Molecular Genetic Characterization of SMAD Signaling Molecules in Pulmonary Arterial Hypertension

M. Talat Nasim,[1]† Takeshi Ogo,[1]† Mohammad Ahmed,[1] Rebecca Randall,[1] Hasnin M. Chowdhury,[1] Katie M. Snape,[1] Teisha Y. Bradshaw,[1] Laura Southgate,[1] Grace J. Lee,[1] Ian Jackson,[2] Graham M. Lord,[2] J. Simon R. Gibbs,[3] Martin R. Wilkins,[4] Keiko Ohta-Ogo,[5] Kazufumi Nakamura,[5] Barbara Girerd,[6] Florence Coulet,[7] Florent Soubrier,[7] Marc Humbert,[6] Nicholas W. Morrell,[8] Richard C. Trembath[1]* and Rajiv D. Machado[1]

†These authors made an equal contribution to this work

[1]Department of Medical and Molecular Genetics, and [2]MRC Centre for Transplantation, King's College London, School of Medicine, Guy's Hospital, London, UK; [3]National Heart & Lung Institute, Imperial College London, UK; [4]Division of Experimental Medicine, Imperial College of Medicine, London, UK; [5]Department of Cardiovascular Medicine, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Japan; [6]Univ Paris-Sud, INSERM U999, Service de Pneumologie, Hôpital Antoine Béclère, Assistance Publique-Hôpitaux de Paris, Clamart, Paris, France; [7]Laboratoire d'Oncogénétique et d'Angiogénétique Moléculaire, Departement de Génétique, Groupe Hospitalier Pitié-Salpêtrière, Paris, France; [8]Department of Medicine, University of Cambridge, Cambridge, UK.

*Correspondence to: Professor Richard C. Trembath, Division of Genetics and Molecular Medicine, King's College London, School of Medicine, 8th Floor Tower Wing, Guy's Hospital, London SE1 9RT, United Kingdom.

E-mail: richard.trembath@kcl.ac.uk; Tel: +44 2071887994; Fax: +44 2071882585

ABSTRACT

Heterozygous germ-line mutations of *BMPR2* are the major contributor to familial clustering of pulmonary arterial hypertension (PAH). To further explore the genetic basis of PAH in isolated cases we undertook a candidate gene analysis to identify potentially deleterious variation. Members of the BMP pathway, namely *SMADs 1, 4, 5* and *9,* were screened by direct sequencing for gene defects. Four variants were identified in *SMADs 1, 4* and *9* amongst a cohort of 324 PAH cases, each not detected in a in a substantial control population. Of three amino-acid substitutions identified, two demonstrated reduced signaling activity *in vitro*. A putative splice site mutation in *SMAD4* resulted in moderate transcript loss due to compromised splicing efficiency. These results emphasize the central role of *BMPR2* mutation in the pathogenesis of PAH and indicate that variation within the SMAD family represents an infrequent cause of disease.

Key Words: pulmonary hypertension • BMPR2 • Receptor SMADs • SMAD4 • transcriptional regulation.

PAH (MIM# 178600) is a progressive vascular disorder often fatal as a result of right heart failure [Humbert et al., 2010]. Mutations of *BMPR2* (MIM# 600799), encoding a bone morphogenetic type II receptor of the TGF-β family, are the major genetic determinant in familial PAH (FPAH). Idiopathic PAH (IPAH), defined as arising spontaneously in the absence of a recorded family history of disease, is indistinguishable from the familial form [Machado et al., 2009]. The detection of germ-line *BMPR2* mutation in ~25% of IPAH cases, posing hereditary risk to offspring, has led to the re-classification of mutation carriers as heritable PAH (HPAH) [Deng et al., 2000; Lane et al., 2000; Machado et al., 2009]. Over 70% of HPAH mutations predict premature truncation likely leading to transcript loss via the nonsense-mediated decay (NMD) pathway [Machado et al., 2009]. Deleterious *BMPR2* mutation in the IPAH population indicates a genetic basis to disease and suggests the existence of additional genetic and/or environmental factors in PAH pathogenesis yet to be fully annotated. Indeed, rare disease alleles of the TGF-β receptors, ALK1 (MIM# 601284) and endoglin (MIM# 131195), and members of the SMAD family have recently been shown to underlie susceptibility to HHT and pulmonary hypertension [Harrison et al., 2005; Shintani et al., 2009]. BMPR-II triggers signal transduction through the receptor SMADs (R-SMADs) 1, 5 and 9, upon ligand binding and complex formation with a type I receptor, namely BMPR-1A, -1B or ALK1. R-SMADs translocate to the nucleus in complex with co-SMAD4 to regulate transcription of target genes, a process dependent on the presence of a conserved proximal DNA binding MH1 domain and a trans-activating MH2 domain [Massague et al., 2000]. As impaired responses to BMP-specific SMADs are acknowledged as critical in the development of PAH [Nishihara et al., 2002; Rudarakanchana et al., 2002], we sought to assess the genetic contribution of this pathway to disease susceptibility by conducting an extensive survey of a large PAH cohort, excluded for *BMPR2* mutation. Here we report the identification of four gene defects in *SMAD1* (MIM# 601595), *SMAD4* (MIM#

600993) and *SMAD9* (MIM# 603295). Functional analyses suggest significant, albeit limited consequences on transcript integrity, signaling and target gene regulation.

Ethical approval for these studies was obtained by local ethical committees and all patients provided informed consent. A diagnosis of PAH was confirmed as described previously [Machado et al., 2009]. The European IPAH cohort comprised a total of 158 subjects ascertained by specialist UK and European centers and displayed a gender bias favoring females (1.9:1). The PAH panel with associated disease (APAH) (n=136) comprised cases with HIV infection (n=9), portal hypertension (n=11), congenital heart disease (n=15), thromboembolic disease (n=42) and connective tissue disease (n=59). Japanese IPAH subjects (n=30) were ascertained through a single specialist referral centre. All patients had been screened for *BMPR2* mutation by direct sequencing and/or DHPLC employing primer sets previously described [Machado et al., 2001]. Direct sequencing was performed on ABI377 fragment analyzer. DHPLC was performed using the Transgenomic WAVE Nucleic Acid Fragment Analysis system containing a DNASep column (Transgenomic, Crewe, UK) according to manufacturer's instruction. Due to the inherent limitations of these techniques in detecting large genomic rearrangements the existence of *BMPR2* mutation in a small proportion of samples cannot be categorically excluded. The control population, derived from the 1958 British birth cohort and in-house samples were of self-defined Western European ancestry (n=960) or of Japanese descent (n=340), ranging between the 3rd and 5th decade of life. An additional panel of French controls (n=284) were used to further confirm absence of variation identified in patients of French origin**.** All coding exons, intron-exon boundaries and splice isoforms of the candidate genes *SMAD1*, *SMAD4*, *SMAD5* (MIM# 603110) and *SMAD9* were DNA sequenced and analyzed on an ABI3730xl DNA Analyzer as per standard protocols [Southgate et al., 2011]. Wild-type constructs for *SMADs 1, 4* and *9* were obtained from Addgene. Variation was introduced by performing site-directed mutagenesis with the

QuickChange kit (Stratagene) on the wild-type template according to the manufacturer's instructions. Constructs to determine splicing efficiency, splicing and luciferase assays are described in Supplementary methods. RNA was isolated from mammalian cells using TRI-Reagent (Sigma), and cDNA was synthesized using random primers and MMLV Reverse Transcriptase (Promega) following the manufacturer's protocol. RT-PCR was carried out using Hi-Fidelity Extensor Master Mix (ABgene). Quantitative PCR for determining transcript levels of *SMAD4*, inhibitor of DNA binding 2 (*Id2*), beta actin and GAPDH were performed using TaqMan Gene Expression Assays (Applied Biosystems) on a 7900HT Fast Real-Time PCR system (Applied Biosystems) according to the manufacturer's protocol. Lung tissue from the patient harboring the SMAD9 variant was fixed in 10% buffered formalin and embedded in paraffin as per standard procedures. Five-mm-thick sections were serially cut and stained with hematoxylin and eosin and Elastica van Gieson for histological examination by light microscopy.

Several known polymorphisms but no putative mutations were identified in the APAH samples. Screening of *SMAD1* in the IPAH patient series revealed a novel missense variant (c.8T>C, p.V3A) within exon 1, upstream of the MH1 domain (Fig. 1A). Within *SMAD4* two defects were identified in independent IPAH cases, namely a predicted splice-site mutation (c.1448-6T>C) upstream of the terminal exon 11 and a missense variant (c.38A>G, p.N13S) (Fig. 1A-C). A novel missense variant, (c.127A>G, p.K43E), was observed in a patient of Japanese origin in *SMAD9* impacting upon the MH1 domain (Fig. 1A). The lack of parental material precluded a determination of whether identified variants were *de novo*. All variants were absent in 960 European control samples and an additional 284 French control samples. The *SMAD9* variant was further excluded from a panel of 340 Japanese controls. Of note, the identified variants were absent from the 1000 Genomes database and an in-house repository of 60 exome sequenced control samples. All three of the substituted amino acid residues detected in *SMADs 1*, *4* and *9*

were highly conserved through evolution, to the level of *Takifugu rubripes* (Fig. 1D). High degrees of conservation have traditionally been considered to be a measure of the potential importance of amino acid residues to protein function. The clinical features of the subjects harboring potential disease causing variation within *SMAD* genes are presented in Supp. Table S1. To investigate the impact of the missense variants on SMAD mediated signaling, we adapted a SMAD reporter assay to be selective to BMP stimulation [Nasim et al., 2008] (Supp. Fig. S1). Wild type constructs for SMADs 1 and 4 activated reporter activity as anticipated (Supp. Fig. S1). Over-expression of the SMAD1 variant generated reduced basal activity by comparison to wild type and impaired responses to ligand stimulation (Fig. 2A and B). The SMAD9 p.K43E construct also displayed reduced reporter activity when compared to wild type (Fig. 2C), a diminished response to BMPR-II over-expression (Fig. 2C) and stimulation by BMP4 and BMP9 ligands (Fig. 2D). No difference in the level of activation was seen between the wild type and missense SMAD4 constructs either in the absence or presence of ligand (Fig. 2A and data not shown) which may indicate subtle effects on function not detectable by this reporter system or the possibility that this is an extremely rare population variant with no impact on PAH susceptibility. Notably, this impaired activity was modest relative to classical, functionally null BMPR-II substitutions underlying familial and hereditary PAH [Nasim et al., 2008]. *In silico* analysis of the c.1448-6T>C variant predicted impaired splicing compared to wild-type [Nalla and Rogan, 2005]. By quantitative RT-PCR analyses, we observed a putative reduction in transcript level of the c.1448-6T>C variant in patient derived leukocytes compared to the wild type; however, due to the known variability of mRNA expression in leucocytes this finding was treated as indicative (Supp. Fig. 2D). To investigate this observation further we performed RT-PCR on these samples and excluded the existence of aberrant splice products thereby indicating pre-mRNA splicing was preserved (Supp. Fig. S2A and B) and the NMD surveillance machinery likely not activated

although in the absence of a patient cell line formal exclusion of this process could not be confirmed. We next questioned whether the reduction in *SMAD4* transcript level might be explained by an alteration in splicing efficiency. Using an assay of splicing competency (Supp. Fig. 2C) [Nasim et al., 2002], we observed a significant reduction in reporter activity for the variant construct relative to wild type, consistent with a decrease in splicing efficiency (Fig. 2E). Both *SMAD4* variants were distinct to mutations observed in HHT alone and JP-HHT, the majority of which cluster in the MH2 domain encoded by exons 8-11 [Gallione et al., 2006; Iyer et al., 2010]. Unusually, these variants represent the most proximal amino acid substitution and only defective splice allele associated with vascular disease but the present data set is insufficient to draw meaningful correlations with phenotype. To better elucidate the functional significance of SMAD mutations in PAH, we assessed the expression of the *Id2* gene in pulmonary artery smooth muscle cells (PASMCs) heterozygous for the *SMAD9* mutation p.K43E, derived from explant lung taken at transplantation which exhibited extensive pulmonary vascular remodeling (Supp. Fig S3). The *Id2* transcript, a downstream target of BMP and TGF-β signaling was significantly reduced in the *SMAD9* variant cell line comparative to WT PASMC although responses to ligand were largely preserved (Fig. 2F). A similar level of reduction of *Id2* transcript expression was observed in a PAH cell line with a previously identified, pathogenic *BMPR2* mutation (p.W9X) (Suppl. Fig S4). Western blotting and immunofluorescence revealed no differences in protein stability and subcellular localization between WT and missense constructs of *SMADs 1, 4* and *9* (Supp. Fig S5), suggesting that variant proteins were likely to be present and intact in the cell. In light of the moderate effect size of these variants it is plausible that they may require one or more genetic or environmental triggers to precipitate disease. Alternatively it remains feasible that they may more substantially perturb one or more of the newly emerging SMAD-independent pathways. The systematic analysis of the BMP-specific SMAD family in

large patient cohorts is of central importance in determining the genetic contribution of this pathway, vital in maintenance of the pulmonary vasculature, to the pathogenesis of PAH. The low frequency of potentially pathogenic variation in genes encoding SMAD mediators of BMP would suggest that SMAD mutation makes a small contribution to inherited genetic susceptibility to I/APAH. The application of recent technological advances in high-throughput sequencing of coding and whole genomes may prove effective in gaining a fuller understanding of the genetic factors contributing to PAH, providing an opportunity for rapid advances in the understanding of the molecular mechanisms underlying this disorder.

ACKNOWLEDGMENTS

The authors wish to thank the patients and their families for participating in this study and to Cambridge NIHR Biomedical Research Centre for assistance in patient recruitment and sample collection. This work was supported by grants from the British Heart Foundation (RG/08/006/25302 to R.C.T., FS/07/036 to R.D.M.). M.T.N was supported by a fellowship award from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust

REFERENCES

Deng Z, Morse JH, Slager SL, Cuervo N, Moore KJ, Venetos G, Kalachikov S, Cayanis E, Fischer SG, Barst RJ, Hodge SE, Knowles JA. 2000. Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. Am J Hum Genet 67:737-744.

Gallione CJ, Richards JA, Letteboer TG, Rushlow D, Prigoda NL, Leedom TP, Ganguly A, Castells A, Ploos van Amstel JK, Westermann CJ, Pyeritz RE, Marchuk DA. 2006. SMAD4 mutations found in unselected HHT patients. J Med Genet 43:793-797.

Harrison RE, Berger R, Haworth SG, Tulloh R, Mache CJ, Morrell NW, Aldred MA, Trembath RC. 2005. Transforming growth factor-beta receptor mutations and pulmonary arterial hypertension in childhood. Circulation 111:435-441.

Humbert M, Sitbon O, Chaouat A, Bertocchi M, Habib G, Gressin V, Yaici A, Weitzenblum E, Cordier JF, Chabot F, Dromer C, Pison C, Reynaud-Gaubert M, Haloun A, Laurent M, Hachulla E, Cottin V, Degano B, Jais X, Montani D, Souza R, Simonneau G. 2010. Survival in patients with idiopathic, familial, and anorexigen-associated pulmonary arterial hypertension in the modern management era. Circulation 122:156-163.

Iyer NK, Burke CA, Leach BH, Parambil JG. 2010. SMAD4 mutation and the combined syndrome of juvenile polyposis syndrome and hereditary haemorrhagic telangiectasia. Thorax 65:745-746.

Lane KB, Machado RD, Pauciulo MW, Thomson JR, Phillips JA, 3rd, Loyd JE, Nichols WC, Trembath RC. 2000. Heterozygous germline mutations in BMPR2, encoding a TGF-beta receptor, cause familial primary pulmonary hypertension. Nat Genet 26:81-84.

Machado RD, Eickelberg O, Elliott CG, Geraci MW, Hanaoka M, Loyd JE, Newman JH, Phillips JA, 3rd, Soubrier F, Trembath RC, Chung WK. 2009. Genetics and genomics of pulmonary arterial hypertension. J Am Coll Cardiol 54:S32-42.

Machado RD, Pauciulo MW, Thomson JR, Lane KB, Morgan NV, Wheeler L, Phillips JA, 3rd, Newman J, Williams D, Galie N, Manes A, McNeil K, Yacoub M, Mikhail G, Rogers P, Corris P, Humbert M, Donnai D, Martensson G, Tranebjaerg L, Loyd JE, Trembath RC,

Nichols WC. 2001. BMPR2 haploinsufficiency as the inherited molecular mechanism for primary pulmonary hypertension. Am J Hum Genet 68:92-102.

Massague J, Blain SW, Lo RS. 2000. TGFbeta signaling in growth control, cancer, and heritable disorders. Cell 103:295-309.

Nalla VK, Rogan PK. 2005. Automated splicing mutation analysis by information theory. Hum Mutat 25:334-342.

Nasim MT, Chowdhury HM, Eperon IC. 2002. A double reporter assay for detecting changes in the ratio of spliced and unspliced mRNA in mammalian cells. Nucleic Acids Res 30:e109.

Nasim MT, Ghouri A, Patel B, James V, Rudarakanchana N, Morrell NW, Trembath RC. 2008. Stoichiometric imbalance in the receptor complex contributes to dysfunctional BMPR-II mediated signalling in pulmonary arterial hypertension. Hum Mol Genet 17:1683-1694.

Nishihara A, Watabe T, Imamura T, Miyazono K. 2002. Functional heterogeneity of bone morphogenetic protein receptor-II mutants found in patients with primary pulmonary hypertension. Mol Biol Cell 13:3055-3063.

Rudarakanchana N, Flanagan JA, Chen H, Upton PD, Machado R, Patel D, Trembath RC, Morrell NW. 2002. Functional analysis of bone morphogenetic protein type II receptor mutations underlying primary pulmonary hypertension. Hum Mol Genet 11:1517-1525.

Shintani M, Yagi H, Nakayama T, Saji T, Matsuoka R. 2009. A new nonsense mutation of SMAD8 associated with pulmonary arterial hypertension. J Med Genet 46:331-337.

Southgate L, Machado RD, Snape KM, Primeau M, Dafou D, Ruddy DM, Branney PA, Fisher M, Lee GJ, Simpson MA, He Y, Bradshaw TY, Blaumeiser B, Winship WS, Reardon W, Maher ER, Fitzpatrick DR, Wuyts W, Zenker M, Lamarche-Vane N, Trembath RC. 2011. Gain-of-Function Mutations of ARHGAP31, a Cdc42/Rac1 GTPase Regulator, Cause Syndromic Cutis Aplasia and Limb Anomalies. Am J Hum Genet 88:574-85.

FIGURE LEGENDS

Figure 1. Identified variation in *SMADs 1*, *4* and *9*.

A: Sequence chromatograms of the missense defects *SMAD1* p.V3A, *SMAD4* p.N13S and *SMAD9* p.K42E, indicating the positions and heterozygous base substitutions relative to wild-type sequence below. **B:** Sequence chromatograms displaying the heterozygous splice mutation in the terminal intron of *SMAD4* (c.1448-6T>C). The horizontal arrows refer to the start of exon 11 of the gene. Identified mutations in patient samples were numbered at the nucleotide level on the basis that +1 corresponded to the A of the ATG translation initiation codon in the reference sequence. **C:** Schematic representation of the genomic organization of *SMAD4* and the functional domains encoded by the gene, namely the N-terminal MH1 and C-terminal MH2 domains shaded in dark grey, separated by the linker region shaded in light grey. Variants identified in this study are indicated below the gene structure. Mutations detected in patients with juvenile polyposis syndrome associated with HHT (JP-HHT) or HHT alone are depicted by the filled symbols where squares represent substitutions, asterisks are small deletions and polygons are nonsense mutations. **D:** Homology comparison of human SMAD1, 4 and 9 illustrating a high level of amino acid conservation across species to the level of the Pufferfish (*Takifugu rubripes*). Amino acids substituted in the IPAH patient cohort are highlighted in bold. Non-conserved residues are shown in italics. The grey shaded boxes depict the DNA binding MH1 domains.

Figure 2. Effect of SMAD substitutions on downstream signaling.

A: SMAD-mediated signaling measured by the 3GC2-Lux reporter in the presence of plasmids encoding WT and variant SMAD1 (n=13) and SMAD4 (n=8). The basal activity of the reporter was set as 100. **B:** Effect of variant SMAD1 on reporter activity was measured in the presence of

BMP4 (n=5) and BMP9 (n=8) ligand stimulation. The activity of the reporter with BMP4 and BMP9 ligand stimulation is normalized to 100. **C:** Basal activation of the reporter by WT and variant SMAD9 with (n=7) and without stimulation by BMPR-II (n=9). **D:** Activation of the reporter by WT and variant SMAD9 in the presence of BMP4 (n=4) and BMP9 (n=4) ligand stimulation. HEK293 and 293T cells were used in experiments outlined in panels A to B and C to D, respectively. An independent plasmid containing the *lacZ* gene was transfected to normalize transfection efficiency. **E.** The ratio of luciferase and β-galactosidase activities represents mRNA splicing efficiency in HEK293 cells (n=12). **F.** The expression of the *Id2* gene was greatly reduced in PASMC harboring *BMPR2* (p.W9X) (n=4) and *SMAD8* (p.K43E) (n=4) variation compared to the wild type. **G.** BMP9 stimulation increased *Id2* expression both in wild type PASMC (n=4) and in the patient cell line (n=4) although transcript levels in the patient remain significantly lower than WT likely due to the lower basal levels of *Id2* stimulation. Data represent mean ± SD. Data between groups were compared using a paired Student's *t*-test. One way analysis of variance (ANOVA) followed by Turkey's HSD tests were carried out for comparison of multiple means. A P-value of $\leq$ 0.05 indicated statistical significance. *P$\leq$ 0.05 compared with corresponding wild types..

Figure 2 A

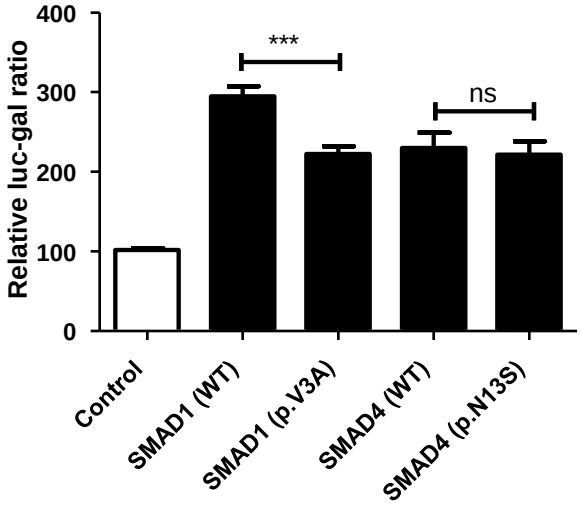

Relative luc-gal ratio
400
300
200
100
0
***
ns
Control
SMAD1 (WT)
SMAD1 (p.V3A)
SMAD4 (WT)
SMAD4 (p.N13S)

Figure 2 B

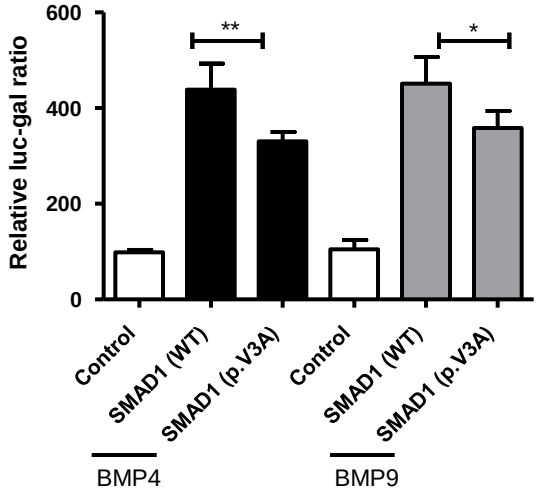

Relative luc-gal ratio
600
400
200
0
**
*
Control
SMAD1 (WT)
SMAD1 (p.V3A)
Control
SMAD1 (WT)
SMAD1 (p.V3A)
BMP4
BMP9

Figure 2 C

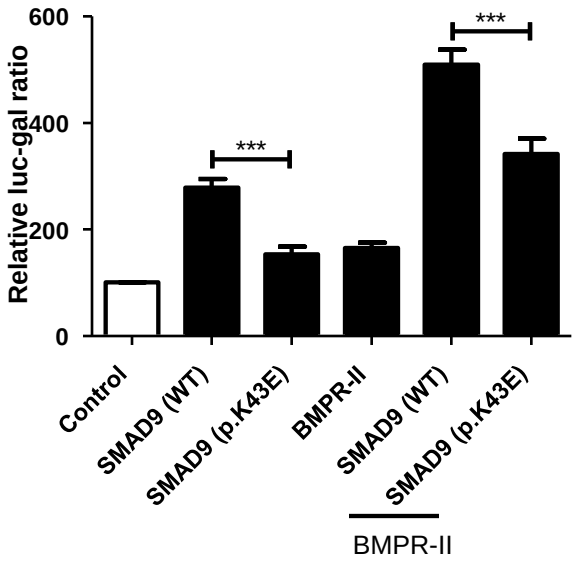

Relative luc-gal ratio
600
400
200
0
***
***
Control
SMAD9 (WT)
SMAD9 (p.K43E)
BMPR-II
SMAD9 (WT)
SMAD9 (p.K43E)
BMPR-II

Figure 2 D

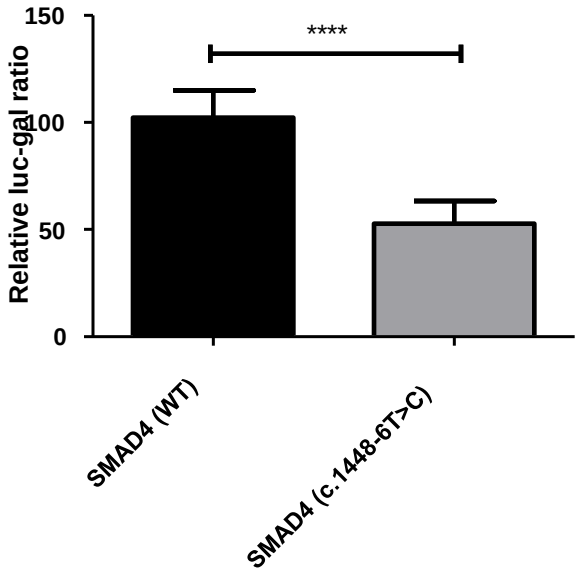


Figure 2 E

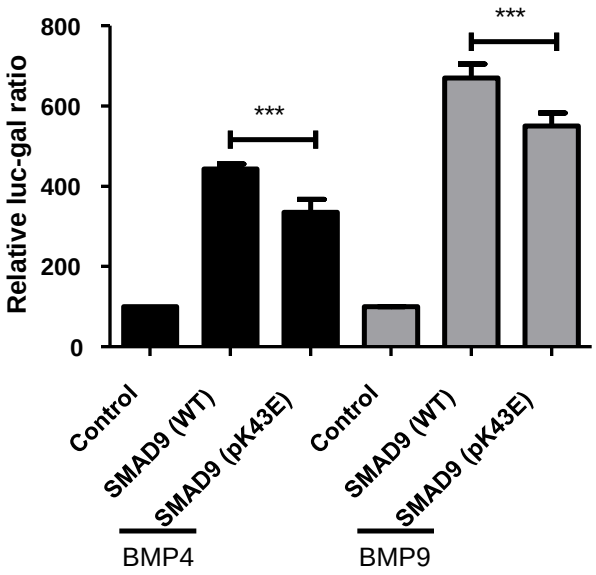
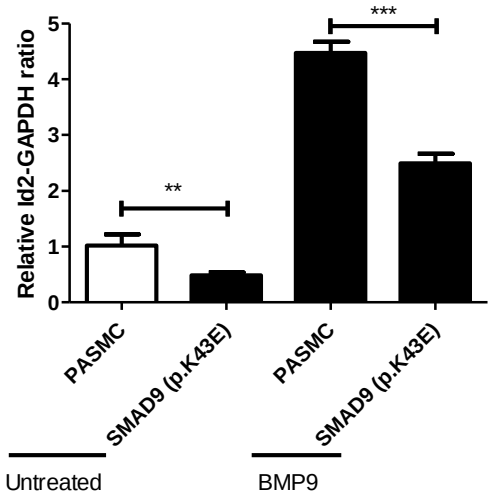
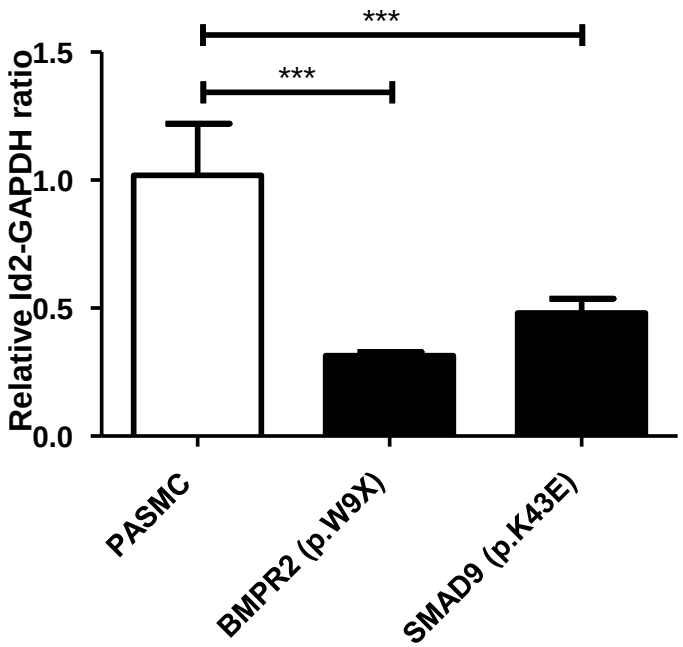

Relative luc-gal ratio
800
600
400
200
0
***
***
Control
SMAD9 (WT)
SMAD9 (pK43E)
Control
SMAD9 (WT)
SMAD9 (pK43E)
BMP4
BMP9

Figure 2 F


Relative Id2-GAPDH ratio
5
4
3
2
1
0
**
***
PASMC
SMAD9 (p.K43E)
PASMC
SMAD9 (p.K43E)
Untreated
BMP9


***
***
1.5
1.0
0.5
0.0
Relative Id2-GAPDH ratio
PASMC
BMPR2 (p.W9X)
SMAD9 (p.K43E)

Supporting information

Supplementary methods

Generation of constructs, cell culture, transient transfection and enzymatic assay

To determine splicing efficiency, constructs were developed based on pTN23 which contains a recombinant splicing unit flanked by two reporter genes encoding luciferase and β-galactosidase proteins [Nasim et al., 2002]. The original splicing unit was further modified by replacing the exon SK (pTN23) by exon 11 (120 bp) along with the preceding 40 bp intron of *SMAD4* using PCR mutagenesis and the resulting fragment was subcloned into *Sal*I/*Xho*I and *Bam*HI sites of pBPLUGA. All plasmids were verified by restriction analyses and sequencing. Cell culture and transfections were performed as previously described [Nasim et al., 2008]. Luciferase and β-galactosidase activities were determined using Dual-light Reporter Assay systems (Applied Biosystems) using an ORION-II Plate Luminometer (Berthold) according to manufacturers' protocols. The splicing assay in mammalian cells was performed as previously described [Nasim et al., 2002].

Immunoblotting

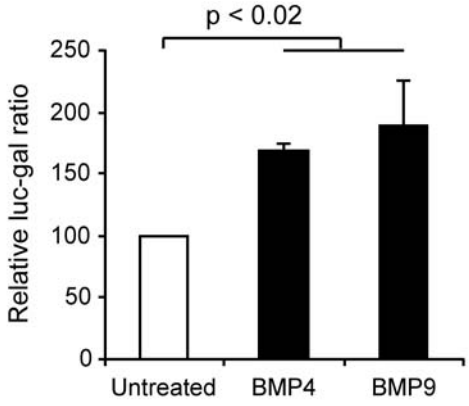
Immunoblotting experiments were conducted on HeLa cells transiently transfected with Flag-tagged constructs expressing wild-type and missense variant versions of SMAD1, 4 and 8 in 10 cm dishes. Cells were lysed in 1 ml lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$, 1% Nonidet P-40, 5 mM EDTA and protease inhibitor cocktail). Proteins were transferred onto nitrocellulose and membranes were probed anti-Flag antibody (Sigma Aldrich, F7425).

diluted 1/500 in blocking buffer (5% milk powder, 0.1% Tween-20 in PBS). Secondary antibodies were HRP conjugated and diluted according to manufacturer's instructions. Proteins were detected by chemiluminescence using the EZ-ECL kit (GE Healthcare).

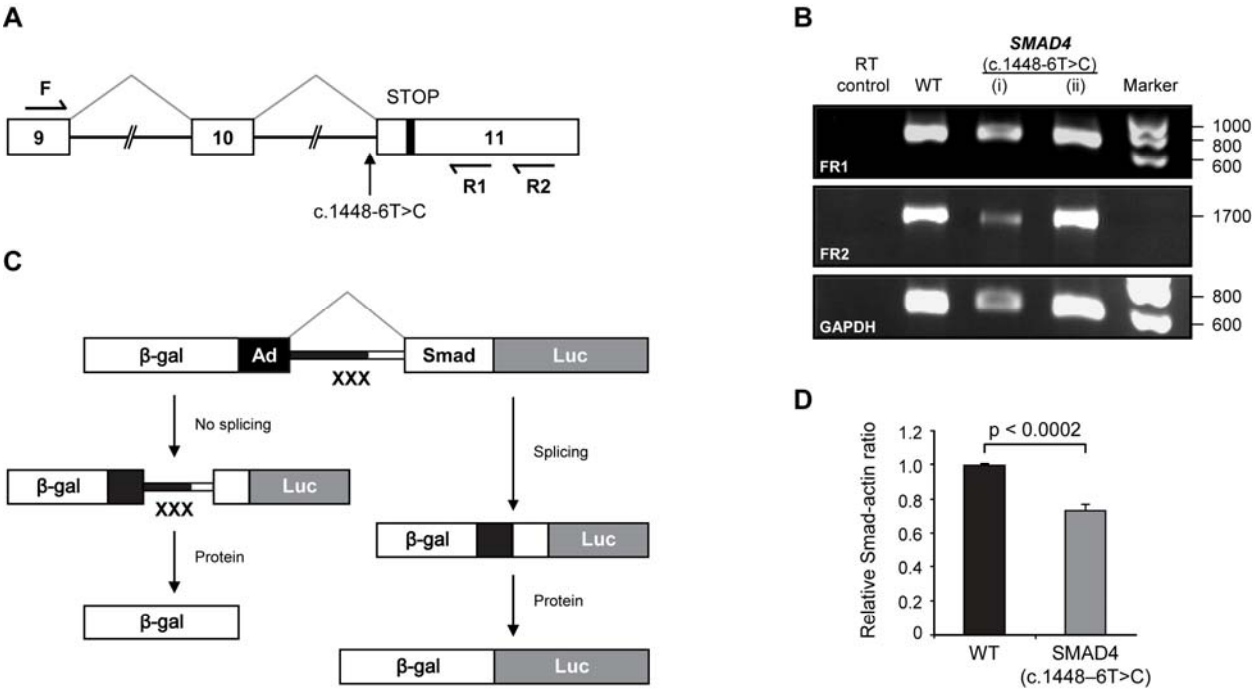
Scanning laser confocal microscopy

Cells were plated on acid treated coverslips in 6-well plates. Following transfection, cells fixed and permeabilised with absolute methanol for 10 min at -20 °C. PBS/1%BSA for 1 h was used as a blocking agent after which cells were stained with anti-Flag antibody diluted according to manufacturer's instructions. Anti-mouse Alexa 488 and was employed as the secondary antibody. Before examination, coverslips were mounted with Vectashield containing DAPI (Vector labs) to facilitate nuclear staining. Confocal microscopy was performed on a Zeiss LSM510 laser scanning microscope (Zeiss, Thornwood, NY) using a Zeiss 63 × 1.4 numerical aperture oil immersion lens.

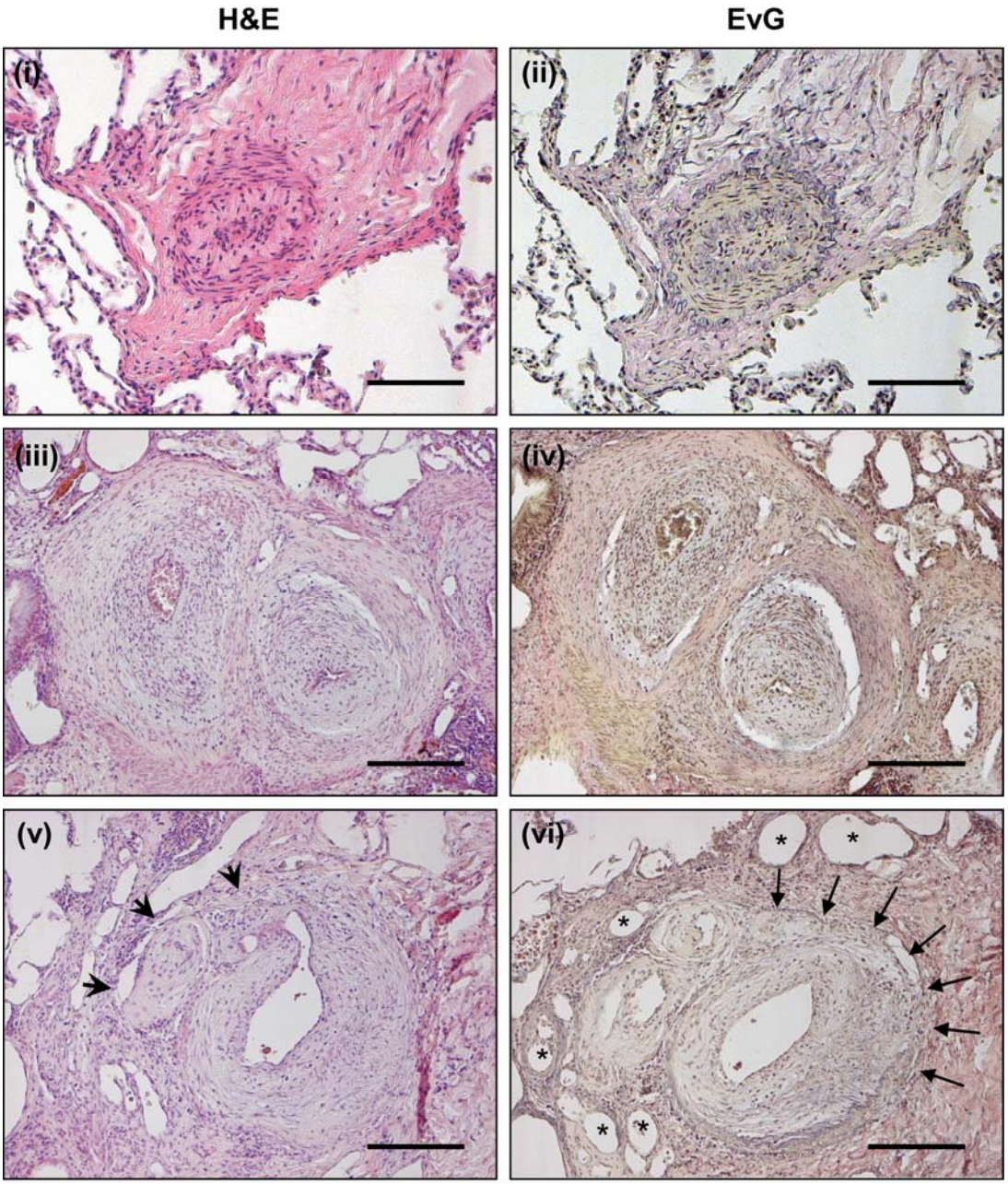
Supplementary Figures



Supplementary Figure 1. Luciferase assay demonstrating SMAD binding element reporter construct responsiveness to the BMP ligands BMP4 and 9 when normalised to untreated cells.

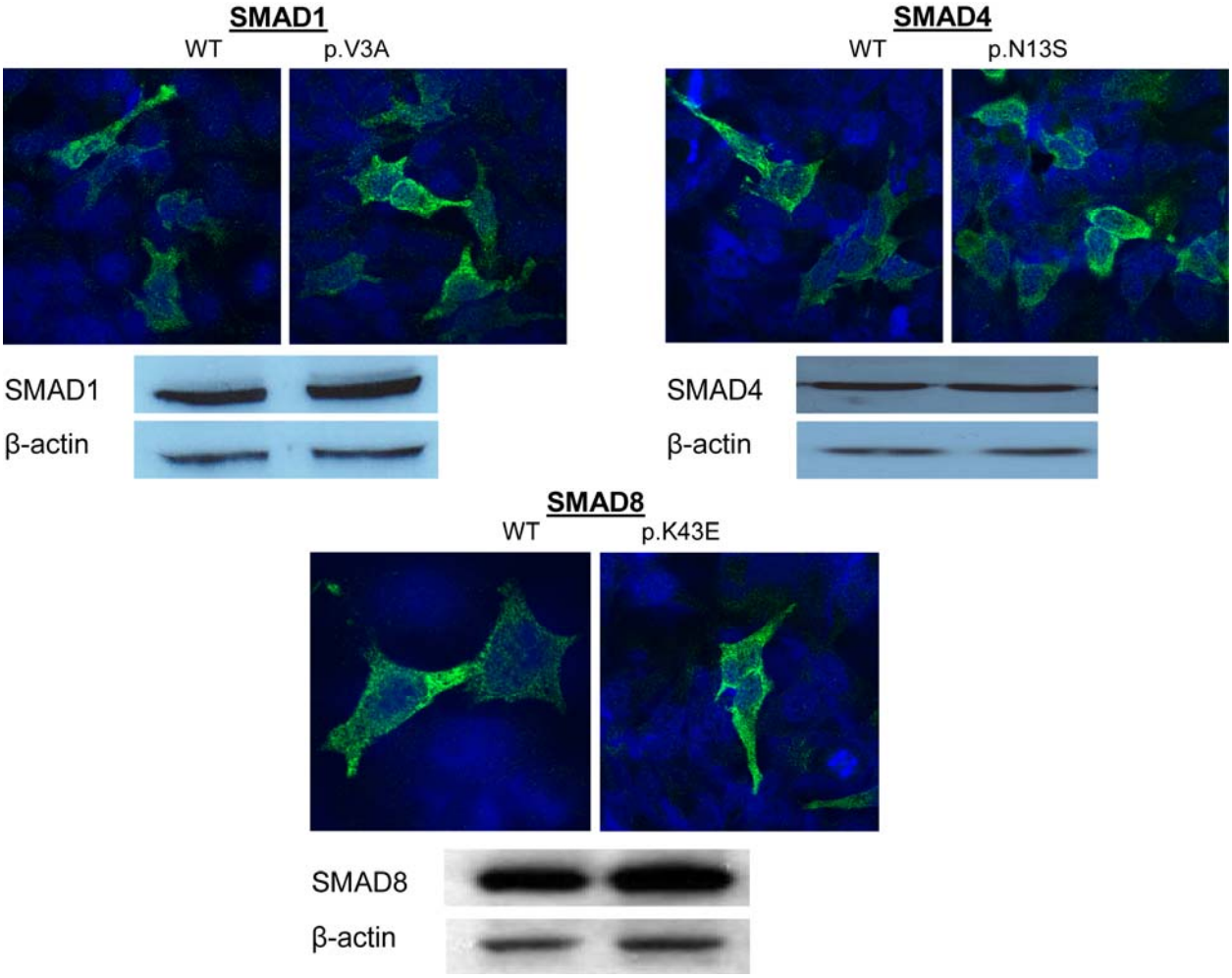


Supplementary Figure 2. RT-PCR on patient leucocytes harbouring the intronic variant c.1448-6T>C. **A.** Patient derived cDNA was amplified using two primer sets, positioned in exon 9 and 11 as indicated in the schematic. The position of the variant is marked. **B.** Panels FR1 and FR2 indicate equivalently sized PCR products between WT and mutant samples (mutant reactions in duplicate). GAPDH, in the bottom panel, was employed as a loading control. **C.** Schematic representation of the construct employed to determine splicing efficiency. **D.** The relative level of SMAD4 transcript in circulating leukocytes is reduced in a subject harboring c.1448-6T>C compared to wild type (WT) as determined by quantitative PCR analysis. The *p* values as calculated by paired student's t-test are indicated.



Supplementary Figure 3. Histopathology of pulmonary vascular lesions in the patient with *SMAD8* substitution stained with Hematoxylin and eosin (H&E) (i, iii and v) and Elastica van Gieson (EvG) (ii, iv and vi). (i) and (ii) intra-acinar pulmonary artery (bar=100µm). (iii) and (iv) two preacinar arteries (bar=200µm). (v) and (vi) Plexiform lesion (arrowhead) with inflammatory infiltrates arising from a stenotic muscular pulmonary artery (bar=200µm). The

destroyed arterial wall and surrounding thin-walled vessels and the dilatation lesions are indicated by arrows and asterisks, respectively.



Supplementary Figure 4. Immunofluorescence and immunoblotting of transiently transfected Flag-tagged WT and variant constructs for each observed amino acid substitution. Confocal microscopy indicates a cytoplasmic distribution with no evidence of mislocalisation or protein degradation in the variants. Western blots indicate the stability of variant proteins is preserved for each variant comparative to WT. β-actin was used as a loading control in each experiment.

Supplementary Table 1.

Table 1. Table of patients' clinical characteristics

	Case 1	**Case 2**	**Case 3**	**Case 4**
SMAD mutations	SMAD1	SMAD4	SMAD4	SMAD8
BMPR2 mutations	None	None	None	None
ALK1 mutations	None	None	None	None
DNA mutation	c.8T>C	c.38A>G	c.1448-6T>C	c.127A>G
Amino-acid change	p.V3A	p.N13S	N/A	p.K43E
Country of origin	France	France	France	Japan
Familial history of PAH	No	No	No	No
Clinical features of HHT	No	No	No	No
Gender	Female	Female	Male	Female
Age at PAH diagnosis	47	35	35	7
NYHA at diagnosis	NYHA IV	NYHA III	NYHA III	NYHA IV
Six-minute walk distance (meters)	75	293	235	Not done
Mean pulmonary artery pressure (mmHg)	52	62	100	87
Mean pulmonary artery wedge pressure (mmHg)	12	12	15	6
Cardiac index (L/min/m^2)	1.53	2.30	2.77	2.17
Pulmonary vascular resistance (mmHg/L/min/m^2)	26.1	21.7	30.7	37.3
Acute vasodilator challenge	Negative	Negative	Negative	Not done
Treatment	Intravenous epoprostenol	Combination of bosentan and tadalafil	Intravenous epoprostenol	Intravenous epoprostenol Lung transplantation
Outcome	Death	Alive	Death	Alive
Cause of death	Right heart failure	N/A	Right heart failure	N/A

Definition of abbreviations: N/A = Not applicable; NYHA = New York Heart Association (functional class of disease severity); mmHg = millimeters of mercury; L/min/m^2 = Liters per minutes per meter squared. Identified mutations in PAH patient samples were numbered at the nucleotide level on the basis that +1 corresponded to the A of the ATG translation initiation codon in the reference sequence.