutes,

supernatants containing the soluble chromatin were transferred to new tubes for immunoprecipitation. Chromatin was precipitated with 5 µg of anti-trimethylated histone 3 lysine 4 antibody (H3K4me3, Abcam, #ab8580) or non-immune rabbit IgG (Diagenode, #C15410206) and Protein A/G Magnetic Beads (Pierce magnetic ChIP kit). DNA was recovered from the protein-chromatin complex by phenol-chloroform extraction and ethanol precipitation, followed by RT-qPCR using primers (**Table S9. Key Resources Table**) generated for predicted H3K4me3 binding in the specific miR210hg promotor site, HIF-2α binding at SNV region.

Prediction and scoring of HIF-2α binding sites

HIF-2α chromatin immunoprecipitation sequencing peaks were published previously (*19*). These are inclusive of sites where HIF-2α directly binds canonical HIF-response elements (HREs) as well as sites where HIF-2α is associated via non-HREs in transcription factor complexes. Primary sequences, consisting of the 300 centrally-located bases of each ChIP-seq peak, were scanned in the MEME suite (*53*) using the default parameters to generate *de novo* HIF-2α association motifs and position weight matrices (PWMs) between 6 and 15 nucleotides in length. Oligonucleotide sequences centered on either the major or minor allele with 5- and 3-prime flanks of the motif length minus one – such that any potential binding site would necessarily include a given SNV – were scanned against each motif PWM (degenerated sequences shown below as degenerate sequences in International Union of Pure and Applied Chemistry [IUPAC] notation and including a canonical HIF-response element [Motif 2] with E-value cutoff of 0.05). The maximum log-odds scores for each SNV allele were computed in Biopython v1.78 (*54*). Score thresholds corresponding to Bonferroni-corrected *P*-value thresholds (0.05/[883 SNPs x 2 alleles x 2 strands x (8 + 8 + 15 nucleotide Motif lengths)] = $0.05/109492 = 4.57 \times 10^{-7}$) were calculated using the *TFMPvalue* (*55*) module in R v4.0.1. SNV alleles were considered to have predicted differential HIF-2α binding if the maximum log-odds score calculated from any PWM was above the score threshold in either the major or minor allele and below the score threshold in the other allele.

Motif 1: GTGACTCA, E-value=$8.1x10^{-28}$

Motif 2: VKACGTGC, E-value=$3.6x10^{-21}$

Motif 3: TTDTTTWYTTTKTTT, E-value = $2.5x10^{-20}$

Identification of long-range chromosome interactions by Hi-C mapping

SNVs in regions with prior-capture Hi-C evidence of interaction with the shared KMT2E/*KMT2E-AS1* promoter (chromosome 7:104650441-104656168) in lung tissue were obtained from the 3DIV database (*29*).

Allele-imbalanced DNA pulldown Western blot (AIDP-Wb)

AIDP-Wb to detect allele specific binding of HIF-1α or HIF-2α to SNV rs73184087 was performed as described (*56*). In brief, a 31 bp biotinylated SNV sequence centered with either A or G allele of SNV rs73184087 was generated by annealing two biotinylated primers (IDT). Approximately 1 µg DNA was then attached to 40 µl of Dynabeads™M-280 Streptavidin (Thermo Fisher Scientific). DNA-beads were mixed with ~100 µg of nuclear protein extracts isolated from hypoxia-treated human PAECs at RT for 1 hour with rotation. After washing off the unbound proteins, the DNA bound proteins were eluted by sample buffer and resolved on an SDS-PAGE gel for Western blot analysis using an antibody directed against HIF-1α or HIF-2α (Novus Biologicals, #NB100-134 and 122). For internal control, the same blot was probed using an antibody directed against PARP-1 (Novus Biologicals, #NBP2-13732).

Chromatin conformation capture (3C) assay

1×10^7 human PAECs or lymphocytes were cross-linked with 1% formaldehyde at RT for 10 min. Nuclei were isolated and genomic DNA was digested with 400U EcoRI (New England Biolabs, #R0101L) overnight at 37°C, 950 rpm. Digested DNA was diluted to 8 ml and ligated with and without 4000U T4 DNA ligase (New England Biolabs, #M0202M) for 4 hours in a 16°C water bath. Then the cross-links were reversed by adding 100 µg Proteinase K (Thermo Fisher Scientific, #EO0491) and incubation

at 65 °C overnight. DNA was then isolated and purified. PCR was performed with primers chosen close to the EcoRI sites (62bp from the EcoRI site at the promoter, 5'-GAC TGA AAA TTA AAT GGT GT-3', and 68bp from EcoR1 site at SNV segment, 5'-AGT TTC AAG CTA GTA TCT GT-3', and amplification of the ligated product resulted in a 130-bp fragment. The primers were also chosen for the EcoRI sites between the promoter and its upstream and downstream segments, to serve as a control. PCR products were separated by gel electrophoresis and analyzed by DNA sequencing.

Transformed lymphocyte analyses

Male and female lymphocytes from PAH patients with A/A or G/G genotypes were separated from whole blood by density gradient centrifugation. Immortalization of lymphocytes was performed by treating 2 x 10^6 lymphocytes with Epstein–Barr Virus stock from B95-8 marmoset cell line and 2 μg/mL of cyclosporin for a week until colonies formed. Transformed lymphocytes were passaged and expanded for experiments. A/A and G/G lymphocytes were treated with cobalt (II) chloride (50 μM, Sigma-Aldrich, #232696) for 24 hours and then harvested for RNA and RT-qPCR analysis.

Reprogramming transformed lymphocytes into induced-pluripotent stem cells

Lymphoblastoid cell lines (LCLs) carrying homozygous A or G at rs73184087 were cultured in RPMI 1640 (Gibco) supplemented with 15 % fetal bovine serum (FBS), 1% MEM nonessential amino acids, 1 mM sodium pyruvate, and 10 mM HEPES buffer at 37 °C and 5 % CO2 in a humidified incubator. The LCLs were electroporated with the Neon™ Transfection System 10 μL Kit (Thermo Fisher Scientific, #MPK1096) using 1.0 μg of each plasmid (pCXLE-hOCT3/4-shp53, pCXLE-hSK and pCXLE-hUL, Addgene) expressing OCT4, SOX2, KLF4, l-MYC, LIN28 and p53 shRNAs (*30*), according to the manufacturer's instructions. The transfected LCLs were transferred to a 12-well plate and incubated for 24 h. At 24 h after electroporation, cells were transferred to a Matrigel-coated 12-well plate and supplemented with iPSC reprogramming medium TeSR-E7 (Stemcell Technologies). When iPSC colonies started to appear, cells were then cultured in mTeSR1 media (Stemcell Technologies) and maintained in a hypoxic incubator (5% O_2). Colonies were manually picked for

further expansion. After 3-4 passages, cells were cultured in Essential 8 media (Thermo Fisher Scientific). iPSCs were characterized by immunofluorescent staining for pluripotency markers NANOG (Cell Signaling Technology, #4903S, 1:200), SSEA4 (Cell Signaling Technology, #4755S, 1:500), OCT4A (Cell Signaling Technology, #2840S, 1:400), and SOX2 (Cell Signaling Technology, #4900S, 1:400). Three germ layer differentiation using STEMdiff Trilineage differentiation kit (Stemcell Technologies, #05230) was done following manufacturer protocol. Cells were fixed and stained for ectoderm (Nestin, Santa Cruz Biotechnology, #sc-23927, 1:200), mesoderm (Brachyury, Abcam, #ab209665, 1:1000) and endoderm (SOX17, Cell Signaling Technology, #81778S, 1:3000) specific markers.

Differentiation of iPSCs into endothelial cells

As described (*31*), starting at passage 10, iPSCs were chemically induced with 6 μM CHIR99021 (Selleckchem, #S2924) for 2 days, followed by 3 μM CHIR99021 for 2 days in RPMI media supplemented with B27 minus insulin (Thermo Fisher Scientific). From day 5, cells were cultured in endothelial differentiation medium: EGM2 (Lonza) supplemented with 50 ng/mL VEGF (Stemcell Technologies, #78073.1), 25 ng/mL FGF (Peprotech, #AF-100-18B) and 10 μM of the TGFβ inhibitor SB431542 (Selleckchem, #S1067). The medium was changed every 2 days. iPSC-Ecs were then purified using magnetic-activated cell sorting (MACS) against CD144 microbeads (Miltenyi Biotec, #130-097-857). Purified iPSC-Ecs were characterized by flow cytometry against endothelial-specific markers, CD31 and CD144.

In vitro angiogenesis assays

Capillary tube formation was performed using a commercial kit (In vitro angiogenesis assay kit, Cultrex, #3470-096-K). Briefly, Matrigel with reduced growth factors was pipetted into pre-chilled 96-well plate (50 μl Matrigel per well) and polymerized for 30 min at 37°C. iPSC-Ecs were resuspended in 100 μl of basic media and seeded in Matrigel coated 96-well plate. After 4-6 h of incubation, tubular structures were photographed using Olympus inverted fluorescent microscope.

Endothelin-1 ELISA and LDH activity assays

Assays were performed according to the manufacturers' instructions using endothelin-1 ELISA kit (Enzo Life Sciences, #ADI-900-020A) and lactate dehydrogenase activity colorimetric assay kit (BioVision, #K726-500).

Seahorse assay

Oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) in cultured PAECs (20,000 cells /well) were measured using Xfe96 Extracellular Flux Analyzer (Seahorse Bioscience) following manufacturer's instructions, as previously described (*40*). Briefly, pre-treated cells were washed and treated with XF medium (Seahorse Bioscience) containing 1mM Sodium-Pyruvate, 2mM L-Glutamine and 10mM glucose (pH = 7.4). OCR was measured at baseline following sequential addition of 1μM Oligomycin, 0.5μM FCCP, and 2μM Rotenone plus 0.5μM Antimycin A. For ECAR, cells were treated with glucose-free XF base medium (Seahorse Bioscience) supplemented with 2mM L-Glutamine (pH = 7.35). ECAR was measured following serial addition of 10mM D-Glucose, 1μM Oligomycin and 100mM 2-Deoxyglucose.

BrdU proliferation assay and caspase 3/7 assay

Cell proliferation and apoptosis assays were performed using BrdU Cell Proliferation Assay Kit (Cell Signaling Technology, #6813) and Caspase-Glo 3/7 Assay (Promega, #G8093) respectively, per the manufacturers' instructions. Signals were normalized to protein concentration determined by BCA (Bio-Rad).

Scratch assay

Confluent PAECs were wounded using pipet tips. Over the course of 12 hours, wound bed closures were followed every 4 hours by capturing brightfield images using EVOS XL CORE imaging system (Thermo Fisher Scientific). Wound closures were quantified using the NIH ImageJ software (http://rsb.info.nih.gov/ij/).

Contraction and co-culture assays

PASMC contraction assay in response to conditioned PAEC media was performed as described previously (*17*). Briefly, collagen-I solution (BD Biosciences) was first prepared by neutralizing it to a pH of 7.5 on ice with 0.1 M NaOH in PBS and 0.1 N HCl. The collagen-I solution was then mixed on ice with growth factor reduced Matrigel (BD Biosciences) to obtain a final concentration of 1.5 mg/ml. PASMCs (50,000 cells/well) were resuspended in 100 µl of the ECM mixture in growth media and plated onto a 96-well plate. After 1 h of incubation at 37 °C, PASMCs-embedded matrices were overlaid with 100 µl of conditioned PAEC serum-free medium (transfected with siRNA or transduced with lentivectors). Medium was changed every 12 h, and at day 4, the gels were imaged, followed by measurement of the relative diameter of the well and the gel using ImageJ software. The percentage contraction was calculated as 100 × (well diameter - gel diameter)/well diameter.

lncRNA structural analysis

Predictions of secondary structures of human *KMT2E-AS1* and mouse *5031425E22Rik* conserved regions were performed by using the RNA Folding form version 2.3 of the mFold server (https://www.unafold.org/mfold/applications/rna-folding-form-v2.php). We concentrated on the highly conserved *KMT2E-AS1* and mouse *5031425E22Rik* regions (1281 to 1397 region for human and 1248 to 1354 for mouse region) for computational predictions of the most stable RNA conformation.

Generation and delivery of AAV6-*E22* vector in vivo

The plenti-*E22* plasmids containing 2500bp full-length *E22* served as templates for PCR to obtain the *E22* gene cloning sequence. The purified PCR products were electrophoresed as 2.5 kb bands, purified, and treated with PNK enzyme (NEB, MA, USA) to add a phosphate to the N-terminus. The fragments were inserted separately into a single-stranded recombinant adeno-associated virus (AAV) backbone carrying a constitutive CMV promoter, as previously described (*32*). The transduction and expression efficiencies of recombinant AAV serotypes 2, 5, 6, 8, 9 were tested in culture mouse pulmonary endothelial cells. Recombinant AAV6 viruses were generated by triple plasmid co-

transfection of 293 cells and purified twice with cesium chloride gradient ultracentrifugation. Vector titers [5×10^{12} genome copy number per ml (v.g./mL)] were determined using DNA dot-blot hybridization. Male C57BL/6J mice were anesthetized by inhaled isoflurane with oxygen using a precision vaporizer (induced at 3%, maintained at 1.5–3%) in a closed ventilation chamber, followed by orotracheal instillation of 1×10^{11}vg/mouse in 100 µl aliquots of AAV6-*E22* or AAV6-*GFP* control.

Generation of *5031425E22Rik* knockout mice

As previously described (*57*), pronuclei of fertilized embryos (C57BL/6, Jackson Laboratory) produced by natural mating, were microinjected with a mixture of EnGen Cas9 protein (New England Biolabs, M0646T) and Cas9 guide RNAs: 5’ guide 2 “GTGATTTCTGTCGGGCGCGTAGG”, 3’ guide 2 “CTGAAAGGACACTCCTTTCGGGG”, cons guide 5 5’ “CAGGAGAAATTGCGCCTCCACGG”, cons guide 19 “GTGGGGTTTGGCGGGGACCCAGG”. The injected zygotes were cultured overnight and transferred to pseudo-pregnant female recipients the following day. Founder mice and their generations were genotyped using designed primers from IDT: 5’P1“CCGCCGCCTCCTATACTTCTTAGC” and 3’P3 “GTGTTCCTGGGTGTTATTCACTTGC” for AD deletion and short_F “GACCTTTTGCTCCTCTCCCT” and 3’P1 “GCGTCAAACCTTCTCCTCCACC” for BD deletion. Lines were backcrossed to C57Bl/6 mice for 7-10 generations. To generate a severe and inflammatory model of Group 1 PAH, AD deletion mice were crossed with IL-6 Tg transgenic mice (*33*). Male mice (10-12 weeks old) were used for experimentation, and C57BL/6 wildtype or IL-6 Tg age/sex-matched littermate mice were used as controls. Knockout mice did not show any developmental or fertility defects, displaying normal litter sizes and no obvious off-target abnormalities.

Mouse model of PH via chronic hypoxia

For modeling hypoxic PH in mice, male C57BL/6 WT littermate male mice (10-12 weeks old, Jackson Laboratory) were subjected to 28 continuous days of normobaric hypoxia in a temperature-humidity controlled chamber (10% O_2, OxyCycler chamber, Biospherix Ltd.) as compared to normoxia (21% O_2). At 4 weeks post-exposure, systemic blood pressure was measured by non-invasive tail-cuff

plethysmography (CODA system, Kent Scientific), according to manufacturer's instructions. Right heart catheterization was performed followed by harvest of lung tissue for RNA and protein extraction and OCT embedding, as previously described (*31*).

Rat model of PAH via SU5416-chronic hypoxia

Male Sprague-Dawley rats (Jackson Laboratory, 10 weeks old) were administered 20mg/kg of SU5416 (Sigma) i.p. and exposed to normobaric hypoxia at 10% O_2 for 3 weeks, followed by 2 weeks of normoxia. DMSO vehicle control or chaetocin (0.25 mg/kg, N = 4 per group) was administered daily via i.p injection throughout the 2-week course of normoxia. At the end of week 5, hemodynamic measurement was performed. Systemic blood pressure was measured by invasive catheterization of the abdominal aorta, followed by tissue harvest.

Rodent echocardiography

Echocardiography was performed using the Vevo 3100 ultrasound machine with the MX400 linear transducer capable of 40 MHz (Visualsonics). The rodent was anesthetized using 3% isoflurane mixed with 1 L/min of 100% oxygen in an induction chamber. Once anesthetized the rodent was then transferred to a warming table and positioned for imaging using 1.5% isoflurane. B-Mode and M-mode imaging was obtained, of the parasternal short axis view of the left ventricular chamber at mid-papillary muscle. Digital echocardiograms were analyzed off-line for quantitative analysis as previously described (*31*). Briefly, Vevo Lab software (VisualSonics) was used to obtain measurements of ejection fraction, fractional shortening and left ventricular wall thickness, by averaging 3 consecutive cardiac cycles. Ejection fraction was calculated based off the M-mode imaging using the formula: 100 (LV Vol;d - LV Vol;s/LV Vol;d). Fractional shorting was calculated by M-mode imaging based on the following formula: 100 (LVID,d – LVID,s/ LVID,d). The left ventricular wall was measured three times via a linear line measuring the depth of the wall.

Isolation of mouse pulmonary vascular endothelial cells

Pulmonary vascular endothelial cells were isolated as we reported (*17*). Briefly, mouse lungs were digested with collagenase D and DNase in 4mL HBSS at 37C for 30 min. FACS buffer (10% FBS in PBS) was added to neutralize collagenase. The sample solution was filtered through 100μm and 400μm cell strainers, followed by centrifugation at 1700rpm for 5 min. Red blood cells were lysed on ice using ACK lysis buffer for 3 min, followed by neutralization and centrifugation. Cell pellets were resuspended and incubated with CD31 mouse microbeads (Miltenyi Biotec, #130-097-418) for 15 min on ice. CD31+ endothelial cells were then collected by magnetic columns (Miltenyi Biotec, #130-122-727) and extruded for RNA extraction.

Immunohistochemistry and immunofluorescence of lung sections

Cryostat sections were cut from OCT embedded lung tissues at 5-10 μm and mounted on gelatin-coated histological slides. Slides were thawed at room temperature for 10-20 min and rehydrated in wash buffer for 10 min. All sections were blocked in 10% donkey serum and exposed to primary antibody and Alexa 488, Cy3 and Cy5-conjugated secondary antibodies (Thermo Fisher Scientific) for immunofluorescence. Primary antibodies included KMT2E (Santa Cruz Biotechnology, C-10, 1:100), H3K4me3 (Abcam, ab8580, 1:100), H3K9me3 (ab8898, 1:100), H3K27me3 (ab6002, 1:100), VEGF (Abcam, ab183100, 1:100), EDN1 (Abcam, ab117757, 1:100), Ki67 (Abcam, ab15580, 1:100), CD31 (Biotechne, AF3628, 1:50) and α-SMA (Abcam, ab32575, 1:250). Images were obtained using Nikon A1 confocal microscopy. Small pulmonary vessels (<100 μm diameter, >10 vessels/section) that were not associated with bronchial airways were selected for analysis (N = 4-5 animals/group). Staining intensity was quantified using ImageJ software (NIH). Degree of pulmonary arteriolar muscularization was assessed in OCT lung sections stained for α-SMA by calculation of the proportion of fully and partially muscularized peripheral (<100 μm diameter) pulmonary arterioles to total peripheral pulmonary arterioles, as previously described (*31*). All measurements were performed blinded to condition.

Genetic analysis of the *KMT2E* gene

Discovery cohort (PAH Biobank)

The PAH Biobank and three European PAH cohorts have been previously described in detail (*28*). Briefly, European-descent adult (age>18 years) individuals with predominantly idiopathic PAH were included from the PAH Biobank (N = 694), a national, multi-center biobank and registry of PAH patients (**Table S5**). The control group included 1,560 European-descent adult individuals without PAH from Vanderbilt University, previously described in detail (*28*). Samples were genotyped on Illumina Omni5-Quad BeadChip array, and samples with low genotyping rate and mismatched clinical and genomic data records were excluded as previously described (*28*). All individuals provided written consent and the cohorts were IRB-approved by each institution.

SNV association analysis in the discovery cohort

Briefly, using available SNV genotypes in the PAH Biobank (*28*) within and flanking (+/-200kb) the lncRNA-KMT2E locus, 59 SNVs were identified with predicted HIF-2α binding to one of either the minor or major SNV alleles, as described above (**Table S6**).

Across this cohort of SNVs, Firth's penalized logistic regression analysis was performed to identify associations with disease risk (PAH cases vs. controls), adjusting for sex, age and two Principal Components (PCs) using PLINK 2 (*58*). Genotypes were coded additively for the number of copies of the minor allele and odds ratios were calculated for each copy of the minor allele. Rsq values were used to generate a number of independent SNVs by doing matrix algebra, resulting in an effective SNV test count of 53.84. After Bonferroni correction, a P value less than (<) 0.00093 (= 0.05/53.84 effective SNVs) was considered statistically significant in the discovery analysis. Firth's penalized logistic regression was then performed to define statistical association of each candidate SNV. Furthermore, LDlink (ldlink.nih.gov) was used to assess LD pattern near SNV rs73184087. Using European samples, hg38 high coverage map, and 500Kb+/- search window, a LD heatmap for SNV rs73184087 was generated.

Validation cohort (UPMC)

Self-identified European-descent individuals with PAH were recruited into a biobank and registry at the UPMC Pulmonary Hypertension Comprehensive Care Center (Pittsburgh, PA, USA) between 2016-2021 and independent of patients recruited for PAH Biobank (discovery cohort). A total of 96 patients diagnosed with WSPH Group 1 PAH were enrolled, including predominantly idiopathic PAH along with hereditary PAH, congenital heart disease-associated PAH, and toxin-associated PAH (**Table S7**). The diagnosis of PAH was confirmed by an expert PAH physician, based on recent WSPH criteria including an elevated mPAP > 20 mmHg, pulmonary capillary wedge pressure $\leq$15 mm Hg, and pulmonary vascular resistance $\geq$ 3 Wood Units (PVR) by right heart catheterization. Controls included 401 non-PH, healthy adult individuals of self-identified European-descent were recruited. Notably, individuals were excluded if suffering from hypertension, diabetes, heart disease, lupus, scleroderma, polymyositis, asthma, emphysema, and lung fibrosis.

SNV association analysis for validation cohort

Human peripheral blood mononuclear cell (PBMC) DNA was genotyped in PAH and control cohorts for the *KMT2E* intronic SNV rs73184087 using TaqMan SNV genotyping assay (Thermo Fisher Scientific, #C__97341322_10) on an Applied Biosystems Quantstudio 6 Flex Fast Real Time PCR device. MAF and frequency of SNV were calculated in patients with PAH and matched controls. Association between SNV and PAH was calculated using Firth's penalized logistic regression, adjusting for age and sex. Crude and age-adjusted odds ratios were calculated for each sex, separately. 95% C.I. and P-values were calculated using a robust sandwich variance estimator. P value less than 0.05 was considered statistically significant. Given the absence of the homozygous G/G genotype in the cohort, odds ratios were derived from comparisons of A/A vs. A/G genotypes.

SNV association analysis for All of Us

Using the All of Us (dataset v6) platform, individuals who underwent Whole Genome Sequencing (WGS) and were diagnosed with Heritable Pulmonary Arterial Hypertension (AoU Concept Identifier: 44783618) or Idiopathic Pulmonary Arterial Hypertension (AoU Concept Identifier: 44782560) or

carried the ICD 10 code for Secondary Pulmonary Arterial Hypertension I27.21 (AoU Concept Identifier: 1326593). Out of these patients, only the ones receiving at least one of the following pulmonary hypertension medications were included in the analysis: Bosentan, Ambrisentan, Macitentan, Sildenafil, Tadalafil, Epoprostenol, Iloprost, Treprostinil, Riociquat, Nifedipine, Diltiazem, Nicardipine or Amlodipine while excluding those diagnosed with erectile dysfunction (AoU Concept Identifier: 3655355). Then, using the ancestry code provided by All of Us, only individuals from European descent were retained in the case cohort (N = 52 individuals) (**Table S8**). For the control cohort, individuals with WGS, European descent and ICD 10 code for "Encounter for general examination without complaint, suspected or reported diagnosis" Z00 (AoU Concept Identifier: 1575980) were selected while excluding any individuals diagnosed with Pulmonary Hypertension (AoU Concept Identifier: 4322024) or erectile dysfunction (AoU Concept Identifier: 3655355). This control cohort was made up of N=11,821 individuals. To assess for association between rs73184087 and PAH risk, a Firth's penalized logistic regression methodology was used, adjusting for age and sex. P value less than 0.05 was considered statistically significant.

Meta-analysis of PAH cohorts

Three additional European-descent cohorts were available for inclusion and testing of the target SNV, rs73184087. Briefly, these cohorts, their demographics, and their clinical characteristics have been described in detail in a prior GWAS study (*28*) and include adult patients with WSPH Group 1 PAH subjects who met the same criteria as the PAH Biobank and UPMC cohorts for the diagnosis of PAH. A meta-analysis of the UPMC cohort, the PAH Biobank, and these three available European-descent cohorts was conducted in R using the *rmeta* package. A fixed effect mode was used when heterogeneity P value was greater than 0.05 to ensure similar variance across the cohorts. P value less than 0.05 was considered statistically significant.

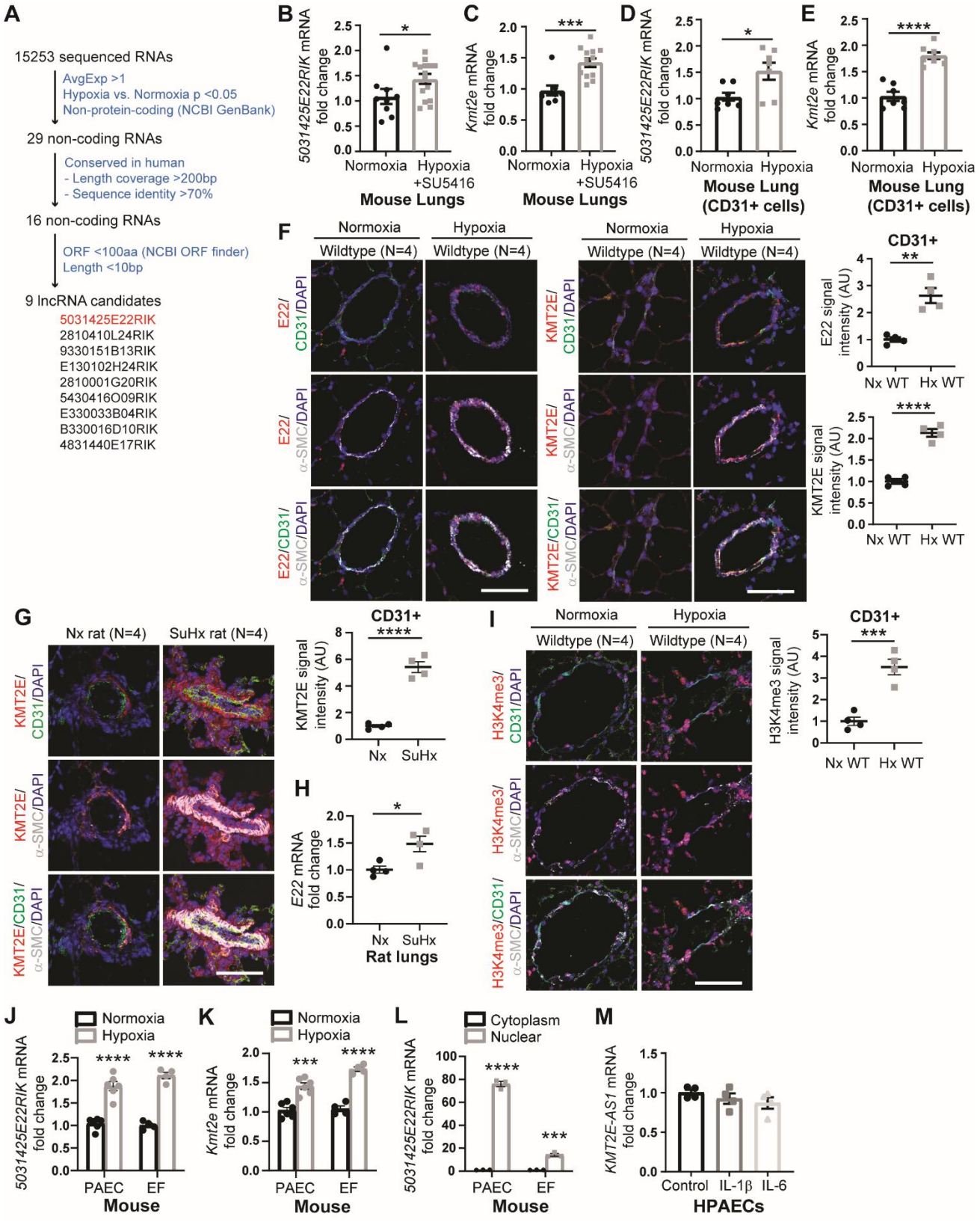

Figure S1. Response of lncRNA *E22/KMT2E-AS1* and KMT2E in hypoxic PH mouse, SU5416-hypoxic PAH rats, cultured mouse cell lines, and human PAECs. (**A**) From RNA-sequencing of SU5416-hypoxic PAH vs. control mouse lung tissues (GEO: GSE61828), differentially expressed long

non-coding RNAs (lncRNAs) are identified. (**B-E**) Mouse lncRNA *5031425E22Rik* (referred to *E22* hereafter) and *Kmt2e* transcripts are quantified in tissue extracts (B-C) (*n = 9-15*; *p<0.05, ***p<0.001, unpaired Student's t-test for *E22* (B), Mann-Whitney test for Kmt2e (C); data represent mean ± SEM), and CD31+ endothelial cells (D-E) of PH mouse lungs by RT-qPCR (*n = 7*; *p<0.05, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). (**F**) Representative fluorescent in-situ hybridization (FISH) and immunofluorescence (IF) staining and quantifications are shown for mouse lncRNA *E22* and KMT2E in lung CD31+ PAECs (green) of chronic hypoxic wildtype (Hx WT) mice as compared to normoxic wildtype (Nx WT) controls (*n = 4*; **p<0.01, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). Scale bars, 50 μm. (**G-H**) Expression of rat homolog of Kmt2e (red; G) and *E22* (H) is shown in SU5416-hypoxic (SuHx) PAH rats vs. normoxic controls by IF staining and RT-qPCR, respectively (*n = 4*; *p<0.05, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). Scale bars, 50 μm. (**I**) IF staining and quantification are shown for H3K4me3 in CD31+ endothelium of hypoxic mice vs. controls (*n = 4*; ***p<0.001, unpaired Student's t-test; data represent mean ± SEM). Scale bars, 50 μm. (**J-K**) *E22* (J) and Kmt2e (K) transcripts are quantified in cultured mouse pulmonary arterial endothelial cells (PAECs) and mouse embryonic fibroblasts (EFs) by RT-qPCR (*n = 4-6*; ***p<0.001, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). (**L**) Expression of *E22* in cytosolic and nucleic fractions of mouse PAECs and EFs is quantified by RT-qPCR (*n = 3*; ***p<0.001, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). (**M**) *KMT2E-AS1* transcript expression is shown in human PAECs exposed to interleukin-1β (IL-1β) or interleukin-6 (IL-6) vs. vehicle control (*n = 4*; ***p<0.001, ****p<0.0001, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM).

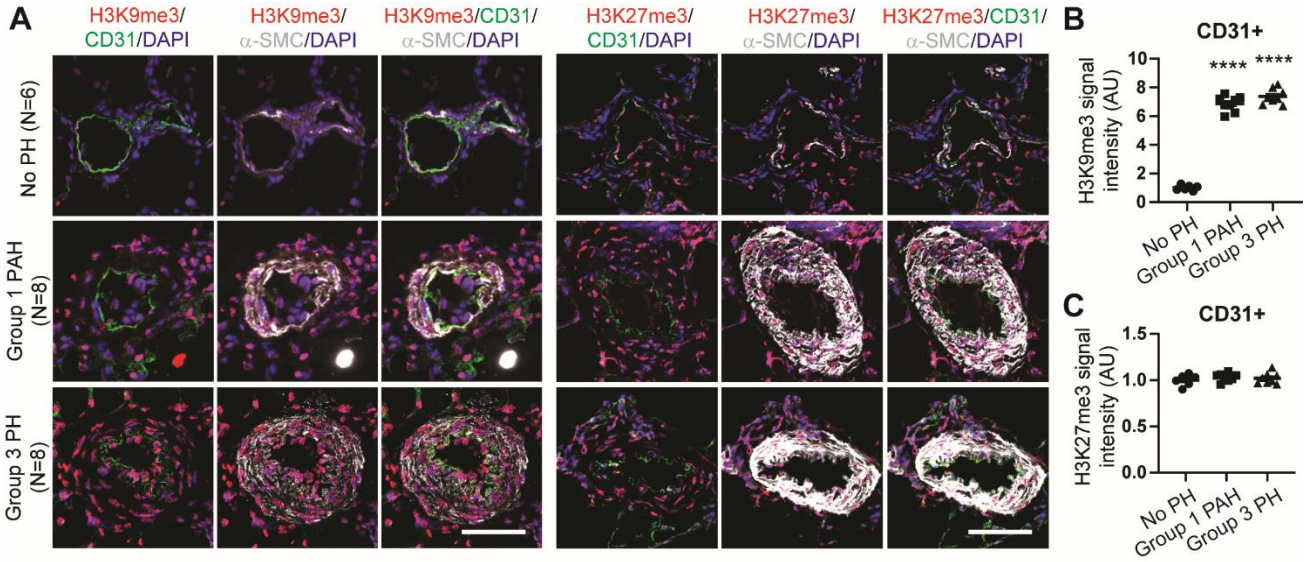

Figure S2. Vascular quantification of other known hypoxia-dependent histone methylation marks in human Group 1 and Group 3 PH. (**A-C**) Representative IF staining and quantifications of H3K9me3 (red; A,B) and H3K27me3 (red; A,C) are shown in CD31+ vascular endothelium of human Group 1 and Group 3 PH vs. control (*n = 6-8*; ****p<0.0001 vs. No PH, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). Scale bars, 50 μm.

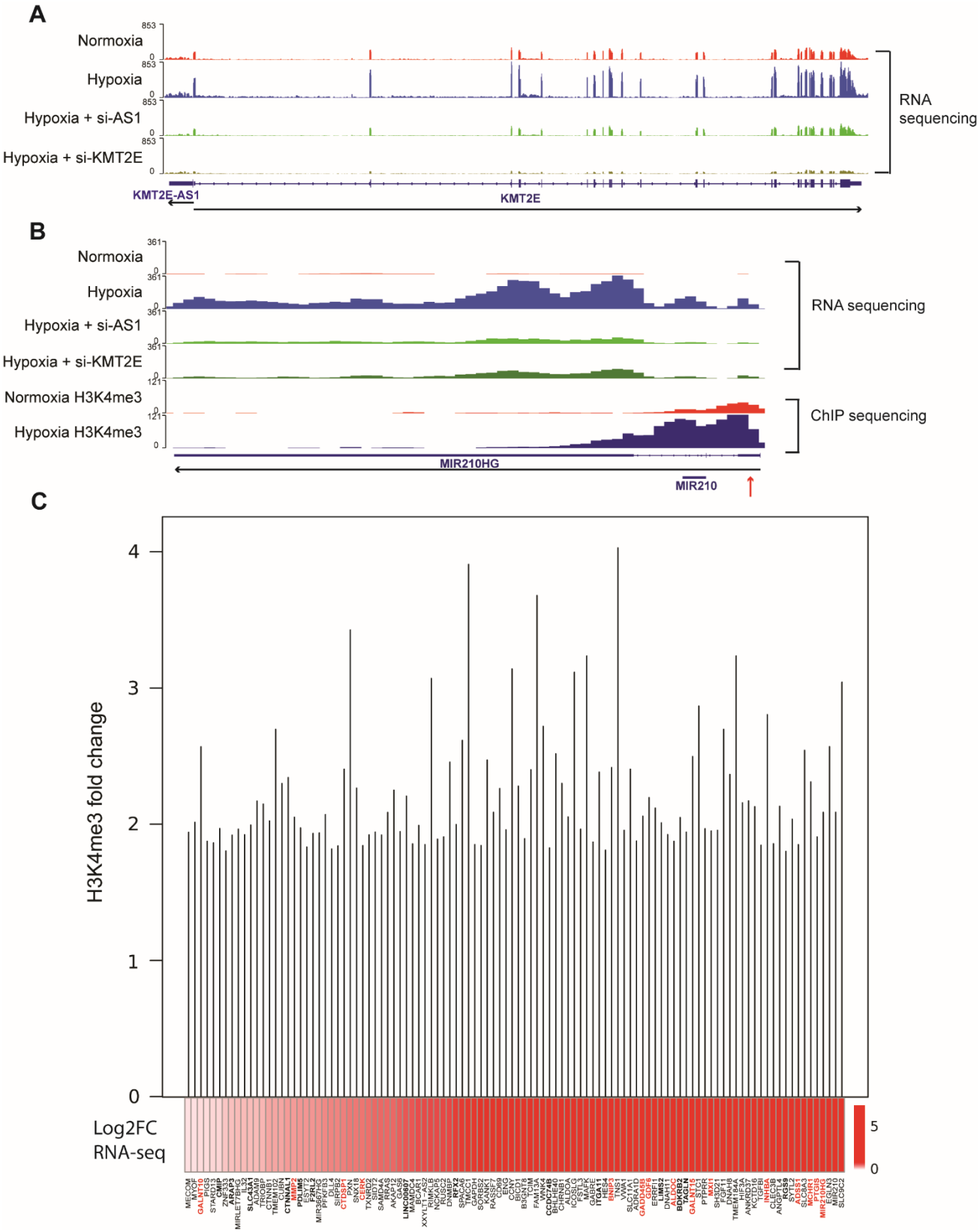

Figure S3. HIF-dependent and metabolism-specific transcripts are controlled by the KMT2E/*KMT2E-AS1* axis and by H3K4me3. (**A-B**) From human PAECs after knockdown of either *KMT2E-AS1* or *KMT2E*, IGV plots of RNA sequencing are shown at the *KMT2E-AS1*/*KMT2E* locus

(A) and at the lncRNA *miR210hg* locus (B, top 4 lanes). From human PAECs under hypoxia vs. normoxia, the IGV plot for H3K4me3-specific ChIP-sequencing is also displayed (B, bottom 2 lanes). Red arrow indicates *miR210hg* gene promoter. (**C**) Via co-analyzing ChIP-sequencing and RNA sequencing data in PAECs, a sub-cohort of genes that carries increased H3K4me3 promoter marks (adjusted $p<0.05$) and increased transcript expression in hypoxia (adjusted $p<0.05$) is shown. Transcripts categorized as “hypoxia” or “metabolism” genes that are reversed by KMT2E/*KMT2E-AS1* knockdown are labeled in red (bold); remaining reversed transcripts are labeled in black (bold).

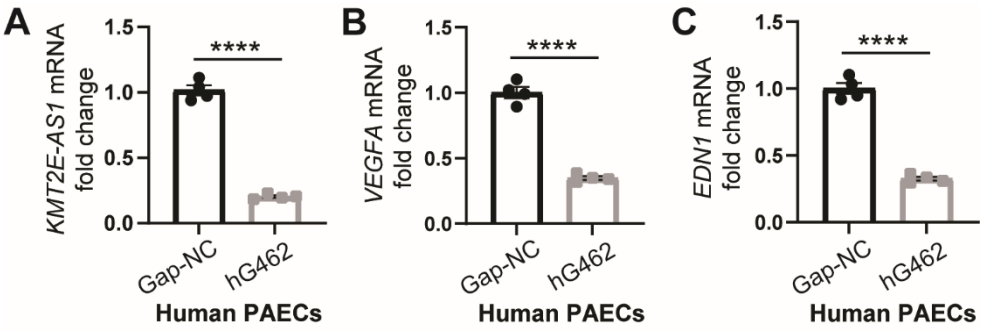

Figure S4. Antisense oligonucleotide gapmer knockdown of *KMT2E-AS1* and downstream effects in human PAECs. (**A-C**) In human PAECs treated with an antisense oligonucleotide gapmer specific for *KMT2E-AS1* (hG462) vs. a gapmer control (Gap-NC), *KMT2E-AS1* (A), *VEGFA* (B), and *EDN1* (C) transcripts are quantified by RT-qPCR (*n* = *4*; ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM).

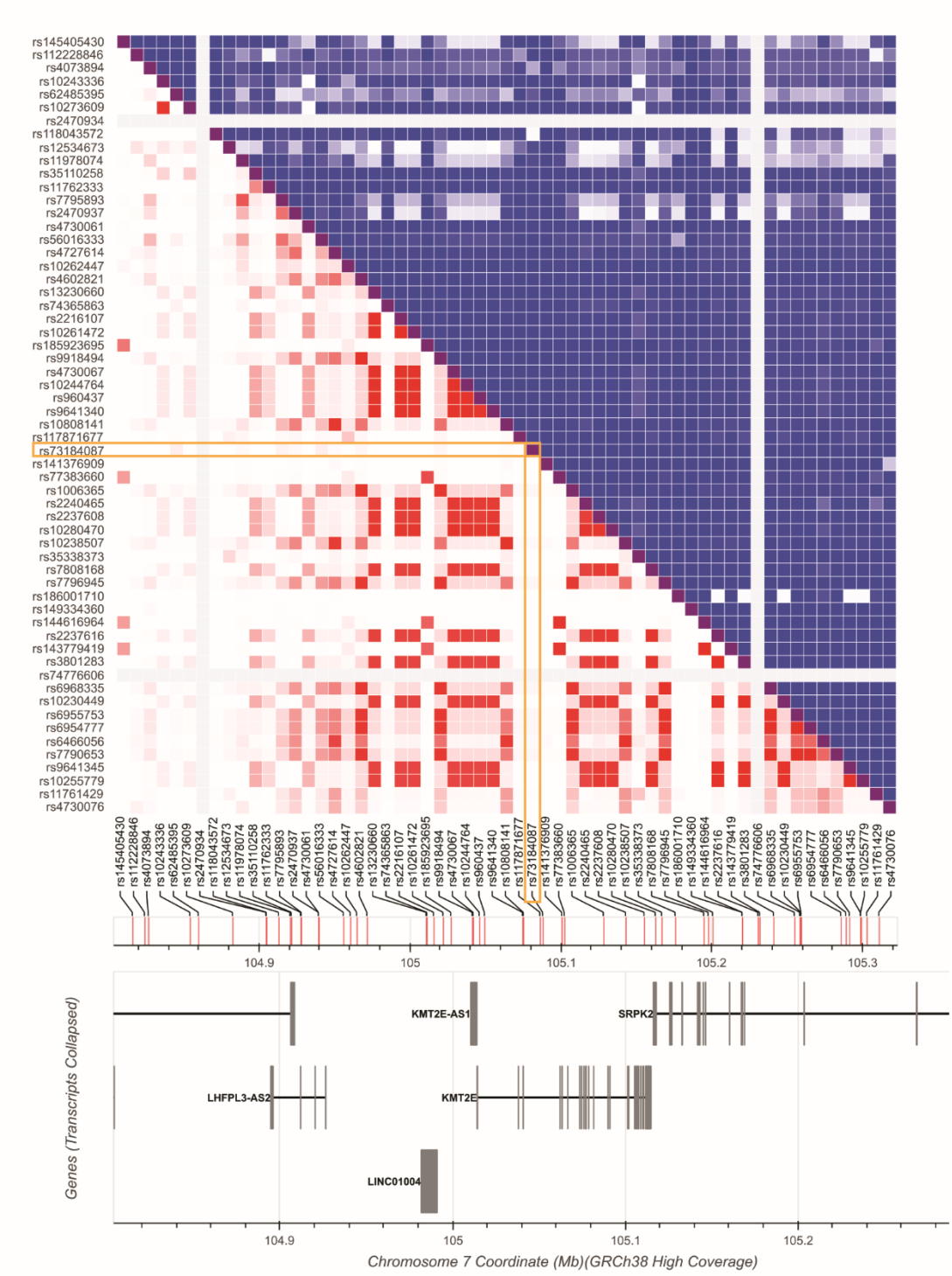

Figure S5. Analysis of linkage disequilibrium of SNV rs73184087. Via LDlink, a linkage disequilibrium (LD) heatmap is displayed. Red indicates strong LD.

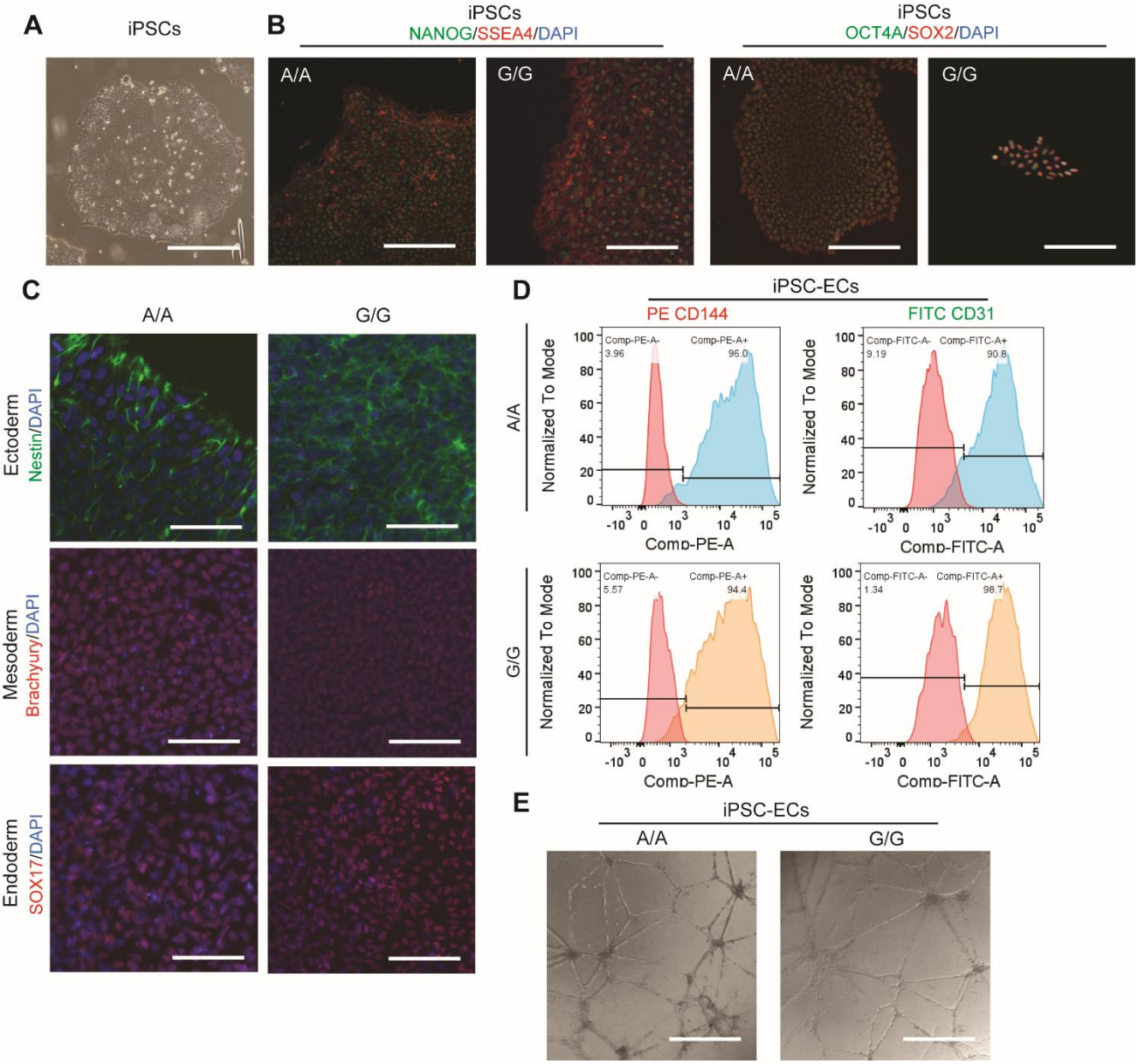

Figure S6. Characterization of induced-pluripotent stem cells and differentiated endothelial cells from transformed lymphocytes carrying A/A or G/G alleles of SNV rs73184087. (**A**) Morphology is shown of induced-pluripotent stem cells (iPSCs) reprogrammed from transformed lymphocytes. Scale bar, 100 μm. (**B**) Expression of 4 pluripotency markers NANOG, SSEA4, OCT4A, and SOX2 is shown in iPSCs with A/A or G/G alleles by IF. Scale bar, 100μm. (**C**) Expression of germ layer-specific markers (Nestin for ectoderm, Brachyury for mesoderm, and SOX17 for endoderm) is shown in iPSCs by IF staining one week after three-germ layer differentiation. Scale bar, 100μm. (**D**) Flow cytometry analysis of iPSC-ECs with CD31 and CD144 markers is shown. (**E**) By tube formation on Matrigel, angiogenic potential is assessed in iPSC-ECs. Scale bar, 100 μm.

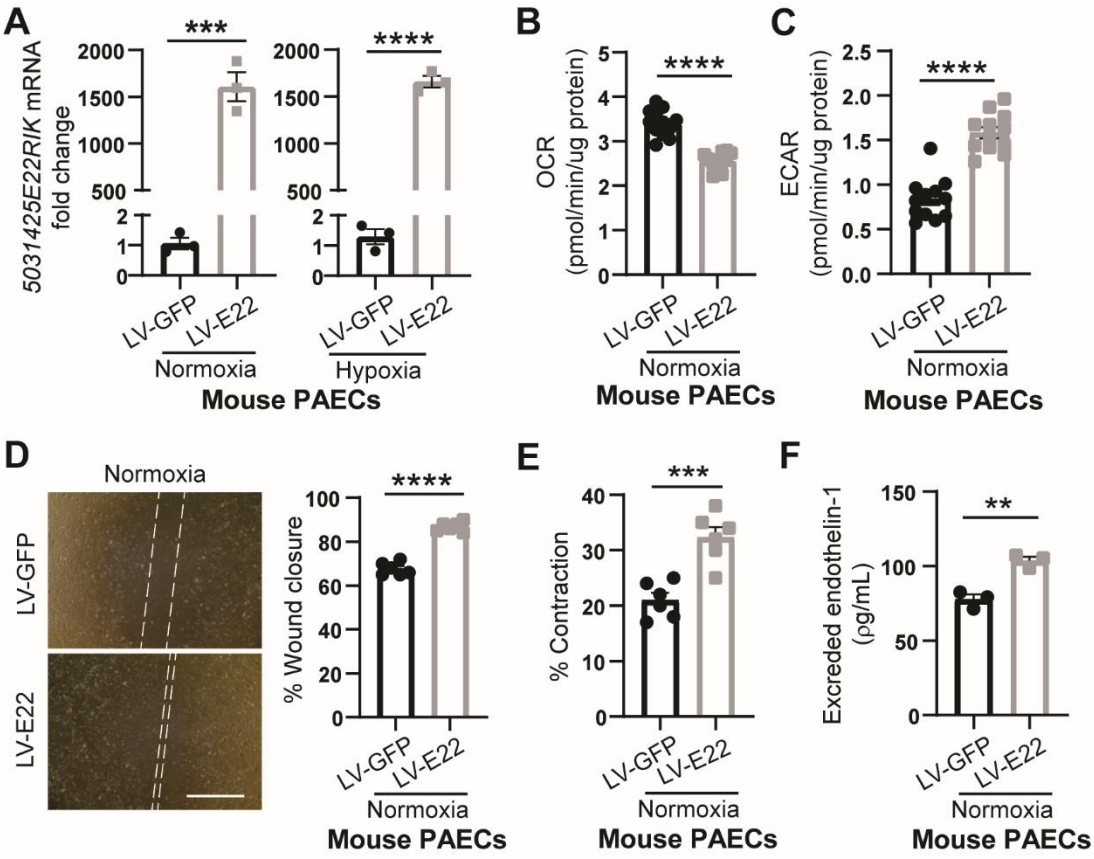

Figure S7. *E22* regulates downstream endothelial pathophenotypes in mouse PAECs. (**A**) *E22* expression is quantified in mouse PAECs after forced expression of *E22* under normoxia (left graph) and hypoxia (right graph) *(n = 3*; ***p<0.001, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). (**B-C**) Seahorse assay quantifies baseline oxygen consumption rate (OCR, B) and extracellular acidification rate (ECAR, C) of mouse PAECs after forced expression of *E22* under normoxia *(n = 12*; ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). (**D**) Scratch wound healing assay measures migration (quantified by percent wound closure) of mouse PAECs after forced expression of *E22* under normoxia *(n = 6*; ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). Scale bars, 200 μm. (**E**) PASMC contraction, quantified by % contraction, is measured in conditioned media generated from mouse PAECs with forced *E22* expression *(n = 6*; ***p<0.001, unpaired Student's t-test; data represent mean ± SEM). (**F**) Endothelin-1 is quantified in conditioned media from mouse PAECs overexpressing *E22* by ELISA *(n = 3*; **p<0.01, unpaired Student's t-test; data represent mean ± SEM).

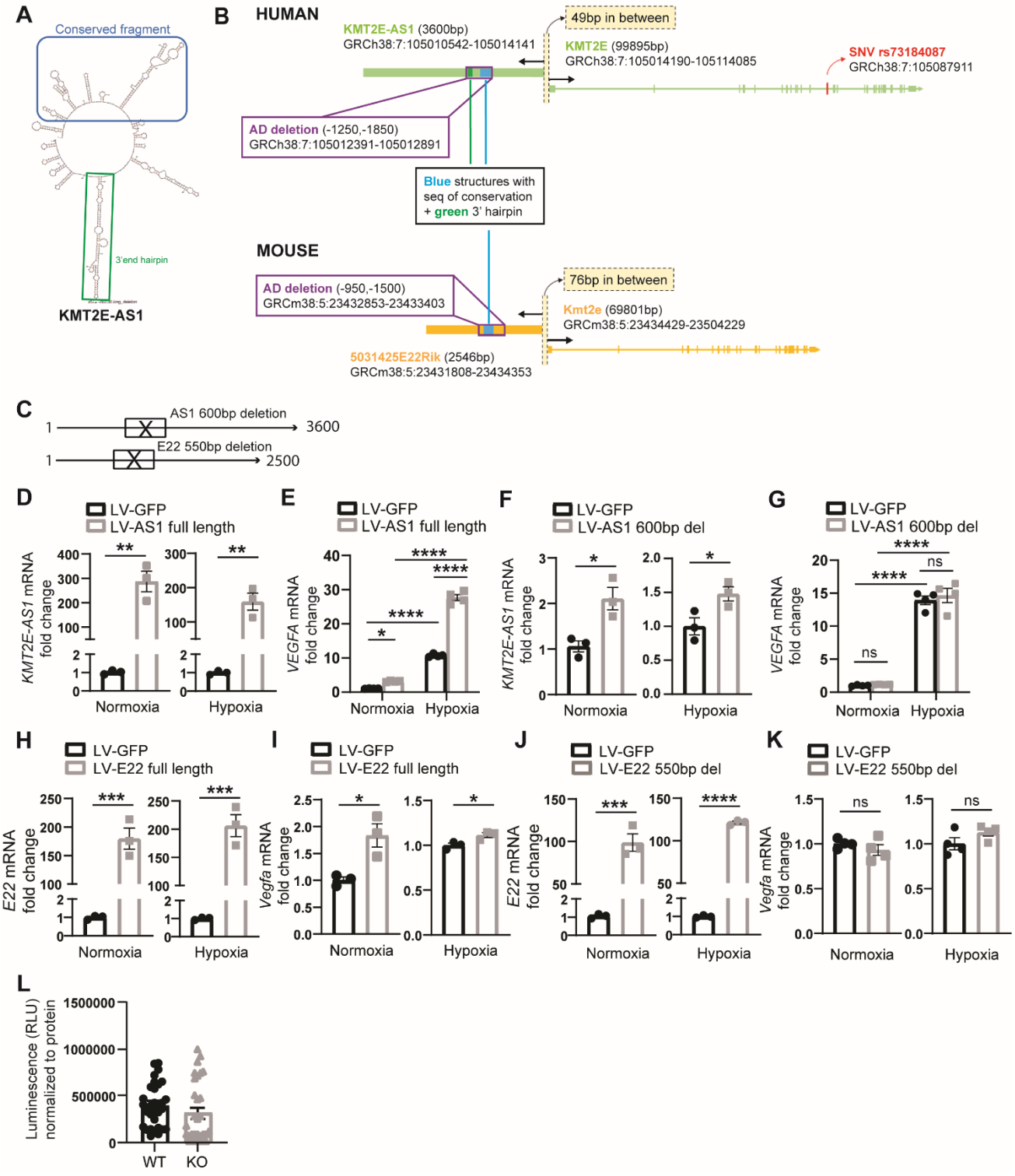

Figure S8. Deletion analysis of *KMT2E-AS1* and *E22* demonstrates that a highly conserved region is required for activity. (**A**) Secondary structure of human *KMT2E-AS1* is shown with highlighted area (blue) displaying a conserved sequence in mouse *E22*. (**B**) Overall schematic shows the chromosomal location and partial sequence conservation across human *KMT2E-AS1* and mouse *E22*. (**C-G**) *KMT2E-AS1* and *VEGFA* transcripts are quantified in human PAECs after lentiviral

delivery of *KMT2E-AS1* full length (D-E) or deletion mutant gene (F-G) missing the conserved sequence (600 bp, C) (*n = 3-4*; *p<0.05, **p<0.01, ****p<0.0001, ns denotes not significant, unpaired Student's t-test for *KMT2E-AS1* (D,F), two-way ANOVA followed by Bonferroni's *post-hoc* analysis for *VEGFA* (E,G); data represent mean ± SEM). (**H-K**) *E22* and *VEGFA* transcripts are measured in mouse PAECs after lentiviral delivery of *E22* full length (H-I) or deletion mutant gene (J-K) missing the conserved sequence (550 bp, C) (*n = 3-4*; *p<0.05, ***p<0.001, ****p<0.0001, ns – not significant, unpaired Student's t-test; data represent mean ± SEM). (**L**) Luciferase activity is quantified in HEK293T cells transfected with plasmid containing the mouse *E22-Kmt2e* promoter carrying a deletion vs. wildtype conserved *E22* sequence upstream of a *NanoLuc* luciferase reporter gene. Data are normalized to protein lysate content (*n = 31*; unpaired Student's t-test; data represent mean ± SEM).

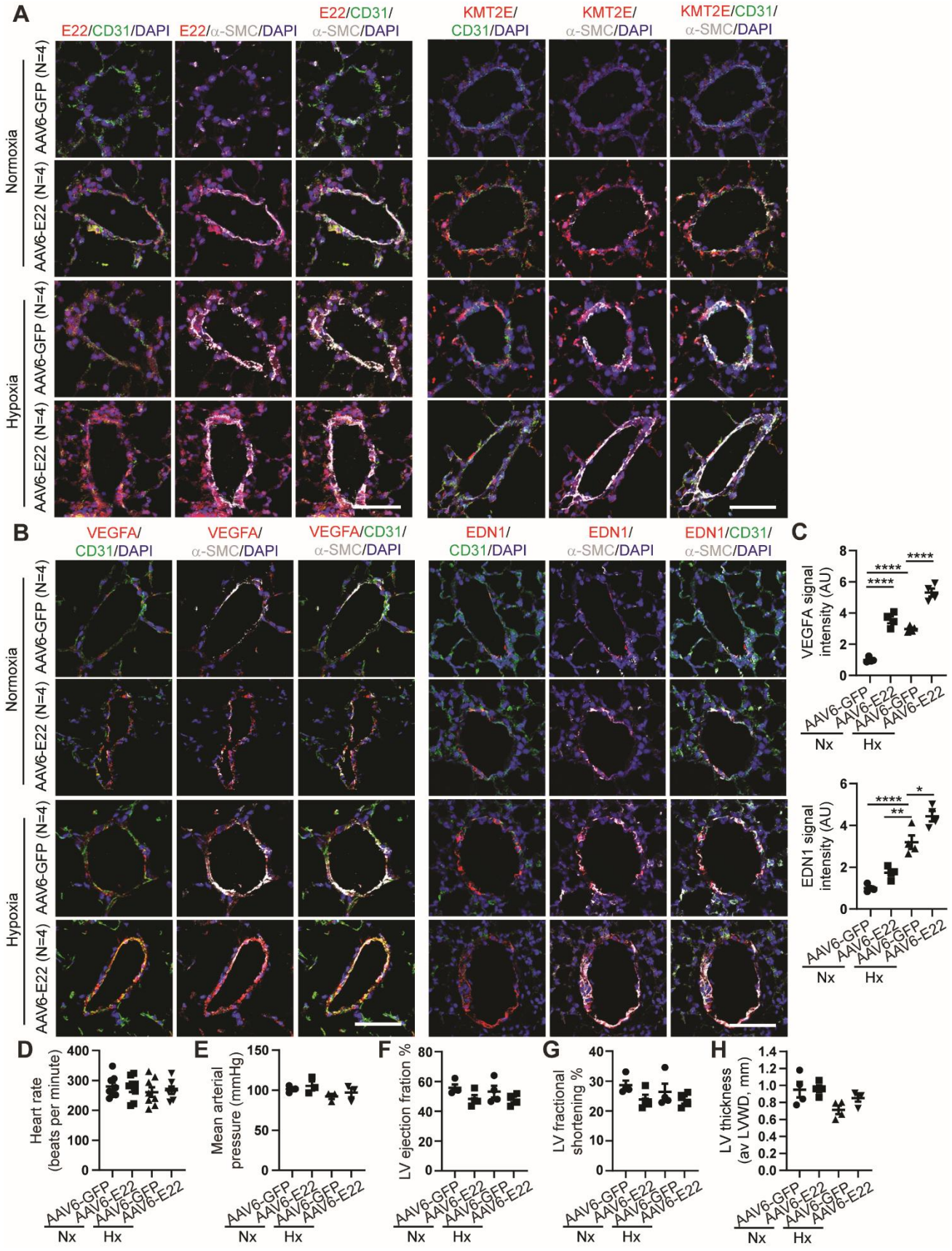

Figure S9. Forced expression of pulmonary vascular *E22* promotes PH pathophenotypes in mice. (**A**) Representative FISH and IF images are shown for *E22* (red) and KMT2E (red) in pulmonary

CD31+ endothelium of mice transduced with AAV6-*E22* vs. AAV6-*GFP* in normoxia and hypoxia. Scale bars, 50 µm. (**B-C**) Representative IF staining (B) and quantifications (C) are shown for VEGFA (red) and EDN1 (red) expression in endothelium of mice transduced with AAV6-*E22* vs. AAV6-*GFP* in normoxia and hypoxia (*n = 4*; *p<0.05, ****p<0.0001, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). Scale bars, 50 µm. (**D**) Heart rates are compared between AAV6-*E22* mice vs. AAV6-*GFP* controls (*n = 7-8*; two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). (**E**) Mean arterial pressure is shown of AAV6-*E22* vs. AAV6-*GFP* cohorts, measured by non-invasive tail-cuff plethysmography (*n = 3*; Kruskal-Wallis test followed by Dunn's *post-hoc* analysis; data represent mean ± SEM). (**F-H**) Echocardiographic measures of left ventricular (LV) ejection fraction (EF, F), fractional shortening (FS, G), and average left ventricular wall dimensions (LVWD, H) are shown for AAV6-*E22* vs. AAV6-*GFP* cohorts (*n = 4*; two-way ANOVA followed by Bonferroni's *post-hoc* analysis for LVEF (F) and LVWD (H), Kruskal-Wallis test followed by Dunn's *post-hoc* analysis for LVFS (G); data represent mean ± SEM).

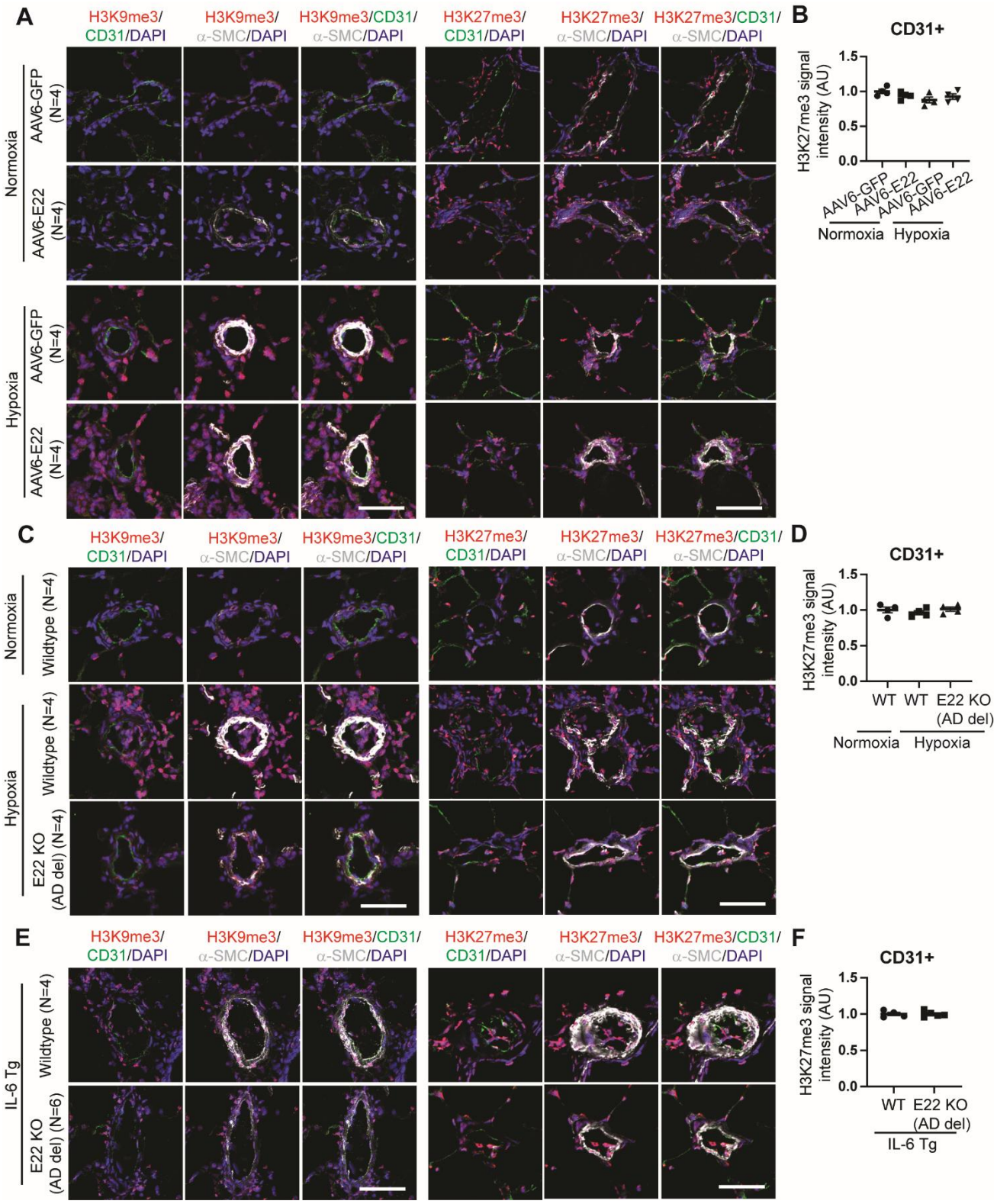

Figure S10. *KMT2E-AS1* partially controls histone 3 lysine 9 trimethylation. (**A-B**) Representative IF images are shown of H3K9me3 (red) and H3K27me3 (red) in AAV6-*E22* mice vs. AAV6-*GFP* controls (A). IF quantification of H3K27me3 expression is displayed for mouse CD31+ vascular endothelium (*n = 4*; two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data

represent mean ± SEM). Scale bars, 50 μm. (**C-D**) Representative IF images of H3K9me3 (red; C) and H3K27me3 (red; C,D) are shown with quantification in CD31+ vascular endothelial cells of *E22* deletion mice in hypoxia vs. wildtype controls in normoxia or hypoxia (*n = 4*; one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). Scale bars, 50 μm. (**E-F**) Representative IF images of H3K9me3 (red; C) and H3K27me3 (red; C,D) are shown with quantification in lung endothelium of hypoxic IL-6 Tg mice carrying *E22* deletion vs. wildtype control (*n = 4-5*; unpaired Student's t-test; data represent mean ± SEM). Scale bars, 50 μm.

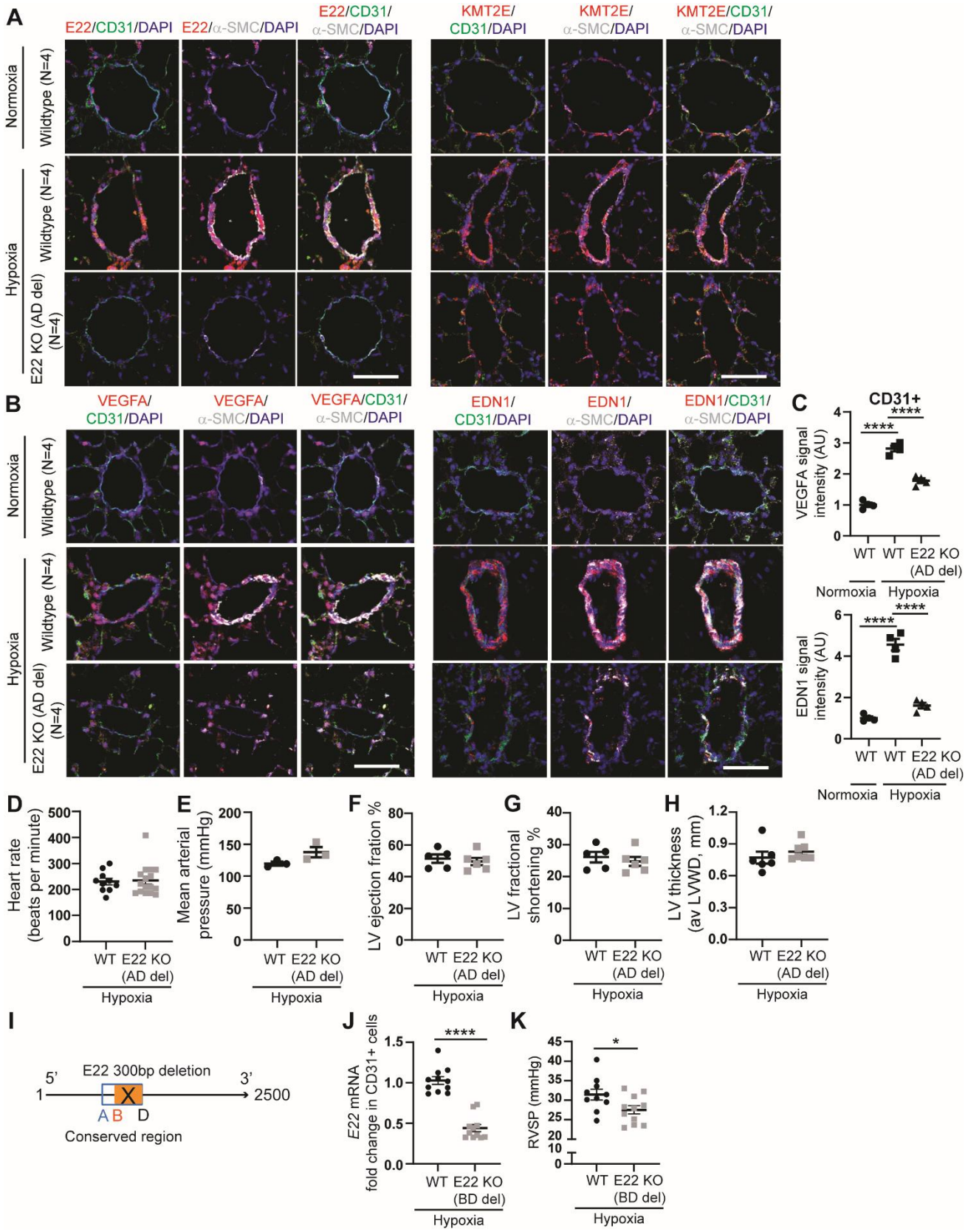

Figure S11. *E22* knockout mice display less hypoxic pulmonary vascular remodeling without affecting heart rate, systemic blood pressure, or left ventricular function. (**A**) Representative

FISH and IF staining with quantifications are shown in CD31+ vascular endothelial cells of *E22* deletion mice in hypoxia vs. wildtype controls in normoxia or hypoxia. Scale bars, 50 μm. (**B-C**) Representative IF staining (B) and quantifications (C) are shown for VEGFA and EDN1 expression in these same mouse cohorts (*n = 4*; ****p<0.0001, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). Scale bars, 50 μm. (**D**) Heart rates of E22 KO mice are compared to wildtype controls (*n = 10-16*; Mann-Whitney test; data represent mean ± SEM). (**E**) Mean arterial pressure of *E22* KO mice are compared with wildtype, as measured by non-invasive tail-cuff plethysmography (*n = 3*; unpaired Student's t-test; data represent mean ± SEM). (**F-H**) Transthoracic echocardiography measures of left ventricular (LV) ejection fraction (F), fractional shortening (G), and average left ventricular wall dimensions (av LVWD, H) in *E22* KO mice are compared with wildtype (*n = 5-6*; unpaired Student's t-test; data represent mean ± SEM). (**I**) A mouse line carrying a separate, slightly smaller deletion of the *E22* conserved region (denoted BD) was generated. (**J**) *E22* transcript expression is quantified in mouse CD31+ endothelial cells isolated from the lungs of hypoxic *E22* KO (BD deletion) mice vs. wildtype controls by RT-qPCR (*n = 11*; ****p<0.0001, Mann-Whitney test; data represent mean ± SEM). (**K**) Right ventricular systolic pressure of hypoxic *E22* mice carrying the BD deletion is compared with hypoxic wildtype PH mice (*n = 10-11*; *p<0.05, unpaired Student's t-test; data represent mean ± SEM).

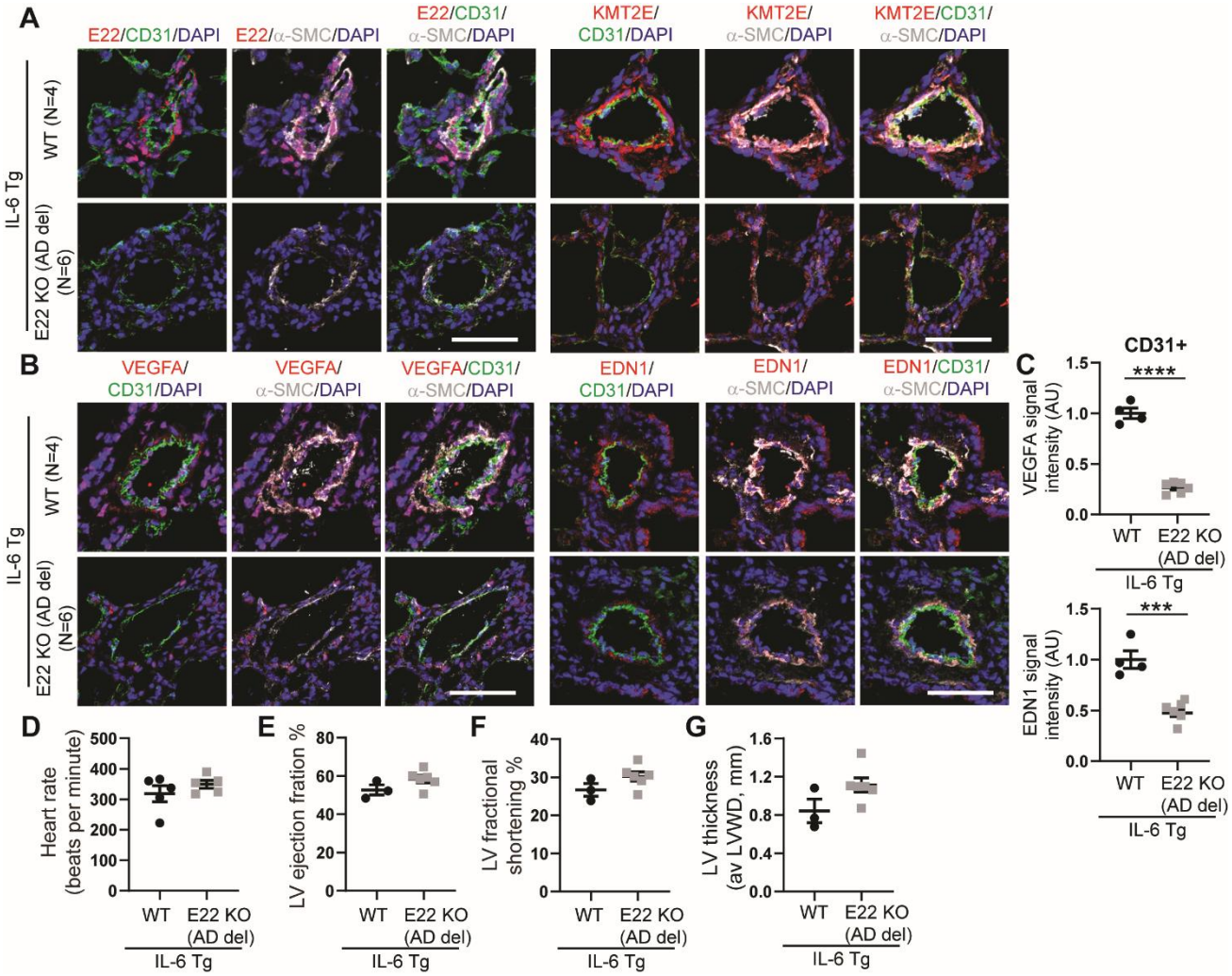

Figure S12. *E22* deletion alleviates PAH in hypoxic IL-6 Tg mice. (**A**) Representative FISH and IF staining of full-length *E22* (red) and KMT2E (red) are shown in hypoxic IL-6 Tg; *E22* KO mouse lungs vs. hypoxic IL-6 Tg wildtype (WT) controls. Scale bars, 50 μm. (**B-C**) Representative IF images (B) and quantifications (C) are displayed for VEGFA (red) and EDN1 (red) in CD31+ endothelial cells of these same mouse cohorts (*n = 4-6*; ***p<0.001, ****p<0.0001, unpaired Student's t-test; data represent mean ± SEM). Scale bars, 50 μm. (**D**) Heart rates of hypoxic IL-6 Tg; *E22* KO mice are compared with hypoxic IL-6 Tg WT controls (*n = 5*; unpaired Student's t-test; data represent mean ± SEM). (**E-G**) Echocardiographic measures of left ventricular ejection fraction (E), left ventricular fractional shortening (F), and average left ventricular wall thickness (G) are shown for hypoxic IL-6 Tg; *E22* KO mice vs. hypoxic IL-6 Tg WT controls (*n = 3-6*; unpaired Student's t-test; data represent mean ± SEM).

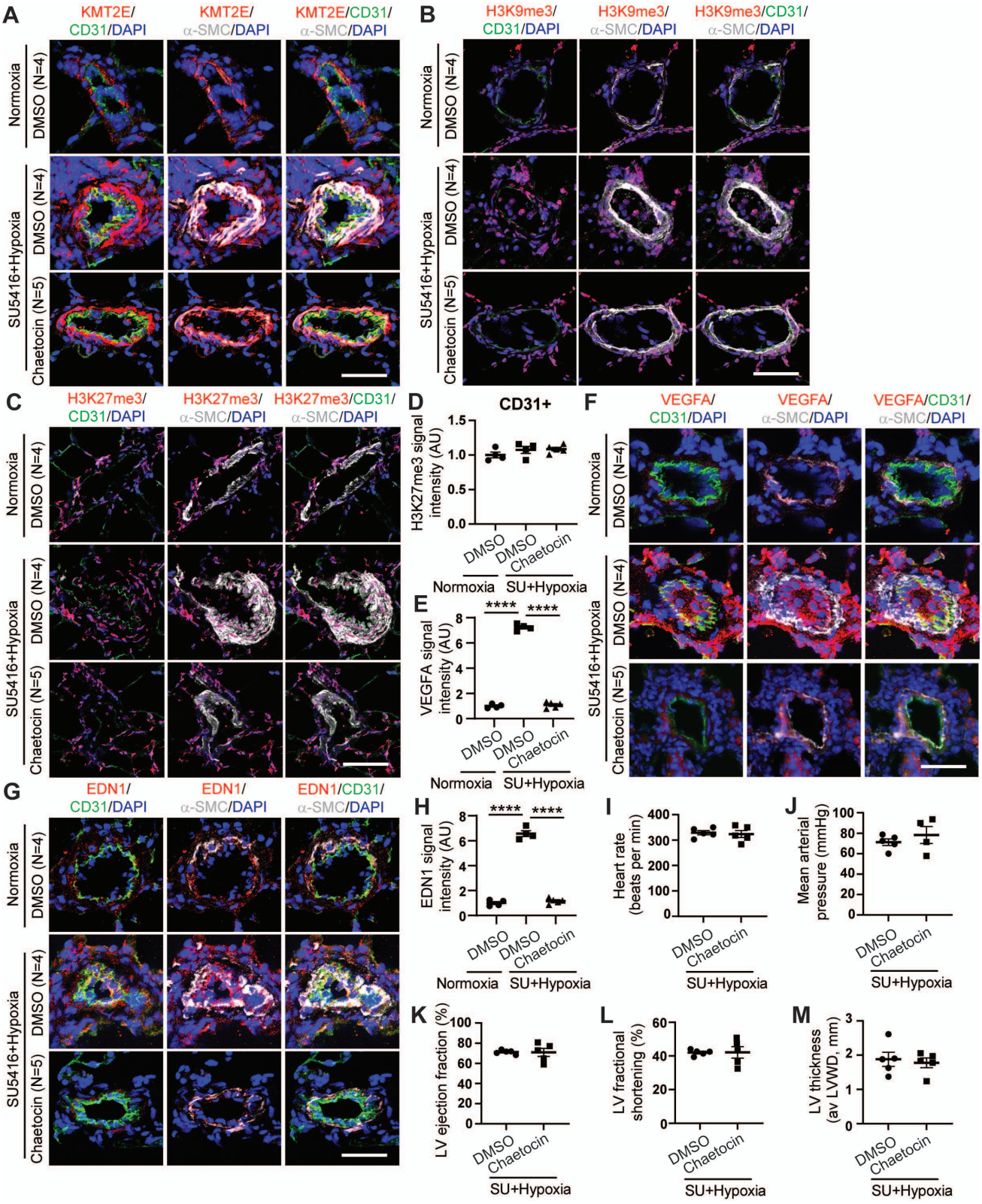

Figure S13. Chaetocin improves PAH in SU5416-hypoxic rats with existing disease. (**A**) Representative IF staining of pulmonary vascular KMT2E (red) is shown in chaetocin or vehicle-

treated SU5416-hypoxic PAH rats vs. normoxic controls. Scale bars, 50 μm. (**B-D**) Representative IF images of H3K9me3 (red; B) and H3K27me3 (red; C,D) are quantified in the same rat cohorts (*n = 4-5*; one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). Scale bars, 50 μm. (**E-H**) Representative IF images and quantifications of VEGFA (red; E,F) and EDN1 (red; G,H) in CD31+ endothelial cells are displayed (*n = 4-5*; ****p<0.0001, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean ± SEM). Scale bars, 50 μm. (**I-J**) Heart rates (I) and mean arterial pressures (J) are compared between chaetocin SU5416-hypoxic PAH rats vs. controls (*n = 4-5*; unpaired Student's t-test; data represent mean ± SEM). (**K-M**) Echocardiographic measures of left ventricular ejection fraction (K), left ventricular fractional shortening (L), and average left ventricular wall thickness (M) are displayed in chaetocin vs. vehicle-treated SU5416-hypoxic PAH rats (*n = 5*; unpaired Student's t-test; data represent mean ± SEM).

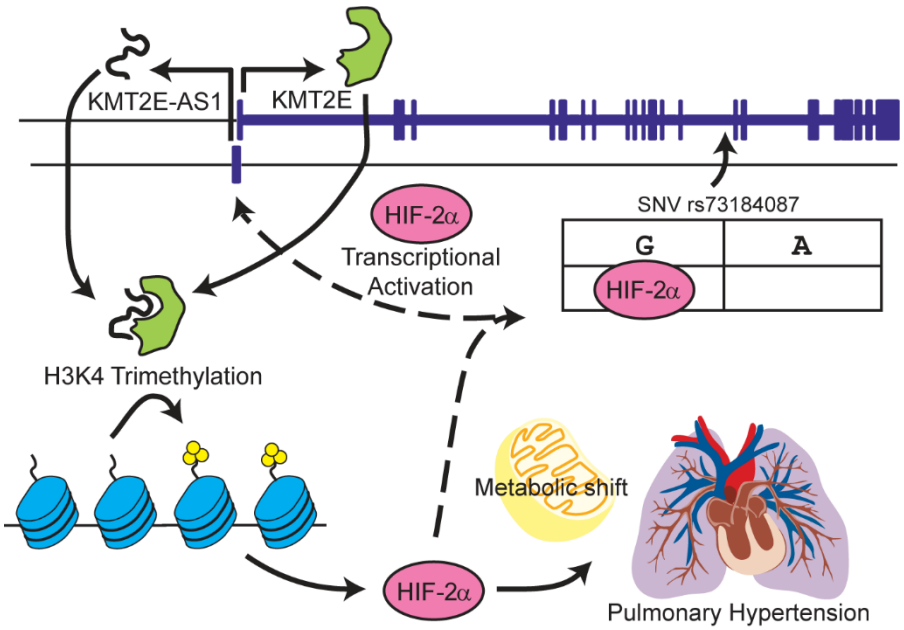

Figure S14. Model of the *KMT2E-AS1*/KMT2E gene tandem that forms a lncRNA-protein complex controlling HIF-2α-dependent endothelial reprogramming in pulmonary hypertension. Proposed mechanism employs a positive feedback loop whereby allele-specific binding of HIF-2α to rs73184087 drives *KMT2E-AS1*/KMT2E transcription and in turn promotes HIF-2α expression.

SUPPLEMENTAL TABLES

Table S1. Clinical characteristics of Group 1 and Group 3 PH patients

I. Group 1 PAH patients *Mean pulmonary arterial pressure (mPAP) #Congenital heart disease (CHD)

Age	Sex	mPAP* (mmHg)	Diagnosis	Clinical description
34	Female	50	Idiopathic	Cardiopulmonary arrest (autopsy)
64	Female	55	Idiopathic	Cardiopulmonary arrest (autopsy)
68	Female	44	Scleroderma	Bilateral lung transplant
12	Male	53	BMPR2 mutation	Bilateral lung transplant
16	Male	62	Idiopathic	Bilateral lung transplant
1	Male	50	Trisomy 21	Lung resection
19	Male	48	Idiopathic	Lung resection
42	Female	57	Scleroderma	Bilateral lung transplant
53	Male	44	Scleroderma	Bilateral lung transplant
55	Female	41	Scleroderma	Bilateral lung transplant
43	Female	35	Autoimmune disease	Bilateral lung transplant
66	Female	44	Scleroderma	Rapid autopsy
37	Female	74	CHD#	Bilateral lung transplant

II. Group 3 PH patients *Mean pulmonary arterial pressure (mPAP) #Idiopathic pulmonary fibrosis (IPF)

Age	Sex	mPAP* (mmHg)	Diagnosis	Clinical description
62	Male	28	IPF# and PH	Bilateral lung transplant
58	Male	28	IPF# and PH	Bilateral lung transplant
63	Male	27	IPF# and PH	Bilateral lung transplant
50	Male	30	IPF# and PH	Bilateral lung transplant
61	Male	37	IPF# and PH	Bilateral lung transplant
69	Female	29	IPF# and PH	Bilateral lung transplant
72	Male	46	IPF# and PH	Rapid autopsy
66	Male	34	IPF# and PH	Bilateral lung transplant
65	Male	28	IPF# and PH	Bilateral lung transplant
65	Male	31	IPF# and PH	Bilateral lung transplant
80	Female	31	IPF# and PH	Rapid autopsy

Table S2. Patients with SNV rs73184087 A/A or G/G alleles (Age and sex-matched)

	Age at blood collection	Sex		Age at blood collection	Sex
A/A	77	Male	G/G	81	Male
A/A	62	Female	G/G	66	Female
A/A	72	Female	G/G	72	Female

Table S3. Differentially expressed genes in PAECs altered in hypoxia and reversed by *KMT2E-AS1* or KMT2E knockdown (excel file).

I. All differentially expressed genes reversed by KMT2E-AS1/KMT2E under hypoxia. Values represent log2 fold change (FDR<0.05). Hx = hypoxia; Nx = normoxia; siNC = siRNA scrambled control; siAS1 = siRNA targeting *KMT2E-AS1*; siKMT2E = siRNA targeting KMT2E. See attached spreadsheet.

II. Genes categorized into different gene networks via gene set enrichment analysis. See attached spreadsheet.

III. Genes under "Hypoxia" gene network ranked by log2foldchange. Genes that are also H3K4 trimethylated under hypoxia vs. normoxia were denoted. See attached spreadsheet.

IV. Genes under "Metabolism" gene network ranked by log2foldchange. Genes that are also H3K4 trimethylated under hypoxia vs. normoxia were denoted. Orange genes denoted genes related to energy generation and TCA cycle. See attached spreadsheet.

Table S4. H3K4 methylated genes in PAECs under hypoxia.

***I. H3K4me3 ChIP sequencing of human PAECs under hypoxia vs. normoxia*.** Fold = log2foldchange. FDR = adjusted p value. See attached spreadsheet.

II. Subcohort of hypoxia-dependent metabolic genes under control of H3K4me3.

Gene ID	**H3K4me3 *"Hypoxia vs. Normoxia"* fold change**	**H3K4me3 *"Hypoxia vs. Normoxia"* padj**	**RNAseq *"Hypoxia si-NC vs. Normoxia si-NC"* log2foldchange**
MECOM	1.940972529	0.016362898	0.114793073
MYOF	2.01404911	0.007115888	0.124284833
GALNT10	2.569000321	6.43375E-05	0.127253476
PIGS	1.874313201	0.039014383	0.142786609
STARD13	1.862823069	0.026237791	0.155465015
CMIP	1.967802881	0.016028465	0.165990964
ZNF333	1.803694464	0.047311468	0.218903231
ARAP3	1.919175831	0.018872889	0.233830301
MIRLET7BHG	1.963450968	0.016321915	0.250558672
IL32	1.922865972	0.02309739	0.257397849
SLC43A1	1.993594713	0.010829845	0.262101346
ADAM9	2.169664673	0.002727836	0.266269007
TRIOBP	2.147716224	0.003738049	0.283383459
CTNNB1	2.023390481	0.010829845	0.33689728
TMEM102	2.697042159	1.07619E-05	0.346313477
CUBN	2.298924613	0.000612	0.349954036
CTNNAL1	2.342181116	0.000457832	0.375691025
MMP2	2.051645395	0.006578295	0.39237532
PDLIM5	1.97301974	0.012482382	0.401380339
ESYT2	1.8320173	0.044581072	0.404349096
F2RL2	1.93259312	0.01633751	0.404763727
MIR3667HG	1.935572358	0.017683858	0.430580966
PFKFB3	2.070308436	0.005263457	0.475885676
DLL4	1.817638275	0.044148419	0.490466905
SIRPB2	1.840561014	0.037395293	0.498858468
CTDSP1	2.403971454	0.00036728	0.513250197
PXN	3.42541904	1.13124E-08	0.529357596
SNX18	2.264512444	0.000818972	0.538706961
CERK	1.842833958	0.041440138	0.545718274
TXNRD2	1.921135742	0.016458754	0.581608186
SIDT2	1.941765525	0.018197294	0.60987071
SAMD4A	1.9201671	0.01633751	0.615265264
RRAS	2.086539103	0.004340781	0.615646843
AKAP12	2.250058047	0.000979066	0.61715088
GAS6	1.944798014	0.016458754	0.654999285
LINC00607	2.205604236	0.00093053	0.6979534
MAMDC4	1.855485917	0.037395293	0.704142182
BCAR1	1.990950071	0.009718354	0.747670514
XXYLT1-AS2	1.84960105	0.036677099	0.756580501
RIMKLB	3.069850667	2.97358E-07	0.757292577
NCKAP5	1.890268201	0.030013693	0.770773929

RUSC2	1.906698561	0.020739908	0.785872938
DNMBP	2.455426278	0.000137	0.791516595
RFX2	1.997210472	0.010327492	0.866462142
SPACA6	2.615480612	2.34249E-05	0.913470897
TMCC1	3.90742558	1.21054E-10	0.92738352
GAPDH	1.850240463	0.042826873	0.931100202
SORBS2	1.843326773	0.048403633	0.993967187
KANK1	2.470732397	0.000110484	1.006003355
RASSF7	2.087177305	0.006769324	1.02650165
CD69	2.261364634	0.001343662	1.048730018
ENG	1.958334845	0.012925652	1.087460879
CCNY	3.139266745	1.4473E-07	1.089786155
BICD1	2.278867189	0.000832875	1.091726348
B3GNT8	1.893509342	0.027994164	1.112994655
TCIM	2.399795855	0.000335733	1.19210265
FAM13A	3.678845572	8.096E-10	1.250798371
WNK4	2.718222441	6.2454E-06	1.255518631
CCDC74B	1.825477154	0.044544737	1.391307549
BHLHE40	2.516573249	0.000110484	1.405853268
CHRNB1	2.299768004	0.000612	1.406574266
ALDOA	2.053185334	0.006476724	1.437606819
ICOSLG	3.114888108	1.92708E-07	1.522583018
FSTL3	1.963513324	0.01633751	1.614514425
MAFK	3.2346389	5.30191E-08	1.645237503
GABRE	1.869263035	0.039224828	1.743816994
ITGA11	2.381874643	0.000268294	1.753455612
HES4	1.809134839	0.041440138	1.855901254
BNIP3	2.415801487	0.000294117	1.880592948
TNS1	4.028186695	4.17193E-11	1.905085008
VWA1	1.955090493	0.015212387	1.927199853
SLC11A1	2.403971454	0.00036728	2.039830295
ADRA1D	1.876278561	0.026237791	2.062917835
GADD45B	2.058564983	0.008482586	2.066908087
GDF6	2.195419498	0.002320067	2.067766267
ERRFI1	2.118019766	0.004928203	2.080664491
LIMS2	2.01002197	0.012334995	2.251768854
DNAH11	1.923577351	0.023513406	2.26832412
ALDOC	1.874313201	0.039014383	2.278157185
BDKRB2	2.049086179	0.007115888	2.330153952
TAGLN	1.941765525	0.018197294	2.348212284
GALNT15	2.496783409	0.000144262	2.376944412
STC1	2.867095098	2.33655E-06	2.386345632
PTPRR	1.966826155	0.01633751	2.400576179
MXI1	1.949805302	0.02077683	2.42427292
SH3D21	1.95515657	0.018872889	2.484341246
FGF11	2.697042159	1.07619E-05	2.640472843
DNAH8	2.365095776	0.00047337	2.817128035
TMEM184A	3.2346389	5.30191E-08	2.961144361
HIF3A	2.156686179	0.002767442	3.015668524
ANKRD37	2.171497035	0.00282431	3.197849657

KCTD16	2.128154826	0.004529566	3.326989809
TGFBI	1.845736165	0.043350151	3.441570074
INHBA	2.803941152	6.2454E-06	3.561354963
CLEC3B	1.856386795	0.043553818	3.603792088
ANGPTL4	2.131434956	0.004601954	3.623350556
RGS9	1.799793556	0.04748274	3.69334356
SYTL2	2.036601774	0.007427848	3.805107131
ADSS1	1.849624643	0.047311468	3.845723983
SLC8A3	2.541754308	7.85865E-05	4.408191612
MCHR1	2.309979904	0.000818972	4.409006494
PTGIS	1.906480934	0.029533183	4.973155943
MIR210HG	2.087177305	0.006769324	5.674122869
EGLN3	2.569244611	8.51716E-05	6.332289481
MIR210	2.087177305	0.006769324	6.781817119
SLC9C2	3.042561693	5.4805E-07	8.111519218

Table S5. PAH Biobank Cohort and SNV association.

I. Demographics

Age at consent	58 [45.6 – 67.4]
Age at catheterization	52.3 [38.7 -62.7]
Mean right atrial pressure (mmHg)	8 [5 - 12]
Mean pulmonary artery pressure (mmHg)	51 [42 - 60]
Pulmonary capillary wedge pressure (mmHg)	10 [7 - 13]
Cardiac output (L/min)	4.3 [3.4 - 5.2]
Cardiac index (L/min/kg/m2)	2.3 [1.8 - 2.8]
Pulmonary vascular resistance (Wood units)	9.8 [6.5 - 14.1]
Functional class, n [%]	
I	22 [5.3%]
II	118 [28.6%]
III	243 [58.8%]
IV	30 [7.3%]

count [%] or median [interquartile range]

II. SNV association with PAH risk (694 cases and 1,560 controls) *Adjusted for sex, age and 2 principal components (PCs).

SNV	MAF in cases	MAF in controls	OR (95% C.I.)*	P
rs73184087 (104728358, G)	0.044	0.025	1.86 (1.30-2.67)	7.40×10^{-4}

Table S6. SNVs within +/-200kb of *KMT2E-AS1*/KMT2E with predicted HIF-2α binding to one of either the minor or major SNV alleles. Change in HIF-2α binding score denotes the maximum difference in log-odds scores between the major and minor alleles among the predicted HIF-2α binding motifs; positive values indicate greater predicted HIF-2α binding to the effect allele. P-values via Firth's penalized logistic regression are listed for the association of SNVs with PAH in the discovery cohort. Red font denotes statistical significance with significant threshold of 0.00093 (= 0.05/53.84 where 53.84 is the effective SNV test count).

RSID	Chr	Position (hg38)	Major Allele	Minor Allele	Change in HIF-2α Binding Score	Effect Allele Frequency (Cases)	Effect Allele Frequency (Controls)	P-value
rs73184087	7	105087911	A	G	+5.45	0.0438	0.0247	0.00074
rs62485395	7	104859691	A	G	-4.72	0.2140	0.1830	0.00782
rs56016333	7	104955801	T	C	-19.21	0.3611	0.3413	0.08937
rs11978074	7	104920506	A	T	-15.19	0.4666	0.4503	0.14033
rs117871677	7	105086107	G	A	+3.13	0.0253	0.0333	0.16091
rs11762333	7	104927430	G	A	+2.00	0.0215	0.0247	0.17418
rs10262447	7	104964657	C	G	+0.89	0.1917	0.2048	0.21944
rs141376909	7	105100573	T	C	-8.68	0.0126	0.0096	0.24669
rs35110258	7	104921318	A	T	-3.32	0.0438	0.0494	0.26499
rs145405430	7	104815994	T	A	+4.32	0.0149	0.0192	0.28139
rs74365863	7	105011110	C	G	-1.68	0.0297	0.0343	0.29750
rs35338373	7	105175707	G	A	+7.01	0.0468	0.0410	0.30425
rs112228846	7	104824050	G	A	-10.39	0.0282	0.0234	0.34024
rs12534673	7	104912884	A	G	-8.71	0.1590	0.1494	0.34647
rs7795893	7	104927852	T	C	-7.68	0.3603	0.3503	0.38085
rs4730076	7	105310738	C	A	+12.29	0.1449	0.1506	0.40399
rs4727614	7	104960407	C	A	+13.24	0.3388	0.3529	0.41041
rs10243336	7	104854113	C	A	-11.13	0.0134	0.0160	0.50380
rs149334360	7	105220100	G	C	-9.69	0.0111	0.0115	0.54641
rs143779419	7	105231708	T	C	-6.53	0.0208	0.0224	0.61732
rs144616964	7	105220106	A	C	-4.32	0.0208	0.0224	0.61732
rs1006365	7	105128172	T	A	-1.32	0.4673	0.4731	0.63196
rs186001710	7	105200541	T	A	-3.32	0.0119	0.0160	0.64014
rs10273609	7	104882359	C	T	+10.39	0.0201	0.0183	0.66020
rs4073894	7	104826517	G	A	+10.88	0.1991	0.2083	0.66046
rs185923695	7	105027040	T	C	-2.26	0.0171	0.0183	0.66249
rs4730061	7	104939329	C	A	-11.10	0.0520	0.0478	0.68229
rs10808141	7	105075043	T	G	+14.97	0.3574	0.3657	0.69586
rs7796945	7	105197603	A	G	-7.35	0.4688	0.4747	0.71508
rs10238507	7	105166760	A	C	-3.32	0.3462	0.3535	0.76523
rs6466056	7	105288820	C	T	+3.58	0.3455	0.3494	0.76902
rs10261472	7	105021865	A	G	-2.00	0.1241	0.1186	0.78316
rs2216107	7	105015595	C	G	+9.30	0.1241	0.1186	0.78316
rs7790653	7	105291098	C	A	+3.13	0.4688	0.4737	0.80030
rs2470934	7	104904501	A	G	-6.34	0.0565	0.0583	0.81750
rs11761429	7	105302678	T	A	-9.38	0.1724	0.1788	0.81965

rs9918494	7	105041202	T	A	-20.48	0.4799	0.4837	0.82514
rs9641340	7	105074490	A	C	-8.24	0.1226	0.1183	0.82988
rs6954777	7	105285474	A	G	-7.68	0.4703	0.4737	0.84260
rs10244764	7	105045870	T	A	-7.59	0.1233	0.1189	0.84278
rs4730067	7	105041801	T	C	-13.76	0.1233	0.1189	0.84278
rs960437	7	105049295	T	C	+8.93	0.1233	0.1189	0.84584
rs6955753	7	105259117	G	T	+4.73	0.4710	0.4744	0.85256
rs6968335	7	105258269	T	C	+6.79	0.4710	0.4744	0.85256
rs10255779	7	105298651	G	A	+2.35	0.1293	0.1285	0.85429
rs9641345	7	105298427	A	C	-5.39	0.1293	0.1285	0.85429
rs4602821	7	104971429	C	T	+6.76	0.4755	0.4782	0.87893
rs7808168	7	105194615	G	A	+2.35	0.1256	0.1208	0.87990
rs74776606	7	105254673	T	G	-12.42	0.0713	0.0728	0.89910
rs118043572	7	104904771	C	T	+7.92	0.0126	0.0144	0.90967
rs2470937	7	104939328	T	A	-11.10	0.4599	0.4628	0.91239
rs13230660	7	105010611	A	G	-4.21	0.1218	0.1183	0.91616
rs2237608	7	105155127	C	T	-13.16	0.1159	0.1131	0.91965
rs10280470	7	105162515	G	A	+7.53	0.1233	0.1208	0.93402
rs10230449	7	105258944	A	G	-3.54	0.1293	0.1292	0.95063
rs3801283	7	105240836	G	A	+12.12	0.1293	0.1292	0.95063
rs77383660	7	105102232	A	C	-8.33	0.0215	0.0215	0.96603
rs2237616	7	105230525	G	T	+12.32	0.1293	0.1295	0.97819
rs2240465	7	105142859	G	T	+4.73	0.1315	0.1282	0.98950

Table S7. UPMC Cohort demographics and SNV association.

I. Demographics

Age at catheterization	49 [36 – 60]
Mean right atrial pressure (mmHg)	9 [5 – 12.3]
Mean pulmonary artery pressure (mmHg)	52 [41.8 – 61.3]
Pulmonary capillary wedge pressure (mmHg)	10 [7.3 – 14]
Cardiac output (L/min)	4.3 [3.4 – 5.4]
Cardiac index (L/min/kg/m2)	2.3 [1.9 – 2.8]
Pulmonary vascular resistance (Wood units)	9.8 [6 – 13.1]

Median [interquartile range]

II. SNV association (96 cases and 401 controls)

rs73184087 (104728358, G)	MAF in cases	MAF in controls	Unadjusted OR (95%CI, P value)	Age adjusted OR (95%/CI, P value)	Sex, age adjusted OR (95%CI, P value)
All (N = 497)	0.073	0.034	2.39 (1.21-4.72, 0.012)	2.38 (1.21-4.70, 0.012)	2.53 (1.25-5.13, 0.01)

Table S8. All of Us demographics and SNV association.

I. Demographics *Includes Heritable Pulmonary Arterial Hypertension, Idiopathic Pulmonary Arterial Hypertension, and Secondary PAH.

	Cases* (N = 52)	Controls (N = 11,821)
Age at Diagnosis	70.6 [62.6 – 75.6]	63.6 [46.6 – 72.6]
Sex at birth		
Male	17 (32.7%)	3,709 (31.4 %)
Female	33 (63.5%)	7,954 (67.3 %)
Other	2 (3.8%)	158 (1.3 %)

II. SNV association with PAH risk (52 cases and 11,821 controls) *Adjusted for sex and age.

SNV	MAF in cases	MAF in controls	OR (95% C.I.)*	P
rs73184087 (104728358, G)	0.077	0.034	2.44 (1.25-4.79)	0.0116

Table S9. Key resources.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
HIF-1α	Novus Biologicals	NB100-134
HIF-2α	Novus Biologicals	NB100-122
PARP-1	Novus Biologicals	NBP2-13732
VEGF	ABCAM	ab183100
H3K4me3	ABCAM	ab8580
H3K4me3	Active Motif	39915
H3K9me3	ABCAM	ab8898
H3K27me3	ABCAM	ab6002
KMT2E (C-10)	Santa Cruz Biotechnology	sc-377182
Histone	Cell Signaling Technology	9715s
β-actin	Santa Cruz Biotechnology	sc-47778
Ki-67	ABCAM	ab15580
α-SMA	ABCAM	ab32575
α-SMA	ABCAM	ab21027
α-SMA	Sigma Aldrich	A5228
EDN1	Novus Biologicals	NB300-526
CD31	Santa Cruz Biotechnology	sc-1506
CD31	Novus Biologicals	AF3628
Digoxigenin	Novus Biologicals	BAM7520
Rabbit IgG Isotype Control	Thermo Fisher Scientific	02-6102
Rabbit IgG	Diagenode	C15410206
Mouse IgG Isotype Control	Thermo Fisher Scientific	02-6502
KMT2E-AS1 Custom LNA Detection Probe	Qiagen	N/A
E22 Custom LNA Detection Probe	Qiagen	N/A
NANOG	Cell Signaling Technology	4903S
SSEA4	Cell Signaling Technology	4755S
OCT4A	Cell Signaling Technology	2840S
SOX2	Cell Signaling Technology	4900S
Nestin	Santa Cruz Biotechnology	sc-23927
Brachyury	ABCAM	ab209665
SOX17	Cell Signaling Technology	81778S
FITC Mouse Anti-Human CD31	BD Biosciences	560984
PE Mouse anti-Human CD144	BD Biosciences	560410
AlexaFluor594 Goat anti-Rabbit IgG (H + L)	Thermo Fisher Scientific	A11012
Goat Anti-Mouse IgG H&L (Alexa Fluor® 488)	ABCAM	ab150113
Bacterial and virus strains		
plenti-CDH-*mE22*-copGFP in Stellar Cell	This paper	N/A
plenti-CDH-AS1 in Stellar Cell	This paper	N/A

plenti-CDH-AS1 with 600 bp deletion in Stellar Cell	This paper	N/A
Biological samples		
Human lung samples	University of Pittsburgh Medical Center; the New England Organ Bank; Boston Children's Hospital	N/A
Human transformed lymphocytes with A/A or G/G alleles	Indiana University School of Medicine	N/A
Human PBMC DNA	University of Pittsburgh Medical Center	N/A
Chemicals, peptides, and recombinant proteins		
MG132	Sigma Aldrich	M7449
Actinomycin D	Sigma Aldrich	A9415
Cobalt (II) chloride	Sigma Aldrich	232696
Protease inhibitor cocktail	Sigma Aldrich	P8340
Recombinant human IL-1β	PeproTech	200-01B
Recombinant human IL-6	PeproTech	200-06
Laduviglusib (CHIR-99021) HCl	Selleckchem	S2924
Human Recombinant VEGF-165	Stemcell Technologies	78073.1
Animal-Free Recombinant Human FGF-basic (154 a.a.)	Peprotech	AF-100-18B
SB431542	Selleckchem	S1067
Commercial assays		
Duolink PLA kit	Sigma Aldrich	DUO92102-1KT
EZ-Magna RIP™ RNA-Binding Protein Immunoprecipitation Kit	Millipore Sigma	17-701
Pierce magnetic ChIP kit	Thermo Fisher Scientific	26157
NE-PER nuclear and cytoplasmic extraction kit	Thermo Fisher Scientific	78835
In vitro angiogenesis assay kit	Cultrex	3470-096-K
Caspase-GLO 3/7 assay kit	Promega	G8090
BrdU cell proliferation assay kit	Cell Signaling Technology	6813
Endothelin-1 ELIZA kit	Enzo	ADI-900-020A
Lactate dehydrogenase activity colorimetric assay kit	BioVision	K726-500
miRCURY LNA miRNA ISH Optimization Kit (FFPE) 5	Qiagen	339455
TSA Plus Cyanine 3 System, for 25-75 Slides	Perkin Elmer	NEL744E001KT
Q5® Site-Directed Mutagenesis Kit	New England Biolabs	E0554S
pPACK-ID: Integrase-defective lentiviral packaging mix	System Biosciences	LV520A-ID
Dual-Glo Luciferase assay	Promega	N1610
Secrete-Pair Dual Luminescence Assay kit	GeneCopoeia	LF031
Neon™ Transfection System 10 µL Kit	Thermo Fisher Scientific	MPK1096
STEMdiff Trilineage differentiation kit	Stemcell Technologies	05230
RNeasy mini kit	Qiagen	74106
Deposited data		
RNA sequencing data (SU5416 hypoxic mice vs. control)	Bertero et al, 2015	GEO: GSE61828
RNA sequencing data (si-*KMT2E-AS1* vs. si-KMT2E vs. control)	This paper	GEO: GSE232799
ChIP sequencing data (H3K4me3 IP, Hypoxia vs. Normoxia)	This paper	GEO: GSE232799

Experimental models: Cell lines		
Human pulmonary arterial endothelial cells	Lonza	CC-2530
Human pulmonary arterial adventitial fibroblasts	ScienCell	3120
Human pulmonary arterial smooth muscle cells	Lonza	CC-2581
Mouse pulmonary arterial endothelial cells	Cell Biologics	C57-6059
Mouse embryonic fibroblasts	ATCC	PTA-9386
HEK293FT	Thermo Fisher Scientific	R70007
Experimental models: Organisms/strains		
Mouse: C57BL/6J	Jackson Laboratory	000664
Mouse: C.B6-Tg(H2-L-IL6)1Kish/J	Jackson Laboratory	008372
Mouse: CRISPR-Cas9 *5031425E22Rik* knockout	This paper	N/A
Mouse: IL6-Tg x CRISPR-Cas9 *5031425E22Rik* knockout	This paper	N/A
Rat: Sprague-Dawley	Charles River	CD IGS rat
Oligonucleotides		
miR210hg_Promoter_F: TGGTACCCTGGACACACAAG	This paper	N/A
miR210hg_Promoter_R: GTCATTTGCATACGGGGCTG	This paper	N/A
SNV_rs73184087_ChIP_F: GCAATTCGATTCAGTTCCTCTT	This paper	N/A
SNV_rs73184087_ChIP_R: AAAGGAAAAGGGATTTCACCA	This paper	N/A
Q5_AS1 Del_1150-1750_F: TCAAGGAGGATCGATTACTTTCTG	This paper	N/A
Q5_AS1 Del_1150-1750_R: TGGGTTCCATATCGGCCG	This paper	N/A
Q5_E22 Del_950-1500_F: ACTCCTTTCGGGGTGTGC	This paper	N/A
Q5_E22 Del_950-1500_R: CGCCCGACAGAAATCACC	This paper	N/A
AS1_Lenti_F: AGATTCTAGAGGTACGCCCGTCC	This paper	N/A
AS1_Lenti_R: CCTTTCTCCCCGCAAAGTATACTTAT	This paper	N/A
E22_Lenti_F: ACTCCGTCTTTCCCACAACA	This paper	N/A
E22_Lenti_R: TGGGCAGGAGTATGAGTTCC	This paper	N/A
Rat E22 1252_F: TAACTGCGTCAGGGCAAGAT	This paper	N/A
Rat E22 1252_R: GGAAGCGCTTTCAAACAGTC	This paper	N/A
Rat_Actin_F: AGCCATGTACGTAGCCATCC	This paper	N/A
Rat_Actin_R: CTCTCAGCTGTGGTGGTGAA	This paper	N/A
TaqMan Gene Expression Assay - *KMT2E-AS1* Human (Hs04233055_s1)	Life Technologies	4331182
TaqMan Gene Expression Assay - KMT2E Human (Hs01096121_m1)	Life Technologies	4331182
TaqMan Gene Expression Assay – ELOC Human (Hs00855349_g1)	Life Technologies	4331182
TaqMan Gene Expression Assay – EPAS1 Human (Hs01026149_m1)	Life Technologies	4331182

TaqMan Gene Expression Assay – mir210hg Human (Hs05030902_s1)	Life Technologies	4331182
TaqMan Gene Expression Assay – VEGFA Human (Hs00900055-m1)	Life Technologies	4331182
TaqMan Gene Expression Assay - Actin Human (Hs00157387_m1)	Life Technologies	4331182
TaqMan Gene Expression Assay - *5031425E22Rik* Mouse (Mm02600639_s1)	Life Technologies	4331182
TaqMan Gene Expression Assay – Kmt2e Mouse (Mm01129499_m1)	Life Technologies	4351372
TaqMan Gene Expression Assay – Vegfa Mouse (Mm00437306_m1)	Life Technologies	4331182
TaqMan Gene Expression Assay - Actin Mouse (Mm02619580_g1)	Life Technologies	4331182
Stealth RNAi Negative Control LO GC	Life Technologies	12935200
Human *KMT2E-AS1* siRNA (n263822)	Life Technologies	4390771
Human KMT2E siRNA (132615)	Life Technologies	AM16708
Human HIF-1α siRNA	Santa Cruz Biotechnology	sc-35561
Human EPAS-1/HIF-2 alpha siRNA	Santa Cruz Biotechnology	sc-35316
Control siRNA	Santa Cruz Biotechnology	sc-37007
KMT2E SNP rs73184087 (C__97341322_10)	Life Technologies	4351379
Gapmer hG462: CAACAGAAAGAAAGCA	This paper	N/A
Recombinant DNA		
plenti-CDH-AS1 with 600 bp deletion	This paper	N/A
plenti-CDH-AS1	This paper	N/A
pCDH-CMV-MCS-EF1-copGFP	System Biosciences	CD511B-1
pLenti-HIF-2α	University of North Carolina, Chapel Hill	N/A
pLenti-HIF-1α	Harvard Medical School	N/A
pCDH-CMV-Nluc	Addgene Plasmid	73038
pEZX-PG04.1	GeneCopoeia	HPRM49354-PG04
pCXLE-hOCT3/4-shp53	Addgene Plasmid	27077
pCXLE-hSK	Addgene Plasmid	27078
pCXLE-hUL	Addgene Plasmid	27080
Software and algorithms		
ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/
PRIdictor	Tuvshinjargal et al., 2016	http://bclab.inha.ac.kr/pridictor/
MEME Suite	Bailey et al., 2006	https://meme-suite.org/meme/
Biopython v1.78	Cock et al., 2009	https://biopython.org/wiki/Download
TFMPvalue	Touzet and Varre, 2007	https://cran.rstudio.com/web/packages/TFMPvalue/index.html
Mfold RNA Folding Form V2.3	Zuker M., 2003	https://www.unafold.org/mfold/applications/rna-folding-form-v2.php

Data File S1. Raw data for all experiments where n<20. See attached excel file.