Supplemental Material

Background

Lung evolution is recapitulated in humans in the five organogenesis phases: embryonic, pseudoglandular, canalicular, saccular, and alveolar.[1] Interruption of any of them results in irreversible malformations and lethal lung developmental disorders (LLDDs) in neonates. LLDDs are a rarely diagnosed (~ 1:100,000) but important group of pediatric lung anomalies presenting with severe progressive respiratory failure and persistent pulmonary arterial hypertension (PAH) in neonates, refractory to treatment.[2] Based on the characteristic histopathological features at lung biopsy or autopsy, LLDDs have been traditionally classified as Alveolar capillary dysplasia with misalignment of the pulmonary veins (ACDMPV), Acinar dysplasia (AcDys), Congenital alveolar dysplasia (CAD), and other unspecified primary pulmonary hypoplasias (PHs).[3,4]

ACDMPV affects ~ 1 in 100,000 newborns[3,5,6] and manifests clinically with progressive hypoxemic respiratory failure and severe PAH often associated with additional malformations of the cardiovascular, gastrointestinal, or genitourinary systems. Histopathologically, ACDMPV features are consistent with disturbance of canalicular, saccular, and alveolar phases of lung development[1] and include thickening of intra-alveolar septa, a reduced number of pulmonary capillaries, the majority of which do not make contact with the alveolar epithelium, medial hypertrophy of small peripheral pulmonary arteries and arterioles, and intrapulmonary vascular anastomoses.[7] In the vast majority of ACDMPV patients, heterozygous SNVs within *FOXF1* or CNV deletions involving *FOXF1* and/or its 60 kb non-coding lung-specific enhancer mapping ~ 286 kb upstream to *FOXF1* on chromosome 16q24.1 have been found.[8-12] *FOXF1* in the developing lungs is expressed in peripheral lung mesenchyme where it mediates sonic hedgehog (SHH) signaling from epithelial cells of branching tubular structures and plays a crucial role in pulmonary angiogenesis.[13-17] *Foxf1*$^{-/-}$ mice are embryonic lethal due to defects in mesodermal differentiation and cell adhesion. Haploinsufficiency of *Foxf1* in *Foxf1*$^{+/-}$ mice causes 90% perinatal lethality.[18-19]

Intriguingly, in contrast to point mutations in *FOXF1*, the ACDMPV-causative CNV deletions arise *de novo* almost exclusively on the maternal chr16q24.1.[10-12] Thus far, we and others have described 50 *de novo* CNV deletions that arose on maternal chromosome 16 and only three *de novo* CNV deletions that arose on paternal chr16q24.1 (**Fig. S1**). To explain this statistically significant bias, we proposed a model with the *FOXF1* enhancer on paternal copy of chr16q24.1 acting stronger.[11] Most recently, we showed that deletion of this enhancer (chr16:86,212,040-86,271,919) (found in ~ 25% of ACDMPV infants) results in a complete shutdown of *FOXF1* and the nearby lncRNA *FENDRR* expression in *cis* and the full ACDMPV lethal phenotype.[20] Interestingly, mice deficient in *Fendrr* developed lethal heart defects and hypoplastic lungs.[21,22]

Importantly, using genome sequencing (GS) in LLDD infants, we have identified statistically significant enrichment of non-coding SNVs within the known or putative lung-specific enhancers paired in *trans* with heterozygous pathogenic variants involving *FOXF1, TBX4,* or *FGF10* that imply a complex compound inheritance of LLDDs.[23-25] They may act as hyper- or hypomorphs, dramatically modifying the LLDD phenotypes. These findings have substantially improved the diagnosis of LLDDs[4] and showed that they are more common than previously thought.

Materials and methods

DNA and RNA isolation, genome sequencing, and variant characterization

The study ethical protocols H-8712 and H-46683 were approved by the Institutional Review Board for Human Subject Research at Baylor College of Medicine. Blood sample from the proband (pt 204.3) and blood and sperm samples from the parents were collected after obtaining the informed written consent. Moreover, lung tissues, umbilical cord, and blood samples from eight previously reported ACDMPV patients (28.7, 60.4, 125.3, 155.3, 170.3, 179.3, 180.3, and 205.3) were also analyzed.

Peripheral blood DNA was extracted using Gentra Purgene Blood Kit (Qiagen, Germantown, MD). Sperm DNA was extracted using a Puregene Core Kit A (Qiagen). RNA was isolated from foetal lung fibroblasts IMR-90 (ATCC, Manassas, VA) using miRNeasy Mini Kit (Qiagen), followed by the removal of trace amounts of DNA using Turbo DNA-free Kit (Invitrogen, Waltham, MA).

DNA sample from the proband (pt 204.3) was tested using the CytoSNP 850Kv1.2 array and software BlueFuse Multi 4.4 (Illumina, San Diego, CA) and Nextera Rapid Capture Panel NGS (TruSightTM One Panel Expanded, Illumina) with SeqNext (JSI medical systems, Ettenheim, Germany). Trio-based GS with 30X coverage was performed using a TruSeq Nano DNA HT Library Prep Kit (Illumina) and the HiSeqX platform (Illumina) at CloudHealth Genomics (Shanghai, China) as previously described.[23] The raw sequencing data were processed using bcl2fastq package (Illumina) and Trimmomatic tool followed by read alignment and mapping to the human genome reference sequence with the BWA 0.7.12 tool.

The copy number variant (CNV) deletion junction was PCR amplified using primers 5'-TGCCTGCTTCTACTGCTTAAA-3' and 5'-ATCTGCTATGGGCTGGATTAAG-3' and then Sanger sequenced. Parental origin of the deletion-bearing chromosome 16 was determined using informative SNPs amplified by PCR using primers 5'-ATCCCTGATGTAGTCAGCTGTGACCAG-3' and 5'-AGAAGAGTTCCCGAATCTTGCCTTGAT-3', of which the second mapped within the heterozygous deletion region. All genomic coordinates refer to the human genome GRCh37/hg19 build.

Histopathology studies

Histopathological studies were performed on formalin-fixed paraffin-embedded (FFPE) lung autopsy tissue stained with hematoxylin and eosin.

Luciferase reporter assays

To determine the pathogenicity of the identified CNV deletion within the *FOXF1* enhancer, we have cloned a 1.8 kb fragment (chr16:86,219,133-86,220,915) of this enhancer, mapping to the deleted region and including evolutionary ultra-conserved sequence interval, in the *FOXF1* promoter-containing *luc2* reporter vector and analyzed it for the ability to regulate *FOXF1* promoter activity. The *FOXF1* promoter region (chr16:86,541,532-86,544,295) was amplified from a normal human DNA and cloned into *Nhe*I-*Xho*I site within multiple cloning site (MCS) of a promoter-less vector pGL4.10 (Promega, Madison, WI) as described previously (14). The 1.8 kb-large *FOXF1* enhancer fragment was amplified from a normal human DNA sample using *Kpn*I primer 5′-tatggtaccGGATGTCACCTTCCTTGTTCAAAGTC-3′ (containing the *Kpn*I site) and

*Nhe*I primer 5′-tatgctagcTAACTGGACTTCCTTACATGCCTCCT-3′. As negative controls for the assay, we have PCR amplified two different genomic regions, comparable in size to the analyzed enhancer fragment. One of the control fragments was chosen based on the ChIP-seq data (ENCODE) showing that it does not specifically bind transcriptional regulators. It was amplified from the human genomic DNA control using *Kpn*I primer 5'-tatggtaccGTTCTGGAGAGGTGGGAAAATCAGT-3' and *Nhe*I primer 5'-tatgctagcCTTACAGAAGACCCAGATGGTTGGA-3'. Another control fragment was amplified from patient's DNA using *Kpn*I primer 5'-tatggtaccCTCCAGTATCACAGTTGCGTGTTAGG-3' and *Nhe*I primer 5'-tatgctagcGTGGCCATCTCTGGGATAGTATTCTG-3', flanking the analyzed CNV deletion. PCR was done using Phusion high-fidelity DNA polymerase (NEB, Ipswich, MA), applying 30 cycles of incubation at 98 °C for 10 s, 58 °C for 30 s and 72 °C for 1.5 min. The amplified fragments were digested with *Kpn*I and *Nhe*I and cloned into *Kpn*I-*Nhe*I site of pGL4.10-*FOXF1* promoter vector, upstream of the promoter. To determine whether the *RP11-805I24.3* promoter is regulated by the 1.8 kb enhancer fragment, we have replaced the *FOXF1* promoter in the described above enhancer fragment-bearing and control vectors with the 1.1 kb fragment (chr16:86,232,068-86,233,141) containing the putative *RP11-805I24.3* promoter, amplified from normal human DNA using *Nhe*I primer 5'-atagctagcGAGCTTGGCTAACATGGTCACTCAG-3' and *Hind*III primer 5'-tataagcttTTCCCTACATTTCTCGAGTCTGTGC-3'.

For transfection, human fetal lung fibroblasts IMR-90 were cultured at 37°C in Eagle's Minimal Essential Medium (ATCC), with 2 mM L-glutamine and 10% fetal bovine serum (ATCC), in the presence of 5% CO_2 on 12-well plates. The cells were transfected in serum-free Opti-MEM (GIBCO, Waltham, MA) using Lipofectamine 3000 (Invitrogen) (4 µl/well) and 1 µg per well of the pGL4.10-*FOXF1* promoter plasmid, bearing either control or enhancer fragment, and 0.1 µg of pGL4.75 (Promega), constitutively expressing *Rluc*. Following cell lysis in Qiazol (Qiagen) 48 h after transfection, RNA was isolated and then converted to cDNA using SuperScript III First-Strand Synthesis System (Invitrogen).

The expression of *luc2* and *Rluc* were determined by measuring levels of their cDNA by qPCR. Custom designed TaqMan primers and probes (*luc2*: assay AP7DRTC, amplicon coordinates: 247-312 in pGL4.10, AY738222; and *Rluc*: assay AP47W76, amplicon coordinates: 1,532-1,593 in pGL4.75, AY738231) were obtained from Applied Biosystems (Waltham, MA). qPCRs were done on CFX Real Time thermocycler (BioRad, Hercules, CA) using TaqMan Universal PCR Master Mix (Applied Biosystems). qPCR conditions included 40 cycles of heating the reaction mixtures for 15 s at 95°C followed by 1 min at 60°C. For relative quantification of the cDNAs and thus transcripts, the comparative C_T method was used. *Luc2* cDNA levels were normalized to those of *Rluc*.

siRNA knock-down of long non-coding RNAs (lncRNAs)

Gene knock-down experiments were done in human lung fibroblast IMR-90 cell line using the Silencer or Custom Select siRNAs (Ambion, Foster City, CA) (**Table S1**). Two Silencer Select Negative Controls (Ambion), constituting siRNAs that do not target any human gene transcript, were used. Two (or one in case of *RP11-805I24.1*) siRNAs were applied at a concentration of 35 pmol/ml of cell culture medium each. Transfections were done in 12-well plates using Lipofectamine RNAiMAX (Invitrogen) at a concentration of 3 µl/ml. Cells for RNA preparation were lysed in Qiazol (Qiagen) 48 h after the transfection. RNA purification, cDNA synthesis, and

qPCR were done as described for the luciferase reporter assay except that the transcript levels were normalized to *GAPDH*. Information on predesigned and custom-designed TaqMan assays (Applied Biosystems) is provided in **Table S2**.

We have also analyzed lncRNA *AC040170.1* located within the Unit 2 of the *FOXF1* enhancer (**Fig. 1**) and transcribed from its putative promoter located within the *FOXF1* enhancer and overlapping 3'ends of the longest *LINC01081* isoforms and *LINC02135* expressed divergently from the *LINC01081-202* putatively bi-directional promoter, antisense overlapping *LINC01081-201*, and the other 5'-extended *LINC01081* isoforms.

LINC01082 *knock-in*

The lncRNA *LINC01082* gene, encoding the splicing variant *LINC01082-201* (chr16:86,229,787-86,233,326), was amplified from the normal control human genomic DNA using the primers 5'-atggtaccGGTTTAGATTAGCCGTGGCCTA-3' and 5'-cgcgaattcTTCTGTTTGAGACATATTAAACAAGCT-3', containing the added *Kpn*I and *Eco*RI cutting sites, respectively. It was subsequently cloned into the *Kpn*I-*Eco*RI site within the MCS of pcDNA3.1 expression vector (Invitrogen). For a control, we have amplified the genomic fragment (chr16:86,166,305-86,169,885) of the size comparable to *LINC01082* from the gene and regulatory sequence desert located over 40 kb upstream of *LINC01082*, using the primers 5'-atggtaccCAACCACAGGGAGTATCTTATGTGCAG-3' and 5'-gcgaattcCGTAAGTCCTAGCTCAGAACCCACTGC-3', and cloned it into the *Kpn*I-*Eco*RI site of pcDNA3.1. As an additional control, we used the empty pcDNA3.1 vector. The constructs were transfected into IMR-90 cells using Lipofectamine 3000, as described for the reporter assays. *LINC01082* over-expression and its effect on the *FOXF1* expression were analyzed by qPCR as described for the knock-down experiments.

DNA methylation studies

To determine whether there are differences in parental chromosome methylation at CpG dinucleotides within the *FOXF1* enhancer region, we have analyzed DNA methylation status of the selected enhancer intervals within the region of hemizygosity in five patients with the maternal CNV deletion of the *FOXF1* enhancer (pts 28.7, 60.4, 125.3, 155.3, and 170.3) and four patients with the paternal CNV deletion (pts 179.3, 180.3, 204.3 and 205.3). Two analyzed regions were from the Unit 1 and two from Unit 2 (**Fig. 1**, **Table S3**). Genomic DNA (500 ng) from the ACDMPV patients with maternal or paternal deletion of the enhancer was digested overnight in NEBuffer 1 with 10 units of *Hpa*II (NEB), sensitive to cytosine methylation within CCGG sequence. PCR amplification of selected enhancer regions containing *Hpa*II site was done using 100 ng of undigested or digested DNA, primers listed in **Table S3**, and *Taq* DNA Pol in standard PCR conditions with 30 amplification cycles.

Results

Clinical findings and histopathology

Clinical course and histopathological findings in the proband (pt 204.3) were consistent with ACDMPV (**Fig. S2**).

A female newborn was transferred to an intensive care unit because of respiratory failure. Intubation was required immediately after Caesarean delivery and pneumothorax was drained. Despite of maximum respiratory support sufficient oxygenation could not be achieved. Nitric oxide was administered with little or doubtful benefit. In addition, relevant catecholamine vasopressor doses were needed to stabilize the hemodynamic function. Therefore, a decision was made to support the patient with va-ECMO. In the transthoracic echocardiography examination, a persistent fetal circulation with a small perimembranous atrial septal defect (leading to right-to-left shunt) was found. Suspecting PAH, right heart catheterization was performed which revealed borderline dysplasia of the pulmonary vessels. Following respiratory and haemodynamic stabilisation, several attempts to wean the patient off ECMO support failed. Without clinical improvement, angiography was repeated and the atrial septal defect was transiently occluded with a balloon catheter. As this intervention lead to a rise in arterial oxygen saturation, the atrial septal defect was closed surgically. The clinical status of the patient had imroved to such an extent that va-ECMO support could be weaned. During surgical removal of the ECMO cannulas (a central cannulation technique), the lungs were inspected and lung biopsy was performed. Shortly after removing the patient off ECMO, cardiorespiratory function deteriorated again. Despite of renewed extracorporeal support and in view of the absence of a causative therapy, the therapy goal was changed to palliative care. The patient died at the age of 2.5 weeks.

Histopathological studies showed abnormal lung architecture with deficient terminal alveolar development. The alveolar septa were widened with variable cellularity and paucity of septal capillaries which were predominantly positioned within the center of the alveolar septa. The autopsy was limited in nature with a single bronchovascular bundle which showed venous-like vascular profiles within the connective tissue, intimately associated with pulmonary arteries. In addition, pulmonary hypertensive changes with medial hypertrophy of pulmonary arteries and arterioles was present.

Genomic structure of the deletion region

Single nucleotide polymorphism (SNP) array analysis in the proband revealed no large CNV abnormalities involving the *FOXF1* gene. No pathogenic variants were detected in *FOXF1* in the targeted panel NGS analysis. However, trio GS revealed an apparently pathogenic *de novo* ~ 8.8 kb CNV deletion mapping ~ 322 kb upstream to the *FOXF1* gene within the centromeric portion (Unit 1) of the ~ 60 kb *FOXF1* lung-specific enhancer region (**Fig. 1, S3)**. The breakpoints of this deletion were mapped to the non-repetitive sequences: chr16:86,213,689-86,213,691 and chr16:86,222,496-86,222,498 with a 5'-CTG-3' microhomology) (**Fig. S4**). Deletion junction-specific PCR showed no evidence of parental mosaicism in the blood samples and the father's sperm sample, indicating that this deletion arose de novo (**Fig. S5**). DNA sequencing of two informative SNPs, rs12711495 and rs12711496, centromeric to the deletion region, revealed that the deletion arose on the paternal chromosome 16 (**Fig. S6**). This CNV deletion completely overlaps the previously reported pathogenic 4,115 bp CNV deletion (chr16:86,216,561-86,220,676) and is overlapped by 32 other *FOXF1* enhancer CNV deletions that did not harbor *FOXF1* (**Fig. S1**), defining the smallest deletion overlap region for all ACDMPV-causative CNV deletions. Importantly, both 8.8 kb and 4.1 kb deletion interval encompass the ~ 660 bp-large evolutionarily ultra-conserved region 660UCR (chr16:86,219,697-86,220,358). Of note, 660UCR overlaps also with the ENCODE candidate Cis-Regulatory Element (cCRE), EH38E1835120 featuring distal enhancer-like signature. Based on the ENCODE 3 database of transcription factor

binding sites identified by ChIP-seq experiments in lung-derived cells, it binds EP300 and is adjacent to the RNA PolII and RAD21 binding sites. Notably, 660UCR region is located in the vicinity of another ultra-conserved ~ 1 kb DNA interval (1000UCR; ~ chr16:86,229,400-86,230,400) that overlaps lung-expressed lncRNA *LINC01082* and its antisense *RP11-805I24.3*. This ultra-conserved region contains the binding sites for EZH2 and several other transcription factors (TFs), suggesting that both ultra-conserved elements constitute a functional Unit 1 of the *FOXF1* enhancer.

Functional analyses of the deletion region

FOXF1 regulation by the ultra-conserved ~ 660 bp enhancer interval

Using luciferase reporter assay in lung IMR-90 cells, we have found that the 1.8 kb enhancer fragment, containing the 660UCR, increased the transcription of *luc2* reporter gene from the subcloned *FOXF1* promoter two-fold (p=0.003) (**Fig. 1B**).

Enhancer lncRNA genes adjacent to the deleted region

We have then characterized the expression and the effect of the enhancer lncRNA genes neighboring and overlapping the ultra-conserved enhancer regions in the Unit 1 on *FOXF1*. We have found that the mutually overlapping *LINC01082* and its antisense *RP11-805I24.3* (*AC135012.3*), located ~ 7 and ~ 1 kb from the region deleted in the proband (pt 204.3), respectively, are expressed in IMR-90 fibroblasts at much lower levels in comparison to, e.g., *FENDRR* or *FOXF1* (regulated by the same enhancer region (21), with *LINC01082* being the least abundant (**Fig. S7**). Using siRNA-based gene silencing, we have found that the decrease of *RP11-805I24.3* expression to 16% ($P < 0.001$) reduced the *FOXF1* transcript level to 39% ($P < 0.001$). The expression of *FENDRR* was reduced by 38% ($P = 0.005$). Depletion of *RP11-805I24.3* to 31% ($P < 0.001$) resulted in a 50% decrease of the *FOXF1* transcript level ($P < 0.001$) (**Fig. 1C**). In response to the depletion of *RP11-805I24.3*, the expressions of lncRNAs *LINC01082, LINC01081* (encoded in the enhancer Unit 2), and *LINC02135* were reduced also to 50-60%. On the other hand, the depletion of *LINC01082* to 18% ($P < 0.001$) did not have any significant effects on the expression of *FOXF1* or any analyzed here lncRNAs. However, the overexpression of *LINC01082* to about the level of *RP11-805I24.3* resulted in two-fold decrease of *FOXF1* expression (**Fig. 1D**).

Since both the 660UCR and *RP11-805I24.3* positively regulate expression of *FOXF1*, we sought whether *FOXF1* expression requires their mutual interaction. The luciferase reporter assay with this ultra-conserved region showed that its spatial proximity to the *RP11-805I24.3* promoter does not modify the activity of this promoter (**Fig. S8**). Thus, our data indicate that 660UCR and *RP11-805I24.3* regulate *FOXF1* independently.

We have previously shown that FOXF1 positively regulates the expression of *FENDRR* (21). We thus sought whether *FOXF1* depletion might also affect the expression of lncRNAs transcribed from the *FOXF1* enhancer. We found that FOXF1 positively regulated expression of *LINC01082* but did not regulate expression of *RP11-805I24.3*. Thus, there is no evidence for a regulatory feedback loop interaction between lncRNAs *FENDRR*, *LINC01082*, and *RP11-805I24.3* and *FOXF1*.

In the *FOXF1* enhancer Unit 2, the levels of *LINC01081* isoforms, including *AC040170.1 (LINC01081:14)*, were higher than that of the NCBI-annotated *LINC01081-202* isoform (data not shown), whereas the expression of *LINC02135*, divergently transcribed from *LINC01081* promoter, was six times higher than the expression of *RP11-805I24.3* (**Fig. S7).** Of note, the

putative promoter of *AC040170.1* is located within one of evolutionary conserved intervals of Unit 2. The lncRNA *LINC01081*, whose gene 3'end overlaps with GeneHancer element GH16J086219, only weakly regulates *FOXF1* (22).

Differential methylation within enhancer Unit 1

Consistent with the previous findings (23,24), using PCR following methylation sensitive restriction digestion, we have found that the CpG-rich region overlapping the *LINC01082* and *RP11-805I24.3* genes and including the GLI2 binding sites (chr16:86,232,367-86,232,979) (21) is methylated on paternal (4/5) but not on maternal (0/4) chromosome 16 (**Fig. 1E**). We have also found that both parental alleles of the conserved 660UCR within Unit 1 are methylated (maternal: 3/4; paternal: 5/5), although with a some potential parental methylation bias that we could not determine. Within Unit 2, two intervals were found to be methylated on both parental alleles, although the methylation of the region within *LINC01081:14*, corresponding to the exon 7 of *LINC01081-202,* was weaker on the maternal allele of the enhancer in two out of four ACDMPV samples analyzed (**Fig. S9)**. In summary, we have found stronger methylation of the enhancer CCGG cytosines on paternally inherited chromosome 16 compared to maternal chromosome 16, including primarily a CpG-rich region with GLI2 binding sites within the Unit 1 and another ultra-conserved non-coding interval located within Unit 2, overlapping *LINC01081*.

Supplementary References

1. Kimura J, Deutsch GH. Key mechanisms of early lung development. *Review Pediatr Dev Pathol.* 2007;10(5),335-347.
2. Nogee LM. Interstitial lung disease in newborns. *Semin Fetal Neonatal Med*. 2017;22(4),227-233.
3. Bishop NB, Stankiewicz P, Steinhorn RH. Alveolar capillary dysplasia. *Am J Respir Crit Care Med.* 2011;184(2),172–179.
4. Vincent M, Karolak JA, Deutsch G, et al. Clinical, histopathological, and molecular diagnostics in lethal lung developmental disorders. *Am J Respir Crit Care Med.* 2019;200(9),1093–1101.
5. Langston C. Misalignment of pulmonary veins and alveolar capillary dysplasia. *Pediatr Pathol*. 1991;11(1),163-170.
6. Sen P, Thakur N, Stockton DW, Langston C, Bejjani BA. Expanding the phenotype of alveolar capillary dysplasia (ACD). *J Pediatr*. 2004;145(5):646-651.
7. Galambos C, Sims-Lucas S, Ali N, et al. Intrapulmonary vascular shunt pathways in alveolar capillary dysplasia with misalignment of pulmonary veins. *Thorax*. 2015;70:84-85.
8. Stankiewicz P, Sen P, Bhatt SS, et al. Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of *FOXF1* cause alveolar capillary dysplasia and other malformations. *Am J Hum Genet.* 2009;84(6),780–791.
9. Sen P, Yang Y, Navarro C, et al. Novel *FOXF1* mutations in sporadic and familial cases of alveolar capillary dysplasia with misaligned pulmonary veins imply a role for its DNA binding domain. *Hum Mutat*. 2013;34(6),801-811.
10. Szafranski P, Dharmadhikari AV, Brosens et al. Small non-coding differentially methylated copy-number variants, including lncRNA genes, cause a lethal lung developmental disorder. *Genome Res*. 2013;23(1),23-33.

11. Szafranski P, Gambin T, Dharmadhikari AV, et al. Pathogenetics of alveolar capillary dysplasia with misalignment of pulmonary veins. *Hum Genet.* 2016;135(5),569-586.
12. Szafranski P, Herrera C, Proe LA, at al. Narrowing the FOXF1 distant enhancer region on 16q24.1 critical for ACDMPV. *Clin Epigenet*. 2016;8,112.
13. Murphy DB, Wiese S, Burfeind P, et al. Human brain factor 1, a new member of the fork head gene family. *Genomics*. 1994;21(3),551-557.
14. Pierrou S, Hellqvist M, Samuelsson L, Enerbäck S, Carlsson P. Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J.* 1994;13(20),5002-5012.
15. Gebb SA, Shannon JM. Tissue interactions mediate early events in pulmonary vasculogenesis. *Dev Dyn*. 2000;217:159–169.
16. Hines EA, Sun X. Tissue crosstalk in lung development. *J Cell Biochem*. 2014;115:1469–1477.
17. van Tuyl M, Liu J, Wang J, Kuliszewski M, Tibboel D, Post M. Role of oxygen and vascular development in epithelial branching morphogenesis of the developing mouse lung. *Am J Physiol Lung Cell Mol Physiol.* 2005;288:L167–L178.
18. Kalinichenko VV, Lim L, Stolz DB, et al. Defects in pulmonary vasculature and perinatal lung hemorrhage in mice heterozygous null for the Forkhead Box f1 transcription factor. *Dev Biol*. 2001;235(2),489-506.
19. Mahlapuu M, Enerbäck S, Carlsson P. Haploinsufficiency of the forkhead gene *Foxf1*, a target for sonic hedgehog signaling, causes lung and foregut malformations. *Development*. 2001;128(12),2397-2406.
20. Szafranski P, Gambin T, Karolak JA, Popek E, Stankiewicz P. Lung-specific distant enhancer cis regulates expression of *FOXF1* and lncRNA *FENDRR*. *Hum Mutat*. 2021;42(6),694-698.
21. Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, et al. The tissue specific lncRNA *Fendrr* is an essential regulator of heart and body wall development in the mouse. *Dev Cell*. 2013;24:206–214.
22. Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife*. 2013;2:e01749.
23. Karolak JA, Vincent M, Deutsch G, et al. Complex compound inheritance of lethal lung developmental disorders due to disruption of the TBX-FGF pathway. *Am J Hum Genet.* 2019;104(2),213-228.
24. Szafranski P, Liu Q, Karolak JA, et al. Association of rare non-coding SNVs in the lung-specific *FOXF1* enhancer with a mitigation of the lethal ACDMPV phenotype. *Hum Genet.* 2019;138(11-12),1301–1311.
25. Karolak JA, Gambin T, Honey EM, et al. A de novo 2.2 Mb recurrent 17q23.1q23.2 deletion unmasks novel putative regulatory non-coding SNVs associated with lethal lung hypoplasia and pulmonary hypertension: a case report. *BMC Med Genom*. 2020;13(1),34.

Table S1. siRNA used in knockdown experiments.

siRNA-targeted gene	Targeted exon (GENCODE transcript ID)	siRNA ID / design	Sense and anti-sense sequences for custom-designed siRNA
RP11-805I24.3	3 (ENST00000599841.1)	555500 / ABKAL12	5'-GCAGAGACUUAGAUUUGAAtt-3' 5'-UUCAAAUCUAAGUCUCUGCag-3'
LINC01082	1 (ENST00000669926.1)	s461135	NA
	2 (ENST00000669926.1)	s59760	NA
LINC01081	7 (ENST00000602425.2)	s552842	5'-UGAUCAAGAUUUUUGCAGAtt-3' 5'-UCUGCAAAAAUCUUGAUCAct-3'
	8 (ENST00000602425.2)	s552843	5'-CGCAAAAUCUGCACCUGGAtt-3' 5'-UCCAGGUGCAGAUUUUGCGtc-3'
FOXF1	1	s5221	NA
	2	s5220	NA

Table S2. TaqMan gene expression assays used in knockdown experiments.

TaqMan-targeted gene	Targeted exon (GENCODE transcript ID)	Assay ID	Amplicon coordinates (GRCh37/hg19) for custom-designed assay
RP11-805I24.3	3 (ENST00000599841.1)	AP47VTU	chr16:86,227,530-86,227,586
LINC01082	1/2 (ENST00000669926.1)	Hs01388639	NA
LINC01081	1 (ENST00000602425.2)	AP32Z7W	chr16:86,319,717-86,319,789
	7/8 (ENST00000602425.2)	ARGZGMA	chr16:86,256,198-86,256,222(context seq.)
LINC02135	3 (ENST00000599486.1)	APNKTVY	chr16:86,325,781-86,325,852
FENDRR	1/2 (ENST00000662100.1)	Hs00419733	NA
FOXF1	1/2	Hs00230962	NA
GAPDH	7/8	Hs02758991	NA

Table S3. Primers used in methylation analysis.

Analyzed region	Genomic coordinates of the amplicon (GRCh37/hg19)	PCR primers
Enhancer Unit 1: 650 bp conserved region	chr16:86,219,739-86,220,347	5'-AGACACACTCAGGTGGCTTCTG-3' 5'-TGTTGCCCTATGTCATACCAATG-3'
Enhancer Unit 1: CpG region with GLI2 sites	chr16:86,232,369-86,232,836	5'-CCTGCGCTAATAAATGCTCCTT-3' 5'-GGTGGTCATCCATTAGCAGTCA-3'
Enhancer Unit 2: *LINC01081* exon 7; in proximity to a putative promoter of *LINC01081* isoforms: *TCONS_00025081*, *00024760*, *00024762*, and *00025084*	chr16: 86,256,080-86,256,591	5'-GAGGAAGAAACTGAGACAAACCAAA-3' 5'-CTTGCACGTTTGGCTTCTTCTA-3'
Enhancer Unit 2: TBX4-binding region	chr16:86,257,709-86,258,282	5'-AGACCCAGGAGCAAAGTACAGG-3' 5'-CAGCAACACCAATGGTAACAGC-3'

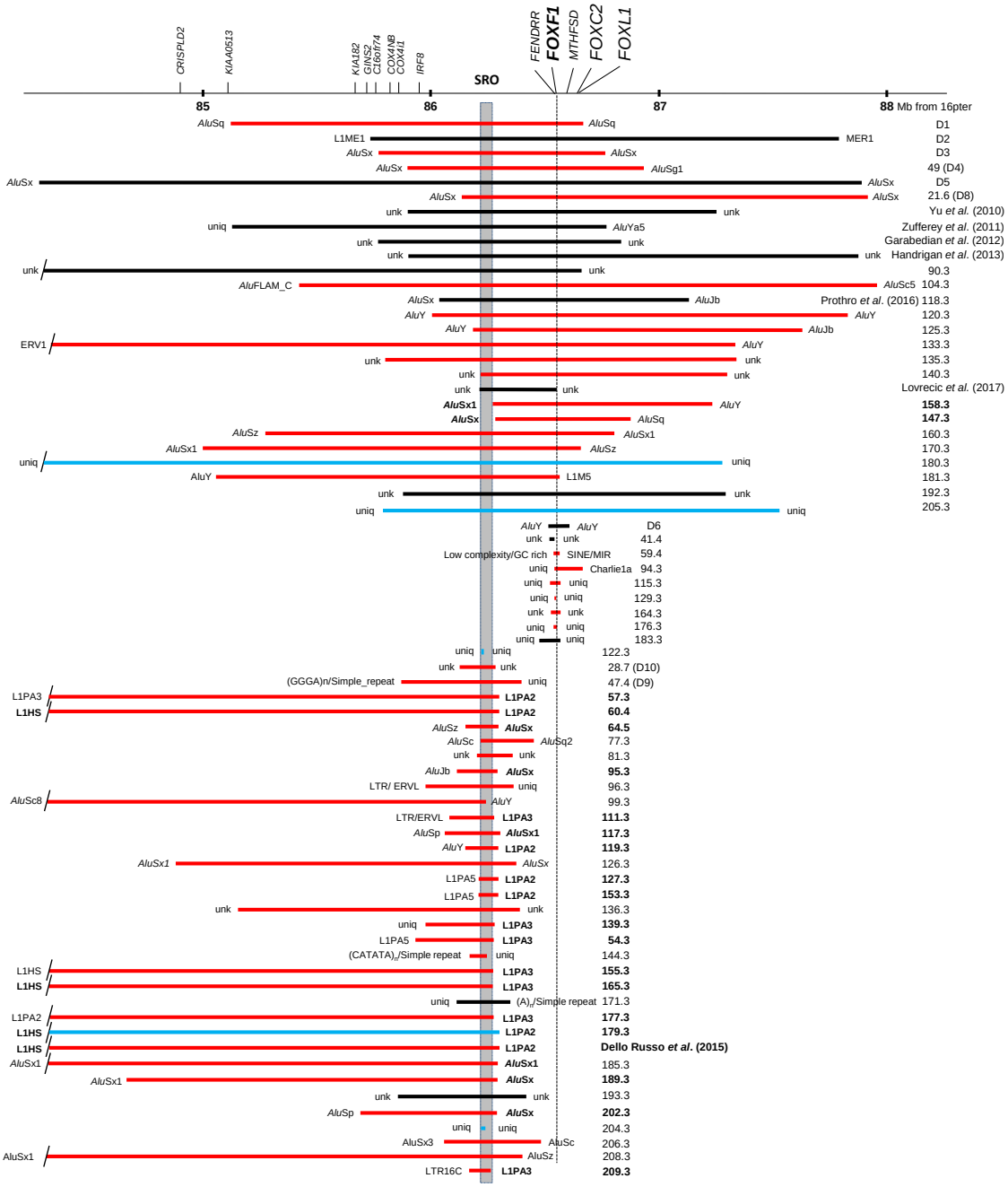

Figure S1. A compilation of the ACDMPV-causative CNV deletions on chromosome 16q24.1. CNV deletions that occurred on maternal chr16q24.1 are shown in red, paternal chromosome in blue, and those for which parental origin could not be determined are in black. The vast majority of CNV deletion breakpoints are mapped within repetitive elements that mediated the rearrangements. SRO, previously determined smallest region of deletion overlap delineating the upstream enhancer region; unk, exact breakpoint sequence unknown; uniq, breakpoint in unique sequence.

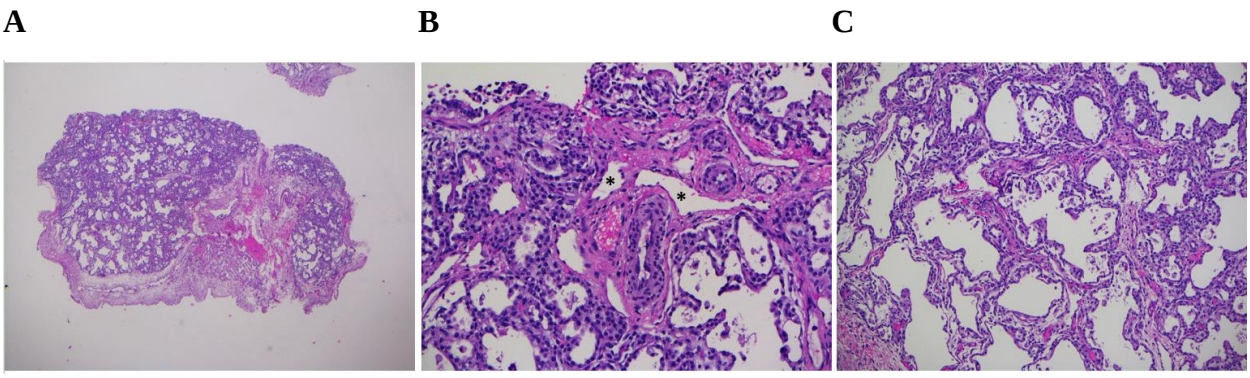

Figure S2. Microscopic findings on lung autopsy. Hematoxylin and eosin stain. (**A**) Low-power view to highlight enlarged and simplified lung architecture with deficient terminal alveolar development. (**B**) Broncho-vascular bundle that shows venous like profiles (black asterisks) within connective tissue, intimately associated with pulmonary artery branches. (**C**) Expanded alveolar septa with loose connective tissue and variably increased cellularity. There is paucity of capillaries which are centrally placed within the alveolar septa. Deficient terminal alveolar development is also appreciated.

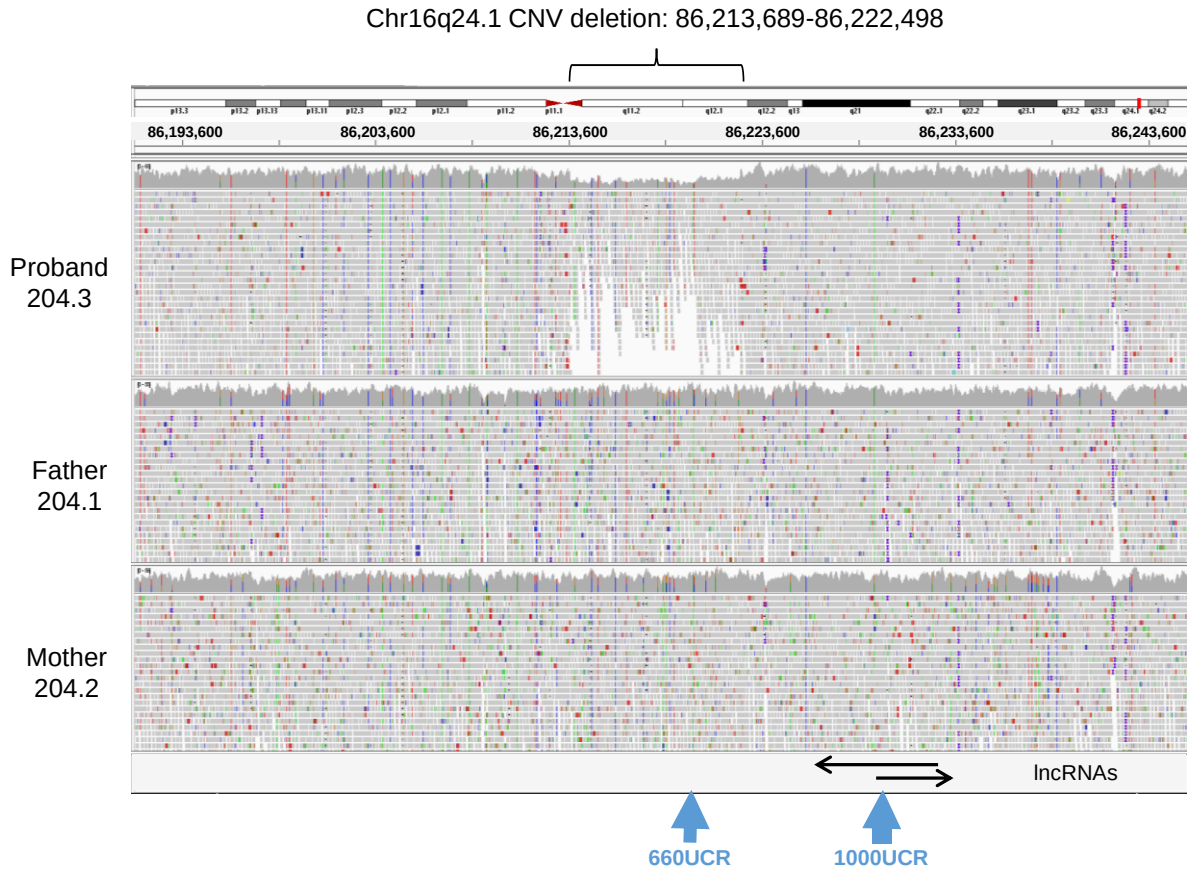

Figure S3. Genome sequencing reads of family trio (ACDMPV case 204) showing a *de novo* 8.8 kb CNV deletion (chr16:86,213,689-86,222,498) within the centromeric portion (Unit 1) of the *FOXF1* lung-specific enhancer.

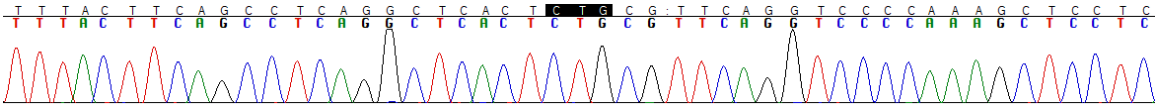

Figure S4. Chromatogram showing the sequence of the chr16q24.1 deletion junction. Microhomology region containing deletion breakpoints is shown on black background.

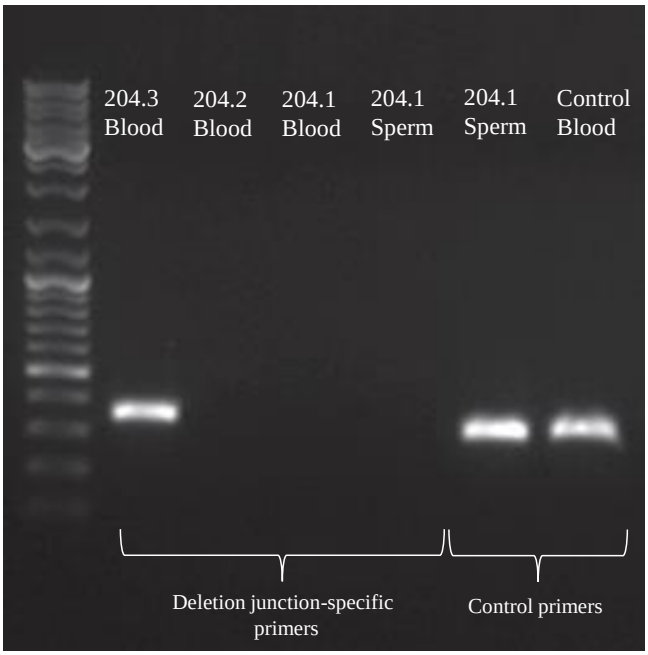

Figure S5. Agarose gel image of the deletion junction-specific PCR in the family 204 trio. Deletion specific amplicon is clearly visible in the proband (204.3) but not in her mother (204.2) or father (204.1), indicating that there is no evidence of parental mosaicism, and the deletion arose *de novo*.

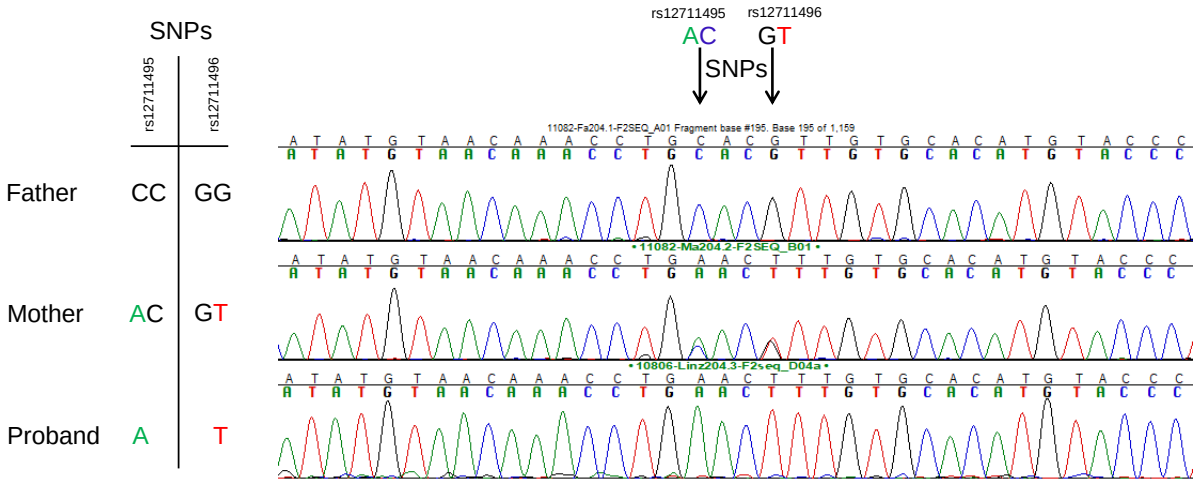

Figure S6. Sequence chromatograms showing parental origin of the deletion. The segregation of informative SNPs (black arrows) indicated that the deletion arose on chromosome 16 inherited from the proband's father.

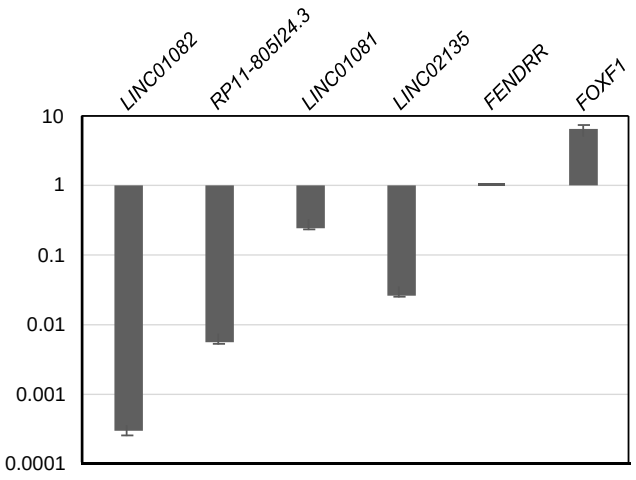

Figure S7. Expression levels in IMR-90 fetal lung fibroblasts of lncRNAs encoded within the *FOXF1* enhancer compared to *FENDRR* (*P*<0.001).

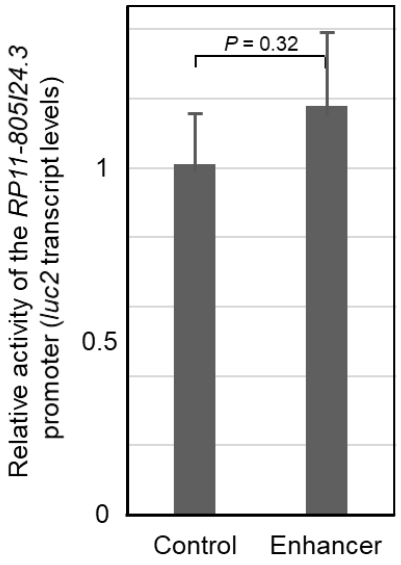

Figure S8. Luciferase reporter assay showing the lack of significant regulation of the *RP11-805I24.3* promoter by the 660UCR interval of the *FOXF1* enhancer.

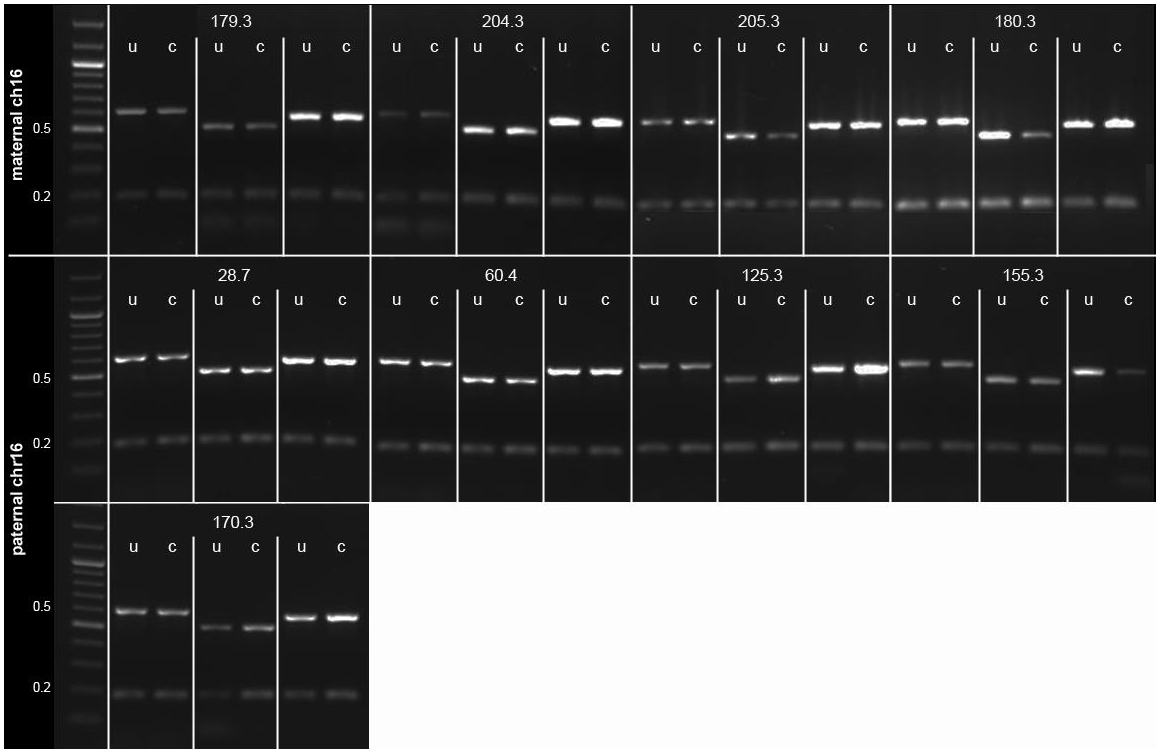

Figure S9. Analysis of the methylation status of 660UCR of *FOXF1* enhancer Unit 1, and two regions of the Unit 2: one overlapping with the 3' end of *LINC01081* and the other located within TBX4-binding region, respectively. DNA was from lungs (pts 179.3, 28.3, 60.4, 125.3, 155.3, 170.3), blood (pts 204.3, 205.3) and umbilical cord (pt 180.3). PCR was done using undigested (u) and digested (c) DNA.