RESEARCH ARTICLE   **Open Access**

A burden of rare variants in *BMPR2* and *KCNK3* contributes to a risk of familial pulmonary arterial hypertension

Koichiro Higasa[1*†], Aiko Ogawa[2†], Chikashi Terao[1], Masakazu Shimizu[1], Shinji Kosugi[3], Ryo Yamada[4], Hiroshi Date[5], Hiromi Matsubara[2] and Fumihiko Matsuda[1]

Abstract

Background: Pulmonary arterial hypertension (PAH) is a severe lung disease with only few effective treatments available. Familial cases of PAH are usually recognized as an autosomal dominant disease, but incomplete penetrance of the disease makes it difficult to identify pathogenic variants in accordance with a Mendelian pattern of inheritance.

Methods: To elucidate the complex genetic basis of PAH, we obtained whole exome- or genome-sequencing data of 17 subjects from 9 families with heritable PAH and applied gene-based association analysis with 9 index patients and 300 PAH-free controls.

Results: A burden of rare variants in *BMPR2* significantly contributed to the risk of the disease ($p = 6.0 \times 10^{-8}$). Eight of nine families carried four previously reported single nucleotide variants and four novel insertion/deletion variants in the gene. One of the novel variants was a large 6.5 kilobase-deletion. In the remaining one family, the patient carried a pathogenic variant in a member of potassium channels, *KCNK3,* which was the first replicative finding of channelopathy in an Asian population.

Conclusions: The variety of rare pathogenic variants suggests that gene-based association analysis using genome-wide sequencing data from increased number of samples is essential to tracing the genetic heterogeneity and developing an appropriate panel for genetic testing.

Keywords: Clinical genetic testing, Gene-based association study, Next generation sequencing, Pulmonary arterial hypertension

Background

Pulmonary arterial hypertension (PAH, MIM #178600) is a rare vascular lung disease presenting increased pulmonary vascular resistance and elevation of mean pulmonary arterial pressure, leading to a grave prognosis of right heart failure without treatment. While survival rates are increasing with a number of recently developed treatments such as epoprostenol, the optimal care of patients with these therapies is unclear due to phenotypic variations and genetic backgrounds.

Despite the complex etiology of PAH due to incomplete penetrance and genetic heterogeneity, multiple genes that cause PAH have been discovered during the last few decades. *BMPR2* variants have been identified in not only >70% of heritable PAH (HPAH), but also 10–40% of idiopathic PAH (IPAH) [1–6]. Variants of other genes including *KCNK3*, *ACVRL1*, *ENG*, *CAV1*, and the *SMAD* family are also rare causes of PAH [7–11]. Furthermore, common variants in *CBLN2* and *KCNA5* were also reported to be associated with the risk of PAH in European Caucasians [12, 13].

Although identification of individuals who carry genetic variants that increase the risk of developing PAH offers an opportunity for earlier diagnosis and finding a therapeutic strategy, the majority of previous studies was

* Correspondence: higasa@genome.med.kyoto-u.ac.jp
†Equal contributors
[1]Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kawara-cho 53, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan
Full list of author information is available at the end of the article

Higasa *et al. BMC Pulmonary Medicine* (2017) 17:57

Page 2 of 8

only focused on the protein coding regions of the most frequently mutated gene *BMPR2* with the use of conventional methods such as Sanger sequencing [6, 14]. Thus, for the patients who have no mutation in *BMPR2*, i.e., around 30% of HPAH and 60–90% of IPAH, another approach, which is practical for multiple genes, is necessary. Considering that current state-of-the-art sequencing technologies allow us to access exome- and genome-wide variants with reasonable costs, unbiased screening of pathogenic variants is beneficial to continue expanding the genetic diagnosis catalogue. However, conventional segregation-based approaches, e.g., linkage analysis, do not have enough power to pinpoint the true pathogenic variants from these genome-wide candidate variants without functional evaluations, especially for diseases with phenotypic and genetic heterogeneity. In this study, applying a gene-based association test, we statistically evaluate the significance of rare variant enrichment in genes responsible for PAH. The strategy we employed here would be useful to unbiasedly elucidate the pathogenicity of multiple rare variants arising from independent founder events in Mendelian diseases as well as common diseases.

Methods

Subjects

We consecutively enrolled five families with PAH and four individual cases with a family history of PAH to this study (Fig. 1). The patients have been diagnosed in the National Hospital Organization Okayama Medical Center between 1996 and 2014. All subjects who participated in our study were approved by the Institutional Review Board of our institutes in which donors gave written informed consent in accordance with institutional and national guidelines.
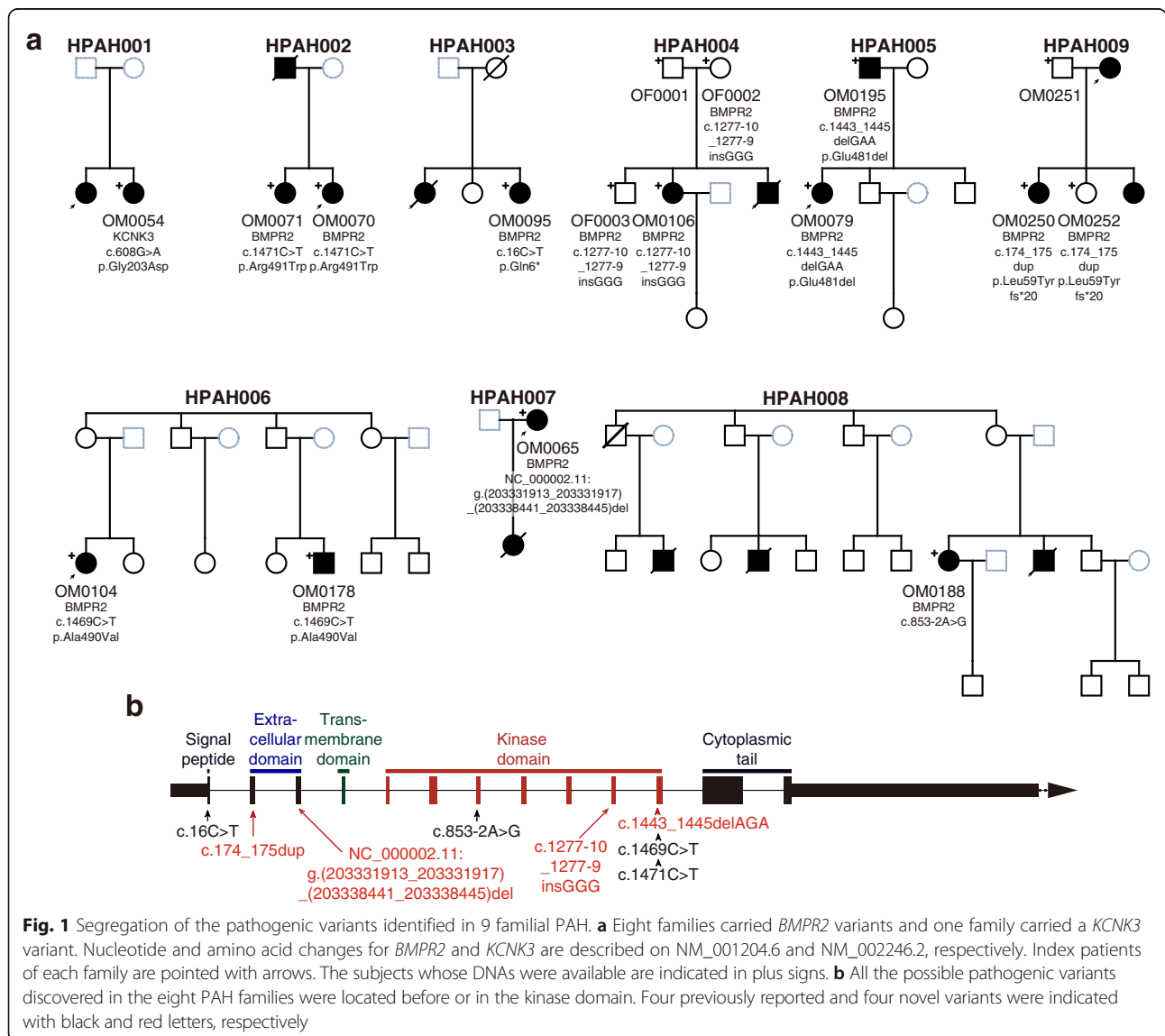

Fig. 1 Segregation of the pathogenic variants identified in 9 familial PAH. **a** Eight families carried *BMPR2* variants and one family carried a *KCNK3* variant. Nucleotide and amino acid changes for *BMPR2* and *KCNK3* are described on NM_001204.6 and NM_002246.2, respectively. Index patients of each family are pointed with arrows. The subjects whose DNAs were available are indicated in plus signs. **b** All the possible pathogenic variants discovered in the eight PAH families were located before or in the kinase domain. Four previously reported and four novel variants were indicated with black and red letters, respectively

Higasa et al. BMC Pulmonary Medicine (2017) 17:57

Page 3 of 8

Next-generation sequencing and data analysis

To understand comprehensive genetic background of these 9 PAH families, we applied whole exome- or genome-sequencing to 12 patients and 5 healthy family members whose DNA were available (Fig. 1 and Additional file 1: Table S1). For the exome sequencing, DNA fragments were enriched by SureSelectXT Human All Exon v4 + UTR (Agilent Technologies, Santa Clara, CA, USA) and then applied to SOLiD™ 5500XL sequencer (Thermo Fisher Scientific inc., Waltham, MA, USA). The whole genome sequencing was conducted with the Illumina HiSeq X sequencer (Illumina Inc., San Diego, CA, USA). After aligning the sequence reads onto the reference genome (NCBI Build 37) using the Burrows-Wheeler Aligner [15], downstream processes including the duplication removal, the recalibration of base quality values, the local realignment, the variant call, and the variant quality score recalibration were analyzed using GATK [16]. The variants were called with an exome sequencing data set of 300 control samples obtained from the Human Genome Variation Database (accession ID: HGV0000004) [17]. The resulting VCF file has been deposited on the same database under accession HGV0000005.

Quality control for association study

After removing variants that were assigned as low quality by the GATK VariantRecalibrator, additional filters were applied to extract high quality variants such as low call rate (<0.9), excessive strand bias (FS > 50), haplotype score (> 5), deviation from Hardy-Weinberg equilibrium (InbreedingCoeff > 0.3), mapping quality of the reads (MQ < 35), excess of zero mapping quality (MQ0 > 100), bias of mapping quality between reference and alternative alleles (MQRankSum < 13), coverage over sample (DP/sample < 10), positional bias of the reads (ReadPosRankSum > 5), quality over depth (QD < 8), and low LOD score (VQSLOD < 0). For the gene-based association analysis, we selected likely protein damaging variants (premature termination, splice site, missense, and indels on exons) to perform the Variable Threshold (VT) test [18] implemented in Variant Association Tools [19].

Annotation and screening of pathogenic variants

All identified variants were annotated using ANNOVAR [20]. Candidate pathogenic variants were screened according to the registrations and frequencies of the variants in the public databases: dbSNP (Build 147) [21], The 1000 Genomes (November 2010 data release) [22], The 10Gen Data Set (version 1.04) [23], NHLBI GO Exome Sequencing Project (ESP6500SI) [24], the Human Genetic Variation Database [17] or ClinVar [25]. For missense variants, PolyPhen-2 [26] and Mutation Taster [27], LRT [28] and PhyloP [29] score were obtained from the dbNSFP database [30]. Damaging effects of splice site variants were evaluated with MaxEntScan [31] and Human Splicing Finder [32].

Results

Gene-based association study

To identify genes responsible for the pathogenesis of PAH, we applied a gene-based association test (Variable Threshold test [18]) to the 60,367 damaging variants extracted from the nine PAH patients and the 300 control samples. These variants were located within 10,744 gene regions. Despite the small sample size, the burden of association between BMPR2 and PAH was highly significant ($p = 6.0 \times 10^{-8}$) compared to the genome-wide significance threshold ($p < 2.4 \times 10^{-6}$) after Bonferroni correction for approximately 21,000 genes (Fig. 2). No other gene was found beyond the threshold.

Identifying the pathogenic variants

The spectrum of rare variants found in BMPR2 is summarized in Table 1. Of the nine families, four carried previously reported single nucleotide pathogenic variants (2 missenses, 1 nonsense, and 1 splice site) [1, 4, 6] and four carried novel insertions/deletions (indels) in this gene (88.9%). One of the novel indels was a large deletion of 6.5 kilobases in length by which one allele lacks the entire region of exon 3 (Additional file 2: Figure S1). Two variants are suspected to be pathogenic although showing incomplete penetrance, since clinically unaffected subjects in the families harbored the same variants found in the patients (Table 1). There were no
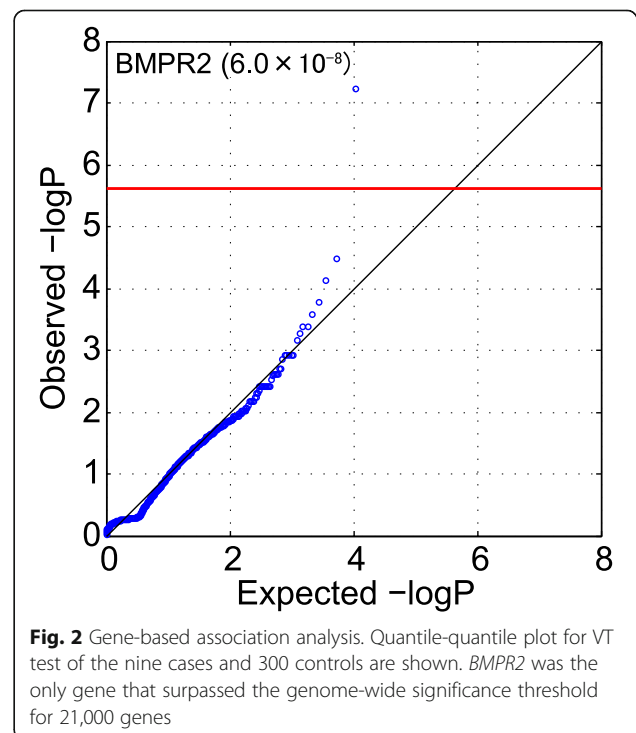

Fig. 2 Gene-based association analysis. Quantile-quantile plot for VT test of the nine cases and 300 controls are shown. BMPR2 was the only gene that surpassed the genome-wide significance threshold for 21,000 genes

Higasa *et al. BMC Pulmonary Medicine* (2017) 17:57

Page 4 of 8

Table 1 Summary of pathogenic variants and clinical features

Family ID	Sample ID	Gender	Diagnosis	Age at diagnosis (y.o.)	mPAP[a] (mmHg)	CI[b] (L/min/m^2)	PVR[c] (dyne·sec/cm^5)	Chromosome	Position	Gene	Exon
1	OM0054	F	HPAH	23	81	2.3	1623	2	26,950,859	KCNK3	2
2	OM0070	F	HPAH	23	55	1.8	1446	2	203,417,496	BMPR2	11
2	OM0071	F	HPAH	27	59	1.8	1371	2	203,417,496	BMPR2	11
3	OM0095	F	HPAH	30	62	2.7	730	2	203,242,213	BMPR2	1
4	OF0001	M	-	-	-			-	-	-	-
4	OF0002	F	-	-	-			2	203,407,024	BMPR2	10
4	OF0003	M	-	-	-			2	203,407,024	BMPR2	10
4	OM0106	F	HPAH	27	47	3.0	609	2	203,407,024	BMPR2	10
5	OM0079	F	HPAH	29	50	1.6	1556	2	203,417,465	BMPR2	11
5	OM0195	M	HPAH	61	67	1.3	2234	2	203,417,465	BMPR2	11
6	OM0104	F	HPAH	23	67	2.1	1257	2	203,417,494	BMPR2	11
6	OM0178	M	HPAH	34	63	3.3	522	2	203,417,494	BMPR2	11
7	OM0065	F	HPAH	26	47	3.5	636	2	203,331,443 ~ 203,338,403	BMPR2	3
8	OM0188	F	HPAH	36	70	1.9	1720	2	203,384,808	BMPR2	7
9	OM0250	F	HPAH	27	50	2.6	720	2	203,329,626	BMPR2	2
9	OM0251	M	-	-	-			-	-	-	-
9	OM0252	F	-	-	-			2	203,329,626	BMPR2	2

[a]mPAP: mean pulmonary artery pressure
[b]CI: cardiac index
[c]PVR: pulmonary vascular resistance
[d]Nucleotide and amino acid changes for BMPR2 and KCNK3 are described on NM_001204.6 or NM_002246.2, respectively, according to the Human Genome Variation Society nomenclature
[e]Quality score from GATK
[f]HPAH, IPAH, H/IPAH: The same pathogenic variants were found either in one of HPAH, IPAH or both HPAH and IPAH patient(s) in the previous reports

Higasa *et al. BMC Pulmonary Medicine* (2017) 17:57

Page 5 of 8

Table 1 Summary of pathogenic variants and clinical features *(Continued)*

Family ID	Nucleotide change[d]	Amino acid change[d]	rs ID	Polyphen	Mutation taster	LRT	PhyloP	MaxEnt scan	Human splice finder	Quality score[e]	Reference[f]
1	c.608G>A	p.Gly203Asp	-	1.000	1.000	0.000	2.407	N.A.	N.A.	179.78	Ma et al.[11], HPAH
2	c.1471C>T	p.Arg491Trp	rs137852746	1.000	1.000	0.000	2.582	N.A.	N.A.	854.18	Deng et al.[1], H/IPAH
2	c.1471C>T	p.Arg491Trp	rs137852746	1.000	1.000	0.000	2.582	N.A.	N.A.	854.18	Deng et al.[1], H/IPAH
3	c.16C>T	p.Gln6*	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	772.78	Momose et al.[6], HPAH
4	-	-	-	-	-	-	-	-	-	-	-
4	c.1277-10_1277-9insGGG	splicing	-	-	N.A.	N.A.	N.A.	235.06	216.94	1932.84	-
4	c.1277-10_1277-9insGGG	splicing	-	-	N.A.	N.A.	N.A.	235.06	216.94	1932.84	-
4	c.1277-10_1277-9insGGG	splicing	-	-	N.A.	N.A.	N.A.	235.06	216.94	1932.84	-
5	c.1443_1445delAGA	p.Glu481del	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	600.26	-
5	c.1443_1445delAGA	p.Glu481del	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	600.26	-
6	c.1469C>T	p.Ala490Val	-	1.000	1.000	0.000	2.582	N.A.	N.A.	598.10	Machado et al.[4], IPAH
6	c.1469C>T	p.Ala490Val	-	1.000	1.000	0.000	2.582	N.A.	N.A.	598.10	Machado et al.[4], IPAH
7	NC_000002.11:g.(203331913_203331917)_(203338441_203338445)del	exon 3 deletion	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	-
8	c.853-2A>G	splicing	-	-	N.A.	N.A.	N.A.	−85.96	−34.14	341.83	Machado et al.[4], IPAH
9	c.174_175dup	p.Leu59Tyrfs*20	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	940.89	-
9	-	-	-	-	-	-	-	-	-	-	-
9	c.174_175dup	p.Leu59Tyrfs*20	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	940.89	-

Higasa *et al. BMC Pulmonary Medicine* (2017) 17:57

Page 6 of 8

BMPR2 variants observed in the remaining one of nine families, but we identified one heterozygous missense variant (p.Gly203Asp) in *KCNK3* by screening the previously reported pathogenic variants [25]. This variant was shown to disrupt the ion-channel function by patch-clump analysis [11]. None of the pathogenic variants we identified was observed in the 300 control samples or in the public database for the Japanese population [17]. Of these, all three missense variants were occurred at highly conserved nucleotides among vertebrates and were assumed to be damaging to the protein function by at least three *in silico* prediction programs [26–29] (Table 1).

Discussion

To our knowledge, this is the first report of gene-based genome-wide association analysis of HPAH. A burden of rare variants in *BMPR2* significantly contributes to risk of the disease ($p = 6.0 \times 10^{-8}$). The approach robustly detected the gene having a large effect on the pathogenesis of PAH, despite the genetic heterogeneity. Eight probands in the nine families harbored possible pathogenic variants in *BMPR2*. Half of these variants were novel indels. One of the novel indels was a large 6.5 kilobase deletion spanning the entire region of exon 3. Another novel indel was a three base insertion (NM_001204.6:c.1277-10_1277-9insGGG) in intron 9 (Additional file 3: Figure S2). Although we could not dismiss that this insertion has no responsibility to the disease pathogenicity, a potential creation of a new splice acceptor site by this insertion was strongly suggested from the multiple splice site prediction tools (Table 1, Additional file 4: Figure S3 and Additional file 5: Figure S4) [31, 32]. The remaining one patient harbored a missense variant in *KCNK3*, which was the first replicative finding of channelopathy in Japanese population. Among the nine families, all variants identified here were mutually exclusive, suggesting that the variants have originated from independent genetic founder events.

Patients who suffered from chronic lung diseases such as chronic obstructive pulmonary disease (COPD) and pulmonary fibrosis are prone to pulmonary hypertension (PH) development. They are categorized as Group 3 in the latest guidelines [33]. Most patients with COPD develop mild PH but 3–5% of them show a further rise in mean pulmonary arterial pressure >35 mmHg. It is unknown how "severe PH-COPD", formerly known as "out-of-proportion PH" is induced. Furthermore, the PAH-approved drugs are yet to be approved for the patients with Group 3 PH. In this study, a male patient (OM0195 in HPAH005) had been treated at another hospital for COPD. He was later diagnosed with PH and referred to our hospital. This patient could be categorized as "severe PH-COPD", if none of his family members developed PAH. Since we had treated his daughter for IPAH at our

hospital, we clinically diagnosed them as HPAH. Genetic testing revealed an in-frame-deletion (c.1443_1445del-GAA) in *BMPR2* in both patients. Underlying genetic predisposition might be one of the reasons for developing "severe PH-COPD". Given our finding and a similarity of morphological appearance of vascular lesions between Group 1 and "severe PH-COPD" patients [33], PAH-approved drug treatments tailored to genetic diagnosis could well be a therapeutic strategy for such patients.

Conclusions

According to the genetic testing registry at the National Institutes of Health, the available panels for clinical genetic testing for PAH do not include *KCNK3* and the detection methods are limited. Considering that pathogenic variants could occur within or spanning non-coding regions with a variety of sizes, the sequencing of the entire region of candidate genes is recommended to further understand the genetic factors relevant to PAH. This strategy will be essential for improving genetic diagnosis and counseling for PAH.

Additional files

Additional file 1: Table S1. Summary of mapping statistics. (PDF 22 kb)

Additional file 2: Figure S1. Large deletion found in the patient in family 8. (a) The large deletion was supported by depth of coverage and discordant paired-end reads. (b) The length of the deletion was approximately 6.5 kilobases and spanned the entire region of exon 3. Breakpoints of the large heterozygous deletion were located in the Alu repeat sequences, which were supported by read depths and paired-end reads spanning the long region. (PDF 529 kb)

Additional file 3: Figure S2. Possible pathogenic variants found in the intron of BMPR2 in family 4. A rare three base insertion (GGG) at 10 bases upstream from exon 10 creates identical sequences around the canonical splicing site (ACAGGG). By the insertion, an out-of-frame protein could be translated due to the new splice acceptor site. (PDF 411 kb)

Additional file 4: Figure S3. Creation of the new splice acceptor site by the three base insertion (NM_001204.6:c.1277-10_1277-9insGGG) predicted from *in silico* programs. (PDF 218 kb)

Additional file 5: Figure S4. Disruption of wild type acceptor site by the single nucleotide variants (c.853-2A > G) predicted from *in silico* programs. (PDF 184 kb)

Higasa *et al. BMC Pulmonary Medicine* (2017) 17:57

Page 7 of 8

Availability of data and materials

The datasets generated and analyzed during the current study are available in the Human Genetic Variation Database repository; http://www.hgvd. genome.med.kyoto-u.ac.jp/repository/HGV0000004.html. http://www. hgvd.genome.med.kyoto-u.ac.jp/repository/HGV0000005.html.

Authors' contributions

RY, HD, HM and FM designed the study. AO, CT, SK and HM collected the DNA and clinical information of pulmonary arterial hypertension patients. MS performed sequencing analysis. KH analyzed the data. KH, AO and FM wrote the manuscript. All authors critically reviewed the final version of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Written informed consent for publication with data was obtained from patients and their family members.

Ethics approval and consent to participate

The study protocol and all subjects who participated in our study were approved by the Institutional Review Board of Kyoto University (approval number G755) and the National Hospital Organization Okayama Medical Center (approval number H24-RINKEN-24/H26-RINKEN-06) in which donors gave written informed consent in accordance with institutional and national guidelines.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kawara-cho 53, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan. [2]Department of Clinical Science, National Hospital Organization Okayama Medical Center, Okayama, Japan. [3]Department of Medical Ethics/Medical Genetics, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [4]Statistical Genetics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan. [5]Department of Thoracic Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan.

References

1. Deng Z, Morse JH, Slager SL, Cuervo N, Moore KJ, Venetos G, et al. Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. Am J Hum Genet. 2000;67:737–44.
2. Girerd B, Montani D, Jais X, Eyries M, Yaici A, Sztrymf B, et al. Genetic counselling in a national referral centre for pulmonary hypertension. Eur Respir J. 2016;47:541–52.
3. Ma L, Chung WK. The genetic basis of pulmonary arterial hypertension. Hum Genet. 2014;133:471–9.
4. Machado RD, Aldred MA, James V, Harrison RE, Patel B, Schwalbe EC, et al. Mutations of the TGF-beta type II receptor BMPR2 in pulmonary arterial hypertension. Hum Mutat. 2006;27:121–32.
5. Machado RD, Southgate L, Eichstaedt CA, Aldred MA, Austin ED, Best DH, et al. Pulmonary arterial hypertension: a current perspective on established and emerging molecular genetic defects. Hum Mutat. 2015;36:1113–27.
6. Momose Y, Aimi Y, Hirayama T, Kataoka M, Ono M, Yoshino H, et al. De novo mutations in the BMPR2 gene in patients with heritable pulmonary arterial hypertension. Ann Hum Genet. 2015;79:85–91.
7. Trembath RC. Mutations in the TGF-beta type 1 receptor, ALK1, in combined primary pulmonary hypertension and hereditary haemorrhagic telangiectasia, implies pathway specificity. J Heart Lung Transplant. 2001;20:175.
8. Chaouat A, Coulet F, Favre C, Simonneau G, Weitzenblum E, Soubrier F, et al. Endoglin germline mutation in a patient with hereditary haemorrhagic telangiectasia and dexfenfluramine associated pulmonary arterial hypertension. Thorax. 2004;59:446–8.
9. Shintani M, Yagi H, Nakayama T, Saji T, Matsuoka R. A new nonsense mutation of SMAD8 associated with pulmonary arterial hypertension. J Med Genet. 2009;46:331–7.
10. Austin ED, Ma L, LeDuc C, Berman Rosenzweig E, Borczuk A, Phillips JA, et al. Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. Circ Cardiovasc Genet. 2012;5:336–43.
11. Ma L, Roman-Campos D, Austin ED, Eyries M, Sampson KS, Soubrier F, et al. A novel channelopathy in pulmonary arterial hypertension. N Engl J Med. 2013;369:351–61.
12. Germain M, Eyries M, Montani D, Poirier O, Girerd B, Dorfmuller P, et al. Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. Nat Genet. 2013;45:518–21.
13. Wipff J, Dieude P, Guedj M, Ruiz B, Riemekasten G, Cracowski JL, et al. Association of a KCNA5 gene polymorphism with systemic sclerosis-associated pulmonary arterial hypertension in the European Caucasian population. Arthritis Rheum. 2010;62:3093–100.
14. Kabata H, Satoh T, Kataoka M, Tamura Y, Ono T, Yamamoto M, et al. Bone morphogenetic protein receptor type 2 mutations, clinical phenotypes and outcomes of Japanese patients with sporadic or familial pulmonary hypertension. Respirology. 2013;18: 1076–82.
15. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.
16. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
17. Higasa K, Miyake N, Yoshimura J, Okamura K, Niihori T, Saitsu H, et al. Human genetic variation database, a reference database of genetic variations in the Japanese population. J Hum Genet. 2016;61: 547–53.
18. Price AL, Kryukov GV, de Bakker PI, Purcell SM, Staples J, Wei LJ, et al. Pooled association tests for rare variants in exon-resequencing studies. Am J Hum Genet. 2010;86:832–8.
19. Wang GT, Peng B, Leal SM. Variant association tools for quality control and analysis of large-scale sequence and genotyping array data. Am J Hum Genet. 2014;94:770–83.
20. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38: e164.
21. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29: 308–11.
22. Clarke L, Zheng-Bradley X, Smith R, Kulesha E, Xiao C, Toneva I, et al. The 1000 genomes project: data management and community access. Nat Methods. 2012;9:459–62.
23. Reese MG, Moore B, Batchelor C, Salas F, Cunningham F, Marth GT, et al. A standard variation file format for human genome sequences. Genome Biol. 2010;11:R88.
24. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012;337:64–9.
25. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016;44:D862–868.
26. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248–9.
27. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11:361–2.
28. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19:1553–61.
29. Siepel A, Pollard K, Haussler D, Apostolico A, Guerra C, Istrail S, et al. New methods for detecting lineage-specific selection. Res Comput Mol Biol Proc. 2006;3909:190–205.
30. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum Mutat. 2013;34:E2393–2402.
31. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol. 2004; 11:377–94.

Higasa *et al. BMC Pulmonary Medicine* (2017) 17:57

Page 8 of 8

32. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37:e67.

33. Seeger W, Adir Y, Barbera JA, Champion H, Coghlan JG, Cottin V, et al. Pulmonary hypertension in chronic lung diseases. J Am Coll Cardiol. 2013; 62:D109–116.