

Allele-specific and hypoxic control of the lncRNA *KMT2E-AS1*/KMT2E axis reprograms the endothelium in pulmonary hypertension

Yi-Yin Tai^{1,3,4,6,#}, Qiujun Yu^{9,#}, Ying Tang^{1,3,4,6}, Wei Sun^{1,3,4,6}, Neil J. Kelly^{1,3,4,6,10}, Satoshi Okawa^{3,4,6,7,8}, Jingsi Zhao^{1,3,4,6}, Tae-Hwi Schwantes-An^{11,12}, Caroline Lacoux¹³, Stephanie Torino¹³, Yassmin Al Aaraj^{1,3,4,6}, Wadih El Khoury^{1,3,4,6}, Vinny Negi^{1,3,6}, Mingjun Liu^{3,4,6}, Catherine G. Corey^{2,3,6,14}, Frances Belmonte^{3,4,6}, Sara O. Vargas¹⁵, Brian Schwartz¹⁶, Bal Bhat¹⁷, B. Nelson Chau¹⁸, Jason H. Karnes¹⁹, Taijyu Satoh^{1,3,4,6,20}, Robert J. Barndt^{1,3,4,6}, Haodi Wu^{1,3,4,6}, Victoria N. Parikh²¹, Jianrong Wang²², Yingze Zhang^{3,5,6}, Dennis McNamara^{4,6}, Gang Li^{3,4,6,23}, Gil Speyer²⁴, Bing Wang⁶, Sruti Shiva^{2,3,6,25}, Brett Kaufman^{3,4,6}, Seungchan Kim²⁶, Delphine Gomez^{3,4,6}, Bernard Mari¹², Michael H. Cho²⁷, Adel Boueiz²⁷, Michael W. Pauciulo²⁹, Laura Southgate^{29,30}, Richard C. Trembath²⁹, Olivier Sitbon³¹, Marc Humbert³¹, Stefan Graf^{32,33,34}, Nicholas W. Morrell^{32,36}, Christopher J. Rhodes³⁵, Martin R. Wilkins³⁵, Mehdi Nouraie^{1,3,5,6}, William C. Nichols²⁸, Ankit A. Desai¹¹, Thomas Bertero¹³, and Stephen Y. Chan^{1,3,4,6,*}

¹Center for Pulmonary Vascular Biology and Medicine, 15213; ²Center for Metabolism and Mitochondrial Medicine, 15213; ³Pittsburgh Heart, Lung, Blood Vascular Medicine Institute, 15213; ⁴Division of Cardiology and ⁵Division of Pulmonary, Allergy, and Critical Care Medicine, 15213; ⁶Department of Medicine, 15213; ⁷Department of Computational and Systems Biology, 15213; ⁸McGowan Institute for Regenerative Medicine, University of Pittsburgh School of Medicine and University of Pittsburgh Medical Center, Pittsburgh, PA, 15219, USA; ⁹Cardiovascular Division, Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO, 63110, USA; ¹⁰Pittsburgh VA Medical Center, Pittsburgh, PA, 15240; ¹¹Division of Cardiology, Department of Medicine and ¹²Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, 46202, USA; ¹³Université Côte d'Azur, CNRS, IPMC, IHU RespiERA, Sophia-Antipolis, 06903, France; ¹⁴Department of Pediatrics, University of Pittsburgh Medical Center Children's Hospital, Pittsburgh, PA, 15224, USA; ¹⁵Department of Pathology, Boston Children's Hospital, MA, 02115, USA; ¹⁶Camp4 Therapeutics, Cambridge, MA, 02139, USA; ¹⁷Translate Bio, Lexington, MA, 02421, USA; ¹⁸Orna Therapeutics, Cambridge, MA, 02139, USA; ¹⁹Division of Pharmacogenomics, College of Pharmacy, University of Arizona College of Medicine, Tucson, AZ, 85721, USA; ²⁰Department of Cardiovascular Medicine, Tohoku University Graduate School of Medicine, Sendai, 980-8575, Japan; ²¹Stanford Center for Inherited Cardiovascular Disease, Stanford University School of Medicine, Stanford, CA, 94305, USA; ²²Department of Computational Mathematics, Science, and Engineering, Michigan State University, MI, 48824, USA; ²³Aging Institute, University of Pittsburgh, Pittsburgh, PA 15219, USA; ²⁴Research Computing, Arizona State University, Tempe, AZ, 85281, USA; ²⁵Department of Pharmacology & Chemical Biology, University of Pittsburgh, Pittsburgh, PA, 15213, USA; ²⁶Center for Computational Systems Biology, Department of Electrical and Computer Engineering, Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, 77446, USA; ²⁷Channing Division of Network Medicine and Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, 02115, USA; ²⁸Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine, Cincinnati, OH, 45229, USA; ²⁹Department of Medical & Molecular Genetics, Faculty of Life Sciences & Medicine, King's College London, London, WC2R 2LS UK; ³⁰Molecular and Clinical Sciences Research Institute, St George's University of London, London, SW17 0RE, UK; ³¹Université Paris–Saclay, INSERM, Assistance Publique Hôpitaux de Paris, Service de Pneumologie et Soins Intensifs Respiratoires, Hôpital Bicêtre, Le Kremlin Bicêtre, 94270, France; ³²Department of Medicine, University of Cambridge, CB2 1TN, UK; ³³NIHR BioResource for Translational Research, Cambridge Biomedical Campus, CB2 0QQ; ³⁴Department of Haematology, University of Cambridge, NHS Blood and Transplant, Long Road, Cambridge, CB2 2PT, UK; ³⁵National Heart and Lung Institute, Imperial College London, London, SW3 6LY, UK; ³⁶Centessa Pharmaceuticals. #Equal contribution

*Corresponding Author:

Stephen Y. Chan, MD, PhD

Pittsburgh Heart, Lung, Blood Vascular Medicine Institute

Division of Cardiology, Department of Medicine, University of Pittsburgh School of Medicine

200 Lothrop Street BST E1240

OVERLINE:

One sentence summary: Underlying a genetic risk association in pulmonary hypertension, hypoxic and SNP-dependent control of the long non-coding RNA *KMT2E-AS1* and its epigenetic protein binding partner KMT2E reprograms HIF-2 α -specific metabolic and gene regulation in pulmonary endothelium.

Editor's summary:

Abstract

Hypoxic reprogramming of vasculature relies upon genetic, epigenetic, and metabolic circuitry, but the control points are unknown. In pulmonary arterial hypertension (PAH), a disease driven by hypoxia inducible factor (HIF)-dependent vascular dysfunction, HIF-2 α promoted expression of neighboring genes, long non-coding RNA (lncRNA) histone lysine N-methyltransferase 2E-antisense 1 (*KMT2E-AS1*) and histone lysine N-methyltransferase 2E (*KMT2E*). *KMT2E-AS1* stabilized KMT2E protein to increase epigenetic histone 3 lysine 4 trimethylation (H3K4me3), driving HIF-2 α -dependent metabolic and pathogenic endothelial activity. This lncRNA axis also increased HIF-2 α expression across epigenetic, transcriptional, and post-transcriptional contexts, thus promoting a positive feedback loop to further augment HIF-2 α activity. We identified a genetic association between rs73184087, a single nucleotide variant (SNV) within a *KMT2E* intron, and disease risk in PAH discovery and replication patient cohorts and in a global meta-analysis. This SNV displayed allele (G)-specific association with HIF-2 α , engaged in long-range chromatin interactions and induced the lncRNA-KMT2E tandem in hypoxic (G/G) cells. In vivo, *KMT2E-AS1* deficiency protected against PAH in mice, as did pharmacologic inhibition of histone methylation in rats. Conversely, forced lncRNA expression promoted more severe PH. Thus, the *KMT2E-AS1/KMT2E* pair orchestrates across convergent multi-ome landscapes to mediate HIF-2 α pathobiology and represents a key clinical target in pulmonary hypertension.

Introduction

Cellular reprogramming by hypoxia relies upon incompletely defined genetic, epigenetic, and metabolic circuitry. Such fundamental biologic concepts are important for pulmonary hypertension (PH) and pulmonary arterial hypertension (PAH) – diseases of lung blood vessels linked to hypoxia and its master transcription factors including hypoxia inducible factor (HIF)- α (1). HIF-2 α in pulmonary arterial endothelial cells (PAECs) is particularly important in promoting this disease (2). However, the broad heterogeneity of disease- and hypoxia-dependent molecular circuitry has bred confusion regarding the development of crucial endothelial pathophenotypes (3).

An integrated understanding of genetic, epigenetic, and metabolic landscapes in the hypoxic endothelium and the pulmonary vasculature is lacking. HIF-dependent pathways regulate metabolic and mitochondrial programs that are dysregulated in hypoxic and diseased pulmonary vasculature in PH (4). Genome-wide molecular profiling in PH (5) has revealed that epigenetic marks of the genome and associated histones are altered in hypoxia across various PH subtypes (6). Furthermore, G9a histone methyltransferase inhibitors targeting primarily histone H3 lysine 9 marks (H3K9) have been reported to improve experimental PAH (7). However, the roles of other hypoxia-driven histone methylation marks in PAH have been poorly described. Histone H3 lysine 4 trimethylation (H3K4me3) is often enriched near promoters of activated genes and drives transcription (8). Although H3K4me3 is increased in hypoxia (9) and controlled by HIF-1 α and HIF-2 α (10), any pathogenic roles of H3K4me3 in orchestrating metabolic reprogramming in hypoxia are still unknown. Moreover, because hypoxia and HIF-2 α constitute crucial triggers of the World Symposium on Pulmonary Hypertension (WSPH) Group 1 PH (PAH) and Group 3 PH (PH due to hypoxic lung disease), hypoxic regulation of H3K4me3 may exert control over key pathogenic pathways in these PH subtypes. Finally, genomic insights are advancing regarding various associations between specific genetic variants with PAH risk, survival, and disease severity (11). However, due to the limited global number of patients with PAH, there are barriers to generating a comprehensive catalogue of genetic variants causatively linked to disease initiation or progression.

Long non-coding RNAs (lncRNAs) can exert regulatory activity across genomic, epigenetic, and metabolic domains (12). lncRNAs are single stranded RNAs that affect cellular function by complexing with chromosomal DNA, RNAs, or proteins, and may prevent microRNA binding to target messenger RNAs (mRNAs). lncRNAs are dysregulated in PH, and certain lncRNAs are controlled by PH triggers, such as hypoxia (13). Characterization of lncRNAs in pulmonary vascular cells has been limited (14), and functional data regarding their roles in PH is just emerging (15, 16). Yet, the majority of lncRNAs do not carry full sequence conservation in mammals, making it challenging to translate in vivo lncRNA biology between rodents and humans.

We hypothesized that specific lncRNAs serve as broad effectors of HIF-dependent pathobiology, orchestrating epigenetic, metabolic, and genomic processes in health and disease. Here, combining insights into genetic epidemiology with molecular mechanism, we identified a lncRNA-protein pair, governed in part by an endogenous human single nucleotide variant (SNV), that carries crucial epigenetic and metabolic functions in endothelial cells and controls PH manifestation in vivo.

Results

Mouse lncRNA *5031425E22Rik* and the human ortholog *KMT2E-AS1* were up-regulated in multiple rodent and human types of PH.

Using a published RNA-sequencing data set (17), we identified differentially expressed lncRNAs in the lungs of PAH mice induced by the small molecule SU5416 and chronic hypoxia compared with controls (**fig. S1A**). After filtering for length, conservation, and non-protein coding status, we identified 9 candidates, one of which was an upregulated lncRNA *5031425E22Rik* (or *E22*) which mapped to a human ortholog histone lysine N-methyltransferase 2E gene-antisense 1 (*KMT2E-AS1*). Across mammals, this lncRNA gene sits on the antisense strand and adjacent (head-to-head) to the histone lysine N-methyltransferase 2E gene (*KMT2E*), a member of a regulatory family controlling H3K4me3 and chromatin remodeling (18) (**Fig. 1A**). Despite its name, this lncRNA does not carry obvious sequence homology with *KMT2E*. Nonetheless, given its conserved chromosomal location and partially conserved sequence with human *KMT2E-AS1*, we analyzed this lncRNA further. Inference from RNA sequencing reads indicated that this lncRNA is most abundantly expressed as a ~2 kb isoform. *E22* and *KMT2E-AS1* transcripts were up-regulated in pulmonary vascular tissue, particularly CD31-positive endothelial cells, in mouse and human cases of PH, accompanied by increased *KMT2E* (**Fig. 1, B to E; fig. S1, B to E; Table S1**). By immunofluorescence microscopy and fluorescence *in situ* hybridization (FISH), increased *KMT2E-AS1* and *KMT2E* were also observed in CD31-positive endothelial cells of remodeled pulmonary arterioles of WSPH Group 1 PAH and Group 3 PH (**Fig. 1, F and G; Table S1**), in hypoxic mice (**fig. S1F**), and in PAH rat models (**fig. S1, G and H**). H3K4me3 was increased in both Groups 1 and 3 PH (**Fig. 1H**), consistent with rodents with PH (**fig. S1I**).

In cultured PAECs, hypoxic exposure drove similar up-regulation of *KMT2E-AS1* and *KMT2E* (**Fig. 1, I and J**). PAAFs also displayed induction of this lncRNA-*KMT2E* pair, but other pulmonary vascular cell types, such as PSMCs, did not (**Fig. 1, I and J**). This specific induction, particularly in PAECs, led to subcellular localization of *KMT2E-AS1* mainly to the nucleus (**Fig. 1K**). In mouse PAECs and fibroblasts, hypoxia also up-regulated *E22* with predominant localization in the nucleus (**fig. S1, J to**

L). Other known inflammatory cytokine triggers of PH, including interleukin-1 beta (IL-1 β) and interleukin-6 (IL-6), did not alter *KMT2E-AS1* (**fig. S1M**) in PAECs, suggesting the central mechanism of hypoxia in controlling this lncRNA axis.

Hypoxic induction of *KMT2E-AS1* and *KMT2E* was HIF-2 α -dependent.

We sought to determine how this tandem lncRNA-KMT2E pair is up-regulated by hypoxia. Previous chromatin immunoprecipitation sequencing (ChIP-Seq) identified interactions between HIF-2 α and the KMT2E promoter (19). Correspondingly, by siRNA knockdown of HIF-1 α and HIF-2 α , we found that HIF-2 α knockdown prevented the increase of *KMT2E-AS1* and *KMT2E* transcripts in hypoxic PAECs (**Fig. 2, A and B**). Conversely, lentiviral overexpression of a constitutively active HIF-2 α gene (20) in normoxia promoted the increased expression of this lncRNA-KMT2E pair (**Fig. 2, C and D**). Thus, HIF-2 α is both necessary and sufficient for the hypoxic up-regulation of *KMT2E-AS1* and the neighboring KMT2E in endothelial cells.

***KMT2E-AS1* interacted with KMT2E to enhance protein stability which increases histone methylation, including H3K4me3.**

Considering the conserved chromosomal neighboring location of *KMT2E-AS1* and *KMT2E*, the functions of lncRNAs to complex with protein partners to affect chromatin remodeling (12), and the known function of KMT2E in H3K4me3 (18), we hypothesized that *KMT2E-AS1* carries epigenetic functions stemming from physical interactions with the KMT2E protein. We screened for and identified an effective siRNA targeting *KMT2E-AS1*. Accordingly, we found that siRNA knockdown of *KMT2E-AS1* in hypoxia (**Fig. 2E**) down-regulated KMT2E protein expression by immunoblot (**Fig. 2F**). Conversely, forced expression of *KMT2E-AS1* in normoxia increased KMT2E protein (**Fig. 2, G and H**). Our results showed transcriptional control of this tandem pair by HIF-2 α (**Fig. 2, A to D**). Yet, in hypoxia, short-term (4h) exposure to the transcriptional inhibitor actinomycin D reduced *KMT2E* transcript (**Fig. 2I**) but not protein expression (**Fig. 2J**), suggesting that KMT2E protein is stabilized in

hypoxia. The in silico PRIdictor algorithm (21), predicted that *KMT2E-AS1* transcript bound to the KMT2E protein (prediction score of 0.534, >0.5 threshold). To validate RNA-protein interaction, RNA-protein immunoprecipitation assay (RIP), using antibodies against KMT2E followed by RT-qPCR of *KMT2E-AS1*, revealed that this lncRNA directly and specifically complexed with KMT2E under hypoxic conditions (**Fig. 2K**). Actinomycin D did not disrupt this physical interaction ($p>0.05$), reflecting the stability of such RNA-protein complexes despite acute reductions in RNA transcript production. Pretreating with the proteasomal inhibitor MG132 reversed the effect of *KMT2E-AS1* knockdown on KMT2E (**Fig. 2L**), indicating that the stabilized complex was responsible for protein expression.

To investigate the downstream consequences of KMT2E stabilization, we performed immunoblots specific for H3K4me3 in hypoxic PAECs. *KMT2E-AS1* knockdown decreased H3K4me3 (**Fig. 2M**). Moreover, forced expression of *KMT2E-AS1* increased the interaction of H3K4me3 with KMT2E by proximity ligation assay (**Fig. 2N**), whereas lentiviral delivery of a deletion mutant of this lncRNA missing a ~600 bp conserved sequence (**fig. S8F**) failed to increase this interaction. Correspondingly, forced expression of full-length *KMT2E-AS1*, but not the deletion mutant lentivirus, increased downstream H3K4me3 as assessed by immunoblot (**Fig. 2O**). Thus, *KMT2E-AS1* is a hypoxia-driven lncRNA that complexes with and stabilizes KMT2E protein to increase H3K4me3 histone trimethylation.

We then sought to determine whether other major histone lysine methylation marks that can be dependent on hypoxia (22), H3K9 and H3K27, were also controlled by this lncRNA-KMT2E tandem. In cultured PAECs, H3K9me3, but not H3K27, marks were increased by hypoxia and modestly reversed by *KMT2E-AS1* knockdown (**Fig. 2M**), a finding consistent with previous studies showing that KMT2E deficiency can induce a histone demethylase LSD1 that specifically reduces H3K9me3 (23, 24). Similarly, immunofluorescence staining of human Group 1 PAH and Group 3 PH demonstrated increased expression of H3K9me3 but not H3K27me3 in diseased pulmonary arterioles

(**fig. S2, A to C**). However, H3K4me3 reversal in PAECs by *KMT2E-AS1* knockdown was more robust than the reversal of H3K9me3 (**Fig. 2M**), consistent with a primary role of this lncRNA axis in H3K4me3.

KMT2E-AS1 drove hypoxic metabolic reprogramming.

Because *KMT2E-AS1* acted in conjunction with KMT2E to regulate primarily epigenetic H3K4me3, we sought to define the landscape of transcriptional alterations under this lncRNA's control during hypoxic endothelial reprogramming. In PAECs, siRNA knockdown of either KMT2E or *KMT2E-AS1* (**fig. S3, A to B**) phenocopied each other by reversing the expression of 2480 genes that were altered by hypoxia (**Fig. 3, A and B; Table S3**). Gene set enrichment analysis [GSEA (25)] of these RNA sequencing results revealed >50% of represented pathways were “hypoxia” and “metabolism” gene networks (**Fig. 3A**), consistent with the major activities of HIF-2 α . Notably, 554 of 1285 reversed metabolism genes (**Fig. 3B**) were related to energy generation and the tricarboxylic acid (TCA) cycle. To determine which of these changes were controlled directly by H3K4me3, H3K4me3 chromatin immunoprecipitation and sequencing (ChIP-Seq) was performed in hypoxic vs. normoxic PAECs (**fig. S3C; Table S4**). The promoters of many of the top dysregulated genes in “hypoxia” and “metabolism” networks regulated by the lncRNA-KMT2E axis carried more H3K4me3 marks in hypoxia, suggesting a direct epigenetic mechanism by which the lncRNA axis modulates HIF-2 α metabolic phenotypes (**Fig. 3B; fig. S3C; Table S3 and 4**). Furthermore, not all H3K4me3-dependent genes in hypoxia were regulated by the lncRNA-KMT2E axis (**fig. S3C**), emphasizing an epigenetic specificity of this regulatory axis. By RNA sequencing and ChIP-Seq overlay (**fig. S3B**) and further independent ChIP-qPCR (**Fig. 3C**), we confirmed the lncRNA *mir210hg* as one of several *KMT2E-AS1*-dependent and H3K4me3-dependent genes.

Consistent with *KMT2E-AS1*'s control over HIF-dependent metabolism, *KMT2E-AS1* knockdown mitigated the hypoxic induction of extracellular acidification rate (ECAR), a representative measure of glycolysis, as well as concomitant lactate dehydrogenase (LDH) enzymatic activity (**Fig. 3, D and H**);

forced *KMT2E-AS1* expression increased ECAR and LDH activity (**Fig. 3, F and I**). Moreover, in cultured PAECs, *KMT2E-AS1* knockdown increased baseline oxygen consumption rate (OCR) (**Fig. 3E**), while forced *KMT2E-AS1* decreased OCR (**Fig. 3G**). Furthermore, vascular endothelial growth factor (VEGFA) was up-regulated by hypoxia, consistent with an increase of H3K4me3 at its gene locus (**Fig. 3J**). Importantly, both *KMT2E-AS1* and *KMT2E* knockdown decreased VEGFA expression. Taken together, these data demonstrated that *KMT2E-AS1* regulates a gene network that decreases oxidative metabolism, increases glycolysis, and controls hypoxic PAEC adaptation.

***KMT2E-AS1* induced HIF via increasing HIF transcription and controlling HIF degradation.**

In addition to the ability of *KMT2E-AS1* to stabilize *KMT2E* protein, our data also indicated that knockdown of *KMT2E-AS1* reduced *KMT2E* transcript (**Fig. 2I; fig. S3A**). Guided by the identification of “hypoxic” response pathway genes regulated by this lncRNA axis that control HIF transcription (*miR210hg*) and HIF protein degradation (*ELOC*) (**Fig. 3B**), we sought to determine whether HIF-2 α -dependent *KMT2E-AS1* reciprocally regulates HIF-2 α expression. siRNA knockdown of *KMT2E-AS1* or *KMT2E* in hypoxic PAECs decreased HIF-2 α expression (**Fig. 4A**), whereas forced expression of this lncRNA increased HIF-2 α in normoxia and hypoxia (**Fig. 4B**). The hypoxia-dependent *miR210hg* gene was specifically induced by this lncRNA axis (**Fig. 3B; fig. S3B**). *miR210hg* is a lncRNA that directly binds the 5'UTR of HIF transcripts (26), leading to increased translation and increased HIF protein expression. Moreover, for both RNA and protein, *KMT2E-AS1* knockdown increased elongin C (*ELOC*), a crucial component of the Von Hippel Lindau complex that controls HIF-2 α protein stability (**Fig. 4, C and D**), indicating that the lncRNA augments HIF-2 α expression via modulation of the hydroxylase-ubiquitin-proteasome pathway. The proteasomal inhibitor MG132 reversed the down-regulation of HIF-2 α by lncRNA knockdown (**Fig. 4E**). Thus, hypoxic induction of *KMT2E-AS1* and *KMT2E* participates in a positive feedback loop to increase HIF-2 α expression via multiple epigenetic, transcriptional, and post-transcriptional mechanisms to activate HIF-2 α transcription and translation as well as to alter HIF-2 α protein degradation.

***KMT2E-AS1* promoted endothelial pathophenotypes of PH.**

KMT2E-AS1 drove endothelial pathophenotypes linked to HIF biology and PH. Knockdown of *KMT2E-AS1* increased PAEC apoptotic potential in hypoxia, while forced expression of *KMT2E-AS1* via lentiviral transduction decreased apoptosis in normoxia (**Fig. 4F**). *KMT2E-AS1* knockdown decreased PAEC proliferation in hypoxia, whereas forced expression increased proliferation in normoxia as observed by BrdU incorporation (**Fig. 4G**). Thus, *KMT2E-AS1* is necessary and sufficient to promote PAEC proliferation. Consistent with these alterations in cell survival and proliferative capacity, *KMT2E-AS1* knockdown decreased PAEC migration in hypoxia, whereas forced expression increased such activity by wound closure assay in vitro, (**Fig. 4, H and I**). Similarly, by quantifying gel contraction as a surrogate of smooth muscle cell contraction when exposed to PAEC-conditioned media, we found that *KMT2E-AS1* knockdown in hypoxic PAECs generated conditioned media that decreased contraction seen in hypoxia, while forced expression of *KMT2E-AS1* in PAECs increased contraction in normoxia (**Fig. 4, J and K**). Consistent with these alterations in vasomotor activity, *KMT2E-AS1* knockdown decreased secreted endothelin-1 (EDN1) in hypoxia, whereas forced expression of *KMT2E-AS1* increased endothelin-1 in normoxia (**Fig. 4L**). Notably, via sequences different from that of the above siRNA, antisense oligonucleotide gapmers targeting *KMT2E-AS1* phenocopied siRNA silencing of *KMT2E-AS1* with knockdown of *KMT2E-AS1* and downstream *VEGFA* and *EDN1* expression (**fig. S4, A to C**). We found that *KMT2E-AS1* inhibited PAEC apoptosis as well as enhanced proliferation, migration, and vasomotor tone, consistent with a role that promotes endothelial dysregulation and PAH.

***KMT2E* SNV rs73184087 (G) allele was associated with WSPH Group 1 PAH risk.**

Genetic control of HIF-2 α activity can be facilitated by single nucleotide variants (SNVs) (27) that alter transcription factor binding sites. SNVs alter HIF binding sites even outside canonical promoter regions (27) with consequent disruptions of long-range genomic interactions with active promoter sites. Thus, we screened for such SNVs relevant to this lncRNA-*KMT2E* locus within a previously reported WSPH Group 1 PAH discovery cohort (“PAH Biobank”) of European-descent (N=694)

individuals vs. individuals without disease (N=1,560) (28) (**Table S5**). Among the 883 genotyped and imputed SNVs (with minor allele frequency>1%) within and flanking (+/-200kb) the *KMT2E-AS1-KMT2E* tandem locus, we prioritized 59 SNVs (**Table S6**) with predicted HIF-2 α binding to one of either the minor or major alleles using position weight matrices (PWMs) derived from HIF-2 α chromatin immunoprecipitation sequencing (19). Among these SNVs, by Firth's penalized logistic regression, we observed a significant association for rs73184087 (GRCh38:7:105087911) with the risk for developing PAH, with the G allele conferring an adjusted odds ratio (OR) of 1.86 (95% CI:1.30-2.67; p=7.40x10⁻⁴) in the discovery cohort (**Fig. 5A**). Linkage disequilibrium pattern analysis revealed that there is no strong association of SNV rs73184087 with other neighboring SNVs (**fig. S5**). PWM scoring predicted more robust HIF-2 α binding to the effect allele (G) of this SNV (Log-odds score 10.01, p<10⁻⁸) vs. the non-effect allele (A) (Log-odds score 4.56). We then replicated the SNV association with disease risk in a second, independent PAH cohort of participants of European-descent from the University of Pittsburgh Medical Center (UPMC cohort, **Table S7**, N=96 cases vs. N=401 non-PAH individuals, adjusted OR 2.53 [95% CI:1.25-5.13]; p=0.01). We also replicated the association of rs73184087 in the All of Us (AoU, dataset v6) cohort, using of ICD10 codes to identify PAH (European descent) patients and other filters (such as pulmonary vasodilator medication use) to increase confidence in diagnostic accuracy (**Table S8**, N=52 cases vs. N=11,821 controls, adjusted OR 2.44 [95% CI: (1.25, 4.79)]; p=0.01). Finally, we replicated the association of this SNV in a global meta-analysis of five cohorts (N=2,181 cases vs. N=10,060 controls; total N=12,241) including the PAH Biobank, UPMC cohort, and three additional European cohorts from a prior study (OR=1.30 [95% CI: 1.08-1.56], p=0.005). Based on the association with disease risk, this SNV was further characterized by functional validation.

SNV rs73184087 displayed allele-specific binding to HIF-2 α and long-range interaction with the shared lncRNA-KMT2E promoter.

SNV rs73184087 is located at a KMT2E intronic site 75 kB downstream of the lncRNA-KMT2E shared promoter. Based on prior-capture Hi-C mapping in lung tissue (29), we found a long-range genomic

interaction between this SNV and the shared promoter (**Fig. 5B**), consistent with the notion that SNV-bound HIF-2 α can readily gain access to the promoter for transcriptional activation. In hypoxic PAEC extracts, oligonucleotides carrying the risk allele SNV rs73184087 (G) displayed preferential and increased binding to HIF-2 α , but not HIF-1 α , as compared with the major allele (A) (**Fig. 5C**). Confirming the functionality of such binding, placement of SNV rs73184087 downstream of a luciferase reporter gene demonstrated increased reporter gene expression with the risk (G) allele vs. ancestral (A) allele in the presence of constitutive HIF-2 α expression (**Fig. 5D**). In transformed lymphocytes from WSPH Group 1 PAH patients carrying SNV rs73184087 (G/G) vs. (A/A) genotypes (**Table S2**), ChIP-qPCR via pulldown of HIF-2 α demonstrated enrichment of binding to the (G/G) vs. (A/A) SNV in hypoxia (**Fig. 5E**). To confirm the long-range interactions of this SNV regardless of its genotype with the shared promoter, using transformed PAH lymphocytes carrying SNV rs73184087 (G/G) or (A/A) genotypes, 3C assays revealed SNV-promoter interaction driven by either the (G/G) or (A/A) genotype (**Fig. 5F**). A 3C assay using PAECs with the A/A genotype confirmed a specific interaction between the SNV and promoter (**Fig. 5G**) but not upstream or downstream of the TSS/promoter (**Fig. 5H**). Furthermore, under cobalt (II) chloride treatment where HIF- α is stabilized in normoxia (**Fig. 5, I and J**), lymphocytes with (G/G) genotype increased *KMT2E-AS1* and *KMT2E* more robustly as compared to those with (A/A) genotype.

To determine whether such HIF-2 α interaction is observed in endothelial cells, we generated inducible pluripotent stem cells (iPSCs) from these transformed lymphocyte lines (30) followed by re-differentiation (31) of those lines into endothelial cells, iPSC-ECs (**fig. S6**). Consistent with our findings in transformed lymphocytes, ChIP-qPCR in these iPSC-ECs revealed that the G/G locus displayed greater HIF-2 α binding vs. the A/A locus (**Fig. 5K**). This increased binding correlated with increased transcript expression of *KMT2E-AS1*, *KMT2E*, and downstream *miR210hg* (**Fig. 5L**). Together, these data define an intronic SNV rs73184087 (G) allele-specific mechanism by which HIF-2 α controls the expression of the lncRNA-KMT2E pair, thus offering a mechanistic explanation underlying the enrichment of SNV rs73184087 (G) allele in WSPH Group 1 PAH.

Mouse lncRNA 5031425E22Rik (E22) phenocopied KMT2E-AS1 via a 500 bp conserved sequence.

We sought to determine if mouse lncRNA *E22* carries similar activity as *KMT2E-AS1* in PAECs. We found that hypoxia upregulated *E22* and *KMT2E* in mouse PAECs, mirroring the regulation of *KMT2E-AS1* in human PAECs. Lentiviral forced expression of *E22* (**fig. S7A**) drove consequent decreased oxygen consumption and increased glycolysis (**fig. S7, B and C**). As with *KMT2E-AS1*, this mouse lncRNA controlled similar endothelial pathophenotypes including migration, contraction, and regulation of vasomotor effectors (**fig. S7, D to F**).

Computational predictions of secondary structures of *E22* and *KMT2E-AS1* revealed putative conserved similarities across these mouse and human isoforms (**fig. S8A**), indicating its importance in this lncRNA's conserved functions. Yet, because of the genomic proximity of this region to the lncRNA-*KMT2E* promoter (**fig. S8B**), it was possible that the chromosomal region encoding this sequence was also important in controlling canonical promoter function in *cis* rather than lncRNA function in *trans*, as reported for other lncRNAs (12). To clarify these roles, we used a deletion mutant analysis to map a sequence responsible for a lncRNA-dependent function, such as VEGFA induction. This approach identified a 550-600 bp region in the 5' end of the lncRNA transcript, conserved in both *KMT2E-AS1* and mouse *E22* (**fig. S8, C to K**). This region corresponded to the same functional domain of *KMT2E-AS1* that controlled the interaction of H3K4me3 with *KMT2E* and thus the expression of H3K4me3 (**Fig. 2, N and O**). Finally, by reporter gene assay, deletion of this region did not affect canonical lncRNA-*KMT2E* promoter activity (**fig. S8L**), emphasizing the specific importance of this region for lncRNA activity. Therefore, these data demonstrated that mouse *E22* and its conserved 550 bp domain can serve as a surrogate to define the in vitro and in vivo pathobiology of human *KMT2E-AS1*.

Overexpression of pulmonary vascular mouse lncRNA 5031425E22Rik worsened PH.

To investigate if forced expression of mouse *E22* in pulmonary vasculature could promote these epigenetic and metabolic alterations in PAECs and thus cause PH, mice received orotracheal administration of a recombinant adeno-associated virus (serotype 6) carrying the 5031425E22Rik transgene (AAV6-*E22*) four weeks before a 3-week exposure to normoxia or hypoxia (**Fig. 6A**). AAV was chosen for long term transgene delivery, and serotype 6 was selective for optimal endothelial delivery based on in vitro infection studies of PAECs, as previously described (32). Compared with AAV6-*GFP*-treated littermate mice, mice that were administered AAV6-*E22* demonstrated a selective upregulation of *E22* in pulmonary vascular CD31-positive endothelial cells, as assessed by *in situ* stain (**fig. S9A; Fig. 6B**). AAV6-*E22* increased KMT2E expression, H3K4me3, and H3K9me3 (but not H3K27me3), and HIF-dependent and vasoconstrictive gene expression with greatest consistency in normoxia and across most of the same indices in hypoxia (**fig. S9A; Fig. 6, C to F; fig. S10, A and B; fig S9, B and C**). AAV6-*E22* increased the presence of proliferation markers (Ki67⁺) in CD31⁺ (endothelial) pulmonary arteriolar cells in comparison with AAV6-*GFP* mice, most robustly in normoxia and with a further trend in hypoxia (**Fig. 6, G and H**). As a result, in both normoxia and hypoxia, AAV6-*E22* increased pulmonary arteriolar muscularization but less robustly vascular thickness (**Fig. 6, I and J**), whereas forced lncRNA expression increased right ventricular systolic pressure (RVSP) (**Fig. 6K**) and right ventricular remodeling (RV to body weight ratio) (**Fig. 6L**) primarily in hypoxia. Delivery did not affect heart rate, systemic mean arterial pressure, or left ventricular function (**fig. S9, D to H**). Thus, the mouse lncRNA *E22* alone was sufficient to promote epigenetic functions and metabolic reprogramming and to induce histologic and hemodynamic indices of PH. Notably, however, the effects of forced *E22* expression in isolation were not equally robust across all indices. In that regard, H3K4me3, H3K9me3, and histologic muscularization all displayed higher signals with hypoxic exposure alone compared with normoxic AAV6-*E22* delivery (**Fig. 6, F, G and J**), potentially indicating the importance of the hypoxic milieu and precise endogenous and tandem transcriptional elevation of KMT2E to drive the full extent of disease.

Mice with a deletion in the conserved 5031425E22Rik/KMT2E-AS1 sequence were protected from PH.

To investigate the full pathobiologic activity of this lncRNA-KMT2E tandem, we sought to determine whether *E22* was necessary to drive PAH under endogenous disease contexts when both the lncRNA and KMT2E are up-regulated. We therefore utilized CRISPR/Cas9 technology to generate a mouse genetically deficient specifically in the conserved 550 bp sequence in *E22* responsible for the control of PAEC activity (**Fig. 7A**). These knockout mice displayed decreased active *E22* and KMT2E in CD31-positive pulmonary vascular endothelial cells (**Fig. 7B; fig. S11A**), consistent with knockdown of *KMT2E-AS1* in human PAECs (**Fig. 2, E and F**). Consequently, under conditions of hypoxic PH to model Group 3 PH, in both lung tissue and CD31-positive PAECs, H3K4me3 and H3K9me3, but not H3K27me3, were downregulated in knockout mice as compared with wild type littermate controls (**Fig. 7, C to E; fig. S10, C and D**). This was accompanied by downstream reduction of VEGFA and EDN1 (**fig. S11, B and C**), consistent with our studies of cultured PAECs (**Fig. 3J, Fig. 4L; fig. S7F; fig. S8, E, G, I and K**) and control of VEGFA by H3K4me3 in hypoxia (**Fig. 3J**). As in cultured cells, *in situ* staining of PAECs displayed lower expression of the proliferation marker Ki67 (**Fig. 7, F and G**). Knockout mice were protected from histologic and hemodynamic manifestations of hypoxic PH, including demonstrating decreased pulmonary remodeling/muscularization, RVSP, and RV/body weight index (**Fig. 7, H to K**) but without other differences in heart rate, blood pressure, or echocardiographic measures of left ventricular (LV) function (**fig. S11, D to H**). To address the possibility of off-target CRISPR/Cas9 editing, a separate mouse line utilizing alternative guide primer pairing was generated with a smaller conserved sequence deletion (300 bp). These mice exhibited similar reductions of *E22* and RVSP (**fig. S11, I to K**).

To model angioproliferative Group 1 PAH, the same mice carrying the 550 bp deletion were crossed with the pulmonary interleukin-6 transgenic (IL-6 Tg) mouse (33) and exposed to chronic hypoxia. We observed decreased *E22* and KMT2E (**Fig. 7L; fig. S12A**), H3K4me3 and H3K9me3 (but not H3K27me3) (**Fig. 7, M to O; fig. S10, E and F**), VEGFA and EDN1 (**fig. S12, B and C**), and Ki67

(**Fig. 7, P and Q**), in CD31-positive pulmonary vascular endothelium. This resulted in a reduction of vascular remodeling, RVSP, and RV/body weight ratio (**Fig. 7, R to U**) but no differences in heart rate, blood pressure, or LV function (**fig. S12, D to G**). These findings demonstrated that this *E22* lncRNA is necessary for promoting experimental Groups 1 and 3 PH in vivo, via epigenetic control of endothelial proliferation.

The histone lysine methyltransferase inhibitor chaetocin ameliorated PAH in vivo.

To determine whether histone lysine methylation, as regulated by the lncRNA-KMT2E pair (**Fig. 2,6,7; fig. S10**), promotes PAH, pharmacologic inhibition of histone lysine methyltransferase activity across H3K4me3 and H3K9me3 was pursued with the small molecule chaetocin (34) in a rat model of PAH (SU5416-hypoxia). A disease reversal protocol was used, where PAH was allowed to develop over 3 weeks of hypoxia before rats were removed from hypoxia and serial i.p. administration of chaetocin was performed over 2 subsequent weeks of normoxia (**Fig. 8A**). In these PAH rats, elevated expression of lncRNA rat homolog of *E22* (**Fig. 8B**) was observed and was not altered by chaetocin. Vascular KMT2E was modestly reduced by chaetocin (**Fig. 8C; fig. S13A**). However, chaetocin decreased downstream H3K4me3 and H3K9me3 (**Fig. 8, D to F; fig. S13, B to D**), VEGFA (**fig. S13, E and F**), EDN1 (**fig. S13, G and H**), as well as Ki67 (**Fig. 8, G and H**) in CD31-positive pulmonary vascular endothelial cells. Chaetocin reduced vascular remodeling, RVSP, and RV to body weight ratio (**Fig. 8, I to L**), without alterations in heart rate, blood pressure, or LV function (**fig. S13, I to M**). When considering these comprehensive gain- and loss-of-function analyses via both pharmacologic and genetic means, we conclude that the mouse *E22* and human *KMT2E-AS1* carry conserved functions as epigenetic mediators of hypoxic metabolic reprogramming, crucial to pulmonary vascular proliferation, remodeling, and overt PH in vivo.

Discussion

Here we defined the *KMT2E-AS1/KMT2E* pair as a lynchpin of HIF-2 α -dependent endothelial pathophenotypes for promoting PH, encompassing robust and convergent genetic, epigenetic, and metabolic processes in this complex disease (**fig. S14**). These findings established epigenetic, metabolic, and RNA-based paradigms that control HIF-2 α -dependent physiology and endothelial reprogramming. Furthermore, by describing the mechanistic underpinnings of large-scale epidemiologic and genetic associations, this work offers rare insight into the genomic functionality of SNVs contributing to clinical manifestations of PH. As a result, these findings introduce the potential of this lncRNA-KMT2E pair as diagnostic markers as well as the lncRNA-KMT2E and downstream H3K4me3 as attractive therapeutic targets.

Our work describes a lncRNA-dependent mechanism by which HIF-2 α controls endothelial pathobiology and PH, with broad implications for other ischemic and hypoxic diseases (1). While HIF-2 α is canonically considered a transcription factor that binds promoter elements directly, our findings add to the emerging roles of genomic architecture, epigenetics (10), and context-specific features (35) that modulate HIF's functions. This is particularly relevant for HIF-specific control of endothelial metabolism in ischemia and hypoxia (36). Our work more specifically identifies poorly defined regulatory paradigms that govern HIF-2 α activity, including allele-specific intronic SNV binding, long-range chromatin interactions, and lncRNA-dependent modulation of epigenetic histone methylation. In the case of SNV rs73184087, HIF-2 α complexing to this locus occurs independently of a canonical HIF-response element. This suggests the role of other factors mediating the interaction of this SNV with HIF-2 α and potentially aiding additional long-range genomic interactions. Whether HIF-2 α utilizes such strategies to exert genome-wide effects (19) and whether other SNVs may control such binding and chromatin contact (27) in additional conditions is not known. Overall, our findings should guide further understanding of context-specific alterations of local hypoxic or HIF-dependent cellular states, such as those recently described for senescent endothelial cells in PH (31) and beyond.

The connected biology between this lncRNA and KMT2E offers guidance in functional understanding of human lncRNAs. Defining the pathobiologic actions of lncRNAs is challenging, due to the lack of full sequence conservation among most human lncRNAs. In this study, partial conservation of lncRNA sequence and secondary structure (**fig. S8A**) coupled with tandem gene chromosomal location guided our investigation. It is possible that other head-to-head chromosomal arrangements among multiple lncRNA-protein coding gene pairs across the genome similarly predict lncRNA activity, based on the protein coding gene's binding or functions. The connected biology of this gene tandem may underlie their cell-type specificity and promoter-specific epigenetic activity (**fig. S3C**). For example, based on our findings that endothelial cells, but not smooth muscle cells, displayed robust *KMT2E-AS1* up-regulation (**Fig. 1, I and J**), the context dependence of KMT2E function may be critically dependent on this lncRNA partner. KMT2E also controls diverse biological processes (37), across hematopoiesis, cancer stem cell renewal, spermatogenesis, autism spectrum disorder (38), and atherogenesis (39). Thus, similar to its binding partner, *KMT2E-AS1* may also carry far-reaching biologic actions in cancer as well as neurologic and cardiovascular conditions. Future studies should also explore if this lncRNA integrates into a broader network of hypoxia-relevant non-coding RNAs – a process that could be aided by computational algorithms applied to pulmonary vascular biology (40).

The identification of a non-coding SNV rs73184087 as a pathogenic mediator of HIF-2 α 's epigenetic activity establishes a fundamental paradigm of a positive feedback loop that augments and sustains hypoxic reprogramming in PH. Notably, while carrying a modest predicted effect on differential HIF-2 α binding, SNV rs73184087 was the sole significant variant associated with PAH risk ($p=0.00074$), perhaps reflecting the importance of genomic location and context in addition to binding. Other predicted HIF binding sites exist in the enhancer regions of the lncRNA-KMT2E gene locus (19), and these would be expected to be active, independent of SNV rs73184087 genotypes. Thus, such genetic control does not define a simple “on-off” switch for hypoxic reprogramming. Rather, our findings demonstrate that the presence of the G allele at rs73184087 augments an already existing ability of HIF-2 α to promote transcription of the lncRNA-KMT2E tandem, with more robust induction

of downstream phenotypes. Such tunable activity could explain cell-type differences, with the actions of the G allele at rs73184087 observed in normoxia in iPSC-ECs but primarily in hypoxia in transformed lymphocytes. These differences may stem from the fact that endothelium can carry augmented expression of HIF-2 α even in normoxia and in PAH disease states (1). Thus, a more obvious display of the effects of the G/G SNP genotype could be seen in endothelium without an exogenous hypoxic trigger.

This positive feedback loop that augments HIF-2 α expression offers a mechanistic explanation for the association between rs73184087 and the risk of PAH. The integral relationship of SNV rs73184087 with HIF-2 α biology may also portend this SNV's role in other hypoxic or ischemic conditions. In contrast, given the complex regulation of H3K4me3, this SNV may carry a variety of contrasting activities in such HIF-dependent conditions. Given the importance of histone methylation in a variety of chronic human diseases (41), well-phenotyped control cohorts without pre-existing chronic disease will be essential for appropriate interpretation of any further targeted SNV rs73184087 association studies. It is notable that the SNV rs73184087 risk allele (G) is known to carry a much lower prevalence in African, Asian, and Latin cohorts (42). Future work should prioritize determining if variations of this allele drive ethnic differences in risk or outcome of PAH.

Lastly, the epigenetic action of *KMT2E-AS1/KMT2E* in hypoxic endothelial metabolism clarifies the pathogenic role of histone methylation in PH and beyond. It also establishes a platform to advance epigenetic and RNA-based diagnostics and therapeutics. Unlike other SET domain-containing methyltransferase family (MLL or KMT2) members, KMT2E promotes H3K4me3 indirectly (43) in a cell-specific manner. While the *KMT2E-KMT2E-AS1* axis primarily controls hypoxic-driven H3K4me3, this axis also modulates H3K9me3, consistent with a therapeutic role of G9a histone methyltransferase inhibitors in PAH (7). Our genetic and pharmacologic data in vivo present an attractive opportunity for pursuing additive or synergistic therapy, with oligonucleotide inhibition of *KMT2E-AS1* coupled with pharmacologic inhibition of both H3K4me3 and H3K9me3. More precise

targeting of specific mediators of HIF-2 α activity rather than HIF-2 α itself may carry higher efficacy for improvement of PAH and may better address the risks of broad endothelial HIF-2 α inhibition in the lung (44). Beyond oligonucleotide technology targeting *KMT2E-AS1* directly, improvement of experimental PAH with chaetocin (**Fig. 8**) offers a foundation for optimizing small molecule inhibitors that can target across both sets of H3K4me3 and H3K9me3 marks. Because KMT2E and downstream histone methylation carry pleiotropic activity beyond the vasculature and lung, simultaneous, lung-specific, and controlled release of drugs would be ideal, as shown with recent inhaled microparticle drug delivery in PH (45).

Our study has several limitations. Although multiple independent PAH vs. control cohorts were used to establish SNV associations with PAH risk, the ability to match clinical characteristics completely across all PAH groups was not possible. Although the PAH Biobank discovery cohort and the UPMC validation cohort were adjudicated by expert physicians to ensure clinical accuracy, phenotypic information in the All of Us dataset relied upon ICD codes and could be prone to error. This study also did not fully define the epigenetic pleiotropy of *KMT2E-AS1*, setting the stage for future work to analyze the link between this lncRNA and other KMT2 factors besides KMT2E. Moreover, the full context of how this lncRNA behaves in hypoxia was not defined. *KMT2E-AS1* may require both the endogenous hypoxic milieu and precise transcriptional elevation of KMT2E, as suggested by our data showing that the pathogenic effects of forced lncRNA expression in isolation (**Fig. 6**) were not as consistently robust as PH improvement after lncRNA knockout (**Fig. 7**).

In summary, by coupling genetic epidemiology with mechanistic exploration, we have identified the *KMT2E-AS1/KMT2E* pair as a lynchpin of hypoxic, genetic, and metabolic endothelial reprogramming for promoting PH. These findings carry broad implications for understanding the fundamental molecular hierarchy driving HIF and hypoxic signaling and endothelial pathobiology. This work offers a roadmap toward more effective diagnostic and therapeutic opportunities focused on RNA-based and epigenetic platforms.

MATERIALS AND METHODS

Study design

This study was designed to investigate the role of long non-coding RNA in pathogenetic development of pulmonary hypertension and explore the potential of RNA-based and epigenetic platforms for early diagnosis and therapeutic opportunities. This objective was addressed by (i) identifying and determining the function of *KMT2E-AS1/KMT2E* lncRNA-protein complex in driving endothelial dysfunction via H3K4 trimethylation (ii) identifying and determining functional single nucleotide variant that mediates pathogenic epigenetic activity in orchestrating hypoxic reprogramming in PH (iii) determining the protective effects with knockdown of *KMT2E-AS1/KMT2E* and inhibition of H3K4me3 in hypoxic PH rodent models, revealing new prospective epigenetic and RNA-based diagnostics and therapeutics in PH.

Cell culture experiments were performed at least three times and at least in triplicate for each replicate. The number of animals in each group was calculated to measure at least a 20% difference between the means of experimental and control groups with a power of 80% and standard deviation of 10%. The number of unique patient samples for this study was determined primarily by clinical availability. Animals sharing same sex, same genotype, and similar body weight were generated and randomly assigned to different experimental groups. No animals were excluded from analyses. Investigators performing hemodynamic data collection and histologic analysis were blinded to the groups.

All animal experiments were approved by the University of Pittsburgh School of Medicine (DLAR). All experimental procedures involving the use of human tissue and plasma and the study of invasive and non-invasive hemodynamics were approved by institutional review boards at the University of Pittsburgh, Brigham and Women's Hospital, and Boston Children's Hospital (Studies 19070274, 990835, 19050364, 19020233, 19050026). Ethical approval for this study and informed consent conformed to the standards of the Declaration of Helsinki.

Human patients

Informed consent was obtained for right heart catheterization and tissue sampling. For tissue samples derived from WSPH Group 3 PH, diagnosis was made by an expert physician based on the criteria of having an elevated mean pulmonary arterial pressure > 20 mmHg (mPAP) by right heart catheterization with PH deemed to be driven by hypoxic lung disease. For tissue samples derived from WSPH Group 1 PAH, diagnosis was made by an expert physician, based on the criteria of having an elevated mPAP > 20 mmHg, pulmonary capillary wedge pressure ≤ 15 mm Hg, and pulmonary vascular resistance ≥ 3 Wood Units (PVR) by right heart catheterization. For PAH cases, the expert physician adjudicated the diagnosis by ruling out left heart disease, hypoxic lung disease, or chronic thromboembolic disease. For OCT fresh-frozen lung samples, human specimens were collected from unused or discarded surgical samples from participants diagnosed with PH (40). Non-diseased human lung specimens were described previously (40).

Statistical analysis

Numerical quantifications for *in vitro* experiments using cultured cells or *in situ* quantifications of transcript expression and physiologic experiments using rodents or human reagents represent mean \pm standard error of the mean (SEM). Immunoblot images are representative of experiments that have been repeated at least three times and quantified using ImageJ. Micrographs are representative of experiments in each relevant cohort. Normality of data distribution was determined by Shapiro Wilk testing. Means of 2 sample groups were compared by an unpaired 2-tailed Student's t-test for normally distributed data, while Mann-Whitney U non-parametric testing was used for non-normally distributed data. For comparisons among different groups, one-, or two-way analysis of variance tests (ANOVA) followed by Bonferroni's *post-hoc* analysis or Kruskal Wallis tests (where appropriate for non-normally distributed data) were performed. A P-value less than 0.05 was considered significant.

List of supplementary materials

Materials and Methods

Figures S1-S14

Table S1-S9

Data file S1

References

1. S. S. Pullamsetti, A. Mamazhakypov, N. Weissmann, W. Seeger, R. Savai, Hypoxia-inducible factor signaling in pulmonary hypertension. *J Clin Invest* **130**, 5638-5651 (2020).
2. Z. Dai, M. Li, J. Wharton, M. M. Zhu, Y. Y. Zhao, Prolyl-4 Hydroxylase 2 (PHD2) Deficiency in Endothelial Cells and Hematopoietic Cells Induces Obliterative Vascular Remodeling and Severe Pulmonary Arterial Hypertension in Mice and Humans Through Hypoxia-Inducible Factor-2alpha. *Circulation* **133**, 2447-2458 (2016).
3. E. D. Michelakis, Spatio-temporal diversity of apoptosis within the vascular wall in pulmonary arterial hypertension: heterogeneous BMP signaling may have therapeutic implications. *Circ Res* **98**, 172-175 (2006).
4. M. K. Culley, S. Y. Chan, Mitochondrial metabolism in pulmonary hypertension: beyond mountains there are mountains. *J Clin Invest* **128**, 3704-3715 (2018).
5. A. R. Hemnes, Using Omics to Understand and Treat Pulmonary Vascular Disease. *Front Med (Lausanne)* **5**, 157 (2018).
6. G. Benincasa, D. L. DeMeo, K. Glass, E. K. Silverman, C. Napoli, Epigenetics and pulmonary diseases in the horizon of precision medicine: a review. *Eur Respir J* **57**, 2003406 (2021).
7. C. Awada, A. Bourgeois, S. E. Lemay, Y. Grobs, T. Yokokawa, S. Breuils-Bonnet, S. Martineau, V. Krishna, F. Potus, J. Jeyaseelan, S. Provencher, S. Bonnet, O. Boucherat, G9a/GLP Targeting Ameliorates Pulmonary Vascular Remodeling in Pulmonary Arterial Hypertension. *Am J Respir Cell Mol Biol* **68**, 537-550 (2023).
8. B. E. Bernstein, E. L. Humphrey, R. L. Erlich, R. Schneider, P. Bouman, J. S. Liu, T. Kouzarides, S. L. Schreiber, Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A* **99**, 8695-8700 (2002).
9. M. Batie, J. Frost, M. Frost, J. W. Wilson, P. Schofield, S. Rocha, Hypoxia induces rapid changes to histone methylation and reprograms chromatin. *Science* **363**, 1222-1226 (2019).
10. B. M. Ortmann, N. Burrows, I. T. Lobb, E. Arnaiz, N. Wit, P. S. J. Bailey, L. H. Jordon, O. Lombardi, A. Penalver, J. McCaffrey, R. Seear, D. R. Mole, P. J. Ratcliffe, P. H. Maxwell, J. A. Nathan, The HIF complex recruits the histone methyltransferase SET1B to activate specific hypoxia-inducible genes. *Nat Genet* **53**, 1022-1035 (2021).
11. N. W. Morrell, M. A. Aldred, W. K. Chung, C. G. Elliott, W. C. Nichols, F. Soubrier, R. C. Trembath, J. E. Loyd, Genetics and genomics of pulmonary arterial hypertension. *Eur Respir J* **53**, 1801899 (2019).
12. L. Statello, C. J. Guo, L. L. Chen, M. Huarte, Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* **22**, 96-118 (2021).
13. M. Sun, W. L. Kraus, From Discovery to Function: The Expanding Roles of Long Non-Coding RNAs in Physiology and Disease. *Endocr Rev* **36**, 25-64 (2015).
14. K. R. Zahid, U. Raza, J. Chen, U. J. Raj, D. Gou, Pathobiology of pulmonary artery hypertension: role of long non-coding RNAs. *Cardiovasc Res* **116**, 1937-1947 (2020).
15. C. M. Zehendner, C. Valasarajan, A. Werner, J. N. Boeckel, F. C. Bischoff, D. John, T. Weirick, S. F. Glaser, O. Roszbach, N. Jae, S. Demolli, F. Khassafi, K. Yuan, V. A. de Jesus Perez, K. M. Michalik, W. Chen, W. Seeger, A. Guenther, R. M. Wasnick, S. Uchida, A. M. Zeiher, S. Dimmeler, S. S. Pullamsetti, Long Noncoding RNA TYKRIL Plays a Role in Pulmonary Hypertension via the p53-mediated Regulation of PDGFRbeta. *Am J Respir Crit Care Med* **202**, 1445-1457 (2020).
16. J. Omura, K. Habbout, T. Shimauchi, W. H. Wu, S. Breuils-Bonnet, E. Tremblay, S. Martineau, V. Nadeau, K. Gagnon, F. Mazoyer, J. Perron, F. Potus, J. H. Lin, H. Zafar, D. G. Kiely, A. Lawrie, S. L. Archer, R. Paulin, S. Provencher, O. Boucherat, S. Bonnet, Identification of Long Noncoding RNA H19 as a New Biomarker and Therapeutic Target in Right Ventricular Failure in Pulmonary Arterial Hypertension. *Circulation* **142**, 1464-1484 (2020).
17. T. Bertero, K. A. Cotrill, Y. Lu, C. M. Haeger, P. Dieffenbach, S. Annis, A. Hale, B. Bhat, V. Kaimal, Y. Y. Zhang, B. B. Graham, R. Kumar, R. Saggarr, R. Saggarr, W. D. Wallace, D. J. Ross, S. M. Black, S. Fratz, J. R. Fineman, S. O. Vargas, K. J. Haley, A. B. Waxman, B. N. Chau, L. E. Fredenburgh, S. Y. Chan, Matrix remodeling promotes pulmonary hypertension through feedback mechanoactivation of the YAP/TAZ-miR-130/301 circuit *Cell Rep* **13**, 1016-

- 1032 (2015).
18. R. C. Rao, Y. Dou, Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* **15**, 334-346 (2015).
 19. J. Schodel, S. Oikonomopoulos, J. Ragoussis, C. W. Pugh, P. J. Ratcliffe, D. R. Mole, High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* **117**, e207-217 (2011).
 20. K. Kondo, W. Y. Kim, M. Lechpammer, W. G. Kaelin, Jr., Inhibition of HIF2alpha is sufficient to suppress pVHL-defective tumor growth. *PLoS Biol* **1**, E83 (2003).
 21. N. Tuvshinjargal, W. Lee, B. Park, K. Han, PRIdictor: Protein-RNA Interaction predictor. *Biosystems* **139**, 17-22 (2016).
 22. J. Kim, H. Lee, S. J. Yi, K. Kim, Gene regulation by histone-modifying enzymes under hypoxic conditions: a focus on histone methylation and acetylation. *Exp Mol Med* **54**, 878-889 (2022).
 23. S. Sebastian, P. Sreenivas, R. Sambasivan, S. Cheedipudi, P. Kandalla, G. K. Pavlath, J. Dhawan, MLL5, a trithorax homolog, indirectly regulates H3K4 methylation, represses cyclin A2 expression, and promotes myogenic differentiation. *Proc Natl Acad Sci U S A* **106**, 4719-4724 (2009).
 24. D. A. Pereira-Martins, I. Weinhauser, J. L. Coelho-Silva, P. L. Franca-Neto, L. Y. Almeida, T. M. Bianco, C. L. Silva, R. F. Franca, F. Traina, E. M. Rego, J. J. Schuringa, A. R. Lucena-Araujo, MLL5 improves ATRA driven differentiation and promotes xenotransplant engraftment in acute promyelocytic leukemia model. *Cell Death Dis* **12**, 371 (2021).
 25. A. Subramanian, P. Tamayo, V. Mootha, S. Mukherjee, B. Ebert, M. Gillette, A. Paulovich, S. Pomeroy, T. Golub, E. Lander, J. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Nat Acad Sci USA* **102**, 15545-15550 (2005).
 26. Y. Du, N. Wei, R. Ma, S. H. Jiang, D. Song, Long Noncoding RNA MIR210HG Promotes the Warburg Effect and Tumor Growth by Enhancing HIF-1alpha Translation in Triple-Negative Breast Cancer. *Front Oncol* **10**, 580176 (2020).
 27. A. Ortiz-Barahona, D. Villar, N. Pescador, J. Amigo, L. del Peso, Genome-wide identification of hypoxia-inducible factor binding sites and target genes by a probabilistic model integrating transcription-profiling data and in silico binding site prediction. *Nucleic Acids Res* **38**, 2332-2345 (2010).
 28. C. J. Rhodes, K. Batai, M. Bleda, M. Haimel, L. Southgate, M. Germain, M. W. Pauciulo, C. Hadinnapola, J. Aman, B. Girerd, A. Arora, J. Knight, K. B. Hanscombe, J. H. Karnes, M. Kaakinen, H. Gall, A. Ulrich, L. Harbaum, I. Cebola, J. Ferrer, K. Lutz, E. M. Swietlik, F. Ahmad, P. Amouyel, S. L. Archer, R. Argula, E. D. Austin, D. Badesch, S. Bakshi, C. Barnett, R. Benza, N. Bhatt, H. J. Bogaard, C. D. Burger, M. Chakinala, C. Church, J. G. Coghlan, R. Condliffe, P. A. Corris, C. Danesino, S. Dobbie, C. G. Elliott, J. Elwing, M. Eyries, T. Fortin, A. Franke, R. P. Frantz, A. Frost, J. G. N. Garcia, S. Ghio, H. A. Ghofrani, J. S. R. Gibbs, J. Harley, H. He, N. S. Hill, R. Hirsch, A. C. Houweling, L. S. Howard, D. Ivy, D. G. Kiely, J. Klinger, G. Kovacs, T. Lahm, M. Laudes, R. D. Machado, R. V. MacKenzie Ross, K. Marsolo, L. J. Martin, S. Moledina, D. Montani, S. D. Nathan, M. Newnham, A. Olschewski, H. Olschewski, R. J. Oudiz, W. H. Ouwehand, A. J. Peacock, J. Pepke-Zaba, Z. Rehman, I. Robbins, D. M. Roden, E. B. Rosenzweig, G. Saydain, L. Scelsi, R. Schilz, W. Seeger, C. M. Shaffer, R. W. Simms, M. Simon, O. Sitbon, J. Suntharalingam, H. Tang, A. Y. Tchourbanov, T. Thenappan, F. Torres, M. R. Toshner, C. M. Treacy, A. Vonk Noordegraaf, Q. Waisfisz, A. K. Walsworth, R. E. Walter, J. Wharton, R. J. White, J. Wilt, S. J. Wort, D. Yung, A. Lawrie, M. Humbert, F. Soubrier, D. A. Tregouet, I. Prokopenko, R. Kittles, S. Graf, W. C. Nichols, R. C. Trembath, A. A. Desai, N. W. Morrell, M. R. Wilkins, U. N. B. R. D. Consortium, U. P. C. S. Consortium, U. P. B. Consortium, Genetic determinants of risk in pulmonary arterial hypertension: international genome-wide association studies and meta-analysis. *Lancet Respir Med* **7**, 227-238 (2019).
 29. I. Jung, A. Schmitt, Y. Diao, A. J. Lee, T. Liu, D. Yang, C. Tan, J. Eom, M. Chan, S. Chee, Z. Chiang, C. Kim, E. Masliah, C. L. Barr, B. Li, S. Kuan, D. Kim, B. Ren, A compendium of promoter-centered long-range chromatin interactions in the human genome. *Nat Genet* **51**, 1442-1449 (2019).
 30. K. Fujimori, T. Tezuka, H. Ishiura, J. Mitsui, K. Doi, J. Yoshimura, H. Tada, T. Matsumoto, M.

- Isoda, R. Hashimoto, N. Hattori, T. Takahashi, S. Morishita, S. Tsuji, W. Akamatsu, H. Okano, Modeling neurological diseases with induced pluripotent cells reprogrammed from immortalized lymphoblastoid cell lines. *Mol Brain* **9**, 88 (2016).
31. M. K. Culley, J. Zhao, Y. Y. Tai, Y. Tang, D. Perk, V. Negi, Q. Yu, C. C. Woodcock, A. Handen, G. Speyer, S. Kim, Y. C. Lai, T. Satoh, A. M. Watson, Y. A. Aaraj, J. Sembrat, M. Rojas, D. Goncharov, E. A. Goncharova, O. F. Khan, D. G. Anderson, J. E. Dahlman, A. U. Gurkar, R. Lafyatis, A. U. Fayyaz, M. M. Redfield, M. T. Gladwin, M. Rabinovitch, M. Gu, T. Bertero, S. Y. Chan, Frataxin deficiency promotes endothelial senescence in pulmonary hypertension. *J Clin Invest* **131**, e136459 (2021).
32. Q. Yu, Y. Y. Tai, Y. Tang, J. Zhao, V. Negi, M. K. Culley, J. Pilli, W. Sun, K. Brugger, J. Mayr, R. Saggarr, R. Saggarr, W. D. Wallace, D. J. Ross, A. B. Waxman, S. G. Wendell, S. J. Mullett, J. Sembrat, M. Rojas, O. F. Khan, J. E. Dahlman, M. Sugahara, N. Kagiyama, T. Satoh, M. Zhang, N. Feng, J. Gorcsan Iii, S. O. Vargas, K. J. Haley, R. Kumar, B. B. Graham, R. Langer, D. G. Anderson, B. Wang, S. Shiva, T. Bertero, S. Y. Chan, BOLA3 Deficiency Controls Endothelial Metabolism and Glycine Homeostasis in Pulmonary Hypertension. *Circulation* **139**, 2238-2255 (2019).
33. M. K. Steiner, O. L. Syrkina, N. Kolliputi, E. J. Mark, C. A. Hales, A. B. Waxman, Interleukin-6 overexpression induces pulmonary hypertension. *Circ Res* **104**, 236-244, 228p following 244 (2009).
34. D. Greiner, T. Bonaldi, R. Eskeland, E. Roemer, A. Imhof, Identification of a specific inhibitor of the histone methyltransferase SU(VAR)3-9. *Nat Chem Biol* **1**, 143-145 (2005).
35. J. L. Platt, R. Salama, J. Smythies, H. Choudhry, J. O. Davies, J. R. Hughes, P. J. Ratcliffe, D. R. Mole, Capture-C reveals preformed chromatin interactions between HIF-binding sites and distant promoters. *EMBO Rep* **17**, 1410-1421 (2016).
36. K. D. Falkenberg, K. Rohlenova, Y. Luo, P. Carmeliet, The metabolic engine of endothelial cells. *Nat Metab* **1**, 937-946 (2019).
37. X. Zhang, W. Novera, Y. Zhang, L. W. Deng, MLL5 (KMT2E): structure, function, and clinical relevance. *Cell Mol Life Sci* **74**, 2333-2344 (2017).
38. S. Dong, M. F. Walker, N. J. Carriero, M. DiCola, A. J. Willsey, A. Y. Ye, Z. Waqar, L. E. Gonzalez, J. D. Overton, S. Frahm, J. F. Keaney, 3rd, N. A. Teran, J. Dea, J. D. Mandell, V. Hus Bal, C. A. Sullivan, N. M. DiLullo, R. O. Khalil, J. Gockley, Z. Yuksel, S. M. Sertel, A. G. Ercan-Sencicek, A. R. Gupta, S. M. Mane, M. Sheldon, A. I. Brooks, K. Roeder, B. Devlin, M. W. State, L. Wei, S. J. Sanders, De novo insertions and deletions of predominantly paternal origin are associated with autism spectrum disorder. *Cell Rep* **9**, 16-23 (2014).
39. J. L. Bjorkegren, S. Hagg, H. A. Talukdar, H. Foroughi Asl, R. K. Jain, C. Cedergren, M. M. Shang, A. Rossignoli, R. Takolander, O. Melander, A. Hamsten, T. Michoel, J. Skogsberg, Plasma cholesterol-induced lesion networks activated before regression of early, mature, and advanced atherosclerosis. *PLoS Genet* **10**, e1004201 (2014).
40. V. Negi, J. Yang, G. Speyer, A. Pulgarin, A. Handen, J. Zhao, Y. Y. Tai, Y. Tang, M. K. Culley, Q. Yu, P. Forsythe, A. Gorelova, A. M. Watson, Y. Al Aaraj, T. Satoh, M. Sharifi-Sanjani, A. Rajaratnam, J. Sembrat, S. Provencher, X. Yin, S. O. Vargas, M. Rojas, S. Bonnet, S. Torrino, B. K. Wagner, S. L. Schreiber, M. Dai, T. Bertero, I. Al Ghouleh, S. Kim, S. Y. Chan, Computational repurposing of therapeutic small molecules from cancer to pulmonary hypertension. *Sci Adv* **7**, eabh3794 (2021).
41. P. Sen, P. P. Shah, R. Nativio, S. L. Berger, Epigenetic Mechanisms of Longevity and Aging. *Cell* **166**, 822-839 (2016).
42. L. Phan, Y. Jin, H. Zhang, W. Qiang, E. Shekhtman, D. Shao, D. Revoe, R. Villamarin, E. Ivanchenko, M. Kimura, Z. Y. Wang, L. Hao, N. Sharapova, M. Bihan, A. Sturcke, M. Lee, N. Popova, W. Wu, C. Bastiani, M. Ward, J. B. Holmes, V. Lyoshin, K. Kaur, E. Moyer, M. Felolo, B. L. Kattman, ALFA: Allele Frequency Aggregator. *National Center for Biotechnology Information, U.S. National Library of Medicine*, (2020).
43. P. Zhou, Z. Wang, X. Yuan, C. Zhou, L. Liu, X. Wan, F. Zhang, X. Ding, C. Wang, S. Xiong, Z. Wang, J. Yuan, Q. Li, Y. Zhang, Mixed lineage leukemia 5 (MLL5) protein regulates cell cycle progression and E2F1-responsive gene expression via association with host cell factor-1 (HCF-1). *J Biol Chem* **288**, 17532-17543 (2013).

44. S. Pasupneti, W. Tian, A. B. Tu, P. Dahms, E. Granucci, A. Gandjeva, M. Xiang, E. C. Butcher, G. L. Semenza, R. M. Tuder, X. Jiang, M. R. Nicolls, Endothelial HIF-2 α as a Key Endogenous Mediator Preventing Emphysema. *Am J Respir Crit Care Med* **202**, 983-995 (2020).
45. A. P. Acharya, Y. Tang, T. Bertero, Y. Y. Tai, L. D. Harvey, C. C. Woodcock, W. Sun, R. Pineda, N. Mitash, M. Konigshoff, S. R. Little, S. Y. Chan, Simultaneous Pharmacologic Inhibition of Yes-Associated Protein 1 and Glutaminase 1 via Inhaled Poly(Lactic-co-Glycolic) Acid-Encapsulated Microparticles Improves Pulmonary Hypertension. *J Am Heart Assoc* **10**, e019091 (2021).
46. D. R. Nyholt, A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* **74**, 765-769 (2004).
47. Y. Zhao, D. Wu, D. Jiang, X. Zhang, T. Wu, J. Cui, M. Qian, J. Zhao, S. Oesterreich, W. Sun, T. Finkel, G. Li, A sequential methodology for the rapid identification and characterization of breast cancer-associated functional SNPs. *Nat Commun* **11**, 3340 (2020).
48. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
49. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
50. M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016).
51. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoutte, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
52. C. S. Ross-Innes, R. Stark, A. E. Teschendorff, K. A. Holmes, H. R. Ali, M. J. Dunning, G. D. Brown, O. Gojis, I. O. Ellis, A. R. Green, S. Ali, S. F. Chin, C. Palmieri, C. Caldas, J. S. Carroll, Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389-393 (2012).
53. T. L. Bailey, N. Williams, C. Misleh, W. W. Li, MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* **34**, W369-373 (2006).
54. P. J. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, M. J. de Hoon, Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422-1423 (2009).
55. H. Touzet, J. S. Varre, Efficient and accurate P-value computation for Position Weight Matrices. *Algorithms Mol Biol* **2**, 15 (2007).
56. N. Hafeez, A. Kirillova, Y. Yue, R. J. Rao, N. J. Kelly, W. El Khoury, Y. Al Aaraj, Y. Y. Tai, A. Handen, Y. Tang, D. Jiang, T. Wu, Y. Zhang, D. McNamara, T. V. Kudryashova, E. A. Goncharova, D. Goncharov, T. Bertero, M. Nouraie, G. Li, W. Sun, S. Y. Chan, Single Nucleotide Polymorphism rs9277336 Controls the Nuclear Alpha Actinin 4-Human Leukocyte Antigen-DPA1 Axis and Pulmonary Endothelial Pathophenotypes in Pulmonary Arterial Hypertension. *J Am Heart Assoc* **12**, e027894 (2023).
57. S. Pelletier, S. Gingras, D. R. Green, Mouse genome engineering via CRISPR-Cas9 for study of immune function. *Immunity* **42**, 18-27 (2015).
58. C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

Acknowledgments

We thank A. Handen (University of Pittsburgh) for input on sequencing analyses; Sanofi for gapmer oligonucleotides; B. McMahon (The University of Pittsburgh Small Animal Ultrasonography Core) for rodent echocardiogram measurements; S. Gingras (University of Pittsburgh) for generation of the lncRNA knockout mouse; and J. Park (University of Pennsylvania) for scientific advice. We thank the NIHR BioResource – Rare Diseases Consortium, UK PAH Cohort Study Consortium, and the US PAH Biobank Consortium as reported in (28) for data access. The *All of Us* Research Program is supported by the National Institutes of Health, Office of the Director: Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 OD026556; 1 OT2 OD026550; 1 OT2 OD026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 OD026551; 1 OT2 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The Participant Center: U24 OD023176; Participant Technology Systems Center: 1 U24 OD023163; Communications and Engagement: 3 OT2 OD023205; 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276. In addition, the *All of Us* Research Program would not be possible without the partnership of its participants.

Funding: This work was supported by NIH grants R01 HL124021 to S.Y.C., HL 122596 to S.Y.C., R00 HL133473 to H.W., HL105333 to W.N. and M.P.; HL137927, HL147148, and HL089856 to M.H.C.; HL141601 to A.B.; and HL143185 to V.N.P.; HL129964 to N.J.K.; AHA grant 18EIA33900027 to S.Y.C.; Plan Cancer 2018 grant 18CN045 to C.L. and B.M.; the French National Research Agency ANR-18-CE14-0025, ANR-21-CE44-0036 and ANR-20-CE14-0006 to T.B.; the French National Cancer Institute INCA-PLBIO 21-094 to T.B.; the BHF Intermediate Basic Science Research fellowship FS/15/59/31839 to C.J.R.; Academy of Medical Sciences Springboard fellowship SBF004\1095 to C.J.R.; and the NSF DBI-1942143 to J.W. The UK National Cohort of Idiopathic and Heritable PAH was supported by the NIHR BioResource; the BHF (SP/12/12/29836); and the UK Medical Research Council (MR/K020919/1). The Advanced High Resolution Rodent Ultrasound Imaging System provided by The University of Pittsburgh Small Animal Ultrasonography Core was supported by NIH

Author Contributions: Y.Y.T., Q.Y., and S.Y.C. conceived and designed the experiments. Y.Y.T., Q.Y., Y.T., W.S., R.J.B., H.W., N.K., J.Z., S.T., Y.A.A., V.N., M.L., F.B., T.S., J.W., D.G., G.L., B.W., S.S., B.K., T.B., and S.Y.C. provided experimental infrastructure and performed the experiments. M.N., M.H.C., A.B., W.C.N., T.S.A., L.S., R.C.T., O.S., M.H., S.G., N.W.M., C.J.R., M.R.W., M.W.P., J.H.K., W.E.K., and A.A.D. performed the SNV association analyses. B.S., B.B., and N.C. provided expertise in lncRNA manipulation. C.L. and B.M. provided structural lncRNA modeling. T.S.A., Y.Z., D.M., W.C.N., S.O.V., and A.A.D. provided human samples. Y.Y.T., Y.T., J.Z. obtained human and rodent PH samples. C.G.C. and S.S. performed Seahorse experiments. S.O., S.K., and G.S. performed sequencing analyses. Y.Y.T., Q.Y., and S.Y.C. wrote the manuscript. All authors participated in interpreting the results and revising the manuscript.

Competing interests: S.Y.C. has served as a consultant for Merck and United Therapeutics. S.Y.C. is a director, officer, and shareholder in Synhale Therapeutics. M.R.W. has served as a consultant for Novartis, Accelerant, MSD, Janssen, MorphogenIX, Chiesi and BenevolentAI. S.Y.C. has held research grants from United Therapeutics and Bayer. M.H.C. has received grant support from Bayer. S.Y.C., S.K., and T.B. have filed patent applications regarding the targeting of metabolism in pulmonary hypertension (Patent: Compositions and Methods for Treating Pulmonary Vascular Disease, US10925869B2). M.R.W. has filed a patent application regarding ZIP12 antibodies (ZIP12 ANTIBODY, WO2022064216A1). The other authors declare no competing interests.

Data and materials availability: All data associated with this study are present in the paper or Supplementary Material. RNA-Seq and ChIP-seq data have been uploaded to the NCBI Gene Expression Omnibus (GEO) repository (GSE232799). De-identified human/patient data are listed in Supplemental Tables. All original code used to process sequencing data has been uploaded to Zenodo (DOI:10.5281/zenodo.10096872). All newly generated cell lines, plasmids or mouse lines will

be shared upon request. Human and rodent samples and data from Cincinnati Children's Hospital Medical Center and Brigham and Women's Hospital shared with the University of Pittsburgh are covered by material transfer agreements.

Figure Legends

Figure 1. Human *KMT2E-AS1* and neighboring *KMT2E* are upregulated across *in vivo* and *in vitro* models of PH. (A) Gene structure of mouse lncRNA *5031425E22Rik* (*E22*) located adjacent to protein-coding gene *Kmt2e* is encoded on the opposite DNA strand and positioned in opposite transcriptional direction. Human lncRNA ortholog *KMT2E-AS1* and neighboring gene *KMT2E* show similar genomic architecture; sequence conservation (purple box 500bp region) is shown within mouse *E22* and human *KMT2E-AS1*. (B-E) Human lncRNA *KMT2E-AS1* and *KMT2E* transcripts are quantified in lung tissue (B-C) ($n = 7-10$; $*p < 0.05$, unpaired Student's t-test; data represent the mean \pm SEM), and CD31⁺ cells (D-E) ($n = 3-4$; $*p < 0.05$, $**p < 0.01$, unpaired Student's t-test; data represent the mean \pm SEM) of patients with WSPH Group 1 PAH (Table S1) by real time-quantitative polymerase chain reaction (RT-qPCR). (F-H) Representative fluorescence *in situ* hybridization (FISH) and immunofluorescence (IF) staining and quantifications of *KMT2E-AS1* (red; F), *KMT2E* (red; G), and H3K4me3 (red; H) are shown in CD31⁺ endothelium of human lung from individuals with Group 1 and Group 3 PH vs. non-PH controls ($n = 5-8$; $***p < 0.001$, $****p < 0.0001$ vs. No PH, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent the mean \pm SEM). Scale bars, 50 μ m. (I) RT-qPCR analysis of *KMT2E-AS1* is shown in cultured human pulmonary arterial endothelial cells (PAECs), adventitial fibroblasts (PAAFs), and pulmonary artery smooth muscle cells (PASMCs) ($n = 4-6$; $**p < 0.01$, $***p < 0.001$, unpaired Student's t-test; data represent the mean \pm SEM). (J) RT-qPCR analysis of *KMT2E* is shown in human PAECs, PASMCs, and PAAFs ($n = 4-6$; $****p < 0.0001$, unpaired Student's t-test for PAECs and PAAFs, Mann-Whitney test for PASMCs; data represent mean \pm SEM). (K) *KMT2E-AS1* (right graph) is measured in cytosolic and nucleic fractions of human cell types by RT-qPCR (GAPDH and U1 snRNA served as cytoplasmic and nuclear controls, respectively; left graphs) ($n = 3$; $*p < 0.05$, $****p < 0.0001$, unpaired Student's t-test; data represent mean \pm SEM).

Figure 2. *KMT2E-AS1* interacts with *KMT2E* to enhance protein stability and increase histone 3 lysine 4 trimethylation (H3K4me3). (A-B) *KMT2E-AS1* (A) and *KMT2E* (B) transcripts are quantified in hypoxic human PAECs with knockdown (siRNA) of HIF-2 α and/or HIF-1 α by RT-qPCR

in comparison with normoxic scramble control (NC) ($n = 4-6$; ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (C) *KMT2E-AS1*, *KMT2E*, and *HIF-2 α* transcripts are quantified in normoxic human PAECs with lentivirus-mediated overexpression of a constitutively active HIF-2 α (LV-HIF-2 α vs. LV-GFP control) by RT-qPCR ($n = 4$; **** $p < 0.0001$, unpaired Student's t-test; data represent mean \pm SEM). (D) HIF-2 α and KMT2E proteins are measured in normoxic human PAECs transduced with LV-HIF-2 α vs. LV-GFP control by immunoblot and densitometry ($n = 3$; *** $p < 0.001$, unpaired Student's t-test; data represent mean \pm SEM). (E) RT-qPCR analysis of *KMT2E-AS1* expression is shown in human PAECs under hypoxia and siRNA knockdown of *KMT2E* or *KMT2E-AS1* ($n = 4$; *** $p < 0.001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (F) KMT2E protein is measured in human PAECs under hypoxia and knockdown of *KMT2E* or *KMT2E-AS1* by immunoblot and densitometry ($n = 4$; * $p < 0.05$, ** $p < 0.01$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (G) *KMT2E-AS1* expression is quantified in human PAECs post-transduction with *KMT2E-AS1* lentivirus (LV-AS1) vs. GFP control (LV-GFP) under normoxia (left panel) and hypoxia (right panel) by RT-qPCR ($n = 3$; ** $p < 0.01$, unpaired Student's t-test; data represent mean \pm SEM). (H) KMT2E protein is measured in human PAECs transduced as in (G) by immunoblot and densitometry ($n = 3$; * $p < 0.05$, ** $p < 0.01$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (I) RT-qPCR analysis of *KMT2E* transcript is shown in human PAECs under hypoxia and *KMT2E-AS1* knockdown or 4h exposure to the transcriptional inhibitor actinomycin D (ActD) ($n = 6$; * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (J) KMT2E protein is measured in human PAECs after *KMT2E-AS1* knockdown or actinomycin D (ActD) exposure under hypoxia by immunoblot and densitometry ($n = 3$; * $p < 0.05$, ** $p < 0.01$, ns – not significant, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (K) Via RNA immunoprecipitation (RIP)-qPCR (IP: KMT2E vs. IgG negative control) of KMT2E protein in normoxia and hypoxia (left immunoblot), *KMT2E-AS1* is quantified by RT-qPCR in the IP fraction of human PAECs after *KMT2E-AS1* knockdown or ActD exposure (4h) under hypoxia (right graph) ($n = 4$; * $p < 0.05$, ** $p < 0.01$, Kruskal-

Wallis test followed by Dunn's post-hoc analysis; data represent mean \pm SEM). (L) KMT2E protein is measured in human PAECs after *KMT2E-AS1* knockdown or proteasomal inhibitor MG132 in hypoxia by immunoblot and densitometry ($n = 4$; * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (M) H3K4me3, H3K9me3, and H3K27me3 are quantified in human PAECs after hypoxia and *KMT2E-AS1* knockdown by immunoblot and densitometry ($n = 3$; * $p < 0.05$, ** $p < 0.01$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis for H3K4me3 and H3K9me3, Kruskal-Wallis test followed by Dunn's post-hoc analysis for H3K27me3; data represent mean \pm SEM). (N) Nuclear interaction of H3K4me3 marks and KMT2E protein is measured by proximity ligation assay (PLA, red, left images) and quantified by PLA counts per cell (right graph) in human PAECs after lentiviral expression of *KMT2E-AS1* (full length) or *KMT2E-AS1* deletion mutant (Fig. S8) vs. GFP control ($n = 3-4$; *** $p < 0.001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). Scale bar, 20 μm . (O) H3K4me3 is measured in human PAECs under lentiviral expression of *KMT2E-AS1* (full length) or *KMT2E-AS1* deletion mutant by immunoblot and densitometry ($n = 3$; ** $p < 0.01$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM).

Figure 3. *KMT2E-AS1* regulates a gene network driving hypoxic metabolic reprogramming. (A) RNA sequencing of human PAECs reveals that 2480 genes are altered by hypoxia and similarly reversed by either *KMT2E-AS1* or KMT2E in hypoxia. Gene set enrichment analysis (GSEA) reveals the major biological processes represented by these reversed genes. (B) Heatmaps display genes in "Hypoxia" and "Metabolism" networks that are altered by hypoxia (leftmost column) and are reversed by *KMT2E-AS1* (middle column) and *KMT2E* (rightmost column) knockdown in hypoxia. Adjusted $p < 0.05$ for each gene shown. H3K4me3 chromatin immunoprecipitation and sequencing (ChIP-Seq) was also performed in hypoxic vs. normoxic PAECs. A sub-cohort of these genes displays increased H3K4me3 marks in hypoxia by co-analyzing these ChIP-Seq and RNA Seq data (* indicates methylated genes with adjusted $p < 0.05$). (C) By ChIP-qPCR (IP: H3K4me3 Ab vs. IgG control), H3K4me3 marks are quantified at the promoter site of the lncRNA *miR210hg* ($n = 3$; ** $p < 0.01$, **** $p < 0.0001$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm

SEM). (D-E) Seahorse assay quantifies extracellular acidification rate (ECAR) (D) and baseline oxygen consumption rate (OCR) (E) of human PAECs after *KMT2E-AS1* knockdown and HIF-2 α overexpression ($n = 10-12$; * $p < 0.05$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (F-G) ECAR (F) and baseline OCR (G) are quantified after *KMT2E-AS1* overexpression ($n = 12$; * $p < 0.05$, *** $p < 0.001$, unpaired Student's t-test; data represent mean \pm SEM). (H-I) Lactate dehydrogenase (LDH) enzymatic activity of human PAECs is measured after *KMT2E-AS1* knockdown (H) and *KMT2E-AS1* overexpression (I), a representative measure of glycolysis ($n = 4$; * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (J) VEGF expression is quantified in hypoxic PAECs after knockdown of *KMT2E* and *KMT2E-AS1* by immunoblot and densitometry ($n = 4$; ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM), overlaid with ChIP-seq of H3K4me3 at the VEGFA gene in hypoxia vs. normoxia (fold change of 1.48).

Figure 4. *KMT2E-AS1* induces HIF-2 α activation and prevents HIF-2 α degradation in driving endothelial pathophenotypes. (A-B) HIF-2 α expression is measured in hypoxic human PAECs after knockdown of *KMT2E-AS1* (A), and overexpression of *KMT2E-AS1* (B) by immunoblot and densitometry ($n = 3-4$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, one-way (A) or two-way (B) ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (C-D) Elongin C (*ELOC*) RNA (C) and protein (D) expression are quantified in human PAECs after hypoxia and *KMT2E-AS1* knockdown (**Fig. 3B**) by RT-qPCR and immunoblot, respectively ($n = 3-5$; * $p < 0.05$, ** $p < 0.01$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (E) HIF-2 α expression is measured in human PAECs after *KMT2E-AS1* knockdown and MG132 treatment by immunoblot and densitometry ($n = 3$; ** $p < 0.01$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (F) Apoptotic caspase 3/7 activity is measured in human PAECs after *KMT2E-AS1* knockdown under hypoxia (left), and *KMT2E-AS1* overexpression under normoxia (right) ($n = 3$; ** $p < 0.01$, **** $p < 0.0001$, unpaired Student's t-test; data represent mean \pm SEM). (G) BrdU

proliferative potential is quantified after *KMT2E-AS1* knockdown in hypoxia (left), and forced *KMT2E-AS1* expression in normoxia (right) ($n = 3-5$; * $p < 0.05$, ** $p < 0.01$, unpaired Student's t-test; data represent mean \pm SEM). (H-I) Scratch wound healing assay (H) measures migration (quantified by percent wound closure) of human PAECs after *KMT2E-AS1* knockdown under hypoxia (left, I), and *KMT2E-AS1* overexpression in normoxia (right, I) ($n = 6$; *** $p < 0.001$, **** $p < 0.0001$, unpaired Student's t-test; data represent mean \pm SEM). (J-K) Human PSMC contraction, quantified by % contraction, is measured in gel matrix (J) with conditioned media from PAECs after knockdown of *KMT2E-AS1* under hypoxia (K, left graph) and after forced expression of *KMT2E-AS1* under normoxia (K, right graph) ($n = 6$; *** $p < 0.001$, unpaired Student's t-test; data represent mean \pm SEM). (L) Endothelin-1 is quantified in conditioned media from human PAECs after *KMT2E* knockdown under hypoxia (left), and forced *KMT2E* expression in normoxia (right) by ELISA ($n = 3$; *** $p < 0.001$, unpaired Student's t-test; data represent mean \pm SEM).

Figure 5. G allele of *KMT2E* SNV rs73184087 binds HIF-2 α to control the *KMT2E-AS1/KMT2E* pair. (A) Among 883 genotyped and imputed SNVs in the PAH discovery cohort (Table S5) within and flanking (+/-200kb) the lncRNA-KMT2E locus, SNVs (Table S6) are displayed with predicted HIF-2 α binding to either the minor or major SNV allele. Independent effective SNV test count was calculated at 53.84 (46). Of those, SNV rs73184087 ranks the highest and meets the P-value threshold of 0.00093 (as indicated by the dashed line on the plot). (B) High-throughput chromatin conformation capture (Hi-C) in lung tissue (29) displays long-range interactions between SNV rs73184087 and the transcription start site/promoter region of *KMT2E-AS1/KMT2E* (as indicated by the blue arcs below the graph). A distance-normalized frequency (magenta dots) greater than the threshold of 2.0 by default (green line) defines a significant interaction with a SNV. (C) By biotin-labeled SNV oligonucleotide incubation with hypoxic human PAEC nuclear extracts followed by streptavidin pulldown and immunoblot with densitometry, SNV binding of HIF-2 α and HIF-1 α is compared across the SNV A vs. G allele ($n = 3$; * $p < 0.05$, unpaired Student's t-test; data represent mean \pm SEM). (D) Luciferase activity is quantified in protein lysates of HEK293T cells after transfection with a

constitutively active HIF-2 α plasmid and a luciferase reporter plasmid carrying the lncRNA-*KMT2E* promoter and SNV rs73184087 (A vs. G allele) ($n = 4$; * $p < 0.05$, unpaired Student's t-test; data represent mean \pm SEM). Luciferase activity is normalized to constitutively secreted alkaline phosphatase (GLuc/SeAP). (E) By ChIP-qPCR (IP: HIF-2 α Ab vs. IgG control) where HIF-2 α is pulled down in hypoxic transformed lymphocytes from WSPH Group 1 PAH patients carrying SNV rs73184087 (G/G) vs. (A/A) genotypes (**Table S2**), SNV enrichment is quantified ($n = 4$; * $p < 0.05$, ** $p < 0.01$, Kruskal-Wallis test followed by Dunn's post-hoc analysis; data represent mean \pm SEM). (F) In a chromatin conformation capture (3C) assay (top diagram) with three pairs of matched transformed lymphocytes (A/A vs. G/G) from WSPH Group 1 PAH patients, PCR is used to detect transcription start site (TSS)/promoter+SNV fusion products indicative of an interaction between SNV rs73184087 and the lncRNA-*KMT2E* promoter. (G-H) 3C assay in human PAECs with SNV A/A genotype defines a ligation product indicative of an interaction between the SNV and promoter (G) but not upstream or downstream of the lncRNA-*KMT2E* promoter (H). (I-J) *KMT2E-AS1* (I) and *KMT2E* (J) transcripts are quantified in transformed lymphocytes carrying A/A or G/G genotype with HIF- α induction by cobalt chloride (50 μ M) by RT-qPCR ($n = 3$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (K) By ChIP-qPCR (IP: HIF-2 α Ab vs. IgG control), where HIF-2 α is pulled down in inducible pluripotent stem cell differentiated endothelial cells (iPSC-ECs) carrying G/G vs. A/A genotype, SNV enrichment is quantified ($n = 3$; * $p < 0.05$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (L) *KMT2E-AS1*, *KMT2E*, and *miR210hg* transcripts are measured in iPSC-ECs carrying A/A vs. G/G genotype by RT-qPCR ($n = 3$; * $p < 0.05$, *** $p < 0.001$, unpaired Student's t-test; data represent mean \pm SEM).

Figure 6. Pulmonary vascular delivery of an AAV6-E22 transgene promotes PH in mice. (A) Experimental design is shown for an AAV serotype 6 (AAV6) carrying either *GFP* or *E22* transgene delivered orotracheally to wildtype C57Bl6 mice 4 weeks before exposure to 3 weeks of chronic hypoxia. (B-C) *E22* (B) and *KMT2E* (C) expression is quantified in mouse AAV6-*GFP* or AAV6-*E22* mouse lung CD31+ endothelial cells by FISH and IF staining ($n = 4$; *** $p < 0.001$, two-way ANOVA

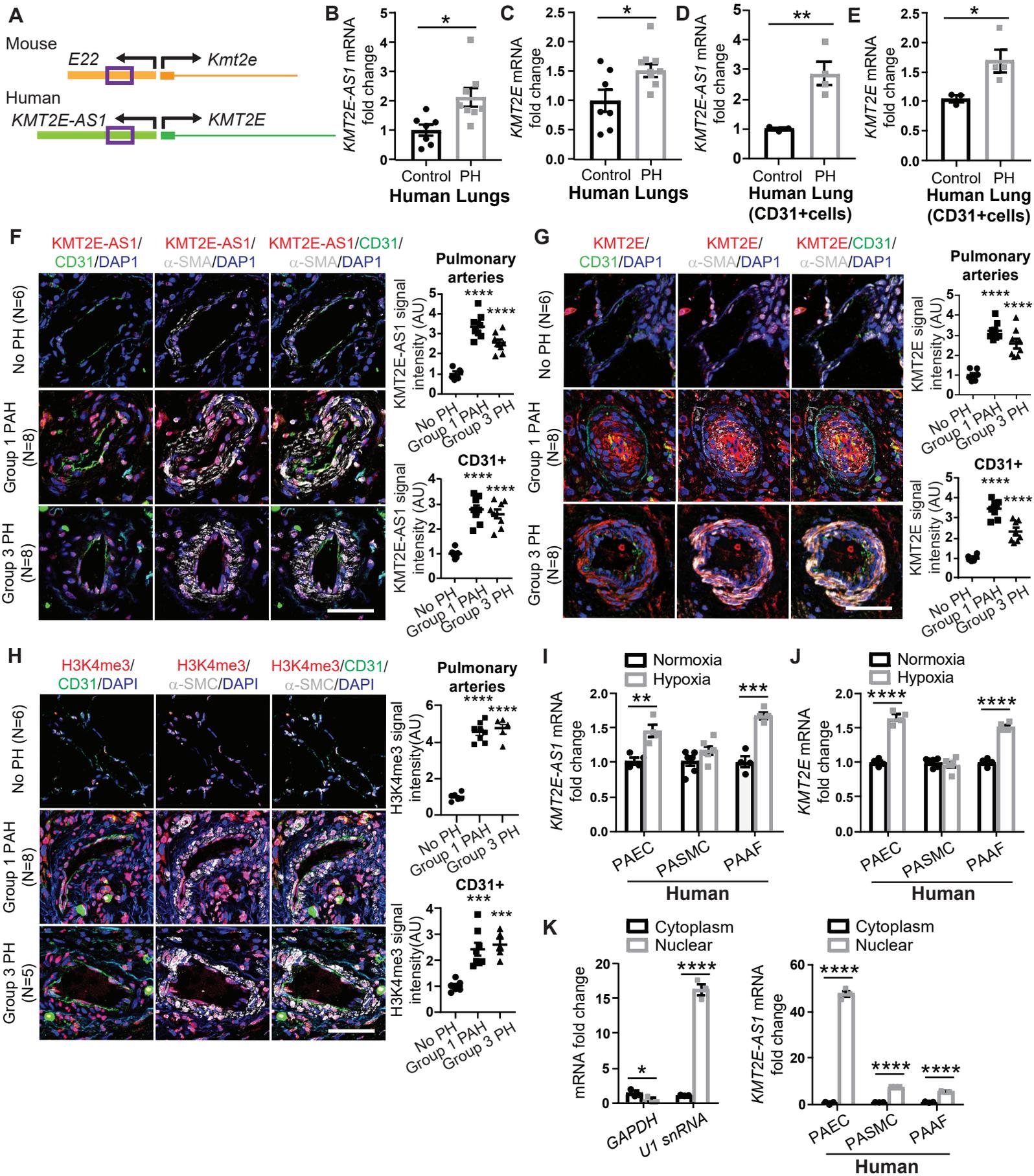
followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (D-H) Representative IF images are shown for H3K4me3 (red; D) and Ki67 proliferation marker stains (red; G) in AAV6-E22 vs. AAV6-GFP mouse lungs. IF quantifications in pulmonary CD31+ vascular endothelium are displayed of H3K4me3 (E), H3K9me3 (F) and Ki67 (H) expression ($n = 4$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). Scale bars, 50 μm . (I-J) Vessel remodeling, indicated by α -SMA staining (white; D,G), is quantified by vessel thickness (I) and muscularization (J) for normoxic and hypoxic AAV6-E22 vs. AAV6-GFP mouse lungs ($n = 4$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (K-L) Right ventricular systolic pressure (RVSP, K) and RV/body weight mass index (L) are measured in AAV6-E22 vs. AAV6-GFP mice ($n = 6-7$; * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$, two-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM).

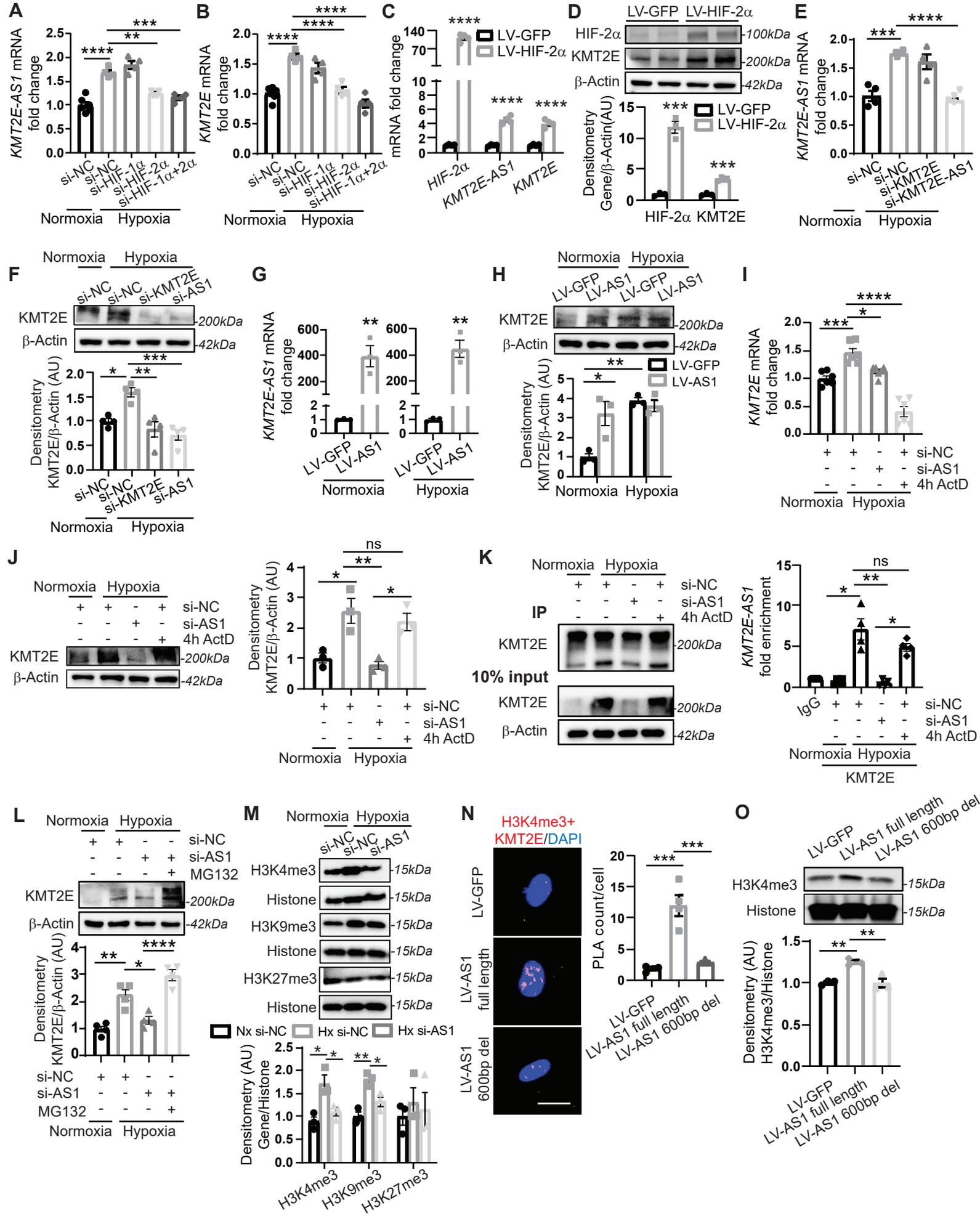
Figure 7. E22 knockout mice display decreased KMT2E and H3K4me3 along with disease improvement in mouse models of WSPH Groups 1 and 3 PH. (A) CRISPR/Cas9 edited mice were generated to be deficient in a conserved 500bp sequence (denoted A-D) shared between human *KMT2E-AS1* and mouse *E22*. (B) FISH and IF staining quantifications are shown for full-length *E22* and KMT2E expression in CD31+ lung endothelial cells of hypoxic *E22* knockout (KO) mice with AD deletion vs. wildtype (WT) controls ($n = 4$; ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (C-G) Representative IF images of H3K4me3 (red; C) and Ki67 (red; F) are displayed in hypoxic *E22* (KO) mice vs. wildtype (WT) controls. IF quantifications in CD31+ PAECs (green) are shown for H3K4me3 (D), H3K9me3 (E) and Ki67 (G) expression ($n = 4$; **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). Scale bars, 50 μm . (H-I) Vessel remodeling, indicated by α -SMA staining (white; C,F), is quantified by vessel thickness (H) and muscularization (I) for *E22* AD KO mice vs. WT controls ($n = 4$; **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (J-K) RVSP (J) and RV remodeling (RV/body weight ratio, K)

are measured in hypoxic *E22* KO mice vs. WT controls ($n = 10-16$; $**p < 0.01$, $***p < 0.001$, unpaired Student's t-test; data represent mean \pm SEM). **(L)** *E22* and KMT2E are quantified in lung CD31+ endothelium of interleukin-6 transgenic (IL-6 Tg) mice crossed onto *E22* KO (AD deletion) mice by FISH and IF staining ($n = 4-6$; $**p < 0.01$, $****p < 0.0001$, Mann-Whitney test for *E22*, unpaired Student's t-test for KMT2E; data represent mean \pm SEM). **(M-Q)** Representative IF images of H3K4me3 (red; M) and Ki67 (red; P) are shown. IF quantifications are displayed of H3K4me3 (N), H3K9me3 (O), and Ki67 (Q) in hypoxic IL-6 Tg *E22* KO mouse lung endothelium as compared to hypoxic IL-6 Tg PAH mice ($n = 4-6$; $*p < 0.05$, $****p < 0.0001$, unpaired Student's t-test for H3K4me3 (N) and Ki67 (Q), Mann-Whitney test for H3K9me3 (O); data represent mean \pm SEM). Scale bars, 50 μ m. **(R-S)** Vessel remodeling (white; M,P) is quantified by vessel thickness (R) and muscularization (S) for IL-6 Tg *E22* KO mice vs. WT controls ($n = 4-6$; $**p < 0.01$, $****p < 0.0001$, unpaired Student's t-test; data represent mean \pm SEM). **(T-U)** RVSP (T) and RV remodeling (U) are measured in IL-6 Tg *E22* KO mice vs. controls ($n = 5-10$; $*p < 0.05$, $**p < 0.01$, unpaired Student's t-test; data represent mean \pm SEM).

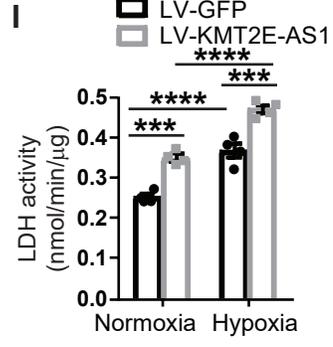
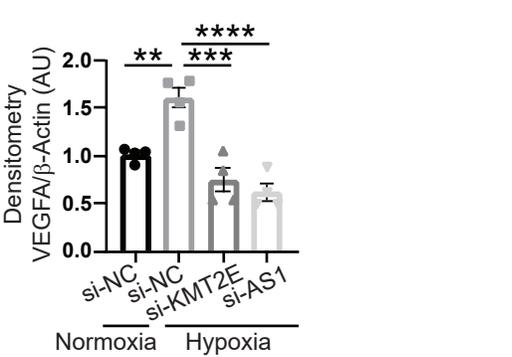
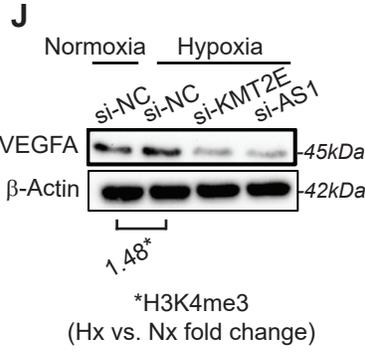
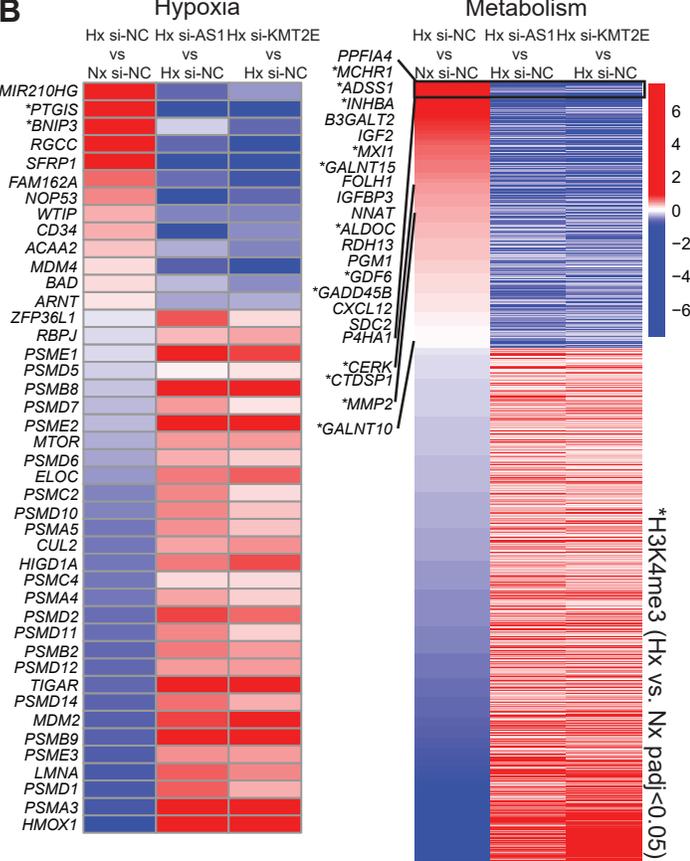
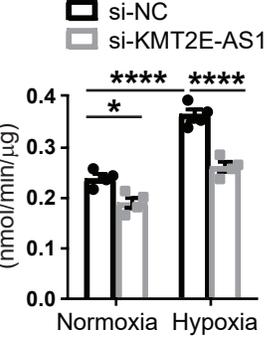
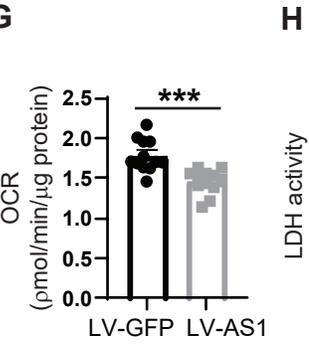
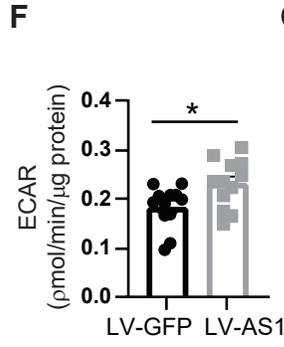
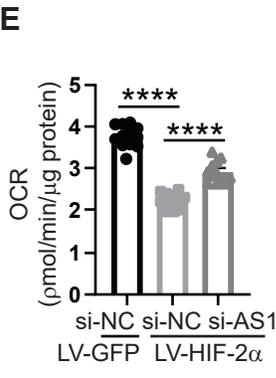
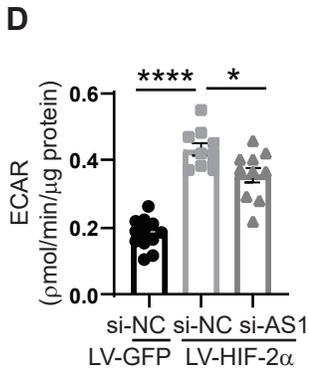
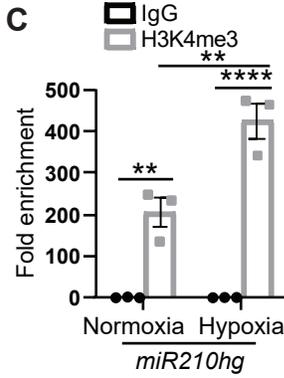
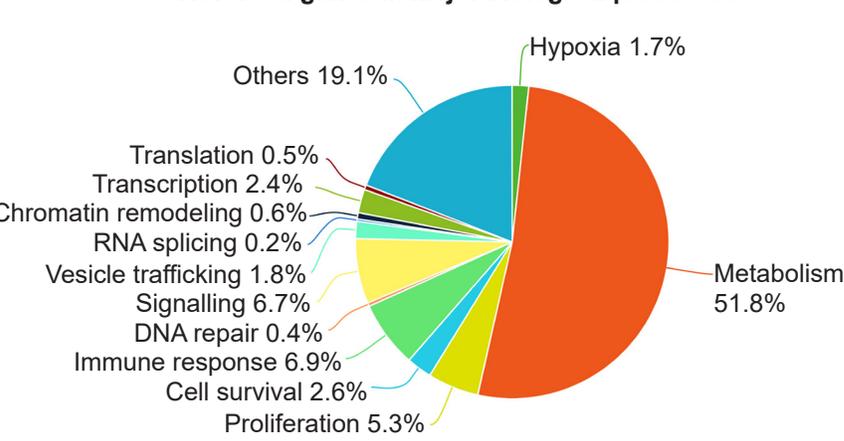
Figure 8. The histone lysine methyltransferase inhibitor chaetocin alleviates PAH in a disease-reversal dosing protocol for SU5416-hypoxic rats. **(A)** Experimental design is shown whereby Sprague-Dawley rats were dosed with SU5416 (SU) and exposed to chronic hypoxia for 3 weeks to generate PAH. Subsequently, PAH rats were dosed with chaetocin vs. DMSO vehicle control by intraperitoneal (IP) injection daily for 2 weeks in normoxia. **(B)** RT-qPCR analysis of *E22* expression in lungs of DMSO- and chaetocin-treated SU+hypoxic rats is compared with normoxic controls ($n = 4-6$; $*p < 0.05$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). **(C)** IF staining quantification of KMT2E expression in CD31+ lung endothelial cells of DMSO- and chaetocin-treated SU+hypoxic rats is compared to normoxic controls ($n = 4-5$; $****p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). **(D-H)** Representative IF images of H3K4me3 (red; D) and Ki67 (red; G) are shown. IF quantifications for H3K4me3 (E), H3K9me3 (F), Ki67 (H) in CD31+ PAECs (green) in SU5415+hypoxic PAH rats are compared to normoxic controls ($n = 4-5$; $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). Scale bars, 50 μ m. **(I-J)**

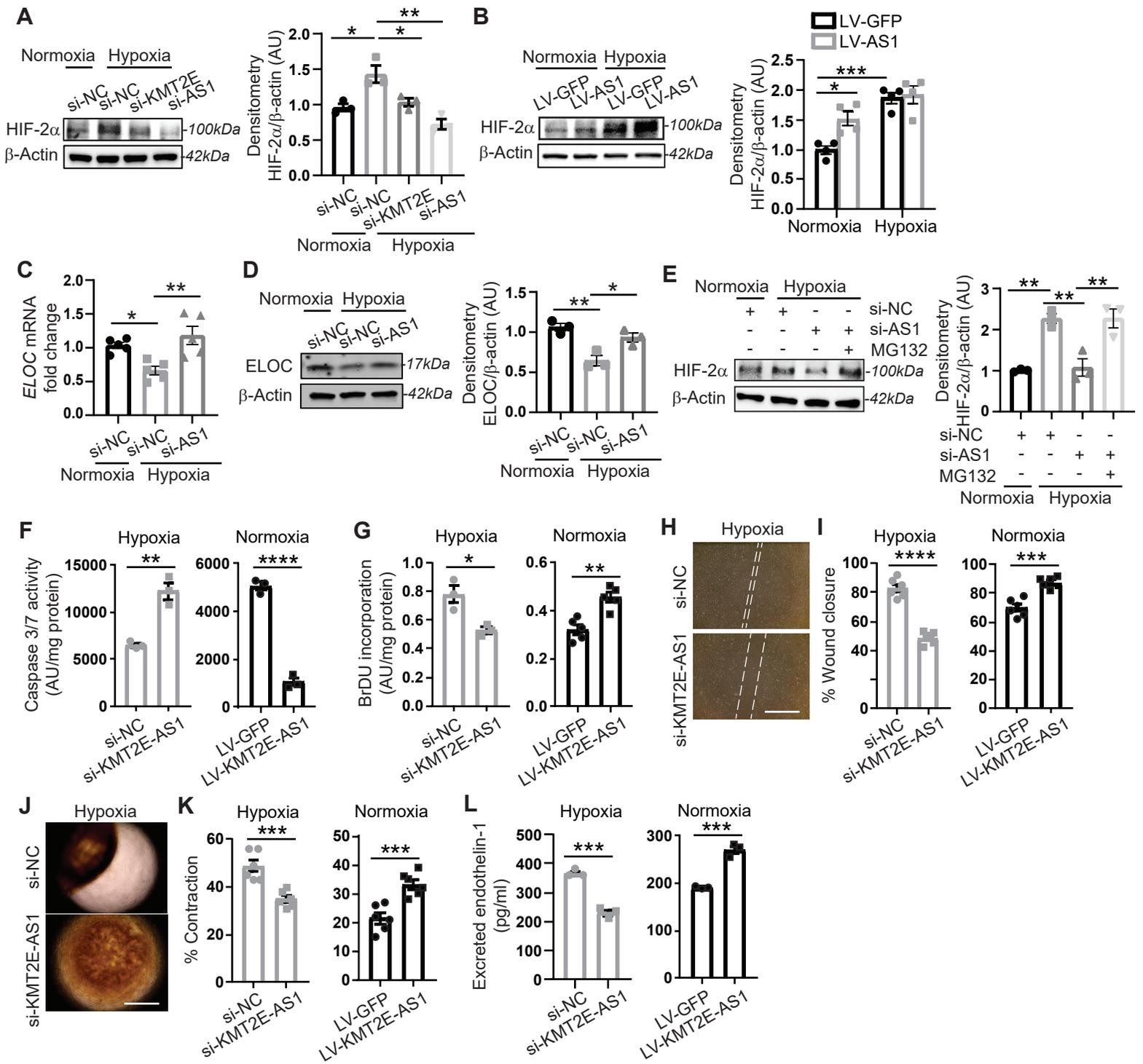
Vascular remodeling, indicated by α -SMA stain (white, D-E), is quantified by vascular thickness (I) and muscularization (J) for chaetocin-treated SU5415-hypoxic PAH rats vs. controls ($n = 4-5$; * $p < 0.05$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM). (K-L) RVSP (K) and RV remodeling (RV/body weight ratio, L) are measured in chaetocin-treated SU5415-hypoxic PAH rats vs. controls ($n = 4-5$; ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, one-way ANOVA followed by Bonferroni's *post-hoc* analysis; data represent mean \pm SEM).

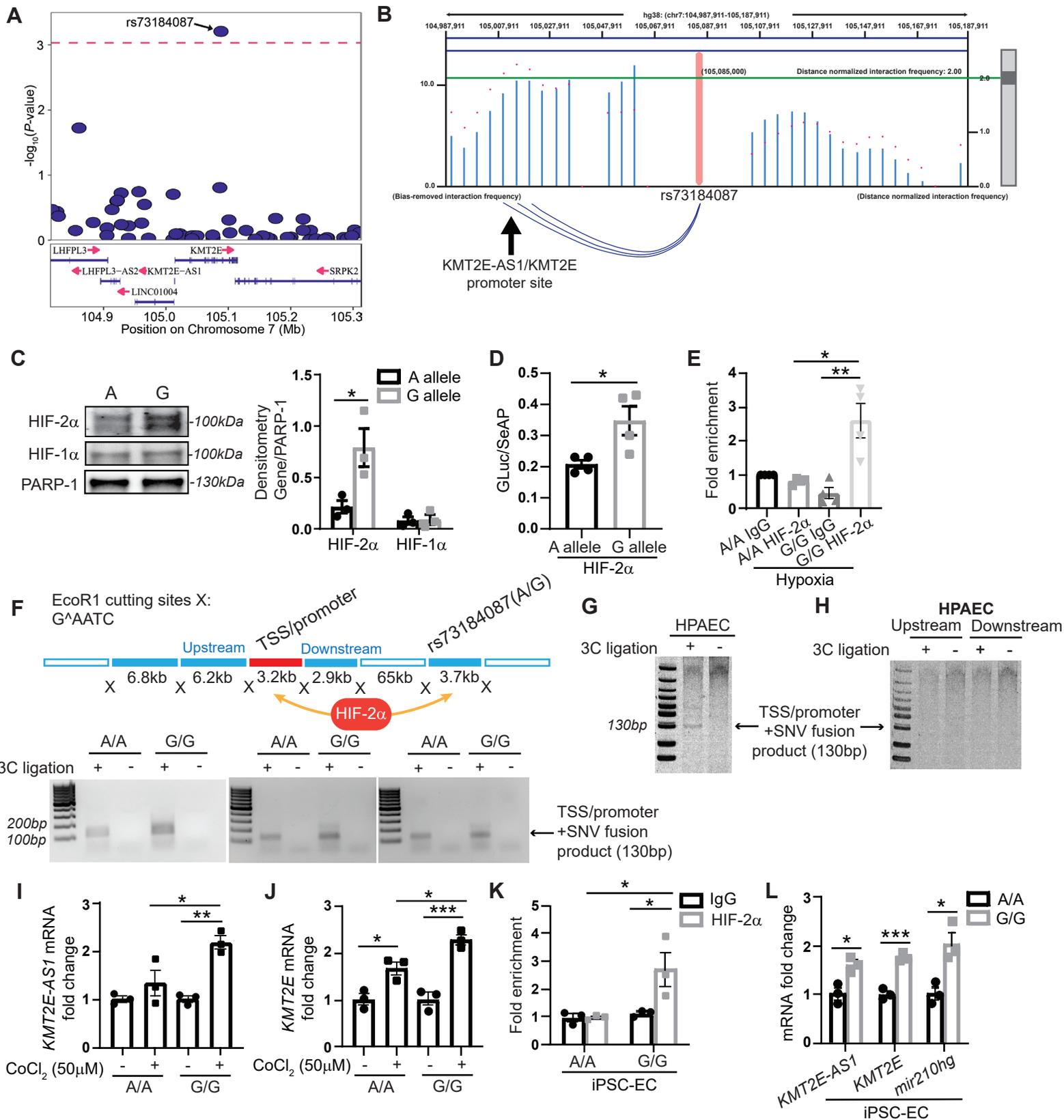


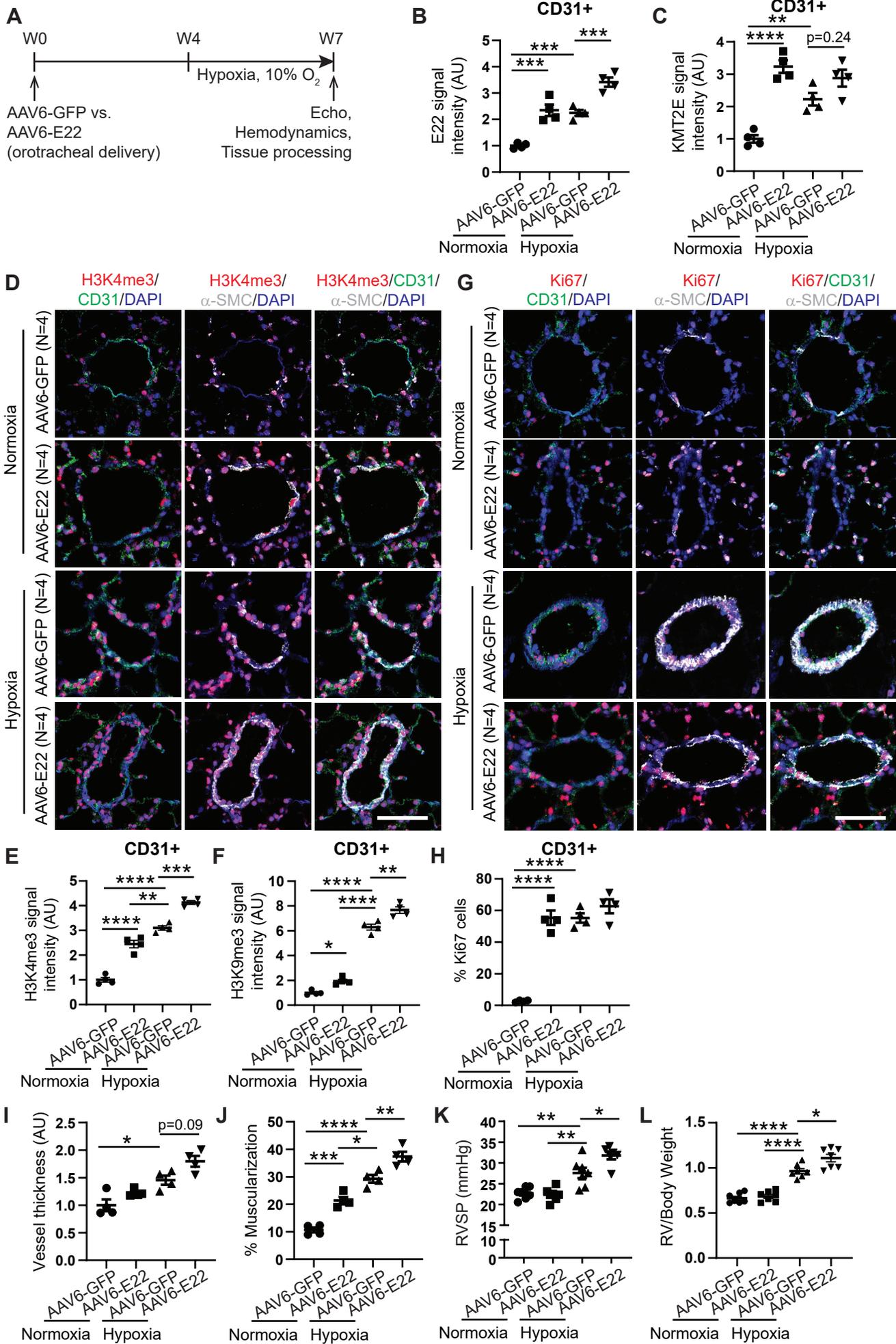


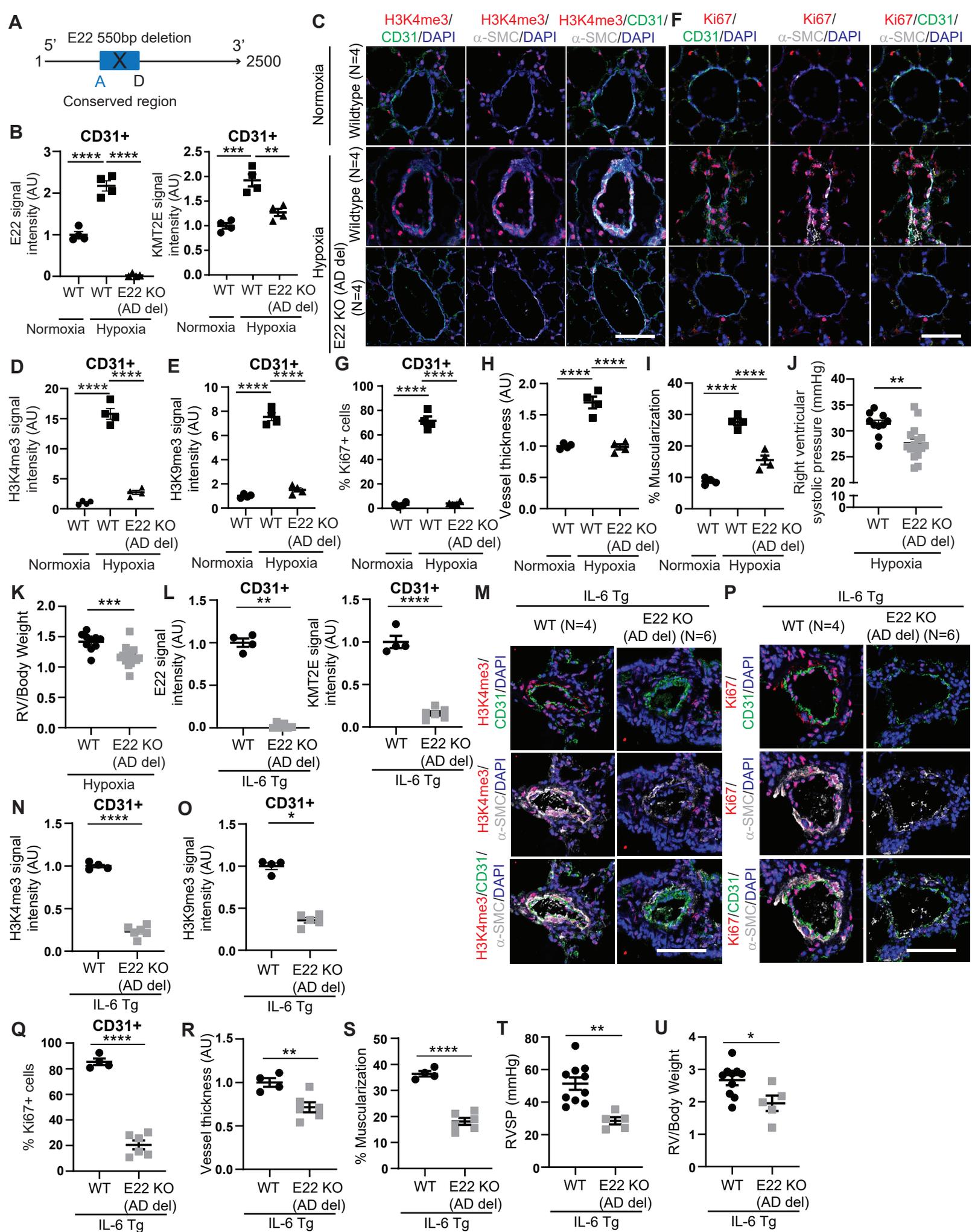
A Reversed genes in major biological processes

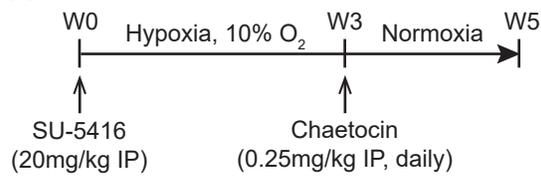
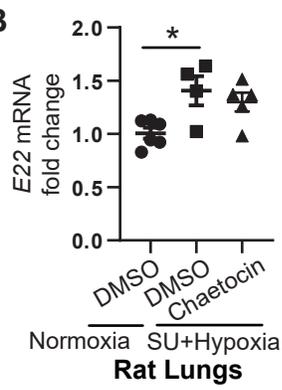
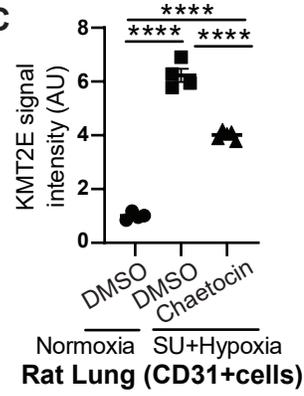
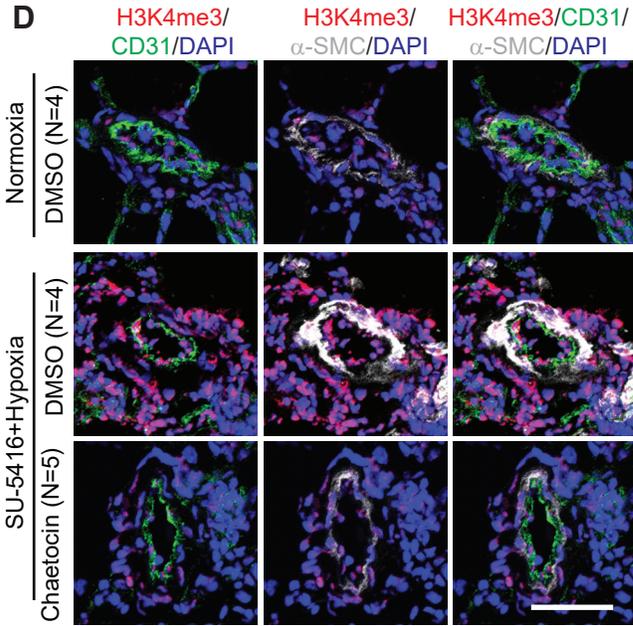
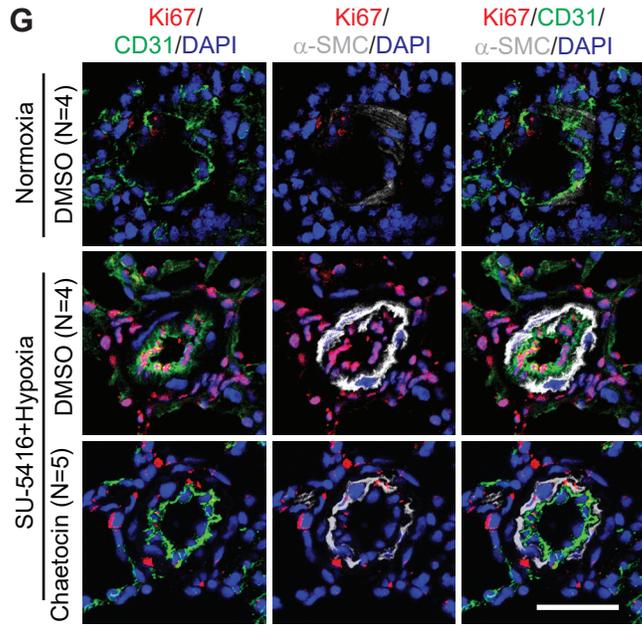
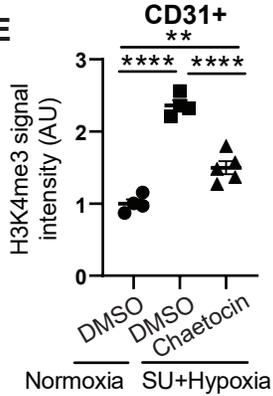
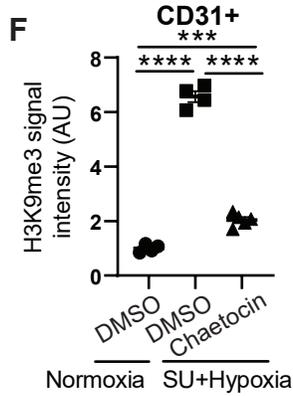
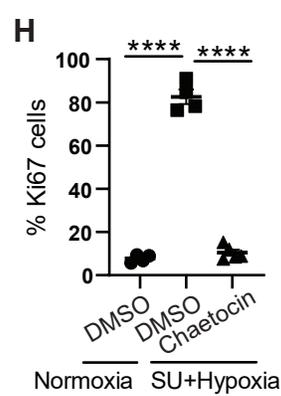
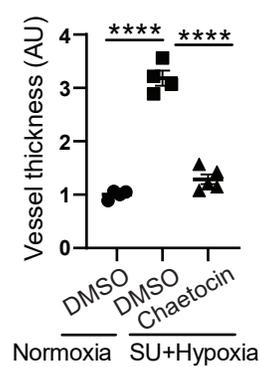
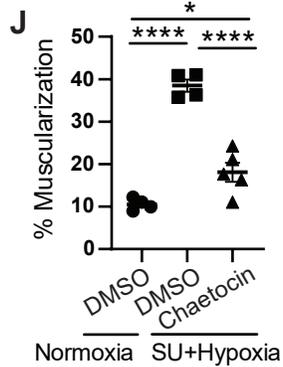
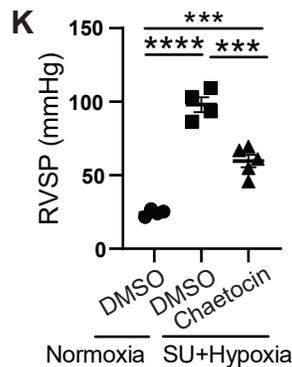
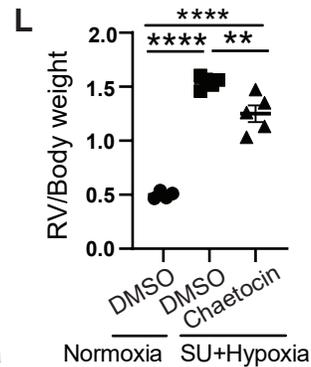










A**B****C****D****G****E****F****H****I****J****K****L**

SUPPLEMENTARY MATERIAL

Allele-specific and hypoxic control of the lncRNA *KMT2E-AS1*/KMT2E axis reprograms the endothelium in pulmonary hypertension

Yi-Yin Tai, Qiujun Yu, Ying Tang, Wei Sun, Neil J. Kelly, Satoshi Okawa, Jingsi Zhao, Tae-Hwi Schwantes-An, Caroline Lacoux, Stephanie Torrino, Yassmin Al Aaraj, Wadih El Khoury, Vinny Negi, Mingjun Liu, Catherine G. Corey, Frances Belmonte, Sara O. Vargas, Brian Schwartz, Bal Bhat, B. Nelson Chau, Jason H. Karnes, Taijyu Satoh, Robert J. Barndt, Haodi Wu, Victoria N. Parikh, Jianrong Wang, Yingze Zhang, Dennis McNamara, Gang Li, Gil Speyer, Bing Wang, Sruti Shiva, Brett Kaufman, Seungchan Kim, Delphine Gomez, Bernard Mari, Michael H. Cho, Adel Boueiz, Michael W. Pauciulo, Laura Southgate, Richard C. Trembath, Olivier Sitbon, Marc Humbert, Stefan Graf, Nicholas W. Morrell, Christopher J. Rhodes, Martin R. Wilkins, Mehdi Nouraie, William C. Nichols, Ankit A. Desai, Thomas Bertero, and Stephen Y. Chan

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-1 β (10 ng/mL) and IL-6 (100 ng/mL) used were from PeproTech. Actinomycin D (Sigma-Aldrich) was used at 1 μ g/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% O₂ at 37 °C or to hypoxic conditions with 0.2% O₂ 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-1 α (sc-35561), HIF-2 α (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™ 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 CCTTTCTCCCGCAAAGTATACTTAT, 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-1 α plasmid and HIF-2 α plasmid (20) were kindly gifted from W. G. Kaelin (Harvard Medical School) and W. Kim (University of North Carolina, Chapel Hill). HIF-1 α and HIF-2 α 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 packaging 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 lentiviral 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'-
gtaGTCCCCTCCCCCTGCCCTGCGGAGCGGGGCGGGCGGGCCCCGTGCGCTTGCCGCCG
CCGCCGCCGCCGCTGCCGCCGCCGCCGCCt and 5'-

ctagaGGCGGCGGCGGCGGCAGCGGCGGCGGCGGCGGCGGCAAGCGCACGGGGCCCCGCC
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 (wmE22lncRNA) or the 2006 bp
truncated version (dmE22lncRNA). 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 1 μ g 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 TTAAAC 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 A CA 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 Secret-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% O₂) vs. normoxia (20% O₂) for 24 hours. 46 hours post-transfection, cells were treated with proteasome 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-1 α (NB100-134, 1:1000) and HIF-2 α (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-2 α binding at SNV region.

Prediction and scoring of HIF-2 α binding sites

HIF-2 α chromatin immunoprecipitation sequencing peaks were published previously (19). These are inclusive of sites where HIF-2 α directly binds canonical HIF-response elements (HREs) as well as sites where HIF-2 α 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-2 α 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 ($0.05/[883 \text{ SNPs} \times 2 \text{ alleles} \times 2 \text{ strands} \times (8 + 8 + 15 \text{ nucleotide Motif lengths})] = 0.05/109492 = 4.57 \times 10^{-7}$) were calculated using the *TFMPvalue* (55) module in R v4.0.1. SNV alleles were considered to have predicted differential HIF-2 α 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: TTDTTTTWYTTTCTTT, 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-1 α or HIF-2 α 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 μ g DNA was then attached to 40 μ l of Dynabeads™M-280 Streptavidin (Thermo Fisher Scientific). DNA-beads were mixed with ~100 μ g 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-1 α or HIF-2 α (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 \times 10⁷ 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 μ g 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 EcoRI 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 2×10^6 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, I-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% O₂). 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 μ M CHIR99021 (Selleckchem, #S2924) for 2 days, followed by 3 μ M 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 μ M of the TGF β 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 μ l Matrigel per well) and polymerized for 30 min at 37°C. iPSC-Ecs were resuspended in 100 μ l 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 1 μ M Oligomycin, 0.5 μ M FCCP, and 2 μ M Rotenone plus 0.5 μ M 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, 1 μ M 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 μ l 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 μ l 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 $100 \times (\text{well diameter} - \text{gel diameter})/\text{well diameter}$.

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 μ l 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 "GTGATTTCTGTCTGGGCGCGTAGG", 3' guide 2 "CTGAAAGGACACTCCTTTTCGGGG", 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 "GTGTTCTGGGTGTTATTCACTTGC" 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% O₂, OxyCycler chamber, Biospherix Ltd.) as compared to normoxia (21% O₂). 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% O₂ 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: $100 \text{ (LV Vol;d - LV Vol;s/LV Vol;d)}$. Fractional shortening was calculated by M-mode imaging based on the following formula: $100 \text{ (LVID,d - LVID,s/ LVID,d)}$. 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 100 μ m and 400 μ m 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 μ m 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 μ m 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 μ m 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-2 α 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.

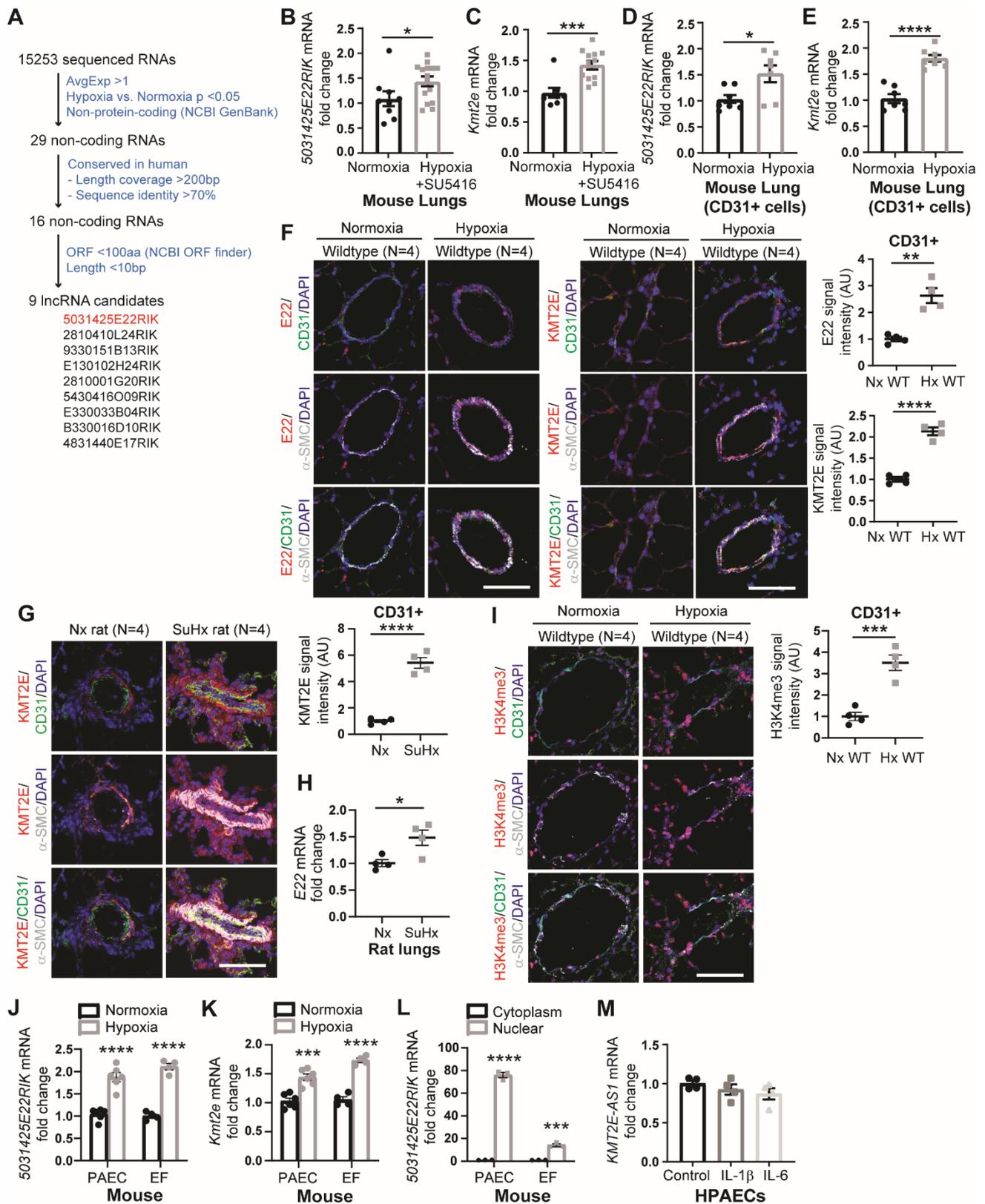


Figure S1. Response of lncRNA *E22/KMT2E-AS1* and *KMT2E* in hypoxic PH mouse, SU5416-hypoxic PAH rats, cultured mouse cell lines, and human PAECs. (A) From RNA-sequencing of SU5416-hypoxic PAH vs. control mouse lung tissues (GEO: GSE61828), differentially expressed long

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 \pm 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 \pm 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 \pm SEM). Scale bars, 50 μ m. **(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 \pm SEM). Scale bars, 50 μ m. **(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 \pm SEM). Scale bars, 50 μ m. **(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 \pm 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 \pm SEM). **(M)** *KMT2E-AS1* transcript expression is shown in human PAECs exposed to interleukin-1 β (IL-1 β) 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 \pm SEM).

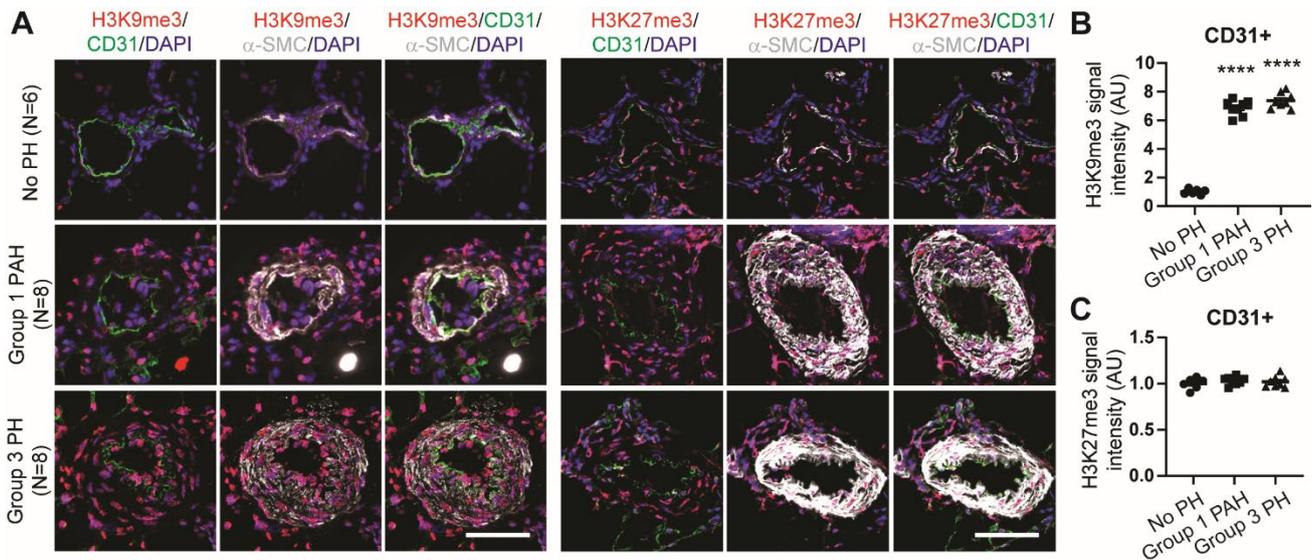


Figure S2. Vascular quantification of other known hypoxia-dependent histone methylation marks in human Group 1 and Group 3 PH. (A-C) Representative IF staining and quantifications of H3K9me3 (red; A,B) and H3K27me3 (red; A,C) are shown in CD31+ vascular endothelium of human Group 1 and Group 3 PH vs.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

(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).

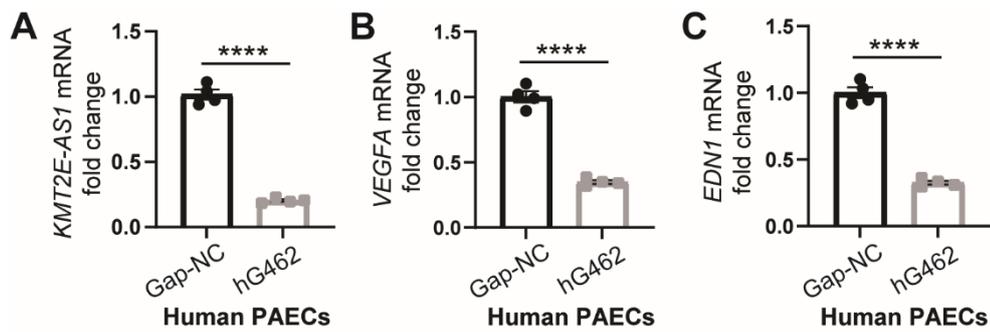


Figure S4. Antisense oligonucleotide gapmer knockdown of *KMT2E-AS1* and downstream effects in human PAECs. (A-C) In human PAECs treated with an antisense oligonucleotide gapmer specific for *KMT2E-AS1* (hG462) vs. a gapmer control (Gap-NC), *KMT2E-AS1* (A), *VEGFA* (B), and *EDN1* (C) transcripts are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

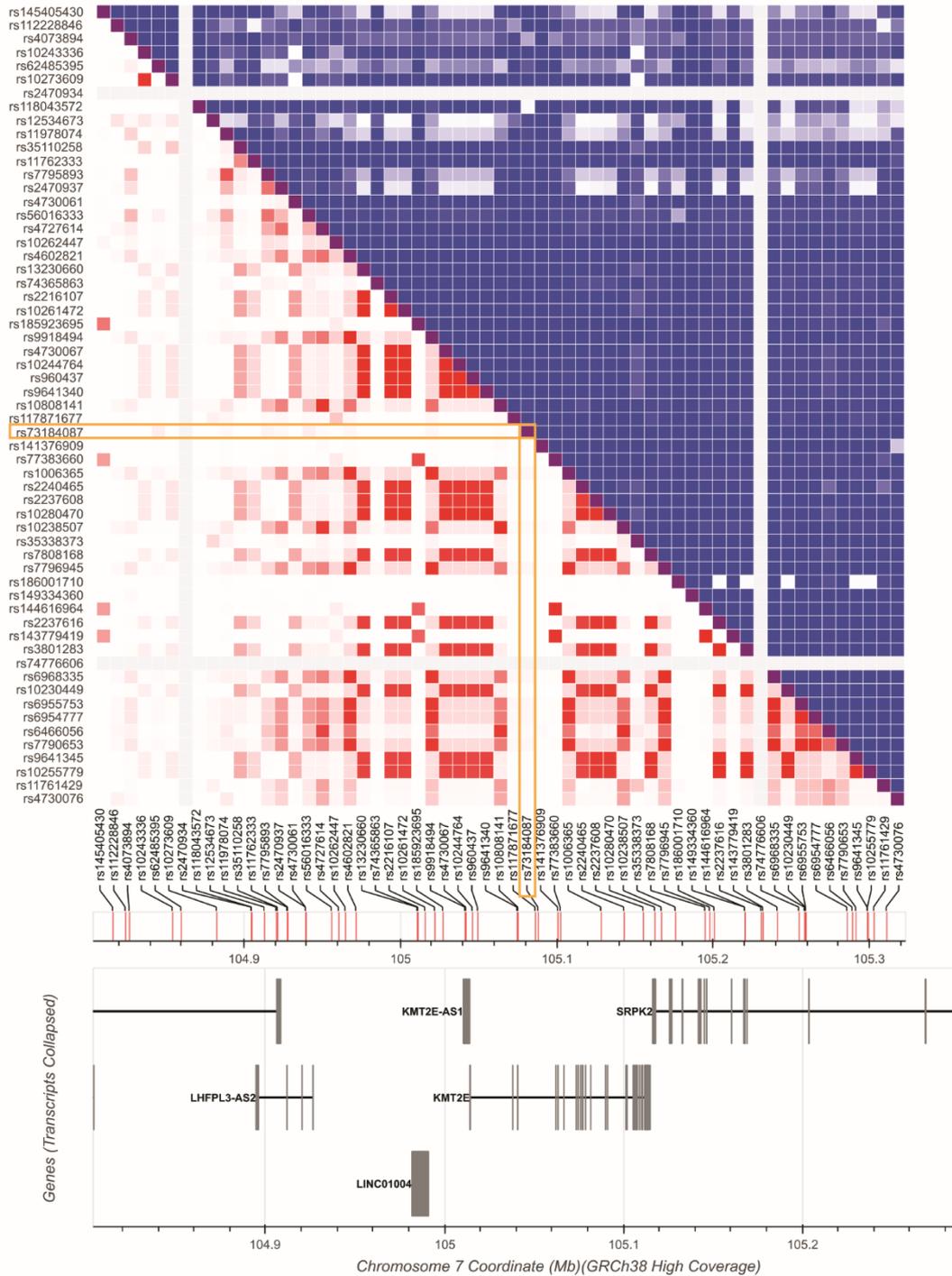


Figure S5. Analysis of linkage disequilibrium of SNV rs73184087. Via LDlink, a linkage disequilibrium (LD) heatmap is displayed. Red indicates strong LD.

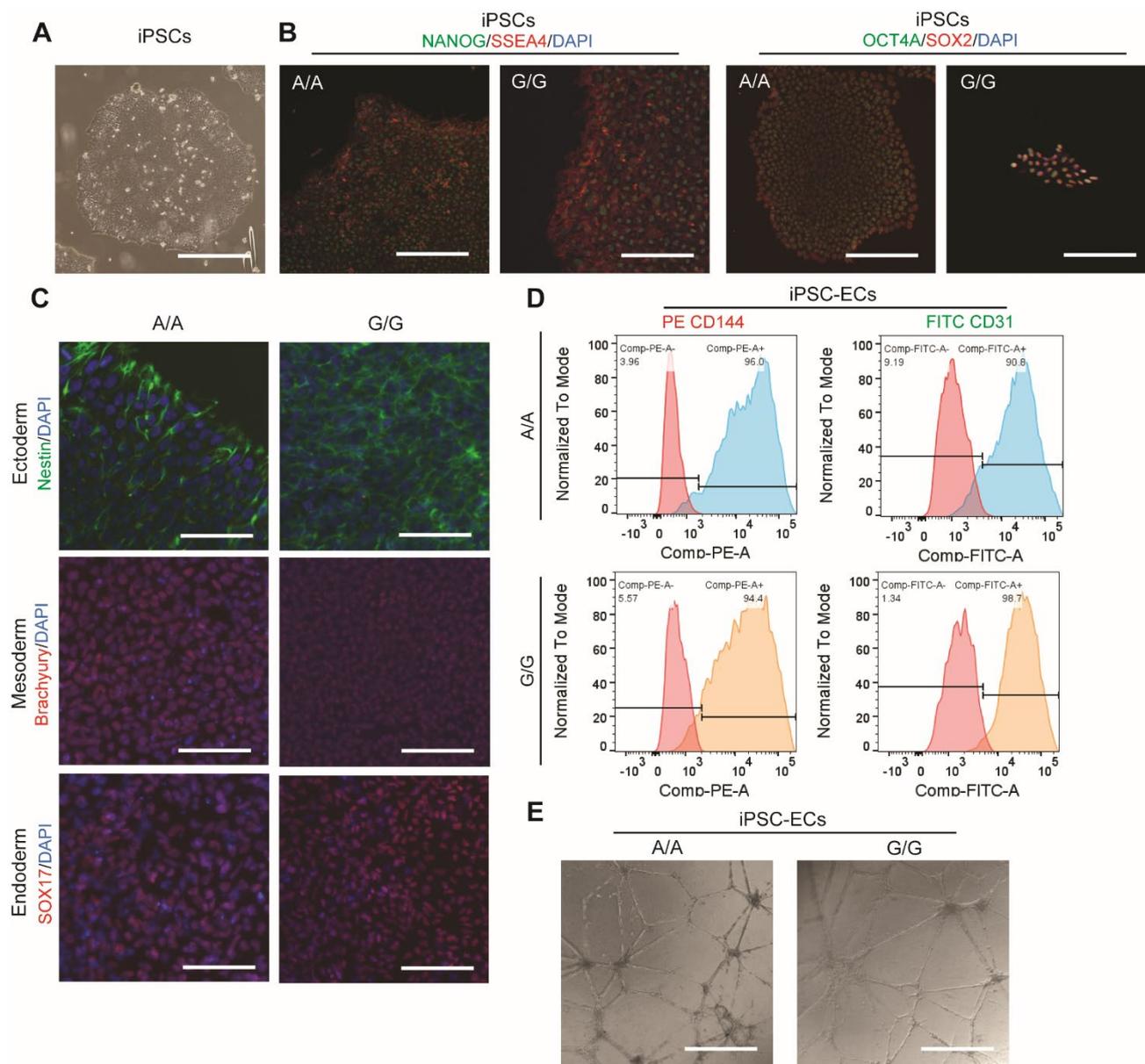


Figure S6. Characterization of induced-pluripotent stem cells and differentiated endothelial cells from transformed lymphocytes carrying A/A or G/G alleles of SNV rs73184087. (A) Morphology is shown of induced-pluripotent stem cells (iPSCs) reprogrammed from transformed lymphocytes. Scale bar, 100 μ m. **(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, 100 μ m. **(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, 100 μ m. **(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 μ m.

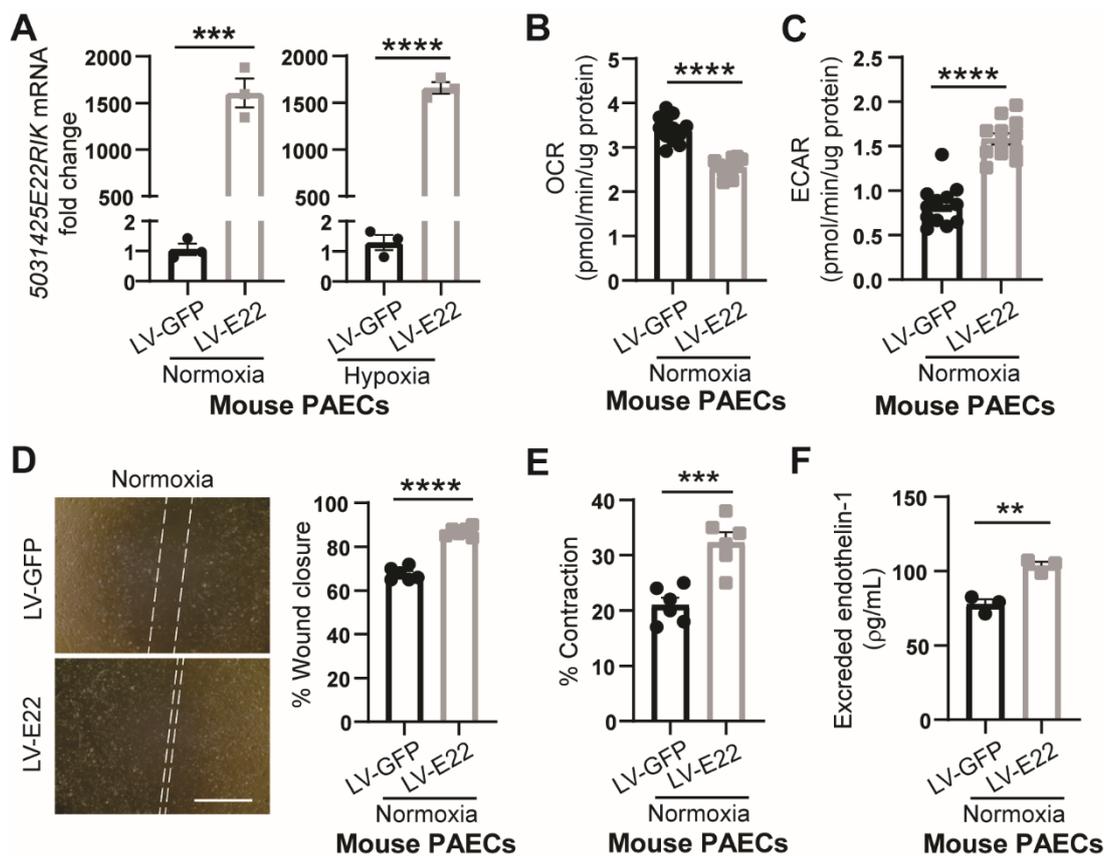


Figure S7. *E22* regulates downstream endothelial pathophenotypes in mouse PAECs. (A) *E22* expression is quantified in mouse PAECs after forced expression of *E22* under normoxia (left graph) and hypoxia (right graph) ($n = 3$; *** $p < 0.001$, **** $p < 0.0001$, unpaired Student's t-test; data represent mean \pm 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 \pm 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 \pm SEM). Scale bars, 200 μ m. (E) PASM 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 \pm 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 \pm SEM).

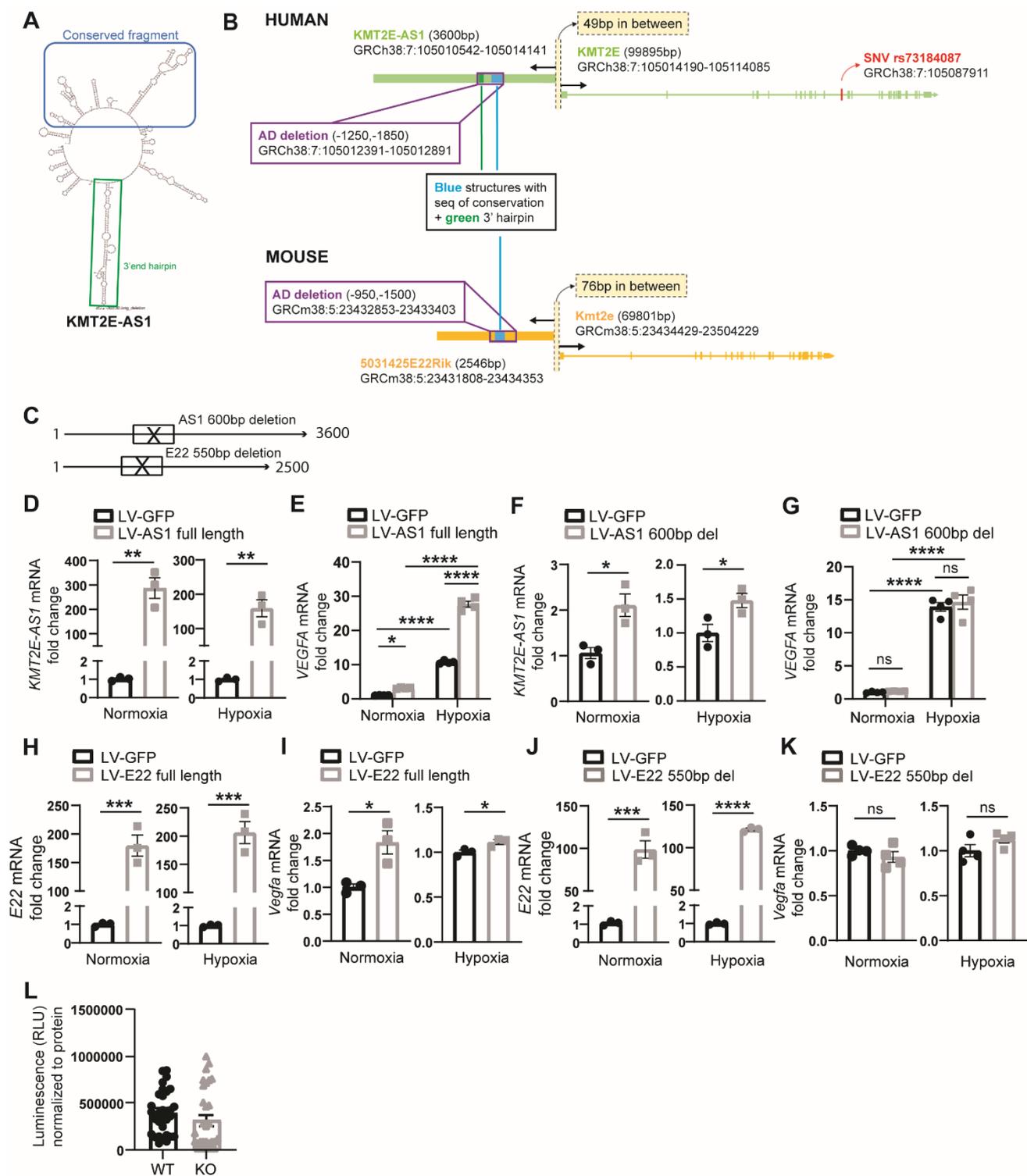


Figure S8. Deletion analysis of *KMT2E-AS1* and *E22* demonstrates that a highly conserved region is required for activity. (A) Secondary structure of human *KMT2E-AS1* is shown with highlighted area (blue) displaying a conserved sequence in mouse *E22*. **(B)** Overall schematic shows the chromosomal 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 \pm 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 \pm 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 \pm SEM).

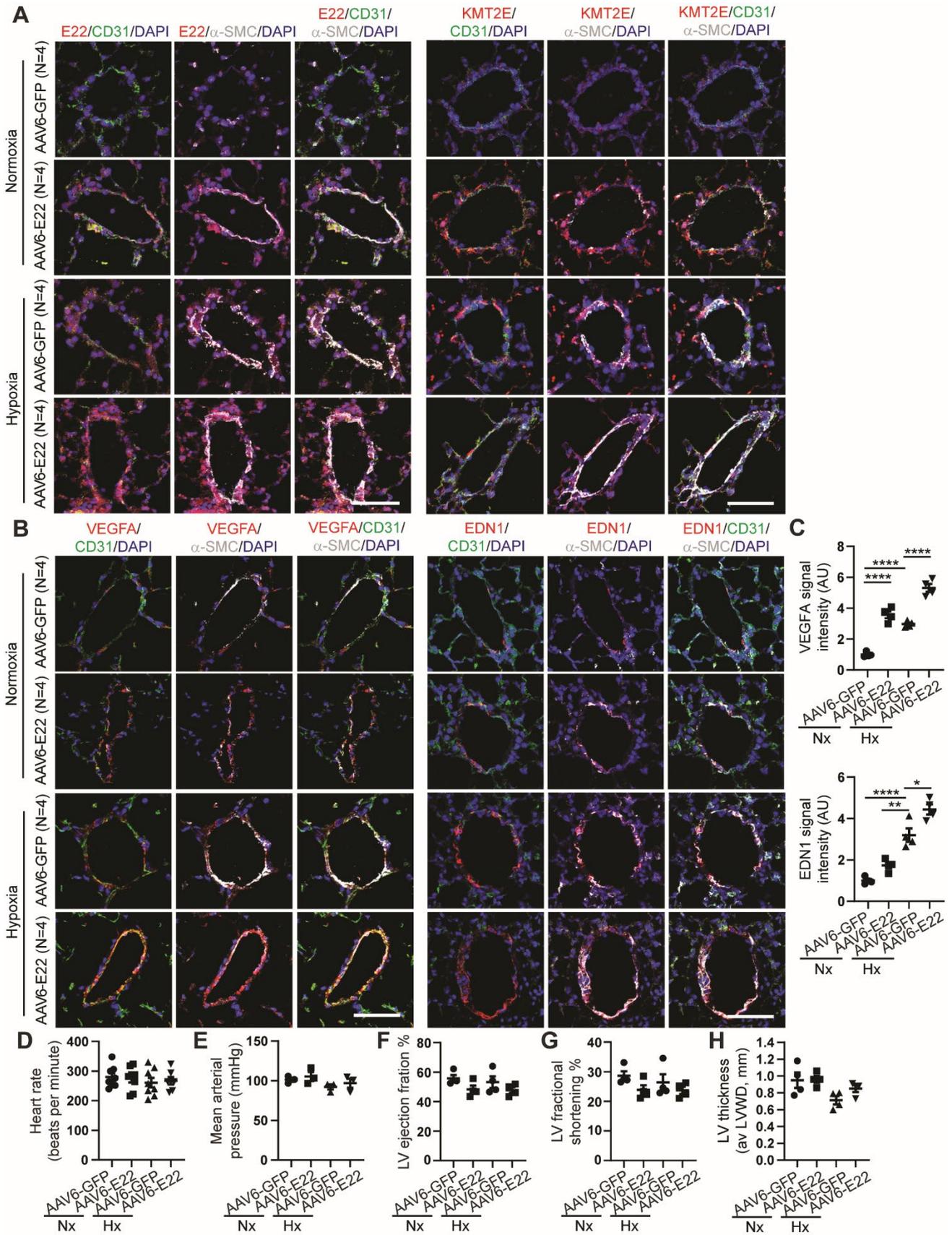


Figure S9. Forced expression of pulmonary vascular *E22* promotes PH pathophenotypes in mice. (A) Representative FISH and IF images are shown for *E22* (red) and KMT2E (red) in pulmonary

CD31+ endothelium of mice transduced with AAV6-E22 vs. AAV6-GFP in normoxia and hypoxia. Scale bars, 50 μ m. **(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 \pm SEM). Scale bars, 50 μ m. **(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 \pm 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 \pm 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 \pm SEM).

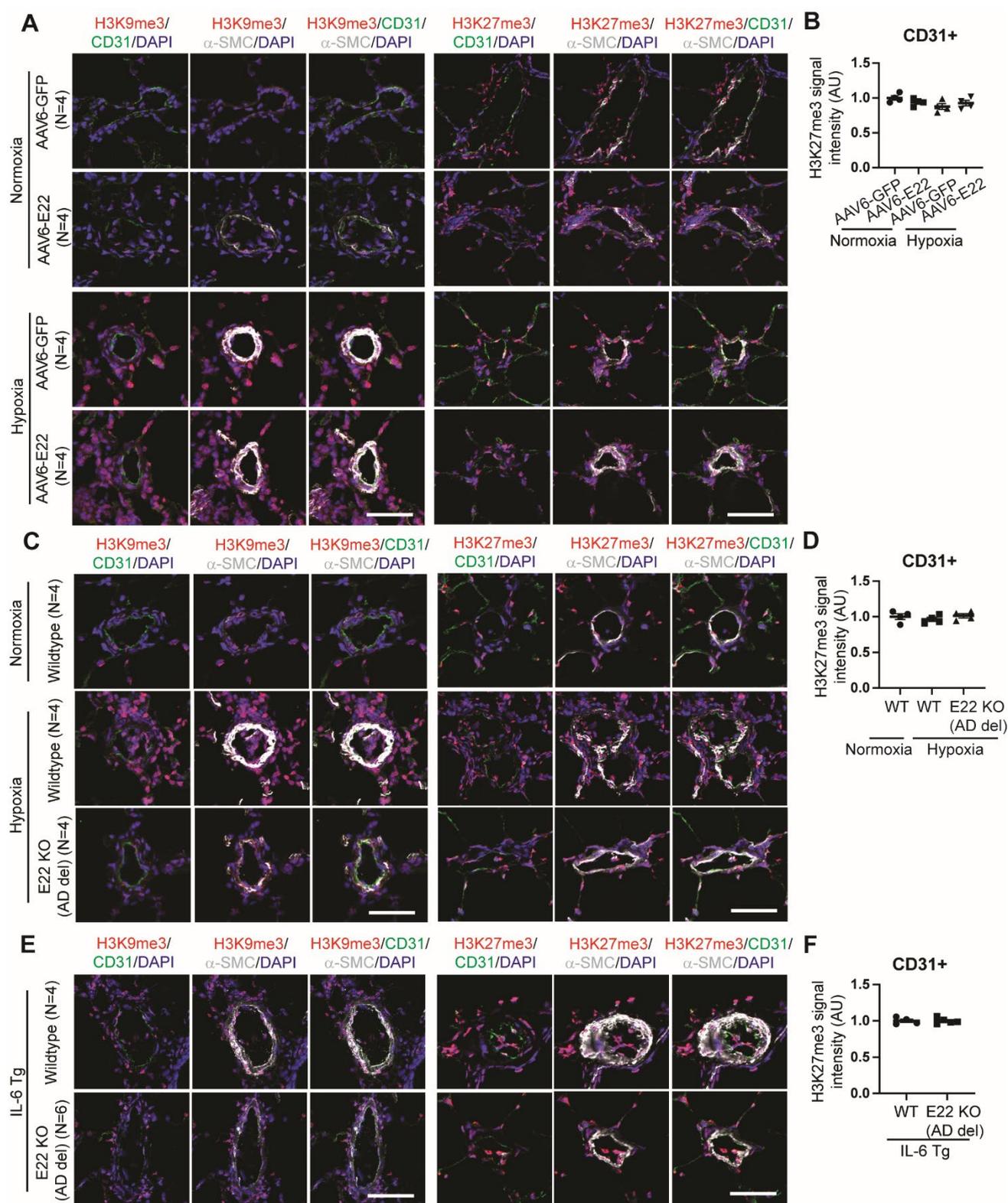


Figure S10. *KMT2E-AS1* partially controls histone 3 lysine 9 trimethylation. (A-B) Representative IF images are shown of H3K9me3 (red) and H3K27me3 (red) in AAV6-E22 mice vs. AAV6-GFP controls (A). IF quantification of H3K27me3 expression is displayed for mouse CD31+ vascular endothelium ($n = 4$; two-way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

represent mean \pm SEM). Scale bars, 50 μ m. **(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 \pm SEM). Scale bars, 50 μ m. **(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 \pm SEM). Scale bars, 50 μ m.

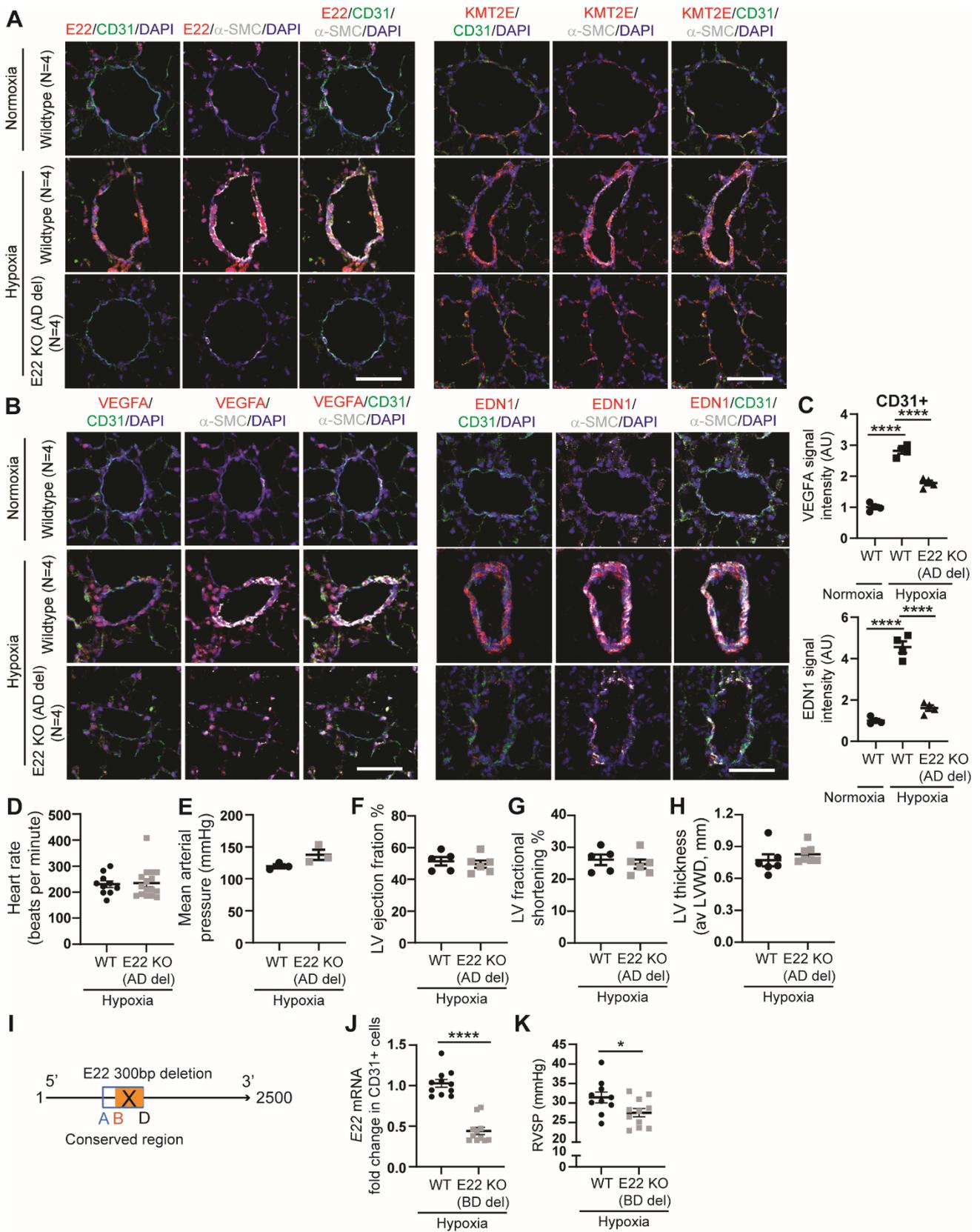


Figure S11. E22 knockout mice display less hypoxic pulmonary vascular remodeling without affecting heart rate, systemic blood pressure, or left ventricular function. (A) Representative

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 μ m. **(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 \pm SEM). Scale bars, 50 μ m. **(D)** Heart rates of *E22* KO mice are compared to wildtype controls ($n = 10-16$; Mann-Whitney test; data represent mean \pm 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 \pm 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 \pm 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 \pm 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 \pm SEM).

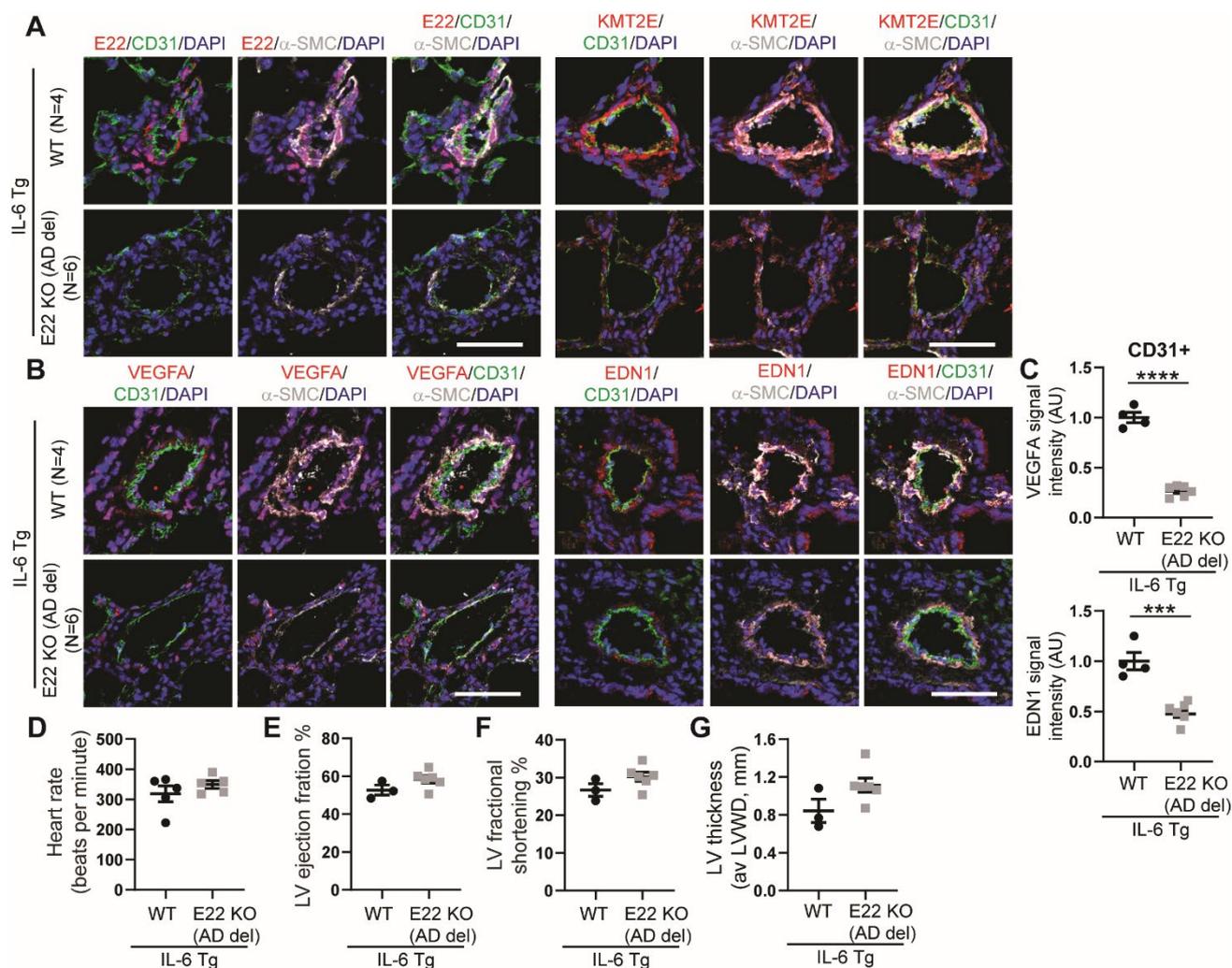


Figure S12. E22 deletion alleviates PAH in hypoxic IL-6 Tg mice. (A) Representative FISH and IF staining of full-length *E22* (red) and *KMT2E* (red) are shown in hypoxic IL-6 Tg; *E22* KO mouse lungs vs. hypoxic IL-6 Tg wildtype (WT) controls. Scale bars, 50 μ m. (B-C) Representative IF images (B) and 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 \pm SEM). Scale bars, 50 μ m. (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 \pm 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 \pm SEM).

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 \pm 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 \pm 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 \pm 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 \pm SEM).

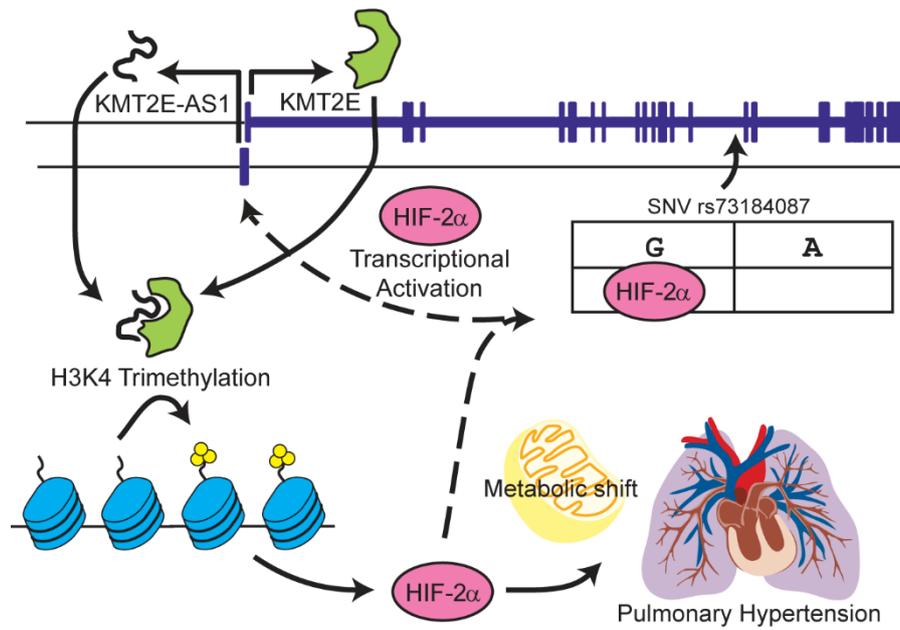


Figure S14. Model of the *KMT2E-AS1/KMT2E* gene tandem that forms a lncRNA-protein complex controlling HIF-2 α -dependent endothelial reprogramming in pulmonary hypertension. Proposed mechanism employs a positive feedback loop whereby allele-specific binding of HIF-2 α to rs73184087 drives *KMT2E-AS1/KMT2E* transcription and in turn promotes HIF-2 α 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
CHRN1	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
ERRF1	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 x 10 ⁻⁴

Table S6. SNVs within +/-200kb of *KMT2E-AS1/KMT2E* with predicted HIF-2 α binding to one of either the minor or major SNV alleles. Change in HIF-2 α binding score denotes the maximum difference in log-odds scores between the major and minor alleles among the predicted HIF-2 α binding motifs; positive values indicate greater predicted HIF-2 α 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-2 α 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/m ²)	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-1 α	Novus Biologicals	NB100-134
HIF-2 α	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
<i>KMT2E-AS1</i> Custom LNA Detection Probe	Qiagen	N/A
<i>E22</i> 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- <i>mE22</i> -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-1 β	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 TM 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 TM Transfection System 10 μ L 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 <i>5031425E22Rik</i> knockout	This paper	N/A
Mouse: IL6-Tg x CRISPR-Cas9 <i>5031425E22Rik</i> 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: GTCATTGCATACGGGGCTG	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: ACTCCTTTCCGGGGTGTGC	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: CCTTTCTCCCCGAAAGTATACTTAT	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 - <i>KMT2E-AS1</i> Human (Hs04233055_s1)	Life Technologies	4331182
TaqMan Gene Expression Assay - <i>KMT2E</i> Human (Hs01096121_m1)	Life Technologies	4331182
TaqMan Gene Expression Assay – <i>ELOC</i> Human (Hs00855349_g1)	Life Technologies	4331182
TaqMan Gene Expression Assay – <i>EPAS1</i> 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 <i>KMT2E-AS1</i> siRNA (n263822)	Life Technologies	4390771
Human KMT2E siRNA (132615)	Life Technologies	AM16708
Human HIF-1 α 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-2 α	University of North Carolina, Chapel Hill	N/A
pLenti-HIF-1 α	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/application/s/rna-folding-form-v2.php

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