Supplements

Bi-allelic loss-of-function variants in *PLD1* cause congenital right-sided cardiac valve defects and neonatal cardiomyopathy

Najim Lahrouchi et al.

Table of contents

Page 1	Overview of Supplementary Tables
Page 2	Overview of Supplementary Videos
Page 3-9	Methods
Page 10-13	**Supplementary Note 1:** Detailed description of the clinical phenotypes in PLD1 families
Page 14	**Supplementary Note 2:** Endocardial cell epithelial to mesenchymal transformation in AV cushion explants
Page 15	**Supplementary Figure 1:** cDNA analysis of c.3000+2T>A *PLD1* splicing variant
Page 16	**Supplementary Figure 2:** Principal component analysis to define ancestry of cases homozygous for I668P in *PLD1*
Page 17	**Supplementary Figure 3:** Western blot of PLD1 wild-type, inactive, and representative missense proteins overexpressed in HEK293T cells
Page 18	**Supplemental Figure 4.** Representative experiment showing PMA-stimulated activation of overexpressed wild-type and inactive PLD1
Page 19	**Supplementary Figure 5.** Placement of homozygous *PLD1* missense variants in gnomAD
Page 20	**Supplementary Figure 6.** Schematic of in vitro AV cushion collagen gel assay
Page 21	**Supplementary Figure 7.** Comparison of in-silico predictions of *PLD1* variants found in cases and gnomAD controls
Page 22-23	**Supplementary Figure 8.** PLD1 inhibition does not affect proliferation or total number of cells on gel
Page 24-26	**References**

SUPPLEMENTARY TABLES (in Microsoft Excel© Workbook file)

Supplementary Table 1: Clinical characteristics of patients with bi-allelic *PLD1* variants

Supplementary Table 2: Overview and annotation of bi-allelic *PLD1* variants found in patients

Supplementary Table 3: Overview and annotation of homozygous *PLD1* variants found in gnomAD individuals

SUPPLEMENTARY VIDEOS (see attached files)

Supplementary Video 1-6: Echocardiography from a selection of congenital heart disease and cardiomyopathy patients

Supplementary Video 7-24: Placement of 18 missense variants on the PLD1 catalytic domain crystal structure with potential local structural effect of residue mutation. **Legend:** a-Helices, cyan; b-sheets, magenta. Pathogenic residue, yellow stick and sphere representations; hydrogen bond, black dashed line; cation-pi interaction, green dashed line; potential structure clashes from mutant residue, red polygons.

METHODS

Case recruitment

The index patient in **Family A** was enrolled at the Amsterdam UMC in Amsterdam, The Netherlands. **Family B** was identified among 75 unrelated patients with severe right-sided valvular congenital heart disease (i.e. Tricuspid atresia or stenosis, Ebstein's anomaly or pulmonary atresia) from the National Registry and DNA bank of congenital heart defects (CONCOR) in The Netherlands.

We subsequently used GeneMatcher[1] to identify other patients with bi-allelic variants in *PLD1* resulting in 13 additional families: **Family C**: Bambino Gesù Children's Hospital IRCCS, Rome, Italy; **Family D**: Department of Clinical Genetics, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus; **Family E**: Research Unit for Rare Diseases, Department of Pediatrics and Adolescent Medicine, 1st Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic; **Family F**: University of Minnesota, Department of Obstetrics, Gynecology & Women's Health, Minneapolis, US; **Family G, N and O** were recruited from the Department of Pediatric Cardiology, Hadassah, Hebrew University Medical Center, Jerusalem, Israel. **Family P** was recruited from the Medical Genetics Department, UPMC Children's Hospital of Pittsburgh, Pittsburgh, Pensylvania, US. **Family Q**: Université Bourgogne-Franche Comté, Dijon, France. **Family R**: University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands. **Family S**: Division of Genetics, Birth Defects and Metabolic Disorders, Ann & Robert H. Lurie Children's Hospital of Chicago, US. **Family T**: Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands and **Family U**: Department of Pediatrics, Ochsner Clinic Foundation, New Orleans, La, US.

Families H-M were recruited to the Congenital Heart Disease Network Study of the Pediatric Cardiac Genomics Consortium (CHD GENES: ClinicalTrials.gov identifier NCT01196182).[2] Subjects were selected for structural CHD. Individuals with either an identified copy number variation (CNV) or chromosomal aneuploidy associated CHD were excluded. For all cases, cardiac diagnoses were acquired from review of all imaging and operative reports and entered as Fyler codes based on the International Paediatric and Congenital Cardiac Codes (http://www.ipccc.net/). All patients underwent evaluation at study entry using a standardized protocol consisting of an interview that includes maternal, paternal and birth history and whether the patient has been examined by a geneticist. An extensive review of the proband's medical record was performed including: height and weight, along with presence or absence of non-cardiac deformities, the availability and results of genetic testing and the presence or absence of a clinical genetic diagnosis. The study protocol was

approved by the local Institutional Review Board and written informed consent was obtained from the patient or their parents.

Next-generation sequencing

In **Family A** whole-genome sequencing was performed in the affected fetus using DNA extracted from chorionic tissue on the Illumina HiSeq X sequencer. FASTQ files were mapped with BWA[3] and calling of variants was conducted in GATK.[4] In the healthy parents, DNA was extracted from peripheral blood and whole-exome sequencing (WES) was carried out. The coding exons were captured using the Agilent SureSelect Human All Exon v.5. The captured fragments were sequenced on the Hiseq2000 sequencer (Illumina, San Diego, CA, USA) to an average depth of >50 reads per target base. The sequence reads were aligned to the human reference genome (UCSC NCBI37.1/hg19) using SOAPaligner (version 2.21)[5], and functional annotation of high-quality variants was performed using SOAPsnp software for the single nucleotide variation. For insertion/deletion detection, sequence reads were aligned by BWA and annotated by GATK for break-point identification. Data analysis was performed in the statistical programming environment R version 3.2.1. Synonymous variations not located at splice sites and variants with a frequency >0.1% in the following publicly available databases were excluded from further analysis using the Exome Aggregation Consortium database.[6] Furthermore, only variants compatible with a recessive pattern of inheritance (i.e. homozygous in the fetus, heterozygous in both parents and heterozygous or absent in the unaffected siblings) were retained for further consideration.

Family B was identified through cohort analysis of 75 patients with severe right-sided valvular congenital heart disease (e.g. tricuspid atresia or stenosis, Ebstein's anomaly or pulmonary atresia) from the National Registry and DNA bank of congenital heart defects (CONCOR). WES and data-analyses were performed as described for **Family A)**.

In **Family C,** clinical trio-exome analysis was performed on fetal DNA extracted from amniotic fluid and genomic DNA extracted from circulating leukocytes of parents using Twist Custom Panel kit (Twist Bioscience) on an Illumina sequencing platform (NovaSeq6000, San Diego, CA). The BaseSpace pipeline and the TGex Knowledge Driven NGS Analysis software (LifeMap Sciences) were used for variant calling and annotation, respectively. Sequencing data were aligned to the hg19 human reference genome. The presence of rare variants (MAF<0.01) involved in heart disease has been evaluated. The variants were filtered and analyzed in silico by using Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotyping v2 (PolyPhen-2) and Mutation Taster for the prediction of deleterious non-synonymous SNVs for human diseases.

In **Family D**, WES was performed (at the Genome Diagnostics Nijmegen, Radboud University Medical Centre, The Netherlands as an outsourced service), on the Illumina NextSeq 500 or NovaSeq. Read alignment (BWA) and variant calling (GATK) were done after exome enrichment (Agilent SureSelectQXT Human All Exon). Variant annotation, selection and prioritizing for pathogenicity was done using an in-house developed strategy. If no causative variants were detected using the gene panel(s), the 'full' dataset was analysed; from a gene panel solely based on frequency. Previous panels were analyzed for missense (PhyloP of base >3.5) and loss-of-function variants. Only those variants that were likely to be causative (by literature, expression profile etc.) were reported. Confirmation of reported variants by an independent technique (such as Sanger sequencing) was performed for low-quality variants (GATK quality scores). This has been validated to have a >99.9% reliability of the non-confirmed variants to be present and to be 'de novo' (if applicable).

In **Family E,** WES was performed in three affected fetuses using DNA extracted from paraffin embedded tissues (II.1, II.3) or cultured chorionic villi (II.5). In the healthy parents, DNA was extracted from peripheral blood and WES was carried out. For DNA enrichment, barcoded DNA libraries and NimbleGen SeqCap EZ Exome v3 (Roche, Madison, USA) were used according to the manufacturer's protocol. DNA sequencing was performed on the captured barcoded DNA library using the Illumina Hiseq 1500 system at the Genomic facility in Motol University Hospital in Prague. The resulting FASTQ files were aligned to the Human Genome Reference (hg19) using Novoalign (3.02.10). Following genome alignment, conversion of SAM format to BAM and duplicate removal were performed using Picard Tools (1.129). The Genome Analysis Toolkit, GATK (3.3)[7] was used for local realignment around indels, base recalibration and variant recalibration and genotyping. Variant annotation was performed with SnpEff[8] and GEMINI[9]. Only the sequence variants present in all three affected fetuses and having a frequency lower than 0.05 in the dbSNP, 1000 Genomes, Exome Variant Server (http://evs.gs.washington.edu/EVS/), and our internal exome database were prioritized for further analysis. Identified genetic variants were filtered according to the expected autosomal recessive model of the disease and evaluated according to the biological relevance of corresponding genes. Candidate variants were visualized in Integrative Genomics Viewer (IGV) version 2.3.32.[10] Selected candidate genetic variants were genotyped in DNA extracted from paraffin embedded tissues obtained from the fetus II.4 and validated in all DNA samples subjected to WES by direct Sanger sequencing using the version 3.1 Dye Terminator cycle sequencing kit (ThermoFisher Scientific) with electrophoresis on an ABI 3500XL Avant Genetic Analyzer (ThermoFisher Scientific). Data were analysed using Sequencing Analysis software (ThermoFisher Scientific).

In **Family F**, genomic DNA was isolated from patient II:3, DNA was barcoded, and enriched for the coding exons of targeted genes using hybrid capture technology. Prepared DNA libraries were then sequenced using a next generation sequencing technology. Following alignment, variants were detected in regions of at least 10x coverage. For this specimen of patient II:3, 100% and 100% of coding regions and splicing junctions of genes have been sequenced with coverage of at least 10x and 20x respectively or by Sanger sequencing. The remaining regions did not have 10x coverage, and were not evaluated. Variants were interpreted manually using locus specific databases, literature searches, and other molecular biological principles. All the variants with quality score less than 500 (roughly 40X of coverage for a heterozygous variant) were confirmed by Sanger sequencing. Only variants classified as pathogenic, likely-pathogenic, or unknown significance which are thought to be related to the patient's phenotype or test indication were considered.

In **Family G**, trio WES was pursued on DNA extracted from whole blood of the patient and his parents. Exonic sequences from DNA were enriched with the SureSelect Human All Exon 50 Mb V5 Kit (Agilent Technologies). Sequences were generated on a HiSeq2500 (Illumina) as 125-bp paired-end runs. Read alignment and variant calling were performed with DNA nexus using default parameters with the human genome assembly hg19 (GRCh37) as reference. Exome analysis yielded 48.9-60.6 million mapped reads with a mean coverage of 80-97×. Following alignment to the reference genome (hg19) and variant calling, variants were filtered out if the total read depth was less than 8× and if they were off-target (>8 bp from splice junction), were synonymous, or had minor allele frequency (MAF) > 0.5% in the gnomAD database (Genome Aggregation Database, Cambridge, MA).[11]

In **Families H-M** genomic DNA from venous blood or saliva was captured using the Nimblegen v.2 exome capture reagent (Roche) or Nimblegen SeqxCap EZ MedExome Target Enrichment Kit (Roche) followed by Illumina DNA sequencing as previously described.[12] WES data for congenital heart disease trios from the Pediatric Cardiac Genetics Consortium (PCGC) were obtained through dbGaP (Study Accession: phs001194.v2.p2). This data initially came as an .SRA file which was converted to .fastq format using the dbGaP inhouse SRAToolkit: fastq-dump (https://github.com/ncbi/sra-tools). All fastq files were aligned to the hg19 human reference genome using BWA, resulting in .sam files.[3] All .sam files were then converted to .bam files using samtools.[13] All resulting .bam files were processed using the Genome Analysis Tool Kit (GATK, version 4.1.4.1.) version and underwent INDEL realignment and base recalibration.[4] Resulting files were combined and jointly genotyped producing GVCF files with the GATK haplotype caller. Variants found were filtered using GATK variant recalibrator, reducing potential false positives calls in the dataset. Tranche values were set to 99.5 and

99.0 for SNPs and INDELs, respectively. Final filtered variant calls were annotated using ANNOVAR.[14] The entire pipeline was parallelized over thousands of computer cores through the Cartesius server of SURFSara.

In **Families N-O** exonic sequences from DNA samples of all available patients in both families (i.e. A-II-3, A-II-4, B-II-1 and B-II-3) were enriched with the SureSelect Human All Exon 50 Mb V.4 Kit (Agilent Technologies, Santa Clara, California, USA), as previously described.[15] Sequences (100-bp paired-end) were generated on a HiSeq2500 (Illumina, San Diego, California, USA). Read alignment and variant calling were performed with DNAnexus (Palo Alto, California, USA) using default parameters with the human genome assembly hg19 (GRCh37) as reference.

In **Family P**, the DNA was enriched for the complete coding regions and splice site junction for most genes of the human genomes using a proprietary capture system developed by GeneDx for next-generation sequencing with CNV calling (NGS-CNV). The enriched targets were simultaneously sequenced with paired-end reads on an Illumina platform. Bi-directional sequence reads were assembled and aligned to reference sequences based on NCBI RefSeq transcripts and human genome build GRCh37/UCSC hg19. The mean depth coverage was 82X and 98.6% of the targeted region was covered by at least 10 sequence reads (10x coverage). Using a custom developed analysis tool (XomeAnalyzer), data were filtered and analyzed to identify sequence variants and most deletions and duplications involving three or more coding exons.[16] Smaller deletions or duplications may not be reliably identified. Reported clinically significant variants were confirmed by an appropriate orthogonal method in the proband and, if submitted, in selected relatives as necessary. Sequence and copy number variants are reported according to the Human Genome Variation Society (HGVS) or International System for Human Cytogenetic Nomenclature (ISCN) guidelines, respectively. Reportable variants include pathogenic variants, likely pathogenic variants and variants of uncertain significance. Likely benign and benign variants, if present, are not routinely reported.

In **Family Q**, WES was conducted on case II-4. The exome was captured using the Twist Human Core Exome Kit (Twist bioscience, South San Francisco, CA, USA) and sequenced on the NovaSeq 6000. Segregation analysis we performed using Sanger sequencing.

In **Family R**, WES was performed as described for Family S (see below). A virtual gene panel of monogenic diseases based on approximately 3,850 genes from the Clinical Genomics Database (CGD) database and Online Mendelian Inheritance in Man (OMIM) were filtered with omission of genes associated with late-onset diseases, as described previously.[17–19]

Variants in this gene panel were analyzed using GAVIN (Gene-Aware Variant Interpretation), minor allele frequencies from gnomAD, and subsequent annotation with OMIM terms, Combined Annotation Dependent Depletion (CADD) scores, and reported modes of inheritance using MOLGENIS.[20–22]

In **Family S**, WES trio with mitochondrial genome sequencing was performed on the proband and unaffected parents using DNA extracted from peripheral blood samples following informed consent. Testing was performed at GeneDx (Gaithersburg, Maryland, USA), a CLIA certified commercial laboratory in the United States. The DNA was enriched for the complete coding regions and splice site junctions for most genes using a proprietary capture system developed by GeneDx for next generation sequencing with CNV calling (NGS-CNV). The enriched targets were simultaneously sequenced with paired-end reads on an Illumina platform, then assembled and aligned to reference sequences based on NCBI RefSeq transcripts and human genome build GRCh37/UCSD hg19. Data analysis was performed using the latest version of a custom-developed analysis tool by GeneDx, XomeAnalyzer. The general assertion criteria for variant classification are publicly available on the GeneDx.

In **Family T**, WES was performed with a parent–offspring trio approach as previously described.[23] Briefly, the exome was captured using the Agilent SureSelect XT Human All Exon V6 kit (Agilent, Santa Clara, CA, USA). Exome libraries were sequenced on an Illumina NextSeq500 instrument (Illumina, San Diego, CA, USA) with 151 bp paired-end reads at an average coverage of 100x and with >95% of the exome covered >20x. The mean coverage of the *PLD1* gene was 123x. Sequence reads were aligned to the hg19 reference genome using BWA version 0.7.5a. Variants were subsequently called by the Genome Analysis Toolkit Haplotype Caller, version 3.7 and annotated using a custom diagnostic annotation pipeline. Then variants were filtered for rarity excluding those with a minor allele frequency (MAF) >1% for heterozygous variants in an autosomal dominant inheritance model and those with MAF >2% for homozygous or compound heterozygous variants in an autosomal recessive inheritance model. After exclusion of (likely) pathogenic variants in 315 known cardiomyopathy-related genes, we selected for homozygous or compound heterozygous variants. Variant pathogenicity was assessed using data from the Agilent Alissa clinical informatics platform (Cartagenia, Leuven, Belgium) and Alamut Visual software (Interactive Biosoftware, Rouen, France). Literature and allele frequencies from an in-house population database were also used for pathogenicity determinations.

In **Family U** whole-exome sequencing was performed using genomic DNA from the patient and his parents. The exonic regions and flanking splice junctions of the genome were

sequenced by massively parallel (NextGen) sequencing on an Illumina sequencing system with 100bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19, and analyzed for sequence variants using a custom-developed analysis tool (Xome Analyzer). Capillary sequencing or another appropriate method was used to confirm all potentially pathogenic variants identified in this individual and relative samples, if submitted. Sequence alterations were reported according to the Human Genome Variation Society (HGVS) nomenclature guidelines.

Genetic variants were submitted to ClinVar (www.ncbi.nlm.nih.gov/clinvar/).

SUPPLEMENTARY NOTE 1 | Detailed description of the clinical phenotypes in a subset of *PLD1* patients

Family A

The index family (**Family A**) presented with a live born child with severe righted sided cardiac anomalies (II-1) and a fetus with a very similar right sided congenital heart disease (II-5) in addition to one miscarriage that had occurred before 16 weeks' gestation (II-2) (**Figure 1a**). The first-born female baby (II-1) had an Ebstein's anomaly with severe tricuspid regurgitation and pulmonary atresia diagnosed at 6 months of age. Her right ventricle was severely hypoplastic with 'paper-thin'-walls and the right atrium was grossly dilated suggestive for Uhl's disease. She died at the age of 11 months shortly after the first palliative surgery towards a Fontan circulation. The severely affected fetus (II-5, also a female) has already been extensively described in another publication (see ref[37]), and was diagnosed at 18 weeks gestation with pulmonary atresia, congenitally unguarded tricuspid orifice and Uhl's anomaly of the right ventricle (**Figure 1C**). The right ventricle was dilated. The tricuspid valve was severely dysplastic with only a rudimentary valvular remnant of the anterior leaflet and no papillary muscles. Free tricuspid regurgitation could be seen on color Doppler examination. The right ventricular wall was extremely thin in keeping with Uhl's disease and the systolic function was poor.[38] The pulmonary valve was dysplastic and atretic and the pulmonary arteries were small. The left side of the heart was structurally normal and there was a small apical ventricular septal defect (VSD) present. A mild mitral valve regurgitation was evident. Because of the fetus' poor prognosis the parents decided to terminate the pregnancy at 18 weeks of gestation. Post-mortem examination of the fetus confirmed the echocardiographic findings (**Figure 1D**).

Family B

The index patient in Family B (**Figure 1a**) was born with tricuspid atresia, atrial septal defect (ASD), VSD and hypoplasia of the pulmonary artery.

Family C

The fetus was the only child of healthy non-consanguineous parents from Bangladesh. The pregnancy was complicated at 19 weeks of gestation by the detection of an apparently isolated congenital heart defect (overriding aorta and ventricular septal defect) during an ultrasound exam. A fetal echocardiogram confirmed the congenital heart defect, including also pulmonary valve stenosis and right ventricular hypertrophy. These findings suggested the diagnosis of a severe form of tetralogy of Fallot.

Family E

In Family E, three spontaneous miscarriages (II-1, II-2, II-3) had occurred at respectively 19, 9 and 9 weeks of gestation. Subsequently two pregnancies (II-4, II-5) were terminated because of cardiac defects diagnosed on prenatal screening (**Figure 1**). In both cases the parents chose to terminate the pregnancies. Stored video images of the fetal echocardiograms were reviewed. In the first fetus (II-4) the echocardiogram was performed at 14 weeks' gestation and it showed a normal atrial situs, levocardia with atrioventricular and ventricular arterial concordance and cardiomegaly. The right ventricular apex appeared hypertrophic, the anterior wall was poorly contractile and the interventricular septum bulged towards the left. The atrioventricular valves were in the normal position and the tricuspid valve motion was reduced. Free tricuspid regurgitation was seen on color Doppler with continuous right to left shunting over the oval foramen. The right atrium was dilated. Retrograde flow was seen in the ductus arteriosus with free pulmonary regurgitation consistent with a functional pulmonary atresia due to poor right ventricular function. The color Doppler images showed evidence for possible coronary artery to right ventricular fistulas. The left ventricle and mitral valve appeared normal. There was a small amount of pericardial fluid seen. In the second fetus (II-5) the echocardiogram was performed at 13 weeks' gestation. There was a normal situs, levocardia, atrioventricular concordance and a small pericardial effusion. Here again a hypertrophied right ventricle was seen with free tricuspid regurgitation and continuous right to left shunting over the oval fossa observed on the color Doppler. The offsetting between the tricuspid and mitral valve was possibly a little more than expected (Ebsteins'-like). An apical muscular VSD and evidence for possible coronary artery to right ventricular fistulas were present. Unfortunately, no images of the pulmonary valve or arterial duct flow were available for evaluation.

Family F

Family F consisted of a triplet affected by a spectrum of congenital heart disease (**Figure 1**). The initial diagnosis of congenital heart disease was made during fetal echocardiography performed at 20 weeks of gestation. Infant A had a normal heart. Infant B had a bicuspid stenotic pulmonary valve, VSD, ASD, and bilateral superior vena cava. The patient underwent balloon dilation of the pulmonary valve at the age of four months and is now growing and developing normally. Infant C had tricuspid atresia, absent pulmonary valve, right ventricular hypoplasia due to right ventricular septal mass, left ventricular non-compaction cardiomyopathy and dual atrioventricular node physiology, in addition to multi-factorial pulmonary hypertension. He underwent a bidirectional Glenn palliation, atrial septectomy and ligation of a persistent ductus arteriosus (PDA).

Family P

This was a full-term male prenatally diagnosed with complex congenital heart disease, postnatally characterized by tricuspid atresia, hypoplastic right heart syndrome (HRHS), absent pulmonary valve, right ventricular outflow tract aneurysm, restrictive ASD and tortuous PDA. Congenital heart disease was diagnosed at 18 weeks gestation. He developed an ectopic atrial tachycardia at 31 weeks gestation successfully treated with maternal flecainide and digoxin therapy. He was born by induced vaginal delivery and therapy with prostaglandin E was started immediately. He required intubation in the delivery room for apnea as well as milrinone, digoxin, epinephrine and vasopressin. Ultimately ECMO support was initiated at 5 days of life due to persistent lactic acidosis and hypoperfusion. His postnatal course was complicated by a persistent *S. aureus* bacteremia with necrotizing pneumonia, right temporal lobe intraparenchymal hemorrhagic stroke, left tension pneumothorax, anasarca, acute kidney injury with continuous renal replacement therapy (CCRT), cavitary left lung lesions, thick secretions and a large right hemithorax clot. Bronchoscopic removal of the hemithorax clot was complicated by hemorrhage due to the friable infected lung tissues and his clinical condition deteriorated. He then developed a systemic-inflammatory-response-syndrome (SIRS)-like response and care was withdrawn at 23 days of age. Autopsy was declined by the family.

Family R

This fetus presented with hydrops fetalis at 19+6 weeks gestation. Fetal ultrasound showed a muscular tricuspid atresia and pulmonary atresia with a hypoplastic and thin walled right ventricle (HRHS). Additionally, a narrow ascending aorta and a wide aortic arch and descending aorta was observed. Left ventricular function was severly impaired and echocardiographic appearance was similar to endocardial fibroelastosis. Due to poor prognosis and development of mirror syndrome in the mother, the pregnancy was terminated at 22+1 weeks gestation. Degenerative changes with fibrosis and calcification were seen on the left ventricular microscopic sections.

Family S

This male patient was born to a healthy mother (20 years of age) at the gestational age (GA) of 39-week with a birth weight 3325 grams. The infant was noted prenatally on the 20-week GA fetal ultrasound to have a complex cardiac defect consistent with Ebstein's anomaly and pulmonary atresia. Amniocentesis was performed at 22 weeks GA for chromosome analysis and prenatal SNP array; both were normal. Upon delivery, he required CPAP and intermittent PPV with FiO2 titrated per NRP. He was then weaned to room air and initially stable with continuous PGE infusion for maintenance of his ductus arteriosus. Initial postnatal

echocardiogram on day of life (DOL) 1 confirmed Ebstein's anomaly with severely dilated right atrium and nearly absent right ventricle, pulmonary atresia, trabeculated left ventricle with depressed systolic function, and a large ASD. His neonatal course was further complicated by severe anemia requiring packed red blood cell transfusion and central line-associated right lower extremity thrombosis. Due to worsening heart function, he was listed for cardiac transplantation with A1 status on DOL 12 and underwent ductal stent placement on DOL 20. On DOL 45, he underwent an ABO-incompatible orthotopic heart transplantation with no significant operative complications, and discharged home two weeks post-transplantation. His post-transplantation course has been complicated by CMV infection with persistent low viral count. His cardiac function remains stable with no evidence of rejection on cellcept and tacrolimus. The patient has been making good developmental progress and at 15 months is crawling, pulling to stand, and taking a few independent steps. He says several single words in both English and Spanish.

The patient's family history is notable for a deceased sister with diagnosis of Ebstein's anomaly, who deceased on DOL10, and maternal grandmother who died of cancer diagnosed in her 30s or 40s. Genetic testing performed on the patient during the initial admission included chromosome microarray analysis and comprehensive cardiomyopathy gene panel with normal results. Subsequently, WES trio analysis with mitochondrial genome sequencing was obtained, which revealed compound heterozygous variants of unknown significance in the PLD1 gene (c.1219C>T/c.2338G>A) with the mother carrying the c.1219C>T variant and the father carrying the c.2338G>A variant.

Family T

The mother of the index patient from Family S presented at 33 weeks' gestation with persistent reduced fetal movements. Fetal ultrasound revealed hydrops fetalis due to severe cardiomyopathy and an emergency delivery by caesarean section was performed. No congenital heart defects were noted during echocardiography. Post-natal cardiac failure persisted despite medical treatment and the child died at the age of 3 days due to circulatory failure. There were no dysmorphic features. No post mortem examination was performed. The parents are healthy, of Dutch descent and non-consanguineous.

Family U

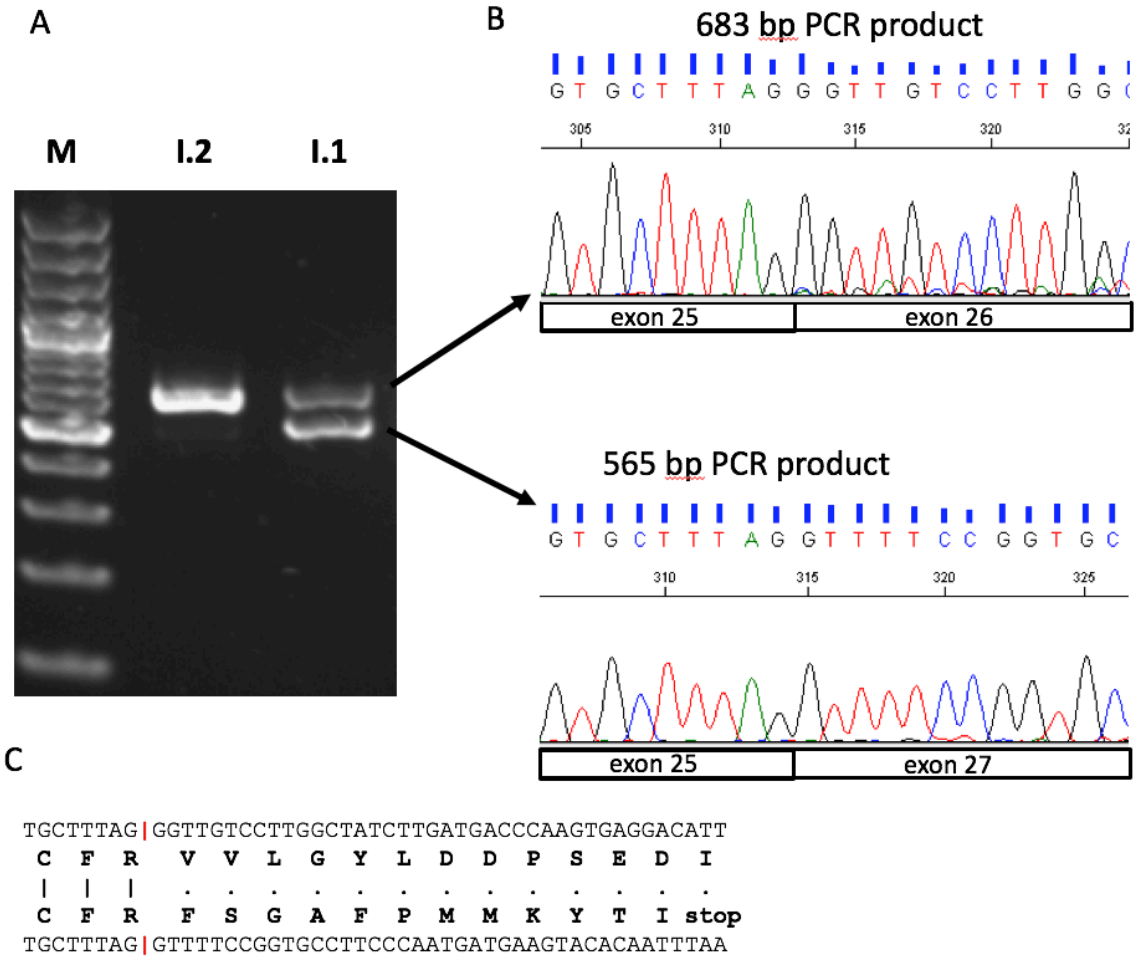
The index patient in Family T was a full-term infant (birth weight ~2.5kg) who developed difficulty breathing and bloody vomiting. The patient died at the age of 8 days. The autopsy showed patchy replacement of the myocardium by histiocytoid cells in a geographic distribution, with occasional plasma cells and eosinophils. This finding was most consistent with a diagnosis of histiocytoid cardiomyopathy.

SUPPLEMENTARY NOTE 2 | Endocardial cell epithelial to mesenchymal transformation in AV cushion explants

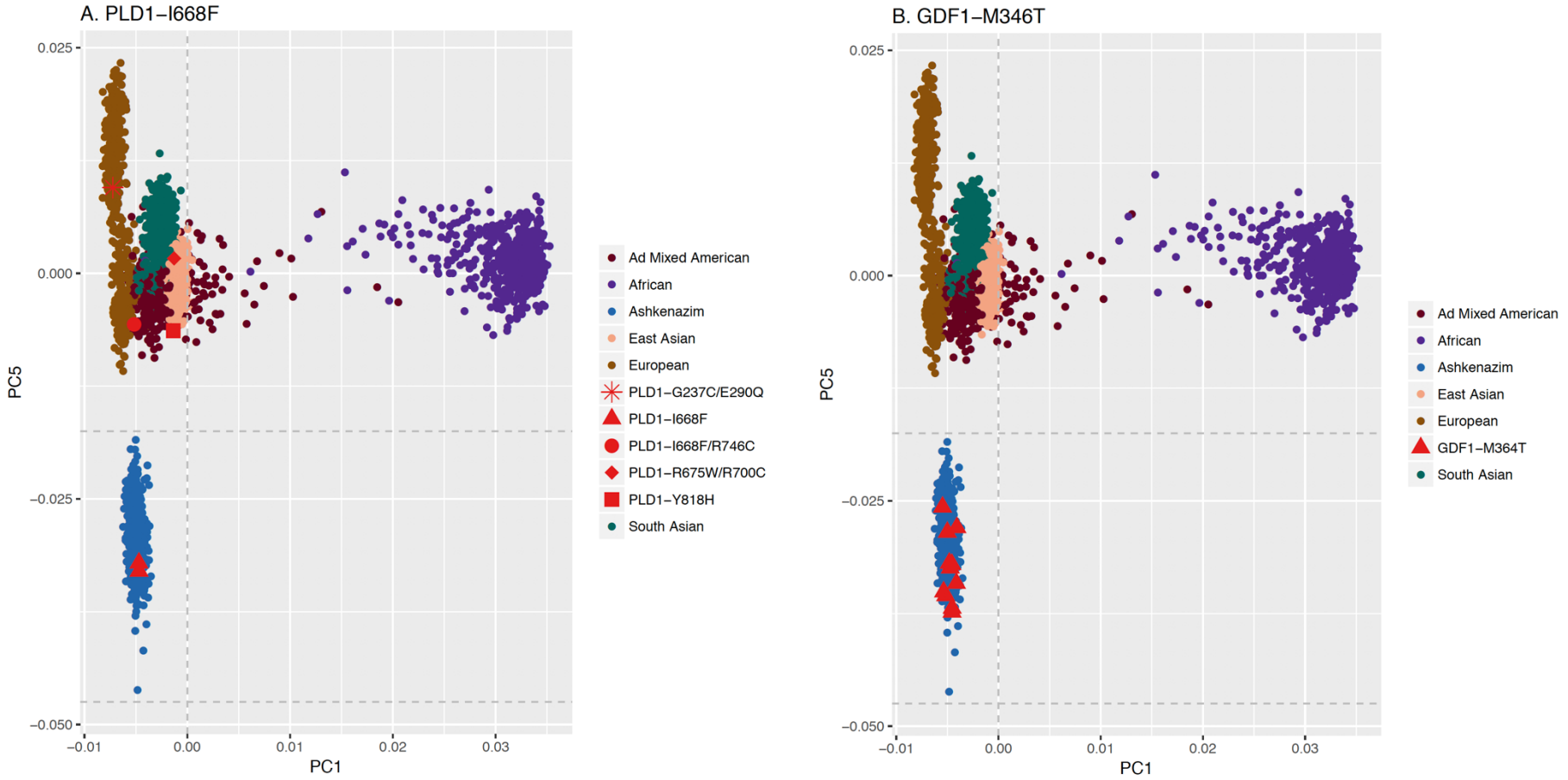
To address the possibility that the decrease in the number of cells in the gel is due to a difference in proliferation we assayed AVC incubated with PLD inhibitors for BrdU incorporation as described (**Supplemental Figure 8c**).[34] As previously reported, basal proliferation rates in AVC explants are low, which was observed in the DMSO control group. A significant decrease in proliferation was observed with PLD1 inhibition, yet this decrease of 1% is insufficient to account for the 80% decrease in the number of transformed cells previously observed. Therefore, the decrease in endocardial EMT is likely not due to differences in proliferation.

Since PLD can regulate cell survival via the mTOR pathway, we also examined the possibility that PLD inhibition results in a decrease in the number of cells on and in the gel, possibly due to cell death or apoptosis. An MTS assay of the endothelial and mesenchymal cells on and in the gel, as previously described[39], was undertaken to assay whether PLD inhibition resulted in a change in cell density (**Supplemental Figure 8d**). If the cells that no longer invade the gel as result of PLD1 inhibition undergo apoptosis or cell death, then the total number of cells on the gel will be lower and thus the cell density will be decreased. We incubated 15 explants with VU0359595 (PLDI -95), VU0155056 (PLD2), or DMSO for 48 hours, long enough to let the endothelial cells migrate from the explants and invade the gel. The myocardium was removed from each explant and the MTS assay was performed. Treatment with PLD1 or PLD2 inhibitor did not result in either a significant decrease or increase in cell density compared to DMSO treated controls (**Supplemental Figure 8d**). Thus, PLD1 treatment does not increase cell death or decrease proliferation.

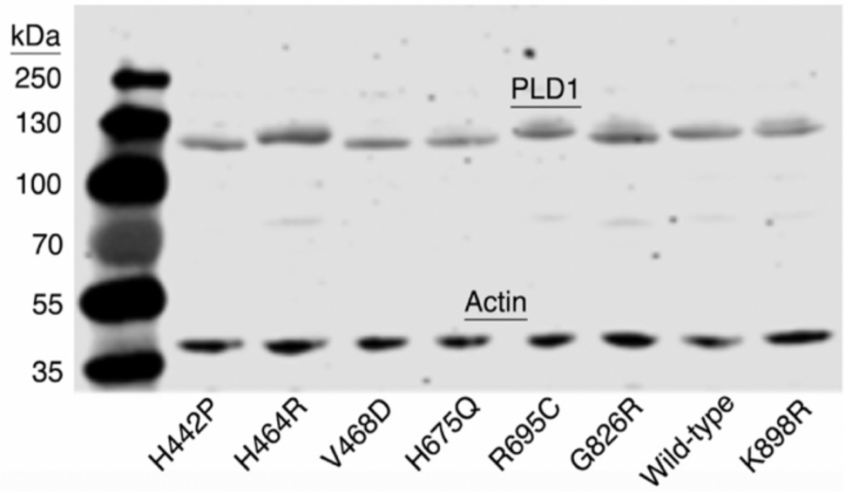
Supplementary Figure 1. cDNA analysis of c.3000+2T>A *PLD1* splicing variant. To assess the effect of the c.3000+2T>A variant on *PLD1* mRNA expression, splicing and stability, we isolated total RNA from peripheral mononuclear blood cells (PMBC) of both parents I.1 and I.2 and performed reverse transcription polymerase chain reaction (RT-PCR) analyses of the *PLD1* mRNA. The analysis of RT-PCR products revealed one PCR product of ~ 683 bp corresponding to that of predicted size of *PLD1* cDNA fragment in a non-carrier (I.2) (**panel A**). In the heterozygous carrier of the c.3000+2T>A variant (I.1) we obtained two PCR products. One PCR product was of a similar size of ~ 683 bp as in the non-carrier. The second PCR product was shorter in size than of the non-carrier. To identify the difference in the splicing pattern of *PLD1* in the carrier we isolated and Sanger sequenced both RT-PCR fragments (**panel B**). This analysis revealed that the larger PCR product corresponds to the expected *PLD1* cDNA fragment. The shorter PCR product of ~565 bp skipped exon 26 of *PLD1*. This deletion encodes for a truncated version of PLD1 that lacks 133 amino acids and contains a short neopeptide composed of 12 amino acids (FSGAFPMMKYTI) on its C-terminal end (NP_002653; p.V962FfsTer11).



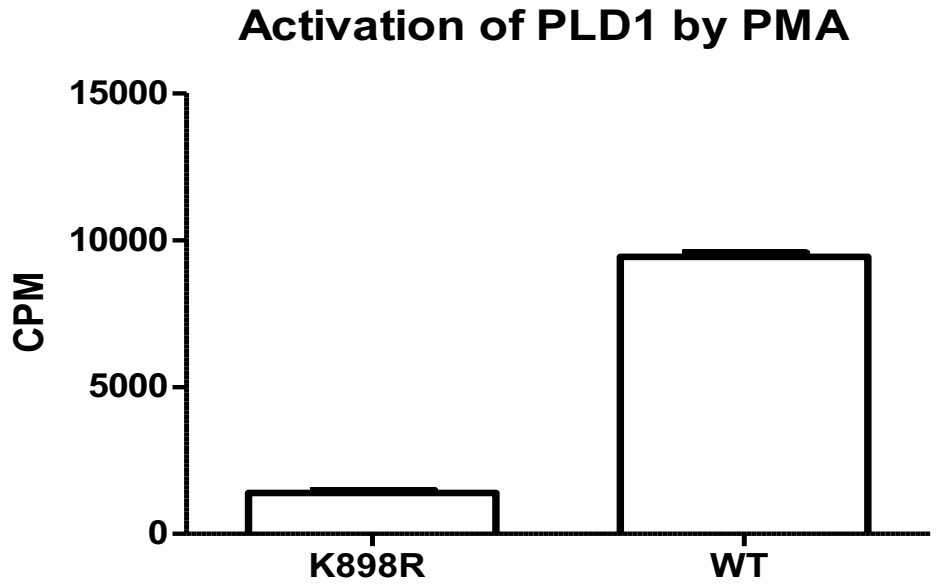
Supplementary Figure 2. Principal component analysis to define ancestry of cases homozygous for p.I668P in *PLD1*. The patients homozygous for p.I668P (defined as red triangles, panel A) cluster best to Ashkenazi Jews (AJS, defined in blue within the grey dashed lines, panel A). Previously identified ASJ patients that carried the GDF1-M346T[40] founder variant in the PCGC dataset (defined as red triangles, panel B) cluster best with AJS (defined in blue within the grey dashed lines, panel B). Principal components analyses were performed using the 1000 Genomes populations (Phase3v5) and the ASJ individuals from the Gene Expression Omnibus database[26] (accession GSE23636) as a reference.



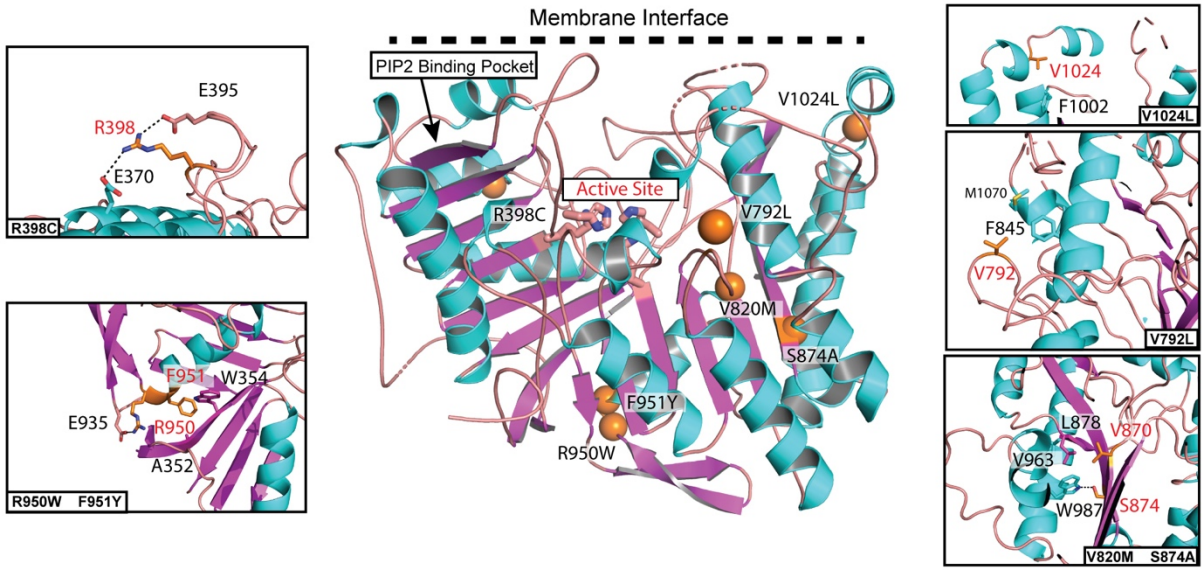
Supplementary Figure 3. Western blot of PLD1 wild-type, inactive, and representative missense proteins overexpressed in HEK293T cells.



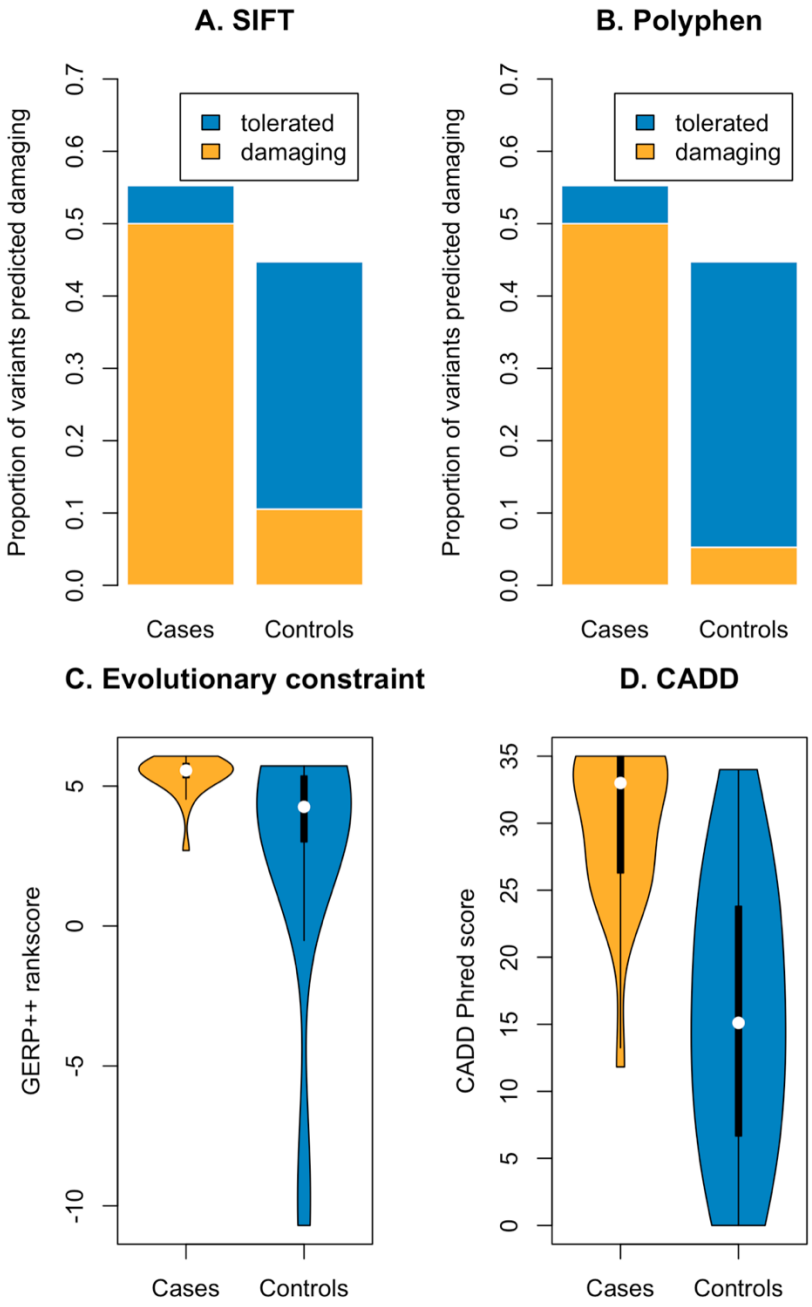
Supplementary Figure 4. Representative experiment showing PMA-stimulated activation of overexpressed wild-type and inactive PLD1. HEK293T cells transfected with wild-type and catalytically-inactive (K898R) HA-tagged human PLD1 and assayed in the presence of PMA, an activator of Protein Kinase C (PKC). ***, P < 0.001. Technical replicates were used for analysis in this representative experiment.



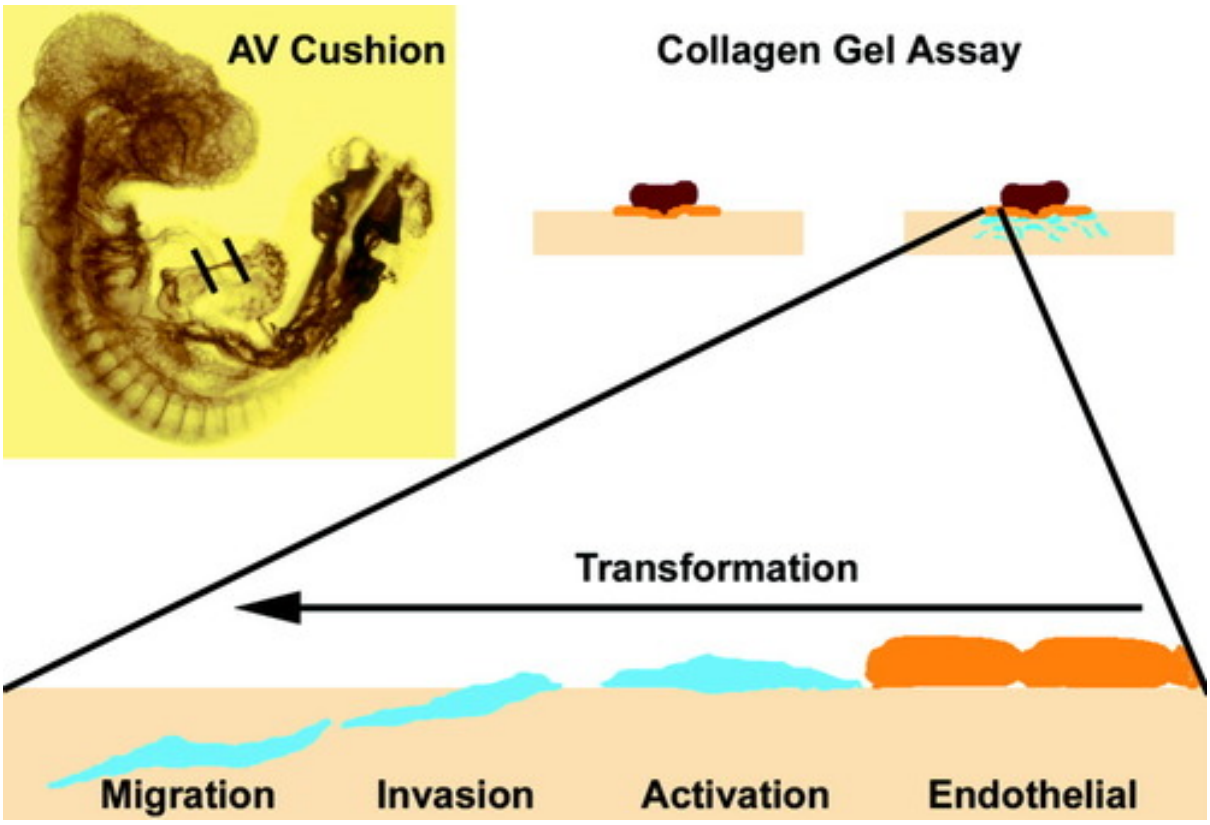
Supplementary Figure 5. Placement of homozygous PLD1 missense variants in gnomAD on the PLD1 structure. Orange spheres, homozygous missense variant found in gnomAD.[11]



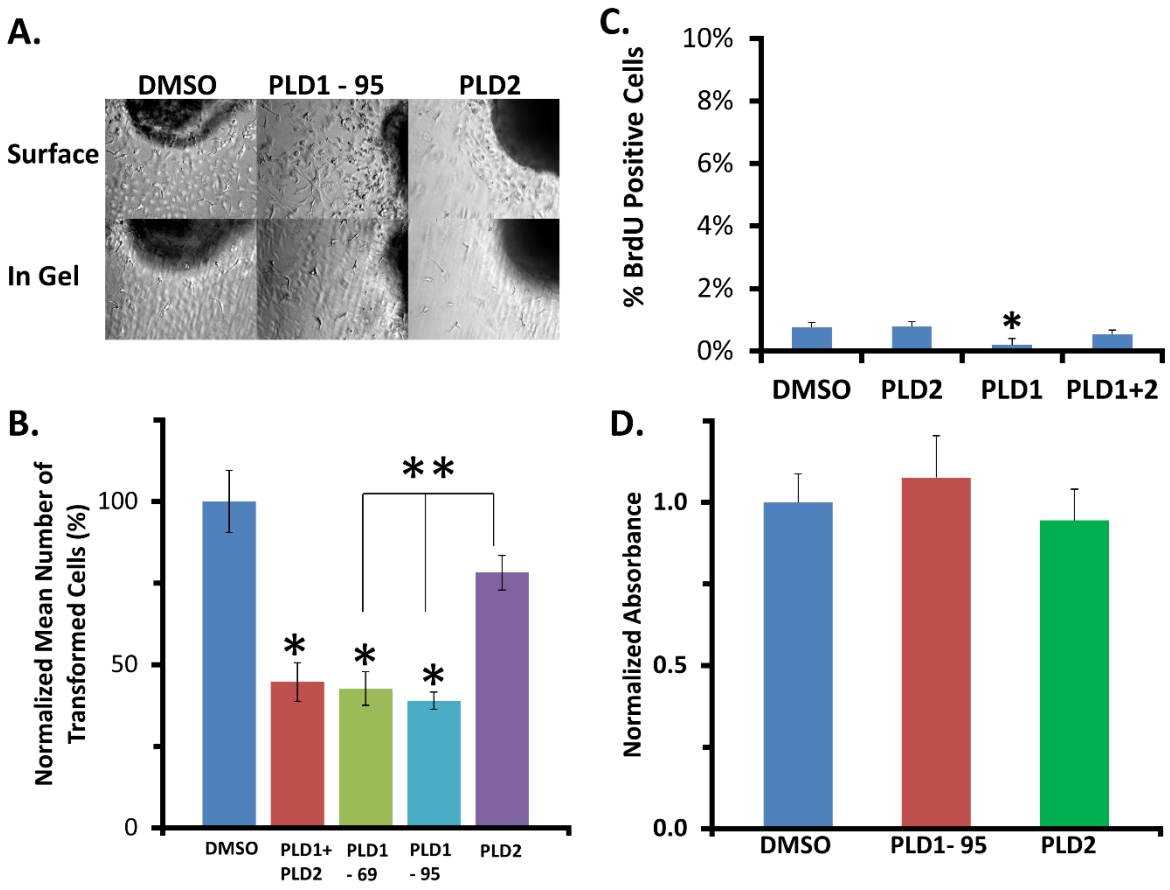
Supplementary Figure 6. Comparison of in-silico predictions of *PLD1* missense variants found in patients and gnomAD controls.[11] Statistical comparisons were performed using Fisher's exact test for categorical variables (SIFT and Polyphen, $P=3.7\times10^{-5}$ and 1.2×10^{-6}, respectively), the Mann-Whitney U test for non-normally distributed continuous variables (GERP++, $P= 7.6\times10^{-4}$) and the Student T-test for normally distributed continuous variables (CADD, $p=3.3\times10^{-5}$). All statistical analyses were performed using R (version 3.2.1), and a p-value of < 0.05 was considered statistically significant. **Abbreviations:** CADD: Combined Annotation Dependent Depletion, GERP: Genomic Evolutionary Rate Profiling, SIFT: Sorting Intolerant From Tolerant, Polyphen: Polymorphism Phenotyping.



Supplementary Figure 7. Schematic of *in vitro* AV cushion collagen gel assay. Stage 18 chick embryo with the AV cushion region demarcated is shown. AV cushion explants placed on a collagen gel form an endothelial cell monolayer and a portion of the endothelial cells undergo epithelial to mesenchymal transformation (EMT). Stages of transformation are indicated. Permission was obtained to reuse figure from.[41]



Supplementary Figure 8. PLD1 inhibition does not affect proliferation or total number of cells on gel. The effect of PLD activity on endocardial EMT was examined using a collagen gel assay. Hamburger-Hamilton (HH) stage 16 chick atrioventricular canal (AVC) cushion explants were incubated with small molecule inhibitors of PLD for 48 hours on a collagen gels. A lower dose of inhibitor, 500nM, yielded a similar result (**A-B**) and demonstrated that the percent inhibition was dose-dependent. (**A**) Representative photomicrographs of AVC explants from (**B**) on collagen gels incubated with small molecule inhibitors of PLD1 (VU0359595), PLD2 (VU0285655-1), or DMSO. (**B**) HH16 AVC explants were incubated with 500nM doses of small molecule inhibitors of PLD activity or DMSO (control, n=29). The number of cells undergoing EMT per explant were counted for each condition and normalized over 3 separate experiments to relevant control explants (100%). PLD1-PLD2 (VU0155056, n=30) inhibits both PLD1 and PLD2 activity. PLD1-69 (VU0155069, n=30) and PLD1-95 (VU0359595, n=30) selectively inhibit PLD1. PLD2 (VU0285655-1, n=30) selectively inhibits PLD2. * and ** - $p<0.01$ compared to control. (**C**) HH16 AVC cushion explants incubated with 5μM of PLD inhibitors were tested for BrdU incorporation to determine if changes in proliferation occurred after PLD inhibitor incubation. After incubating the AVC explants for 48 hours, the total number of BrdU positive cells per explant were counted and normalized to the total number of cells undergoing EMT per condition. n=18 for each condition, spread over 3 independent experiments. * - $p<0.01$ (**D**) HH16 AVC explants were incubated with 500nM doses of small molecule inhibitors of PLD activity for 48 hours, long enough to let the endothelial cells migrate from the explants and invade the gel, after which the myocardium was removed and an MTS assay performed on the remaining endothelial and mesenchymal cells on and in the gel to determine if the decrease in the number of cells undergoing EMT after PLD1 inhibition was the results of altered cell density. Three separate experiments were performed using 15 explants per condition per MTS assay and the mean absorbance was normalized to control (DMSO). No significant difference was observed between conditions. All p-values were calculated using Student's T-test and Bonferroni correction was used to correct for multiple testing.


A.
DMSO
PLD1 - 95
PLD2
Surface
In Gel
C.
% BrdU Positive Cells
10%
8%
6%
4%
2%
0%
*
DMSO
PLD2
PLD1
PLD1+2
B.
Normalized Mean Number of Transformed Cells (%)
100
50
0
**
*
*
*
DMSO
PLD1+ PLD2
PLD1 - 69
PLD1 - 95
PLD2
D.
Normalized Absorbance
1.0
0.5
0.0
DMSO
PLD1- 95
PLD2

REFERENCES

1. Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. *Hum. Mutat.* **36**, 928–930 (2015).
2. Gelb, B. *et al.* The Congenital Heart Disease Genetic Network Study. *Circ. Res.* **112**, 698–706 (2013).
3. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
4. Van der Auwera, G. A. *et al.* From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* (2013). doi:10.1002/0471250953.bi1110s43
5. Gu, S., Fang, L. & Xu, X. Using SOAPaligner for Short Reads Alignment. *Curr. Protoc. Bioinforma.* **44**, (2013).
6. Karczewski, K. J. *et al.* The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res.* **45**, D840–D845 (2017).
7. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–303 (2010).
8. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin).* **6**, 80–92 (2012).
9. Paila, U., Chapman, B. A., Kirchner, R., Quinlan, A. R. & Lander, E. GEMINI: Integrative Exploration of Genetic Variation and Genome Annotations. *PLoS Comput. Biol.* **9**, e1003153 (2013).
10. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–6 (2011).
11. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
12. Zaidi, S. *et al.* De novo mutations in histone-modifying genes in congenital heart disease. *Nature* **498**, 220–3 (2013).
13. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
14. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
15. Ta-Shma, A. *et al.* Congenital valvular defects associated with deleterious mutations in the PLD1 gene. *J. Med. Genet.* **54**, 278–286 (2017).
16. Retterer, K. *et al.* Assessing copy number from exome sequencing and exome array CGH based on CNV spectrum in a large clinical cohort. *Genet. Med.* **17**, 623–629 (2015).
17. van Diemen, C. C. *et al.* Rapid Targeted Genomics in Critically Ill Newborns. *Pediatrics* **140**, e20162854 (2017).

18. Solomon, B. D., Nguyen, A.-D., Bear, K. A. & Wolfsberg, T. G. Clinical genomic database. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 9851–5 (2013).
19. Online Mendelian Inheritance in Man (OMIM) , Morbid Map, (released April 1st 2014, downloaded from OMIM FTP.
20. van der Velde, K. J. *et al.* GAVIN: Gene-Aware Variant INterpretation for medical sequencing. *Genome Biol.* **18**, 6 (2017).
21. Karczewski, K. J. *et al.* Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* 531210 (2019). doi:10.1101/531210
22. Swertz, M. A. *et al.* The MOLGENIS toolkit: rapid prototyping of biosoftware at the push of a button. *BMC Bioinformatics* **11**, S12 (2010).
23. Herkert, J. C. *et al.* Toward an effective exome-based genetic testing strategy in pediatric dilated cardiomyopathy. *Genet. Med.* **20**, 1374–1386 (2018).
24. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
25. Belkadi, A. *et al.* Whole-exome sequencing to analyze population structure, parental inbreeding, and familial linkage. *Proc. Natl. Acad. Sci.* **113**, 6713–6718 (2016).
26. Bray, S. M. *et al.* Signatures of founder effects, admixture, and selection in the Ashkenazi Jewish population. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16222–7 (2010).
27. Gandolfo, L. C., Bahlo, M. & Speed, T. P. Dating Rare Mutations from Small Samples with Dense Marker Data. *Genetics* **197**, 1315–1327 (2014).
28. Morris, A. J., Frohman, M. A. & Engebrecht, J. Measurement of Phospholipase D Activity. *Anal. Biochem.* **252**, 1–9 (1997).
29. Altshuller, Y. M. *et al.* Human ADP-ribosylation Factor-activated Phosphatidylcholine-specific Phospholipase D Defines a New and Highly Conserved Gene Family. *J. Biol. Chem.* **270**, 29640–29643 (1995).
30. Leiros, I., Secundo, F., Zambonelli, C., Servi, S. & Hough, E. The first crystal structure of a phospholipase D. *Structure* **8**, 655–67 (2000).
31. Dixon, J. E. & Stuckey, J. A. Crystal structure of a phospholipase D family member. *Nat. Struct. Biol.* **6**, 278–284 (1999).
32. Sung, T.-C. Mutagenesis of phospholipase D defines a superfamily including a trans-Golgi viral protein required for poxvirus pathogenicity. *EMBO J.* **16**, 4519–4530 (1997).
33. DeLaughter, D. M., Saint-Jean, L., Baldwin, H. S. & Barnett, J. V. What chick and mouse models have taught us about the role of the endocardium in congenital heart disease. *Birth Defects Res. Part A Clin. Mol. Teratol.* **91**, 511–525 (2011).
34. Townsend, T. A. *et al.* Endocardial cell epithelial-mesenchymal transformation requires Type III TGFβ receptor interaction with GIPC. *Cell. Signal.* **24**, 247–56

(2012).

35. Townsend, T. A., Wrana, J. L., Davis, G. E. & Barnett, J. V. Transforming growth factor-beta-stimulated endocardial cell transformation is dependent on Par6c regulation of RhoA. *J. Biol. Chem.* **283**, 13834–41 (2008).
36. Townsend, T. A. *et al.* BMP-2 and TGFβ2 shared pathways regulate endocardial cell transformation. *Cells. Tissues. Organs* **194**, 1–12 (2011).
37. van der Palen, R. L. F., van der Wal, A. C., Robles de Medina, P. G., Blom, N. A. & Clur, S.-A. B. Uhl's anomaly: Clinical spectrum and pathophysiology. *Int. J. Cardiol.* **209**, 118–21 (2016).
38. UHL, H. S. M. A previously undescribed congenital malformation of the heart: almost total absence of the myocardium of the right ventricle. *Bull. Johns Hopkins Hosp.* **91**, 197–209 (1952).
39. Sánchez, N. S. *et al.* The cytoplasmic domain of TGFβR3 through its interaction with the scaffolding protein, GIPC, directs epicardial cell behavior. *Dev. Biol.* **358**, 331–43 (2011).
40. Jin, S. C. *et al.* Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. *Nat. Genet.* **49**, 1593–1601 (2017).
41. Barnett, J. V. & Desgrosellier, J. S. Early events in valvulogenesis: A signaling perspective. *Birth Defects Res. Part C Embryo Today Rev.* **69**, 58–72 (2003).