

## Supplemental Information

### METHODS

#### Clinical Exome Sequencing

DNA extracted from the infant's, maternal, and paternal blood samples was used for WES at the Center for Pediatric Genomic Medicine at Children's Mercy Hospital. Exome libraries were prepared using TruSeq (Illumina, San Diego, CA) and xGen exome enrichment (Integrated DNA Technologies [IDT], Coralville, IA) according to the manufacturers' standard protocols. Briefly, 250 ng of high-quality genomic DNA was sheared by Covaris sonication to an average size of 450 bp. DNA fragments underwent end repair, A tailing, adapter ligation, and associated AMPure bead cleanups on a Hamilton NGS STAR liquid handler (Hamilton Company). After adapter ligation, 10 cycles of standard polymerase chain reaction (PCR) were performed with KAPA HiFi master mix and primers specific to the ligated adapters, followed by a bead cleanup. The resulting libraries underwent quality assessment for appropriate concentration and final library size. Six libraries were pooled together at 1000 ng each for exome enrichment. Pooled libraries were blocked to prevent nonspecific binding and lyophilized with a vacuum concentrator. Pools were reconstituted with IDT enrichment buffer, standard exome baits, copy number variants baits, and custom mitochondrial baits. This exome enrichment reaction underwent a 4-hour hybridization, followed by incubation with streptavidin beads and serial washes with IDT enrichment wash buffers on a Sciclone liquid handler (PerkinElmer). An additional 10 cycles of standard PCR were performed with KAPA HiFi master mix and primers specific to the adapters, followed by a final bead

cleanup. The resulting enriched pools underwent quality assessment for appropriate concentration and final library size, as well as a TaqMan quantitative PCR assay to ensure successful enrichment of target regions before standard free adapter blocking (Illumina) was performed. Cleaned, adapter-blocked pools were loaded on a NovaSeq 6000 sequencing platform with a run configuration of  $151 \times 8 \times 8 \times 151$ . Resulting data were processed with bcl2fastq conversion software (Illumina). Sequence alignment and variant detection was performed with the Illumina DRAGEN Bio-IT Platform versions 3.02 to 3.6.3 (Illumina) against the GRCh37 reference genome.

#### Variant Annotation

Genetic variants with respect to hg19 were identified and imported into a proprietary annotation tool, Rapid Understanding of Nucleotide Variant Effect Software (RUNES). RUNES incorporated data from Ensembl Variant Effect Predictor software, produced comparisons to dbSNP, ClinVar, the Exome Aggregation Consortium, Genome Aggregation Database (<https://gnomad.broadinstitute.org>), and known disease variants from the Human Gene Mutation Database.<sup>19–21</sup> RUNES reported allele frequencies (minor allele frequency) derived from the Center for Pediatric Genomic Medicine Variant Warehouse Database. Potentially deleterious variants were identified using Variant Integration and Knowledge Interpretation in Genomes.<sup>19–21</sup> Because common variants (minor allele frequency >1%) are less likely to be functionally deleterious, our focus was on rare and novel variants.<sup>22</sup> Biological context and prediction of deleterious functional effects were estimated using PolyPhen version 2 (<http://genetics.bwh.harvard.edu/pph2>) and sorting intolerant from tolerant.<sup>23</sup>

#### Cell Culture

Primary HPMECs commercially purchased from ScienCell (Carlsbad, CA) were used to generate an immortalized line (HPMEC-Im) in our laboratory, as reported previously.<sup>10</sup> These cells express EC specification markers and uptake oxidized low-density lipoprotein.<sup>10</sup> Cells were grown using complete EC medium (ScienCell) in a humidified incubator containing 5% CO<sub>2</sub> at 37°C. HEK293 cells were purchased from ATCC (Manassas, VA) and grown according to the manufacturer's instructions.

#### Design of Wild-Type/Variant Construct and Transfection

Human EPHB4 cDNA C-His tag (Sino Biological Inc., USA) was cloned into pcDNA3.1-EGFP (Addgene, Cambridge, MA) between Hind3 and Xba1 sites to generate pcDNA3.1-EPHB4 C-His tag. EPHB43'-fragment was amplified by PCR with primers, 5'-GAGATCGATGTCCTACGTCAAGATTGAAGAGGT-3' and 5'-TGATCTAGATCAGTACTGCGGGGCCGGTCCTCCTGT-3', cut with Bsu15I and Xba1 (Thermo-Fisher) and cloned into the pcDNA3.1-EPHB4 C-His tag to generate pcDNA3.1-EPHB4. EphB4 point mutation was introduced by a 2-step PCR strategy to generate pcDNA3.1-EPHB4 (p.Ala700Thr). HPMEC-Im and HEK293 grown in 6-well tissue culture plates were transfected overnight with 2 µg of the indicated plasmids or empty plasmids (mock) with Lipofectamine 3000 (Thermo Fisher Scientific), according to the manufacturer's protocol. Cells were allowed to recover for 48 hours, and cell lysates were used for protein expression studies.

#### Immunoblotting for Quantifying Changes in Protein Expression

Immunoblotting (Bio-Rad, Hercules, CA) was done following the

manufacturer's standard protocol. Antibodies used were rabbit anti-phospho-EPHB4 (Tyr987, PA5-64792; Thermo Fisher Scientific), anti-PROX1 (#14963S; Cell Signaling Technology, Danvers, MA), anti-phospho p44/42 (ERK) (Thr202/Tyr204, #4370; Cell Signaling Technology), anti-p44/42 (ERK) (#4695, Cell Signaling Technology), goat anti-EPHB4 (AF3038; R&D Systems, Minneapolis, MN), and mouse anti- $\beta$ -actin (A1978; Sigma-Aldrich, St Louis, MO). Densitometry was performed using ImageJ software (National Institutes of Health, Bethesda, MD), and fold-change was normalized to  $\beta$ -actin or the corresponding nonphosphorylated antibody.

#### **Autopsy Histopathological Examination-Methods and Staining of Human Podoplanin and PROX1**

After consent, postmortem examination of the deceased infant was performed at 4 months of age. Evisceration and gross evaluation of all organs was accompanied by detailed histopathological analysis and postmortem microbiologic cultures of the spleen and lung. Histological sections of the lungs and intestines were also stained for PROX1 and podoplanin (D2-40), and the results were compared with postmortem control tissues from 2 other infants (of similar age) who had relatively healthy lungs and intestines.

#### **Immunohistochemistry**

Sections (4  $\mu$ m) from formalin-fixed paraffin-embedded tissue blocks were stained for D2-40 (BioLegend,

San Diego, CA) and PROX1 using automated IHC protocols. D2-40 is an in vitro diagnostic monoclonal antibody that reacts with membranous O-linked sialoglycoprotein and is used in 1:40 dilution to identify lymphatic endothelium. PROX1 monoclonal antibody (Cell Signaling Technology) is produced with a synthetic peptide corresponding to residues surrounding Glu567 of human PROX1 protein and is used at 1:200 dilution to recognize the nuclear protein in lymphatic endothelium. The staining was performed with a Leica Bond Max instrument (Leica Biosystems, Buffalo Grove, IL) that was previously validated for routine clinical diagnostic use. After antigen retrieval at high pH, deparaffinized tumor slides were sequentially incubated with the primary antibody, a horseradish peroxidase secondary antibody conjugated to a dextran polymer and a coloring reagent that yields a brown color for positive staining. Separate positive controls made from postmortem lung tissue from similar infants were simultaneously stained. External negative controls were stained similarly except for omission of primary antibodies.

#### **Parental consent**

Written consent was obtained from parents to perform all proposed studies.

#### **Statistical Analysis**

Data are presented as mean  $\pm$  SD. Experimental data from a minimum of 3 independent experiments with

adequate technical replicates was used for quantification. For all data,  $P < .05$  was considered significant, and we initially examined whether distribution of data was Gaussian by using the D'Agostino-Pearson omnibus normality test. Because all the data were normally distributed, one-way analysis of variance with a post hoc Tukey multiple comparisons test was used for analysis. For most analyses, fold changes were calculated relevant to expression/changes in untreated control patients. Statistical analysis was done using GraphPad Prism 7.0 (GraphPad Software, San Diego, CA).

#### **SUPPLEMENTAL REFERENCES**

19. Cummings L, Tucker M, Gibson M, et al. Rare genetic variants in immune genes and neonatal herpes simplex viral infections. *Pediatrics*. 2021;147(1):e20200687
20. Rajeev N, Soden S, Swanson D, Ciaccio CE, Kingsmore SF, Dinwiddie DL. Utility of next generation sequencing in clinical primary immunodeficiencies. *Curr Allergy Asthma Rep*. 2014; 14(10):468
21. Soden SE, Saunders CJ, Willig LK, et al. Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. *Sci Transl Med*. 2014;6(265):265ra168
22. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011;12(11):745–755
23. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc*. 2016;11(1):1–9