medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

PKHD1L1, A Gene Involved in the Stereociliary Coat, Causes Autosomal Recessive Nonsyndromic Hearing Loss

Shelby E. Redfield[1,+], Pedro De-la-Torre[2,3,+], Mina Zamani[4,5], Hina Khan[6], Tyler Morris[2], Gholamreza Shariati[5,7], Majid Karimi[8], Margaret A. Kenna[1, 2], Go Hun Seo[9], Sadaf Naz[6], Hamid Galehdari[4], Artur A. Indzhykulian[2,3,*,^], A. Eliot Shearer[1, 3,*,^], Barbara Vona[10,11,*,^]

1. Department of Otolaryngology and Communication Enhancement, Boston Children's Hospital, 300 Longwood Avenue, BCH-3129, Boston, MA 02115, USA

2. Mass Eye and Ear, Eaton Peabody Laboratories, Boston, Massachusetts, USA

3. Department of Otolaryngology Head and Neck Surgery, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA

4. Department of Biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran

5. Narges Medical Genetics and Prenatal Diagnosis Laboratory, Kianpars, Ahvaz, Iran

6. School of Biological Sciences, University of the Punjab, Quaid-e-Azam Campus, Lahore 54590, Pakistan

7. Department of Medical Genetics, Faculty of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

8. Khuzestan Cochlear Implantation Center (Tabassom), Ahvaz, Iran

9. 3billion, Inc., Seoul, South Korea

10. Institute of Human Genetics, University Medical Center Göttingen, 37073 Göttingen, Germany

11. Institute for Auditory Neuroscience and InnerEarLab, University Medical Center Göttingen, 37075 Göttingen, Germany

* *Corresponding authors*: Artur_Indzhykulian@hms.harvard.edu, Eliot.Shearer@childrens.harvard.edu, barbara.vona@med.uni-goettingen.de

\+ *Contributed equally to this work*

^ *Contributed equally to this work*

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

ORCID

Shelby Redfield	0000-0002-3485-6654
Pedro De-la-Torre	0000-0002-2434-3345
Mina Zamani	0000-0002-7005-3787
Sadaf Naz	0000-0002-1912-0235
Artur A. Indzhykulian	0000-0002-2076-6818
A. Eliot Shearer	0000-0002-5324-4805
Barbara Vona	0000-0002-6719-3447

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Identification of genes associated with nonsyndromic hearing loss is a crucial endeavor, given the substantial number of individuals who remain without a diagnosis after even the most advanced genetic testing. *PKHD1L1* was established as necessary for the formation of the cochlear hair-cell stereociliary coat and causes hearing loss in mice and zebrafish when mutated. We sought to determine if biallelic variants in *PKHD1L1* also cause hearing loss in humans.

Exome sequencing was performed on DNA of three families segregating autosomal recessive moderate to severe nonsyndromic sensorineural hearing loss. Compound heterozygous missense p.[(Gly129Ser)];p.[(Gly1314Val)], homozygous missense p.(His2479Gln) and nonsense p.(Arg3381Ter) variants were identified in *PKHD1L1* that were predicted to be damaging using *in silico* pathogenicity prediction methods. *In vitro* functional analysis of two missense variants was performed using purified recombinant PKHD1L1 protein fragments. We then evaluated protein thermodynamic stability with and without the missense variants found in one of the families. *In vitro* functional assessment indicated that both engineered PKHD1L1 mutant p.(Gly129Ser) and p.(Gly1314Val) constructs significantly reduced the folding and structural stabilities of the expressed protein fragments, providing further evidence to support pathogenicity of these variants. *In silico* molecular modelling using AlphaFold2 and protein sequence alignment analysis were carried out to further explore potential variant effects on protein folding and stability and exposed key structural features that might suggest PKHD1L1 protein destabilization.

Multiple lines of evidence collectively associate *PKHD1L1* with nonsyndromic mild-moderate to severe sensorineural hearing loss. *PKHD1L1* testing in individuals with mild-moderate hearing loss may identify further affected families.

Keywords: Polycystic Kidney and Hepatic Disease 1-Like 1 (*PKHD1L1*), stereocilia, deafness

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Hearing loss-associated genes are implicated in the function of all parts of the delicate molecular machinery that permits human hearing. The inner hair cells (IHCs) and outer hair cells (OHCs) of the organ of Corti contain an apical bundle of ~100 actin-filled protrusions called stereocilia. Upon sound stimulation, stereocilia bundles are deflected by pressure-induced waves within the fluid-filled organ of Corti. Housing the mechanotransduction apparatus at the tips of stereocilia, these bundles mediate the transformation of the mechanical stimulus into an electrical signal the brain interprets as sound. While IHCs convert sound waves into nerve signals, OHCs allow for non-linear amplification of the sound stimuli by changing their length in response to bundle deflection, a process known as electromotility (Brownell 1990). Although IHCs and OHCs have two separate and distinct functions, both sensory cell types require a properly organized, functional stereocilia bundle. Stereocilia have a transiently expressed surface coat, but little is understood about the function or molecular architecture of this surface specialization. To date, there are over 30 genes reported to be critical for stereocilia bundle morphology that are associated with sensorineural hearing loss (SNHL) in humans (Michalski and Petit 2015; Petit and Richardson 2009).

One such stereocilia protein, polycystic kidney and hepatic disease 1-like 1 (*Pkhd1l1*) is critical to hearing in mice (Wu et al. 2019). The *PKHD1L1* gene in humans encodes a 4,243 amino acid protein, which is predicted to be composed by a large extracellular domain, a 20 amino acid transmembrane domain, and a very short cytoplasmic domain of 9 residues. In mice, PKHD1L1 is highly enriched in both IHCs and OHCs, particularly at the tips of OHC stereocilia bundles (Wu et al. 2019). It is hypothesized that this protein makes up a major component of the transient stereociliary coat. Mice lacking *Pkdh1l1* displayed elevated Auditory Brainstem Response (ABR) and Distortion Product Otoacoustic Emissions (DPOAE) thresholds in response to pure tone stimuli in a progressive fashion (Wu et al. 2019). More recent data from zebrafish (*Danio rerio*) with a double knockout of *pkhd1l1a* and *pkhd1l1b* (orthologs of human *PKHD1L1*) show significant deficits in auditory startle response at the larval stage, consistent with an early-onset auditory phenotype in zebrafish (Makrogkikas et al. 2023). Based on these findings in animal models, we sought to determine whether variants in *PKHD1L1* cause hearing loss in humans.

In this study, we propose *PKHD1L1* as a novel cause of autosomal recessive nonsyndromic hearing loss in humans. We describe three unrelated families with biallelic variants in *PKHD1L1* identified via exome sequencing. All three probands present bilateral congenital SNHL which is

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

moderate to severe. In addition, *in vitro* functional evaluation of two of the missense variants in protein fragments show decreased stabilities, suggesting that they may negatively impact their structures and molecular assembly *in vitro*.

Methods

Recruitment and Clinical Assessment

This study was approved by the institutional review boards of Boston Children's Hospital (IRB P-00031494), University Medical Center Göttingen (No. 3/2/16), and the School of Biological Sciences, University of Punjab, Lahore, Pakistan (IRB No. 00005281). Written informed consent was obtained from participating members of the three families or parents for their minor children.

The proband in family 1 derived from non-consanguineous parents ascertained as part of a large clinically diverse cohort of patients with SNHL in the United States. The proband in family 2 derived from consanguineous parents and was ascertained as part of a large ethnically diverse Middle Eastern population rare disease study. The proband in family 3 was born to consanguineous parents and identified from a special education school. The parents did not participate in the study.

Demographic, otolaryngologic, audiological, and relevant medical data were ascertained from the medical records of probands. Affected individuals underwent a complete otologic evaluation. Routine pure-tone audiometry was performed according to current standards and measured hearing thresholds at 0.25, 0.5, 1, 2, 4, 6 and 8 kHz. Pure tone audiometry for proband 3 was performed in ambient noise conditions due to lack of soundproof testing environment. The probands in families 1 and 2 underwent tympanometry and speech audiometry testing. Proband 2 additionally underwent otoacoustic emissions testing.

Exome Sequencing

Genomic DNA (gDNA) from individuals in families 1 (I:1, I:2 and II:1), 2 (III:1, III:2, IV:1 and IV:2) and 3 (II:1) was isolated from either buccal mucosa or whole blood using standard procedures.

Family 1: Exome sequencing was performed in a Clinical Laboratory Improvements Amendments (CLIA)-certified environment (GeneDx, Gaithersburg, MD, USA). Analysis was performed using the DRAGEN pipeline (Illumina, San Diego, CA, USA). Copy number variants were annotated using genome analysis toolkit (GATK) and a normalized segmented depth of coverage model. Variants were prioritized according to allele frequency, *in silico* predictors,

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

variant segregation, prior observations, and gene-disease association, as previously described (Rockowitz et al. 2020).

Family 2: Exome sequencing for the proband was performed using the SureSelect Target Enrichment v7 (Agilent Technologies, Santa Clara, CA) kit and sequenced with a NovaSeq 6000 (Illumina, San Diego, CA, USA) sequencer following manufacturer's protocols. Bioinformatics analysis of the proband in family 2 was performed as previously described (Vona et al. 2021).

Family 3: Exome sequencing for the proband was carried out at 3billion, Inc., Seoul, South Korea (https://3billion.io/index). Briefly, coding exon regions of human genes (~22,000) were captured by xGen Exome Research Panel v2 (Integrated DNA Technologies, Coralville, IA, USA). The captured regions of the genome were sequenced with NovaSeq 6000 (Illumina, San Diego, CA, USA). The raw genome sequencing data analysis, including alignment to the GRCh37/hg19 human reference genome, variant calling and annotation, was conducted with open-source bioinformatics tools including Franklin (https://franklin.genoox.com/clinical-db/home) and 3billion in-house software.

Variant Assessment and Validation

Variants were prioritized based on population and *in silico* pathogenicity prediction software. Variant minor allele frequencies were derived from gnomAD (v2.1.1 and v3.1.2) (Chen et al. 2022; Karczewski et al. 2020). Various pathogenicity prediction tools were used including SIFT (Ng and Henikoff 2001), PolyPhen-2 (Adzhubei et al. 2010), FATHMM (Shihab et al. 2014), MutationTaster (Schwarz et al. 2014), REVEL (Ioannidis et al. 2016) and CADD (Rentzsch et al. 2019). Visualization of amino acid substitution tolerance was supported by the MetaDome web server (Wiel et al. 2019).

Variants were annotated using the *PHKD1L1* NM_177531.6 accession (ENST00000378402.9). The GTEx Portal (GTEx consortium 2013) was referenced for assessing the location of variants across the annotated *PHKD1L1* isoform (Supplementary Fig. S1).

PKHD1L1 variant segregation in families 1 and 2 was confirmed using Sanger sequencing. Confirmation via Sanger sequencing was not performed for the proband of family 3.

Sequence analyses and structural modeling of PKHD1L1 protein

We compared the PKHD1L1 isoform sequence among 10 different species (see Supplementary Table S1 for more details about the selected species). The sequences were obtained from the

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

National Center for Biotechnology Information (NCBI) protein database (Supplementary Table S1). First, each individual protein sequence was used to predict their signal peptides and domains using the Simple Modular Architecture Research Tool (SMART) (Letunic et al. 2021). Signal peptides were further predicted using the SignalP-5.0 (Almagro Armenteros et al. 2019) and the Prediction of Signal Peptides (PrediSi) online servers (Hiller et al. 2004). AlphaFold2 modelling was used to predict the potential signal peptide cleavage site and accurately inform the start and end of each predicted domain before carrying out the protein sequence alignment (Mirdita et al. 2022). Since the PKHD1L1 IPT domains do not have a clear conservation pattern at their IPT protein start and end sequence and connecting linker domains, AlphaFold2 modeling results were then combined with protein sequence alignment to better predict the signal peptide, IPT domain start and end residue positions, and the location of missense mutations.

The ClustalW algorithm (Larkin et al. 2007) on Geneious (Kearse et al. 2012) was used for the sequence identity analysis using the default parameters. Alignment files from Geneious were imported and color-coded in JalView with 35% conservation threshold, as previously described (Kearse et al. 2012). AlphaFold2 simulations of PKHD1L1 fragments were carried out using the Colabfold v1.5.2-patch server using default parameters (Mirdita et al. 2022).

Cloning, expression, and purification of engineered bacterially expressed PKHD1L1 protein fragments and mutant constructs

Wild-type (WT) *Mus musculus* (*Mm*) PKHD1L1 IPT1-3 and IPT5-6 were subcloned into the *Nde*I and *Xho*I sites of the pET21a+ vector. Site-directed mutagenesis was applied to engineer the *Mm*_PKHD1L1 IPT1-3 p.(Gly129Ser) and PKHD1L1 IPT5-6 p.(Gly1314Val) constructs using the QuickChange lightning kit (Agilent Technologies). All constructs were used for protein expression in Rosetta (DE3) competent cells (Novagen) and cultured in TB as reported previously (De-la-Torre et al. 2018). Expressed proteins were refolded at 4 °C using conditions outlined below. WT *Mm* PKHD1L1 IPT1-3 and IPT1-3 p.(Gly129Ser) were refolded by fast or drop-wise dilution as previously reported for other protein families (De-la-Torre et al. 2018): 20 mL of pure denatured protein (0.5-1 mg/mL) was dropped into 480 mL of refolding buffer containing 20 mM TrisHCl [pH 8.0], 150 mM KCl, 50 mM NaCl, 2 mM $CaCl_2$, 400 mM L-arginine, and 2 mM D(+) glucose. *Mm* WT PKHD1L1 IPT5-6 and IPT5-6 p.(Gly1314Val) were refolded by dialysis of 40 mL of eluted denatured protein at 0.5 mg/mL into 1000 mL of refolding buffer (20 mM TrisHCl [pH 7.5], 150 mM KCl, 50 mM NaCl, 5 mM $CaCl_2$, 400 mM L-arginine, 1 mM of glutathione oxidized). Proteins were concentrated using 10,000 Da Amicon Ultra-15 centrifugal

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

filters and purified using size exclusion chromatography with an Akta Purifier System with the S200 16/600 pg and S200 13/300 increase GL columns (GE Healthcare) in a buffer containing 20 mM Tris-HCl pH 7.5, 150 mM KCl, 50 mM NaCl, and 5 mM $CaCl_2$ to preserve the most abundant endolymphatic cations. Purity of the recombinant and mammalian proteins was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Nanoscale Differential Scanning Fluorimetry

WT *Mm* PKHD1L1 IPT1-3 and IPT5-6 protein fragments and their respective missense IPT1-3 p.(Gly129Ser) and IPT5-6 p.(Gly1314Val) proteins were used for functional evaluation *in vitro*. Thermodynamic evaluation and folding stabilities of these constructs in solution were carried out using nano differential scanning fluorimetry (NanoDSF). Pure proteins were concentrated to 0.5-1 mg/mL for NanoDSF using a Prometheus NT.48 (Nanotemper) and scanned between 20-95 °C with a pre-stabilization phase of 1 min and a temperature slope of 2 °C/min (37 min in total). Data were processed using a PR.ThermControl v2.1.2 software and plotted using GraphPad Prism. At least two biological replicates were used for each experiment. Each protein preparation was independently expressed and refolded at least two times (two biological replicates), and evaluated independently on NanoDSF. Each NanoDSF scan was run in parallel using at least four separate capillary tubes for each biological replicate. Each result per biological replicate represents average values of these measurements. Protein folding analysis results were plotted as a relationship of the normalized F350/F330 (%) ratio intensities (T_{onset}). The first derivative of F350/F330 (%) intensities were plotted to obtain the thermal unfolding transition midpoints (T_m).

Results

Clinical Genetics

Family 1 (Fig. 1): The proband is a 10-15-year-old female born to healthy nonconsanguineous parents. She referred bilaterally on newborn hearing screening. Pure tone audiometry has been performed approximately every 6 months and consistently demonstrated a slowly progressive mild to moderate SNHL bilaterally. Between the ages of 0-5 years and 10-15 years, there is a decrease in pure tone average (PTA) of 5 dB for the right ear and 8 dB for the left ear. PTA was 41.25 and 38.75 in the right and left ears, respectively at age 0-5. Most recent audiometric testing showed PTA of 45 and 48.75 in the right and left ears, respectively. Speech audiometry at most recent evaluation (age 10-15 years) demonstrates a 90%-word recognition at a comfortable listening level. Speech recognition threshold (SRT) is 45 dB bilaterally. There is a

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

history of monthly episodes of benign paroxysmal positional vertigo (BPPV) which resolved after Epley maneuver. Imaging studies of the inner ear were not performed. An electrocardiogram and ophthalmology exam were normal. Cytomegalovirus testing was negative. There were no dysmorphic facial features, neurological or developmental abnormalities, or other pertinent history. Exome sequencing revealed candidate variants only in *PKHD1L1*. Other variants were excluded based on a combination of *in silico* predictors, allele frequency, and prior clinical observation. The proband was compound heterozygous for missense variants, p.(Gly129Ser) and p.(Gly1314Val) (Table 1) that segregated within the trio in a Mendelian recessive manner (residue numbering corresponds to the NCBI human (*Hs*) sequence including the signal peptide, Supplementary Table S1).

The p.(Gly129Ser) substitution resides in exon 4 of 78 in *PKHD1L1* and has a maximum allele frequency (MAF) of 0.001471% in gnomAD (v3.1.2). This variant is predicted to be deleterious to protein structure and function via *in silico* predictors (Table 1). This nonpolar glycine to polar serine substitution occurs at the tip of the PKHD1L1 IPT1 domain (Fig. 2b-d). This locus is predicted to be somewhat tolerant to missense substitution (Supplementary Fig. S2). On the other hand, the p.(Gly1314Val), located in exon 32 of 78 in *PKHD1L1,* has a MAF of 0.07204% (gnomAD, v3.1.2) and is also predicted to be deleterious to protein structure and function by *in silico* predictors (Table 1). This glycine to valine substitution occurs within the PKHD1L1 IPT6 domain region (Fig. 2b-d). This locus is predicted to be intolerant to missense substitution (Supplementary Fig. S2).

Family 2 (Fig. 1): The proband is a 6-10-year-old male born to healthy consanguineous parents. Congenital SNHL was suspected and later clinically diagnosed. It has progressed to a bilateral moderate to severe degree. Pure tone audiometry shows moderate to severe SNHL in all frequencies. Speech audiometry understanding is 100% at a comfortable listening level, and the otoacoustic emissions test is the reference on both ears. His SRT is 60 dB and his speech discrimination score (SDS) is 100% at the intensity level of 80 dB. The proband currently uses hearing aids bilaterally. There have been no vestibular or movement abnormalities. Exome sequencing revealed that the proband was homozygous for the p.(Arg3381Ter) nonsense variant that resided in a ~28 Mb run of homozygosity (Table 1). All other variants were excluded on the basis of segregation (Supplementary Table S2). Sanger sequencing at this locus confirmed the homozygous variant and revealed that the parents were both heterozygous carriers of the *PKHD1L1* variant.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The p.(Arg3381Ter) variant is located in exon 62 of 78 and identified with a MAF of 0.01936% in population databases (gnomAD, v3.1.2). This variant occurs in a region with parallel beta-helix (Pbh1) repeats (Fig. 2b and Supplementary Fig. S3 for residue conservation), introducing a premature stop codon that is predicted to result in the loss of approximately 20% of the transcript (~ 882 amino acids), including those encoding the transmembrane domain, and could potentially cause nonsense mediated decay. However, if expressed in a truncated form, lack of the transmembrane domain is likely to impair proper localization of PKHD1L1 in the cell membrane or might induce unconventional secretion of PKHD1L1 protein fragments.

Family 3 (Fig. 1): The proband was a 10-15-year-old male with congenital SNHL born to healthy consanguineous parents. Audiometric testing demonstrated a bilateral severe SNHL. No further follow up was possible for the affected individual. Exome data analysis revealed two homozygous missense variants of interest; one in *PKHD1L1* and one in *MYO7A* (NM_000260.4:c.1123C>G, p.(Leu375Val)). The homozygous variant in *MYO7A* was deprioritized given uncertain and neutral *in silico* predictions with respect to impact on protein structure and function (Supplementary Table S2).

The p.(His2479Gln) substitution is located in exon 49 of 78 and is identified at a MAF of 0.3107% in population databases (gnomAD, v3.1.2). This positively charged histidine to neutral glutamine substitution is located in a topological region with an unpredicted domain structure when using the SMART prediction tool (Fig. 2b).

Investigating the conservation of the mutated residue positions throughout evolution

All three variants do not appear to cluster in any particular region of the PKHD1L1 isoform that was used for alignment and further analysis (Supplementary Fig. S1). In comparing the longest PKHD1L1 isoform sequences among 10 different species, we uncovered an overall amino acid sequence identity of 79.2 % (Supplementary Fig. S3) across all 10 species. Notably, *Mm* and *Hs* PKHD1L1 share 81.8% of amino acid sequence identity (when comparing for identical sites excluding the signal peptides), suggesting high protein sequence conservation between the two species. Although some previous studies report protein sequence alignments and predictions of PKHD1L1 IPT domains (Hogan et al. 2003), an in-depth analysis was necessary to more accurately predict the signal peptide cleavage sites, as well as the starting and ending residues for each IPT domain. The location of the native Gly126, Gly1314, and His2479 residues, where the reported missense variants were detected, are highly conserved across a diverse set of the 10 different PKHD1L1 orthologues analyzed (Fig. 2d-e, Fig. 3a, and Supplementary Fig. S3 and Fig. S4).

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

AlphaFold2 modeling of PKHD1L1 substitutions

PKHD1L1 has 14 predicted IPT extracellular-domain repeats of similar fold but with non-identical protein sequence and labeled as IPT1 to IPT14 from its N-terminal to C-terminal end, and other key domain features (Fig. 2b). The AlphaFold2 model of *Hs* WT IPT1-2 and its mutant p.(Gly129Ser) shows no apparent differences between their predicted structures (Fig. 2f-g), likely because the small side chain carrying this residue is located on a loop region exposed to the solvent. More specifically, the p.(Gly129Ser) variant is located within the connecting loop between the β-strand 6-7 of IPT1, close to a potential disulphide bond formed by p.Cys100 and p.Cys86, also found in plexin-like domains (Fig. 2g). Changes of the polarity or the electrostatic potential of this loop by p.(Gly129Ser) might cause structural changes or altered loop dynamics in IPT1 (Krieger et al. 2005). We also generated AlphaFold2 models for *Hs* IPT5-6 consistent with the expected IPT plexin-like folding for this structure (Fig. 2h-i). According to this structural model, the *Hs* p.(Gly1314Val) mutation is also located at the connecting loop between the β-strand 6-7 of IPT6 (Fig. 2h-j). Furthermore, the p.(Gly1314Val) variant is located within the connecting area between IPT5 and IPT6, and the AlphaFold2 model suggests a structural change for this specific fragment (Fig. 2i-j).

In a previous study, authors used protein sequence analysis of PKHD1, PKHD1L1 and TMEM2 reporting that PKHD1 and PKHD1L1 share two regions of significant sequence homology with TMEM2 (Hogan et al. 2003). AlphaFold2 modelling of the third variant, p.(His2479Gln) and surrounding PKHD1L1 region, revealed a high structural homology with a *Hs* TMEM2 protein (Fig. 3). We identified that this region features a conserved (throughout 10 different PKHD1L1 species, and Supplementary Fig. S4) p.His2479 residue (p.His552 in *Hs* TMEM2, Fig. 3a; PDB: 8C6I (Niu et al. 2023)) reported to form a nickel-finger binding site, which might mediate catalytic functions in TMEM2. Disruption of this site in PKHD1L1 and TMEM2 might impair cation binding (Fig. 3a-f and Supplementary Fig. S3 and S4) and suggests a potential deleterious effect for this variant on protein structure and function. This locus is predicted neutral in terms of tolerance to missense substitution (Supplementary Fig. S2).

Functional testing of the p.(Gly129Ser) and p.(Gly1314Val) substitutions

Next, we cloned, expressed, and purified WT *Mm* PKHD1L1 IPT1-3 and IPT5-6 protein fragments as well as the respective IPT1-3 p.(Gly129Ser) and IPT5-6 p.(Gly1314Val) mutant constructs using size-exclusion chromatography (SEC) (Supplementary Fig. S5). These protein constructs represent key regions of the complete extracellular domain of PKHD1L1 where these mutations might locally affect the structural assembly of the protein. The thermodynamic and

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

folding stabilities were measured using NanoDSF to assess the protein stabilities in solution for WT PKHD1L1 protein fragments and compared to fragments carrying missense mutations (Fig. 4 and Supplementary Fig. S5). For WT IPT1-3, the T_{onset} (melting temperature at which unfolding begins) was measured at a maximum of 58 °C, while the T_{onset} for IPT1-3 p.(Gly129Ser) variant was 52 °C (a 6 °C decrease, Fig. 4 and Supplementary Fig. S5). In addition, there was a decrease on the T_m (melting temperature or point at which 50% of the protein is unfolded) of ~ 4 °C comparing different thermal transition points between WT and the IPT1-3 p.(Gly129Ser) variant (Fig. 4a). These measurements strongly suggest that the p.(Gly129Ser) variant affects PKHD1L1 protein stability within this region.

Similarly, NanoDSF measurements for WT *Mm* PKHD1L1 IPT5-6 and *Mm* IPT5-6 p.(Gly1314Val) showed a T_{onset} of 35.48 °C and 28.56 °C, respectively (~ 7 °C decrease, Fig. 4 and Supplementary Fig. S5). Additionally, WT *Mm* PKHD1L1 IPT5-6 showed a melting temperature T_m of 45.10 °C, while the mutant IPT5-6 p.(Gly1314Val) displayed a decrease on this T_m to 36.6 °C (decreasing the temperature ~ 8.6 °C) (Fig. 4b). These results confirm that both IPT1-3 p.(Gly129Ser) and IPT5-6 p.(Gly1314Val) variants indeed significantly decrease the thermal and folding stabilities of these PKHD1L1 protein constructs.

Discussion

A majority of congenital SNHL is attributable to a genetic etiology, and clinical genetic testing for known SNHL genes is an established standard of care in the diagnostic evaluation of bilateral SNHL in pediatric patients (Shearer and Smith 2015; Smith et al. 2005). There are over 120 known genetic causes of nonsyndromic hearing loss, and gene panel tests are recommended to facilitate accurate and timely genetic diagnosis of SNHL (https://hereditaryhearingloss.org). However, despite advances in genetic testing for SNHL, the diagnostic yield for SNHL ranges from 22.5% to 55.7% with an average of about 40% (Downie et al. 2020; Perry et al. 2023; Rouse et al. 2022; Sloan-Heggen et al. 2016). The identification of novel hearing loss genes is critical to improving diagnostic rates, thus impacting care and management for individuals with SNHL.

PKHD1L1 is predominantly expressed in the OHC stereocilia by P0 to P12 with a basal-to-apical (decreasing) expression gradient and is a major component of the stereocilia surface coat (Wu et al. 2019). *Pkhd1l1*-deficient mice lack the surface coat at the stereocilia tips and exhibit progressive SNHL by ABR and DPOAE measurements, starting as early as 3 weeks. The two functional hypotheses of PKHD1L1 expression at the stereocilia include that it may be required

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

for the correct localization of other stereocilia proteins or during the development of attachment crowns at the stereocilia to secure the tectorial membrane to the bundle. This immature attachment could manifest as a persisting relaxed tectorial membrane coupling (Wu et al. 2019). Still, it is unknown whether PKHD1L1 is the only component of the stereocilia coat. Recently, *pkhd1l1* was shown to play a critical role in regulating hearing in zebrafish (Makrogkikas et al. 2023). *pkhdl1l* has a ubiquitous expression pattern and is sustained for most of embryonic development (Makrogkikas et al. 2023). Through depletion of the two paralogous genes (*pkhd1l1a* and *pkhd1l1b*), double mutant zebrafish exhibited significant hearing loss from the larval stage (6 days post fertilization) which differs compared to progressive hearing loss in the mouse (Wu et al. 2019).

The degree of hearing impairment in the patients we present is fairly broad: The proband in family 1, with p.[(Gly129Ser)];p.[(Gly1314Val)] compound heterozygous variants, was diagnosed with congenital hearing impairment that remains mild to moderate at the age of 10-15 years; the proband in family 2, with a homozygous p.(Arg3381Ter) variant, already showed moderate to severe SNHL at the age of 6-10 years; and the proband in family 3, at the age of 10-15 years, showed severe hearing impairment attributed to the homozygous p.(His2479Gln) variant. While we have identified individuals in three families with variants in *PKHD1L1*, this study highlights the necessity for an extended case series with longitudinal audiological follow up and functional studies to assess variant effects of patient-specific perturbations on development, maturation, and function of the auditory system, as well as explore the potential of accelerated effects of age, noise, or trauma on progression, which remain as current major limitations. Interestingly, the *PKHD1L1* gene has been suggested to be associated with adult-onset hearing loss (Lewis et al. 2023). Since the studied variants are also located in different residue positions in the PKHD1L1 protein sequence, the broad range of hearing impairment from these patients might suggest that these variants differentially impact the protein expression, folding, and/or the stability and function of PKHD1L1.

Thus, we also investigated the conservation of the mutated residue positions throughout evolution. Multiple sequence alignments of the complete PKHD1L1 amino acid sequence from 10 different species were analysed and found to be highly conserved. This suggests that these native residues might be critical to protein folding and assembly. Therefore, variants in these positions might disrupt protein function and potentially cause hearing impairment *in vivo*.

To further predict how these PKHD1L1 mutant variants might affect the PKHD1L1 protein at the structural level, we modelled the structures of the individual domains carrying reported variants

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

(Fig. 2f-i and Fig. 3b-e). The p.(Gly129Ser) substitution in IPT1 was not predicted to exert an apparent structural difference. We speculate that instead changes of the polarity or the electrostatic potential of the β-strand linker loop by p.(Gly129Ser) might alter loop dynamics in IPT1. Interestingly, both glycine substitutions p.(Gly129Ser) and p.(Gly1314Val) are located within the same connecting loop between β-strand 6-7 in IPT1 and IPT6, respectively. While the AlphaFold2 model of the p.(Gly129Ser) mutant shows no apparent structural changes in the predicted structure (Fig. 2f), the AlphaFold2 model of p.(Gly1314Val) shows a conformation change, likely due to steric hindrance on the structure (Fig. 2i). Finally, we also mapped the location of the p.(His2479Gln) using structural modeling by AlphaFold2 (Fig. 3, Supplementary Fig. 4). Our results indicate that the His2479 position (from 10 different species) is 100% conserved with a *Hs* TMEM2 protein (PDB: 8C6I), a regulator of the hyaluronan metabolism (Fig. 3a and 3e-f) (Sato et al. 2023). Interestingly, experiments suggest that TMEM2 activity is calcium-dependent (Yamamoto et al. 2017) and TMEM2 has been previously studied for its structural similarities with the *CEMIP* deafness gene candidate (Yoshida et al. 2013).

Since the p.(Arg3381Ter) introduces a stop codon that would prevent the translation and expression of the transmembrane domain of PKHD1L1 (Fig. 2b), the proper localization of PKHD1L1 in the cell membrane is expected to be affected. Because the helical peptide of PKHD1L1 transmembrane domain is predicted between p.Ile4211-4231Ala (850 amino acids downstream from the Arg3381 site from the PKHD1L1 N-terminal towards the C-terminal end), this large portion of the protein will not be expressed (Fig. 2b-c, Supplementary Fig. S3). This is likely to impair the proper folding, trafficking, and insertion of PKHD1L1 in the stereocilia-plasma membrane, or even result in secretion of PKHD1L1 extracellular fragments that could progressively affect hearing function. Interestingly, secreted versions of extracellular PKHD1L1 have been found in supernatant solution from platelet cells (Maynard et al. 2007) and their soluble concentrations could be modulated by protease inhibitors (Fong et al. 2011), suggesting potential cleavage sites in *Hs* PKHD1L1. However, the role of these cleaved extracellular PKHD1L1 protein fragments is still unknown.

Our NanoDSF thermal folding analysis showed that both T_{onset} and T_m values decreased in the presence of the p.(Gly129Ser) and p.(Gly1314Val) variants. The T_m measurements using NanoDSF showed p.(Gly129Ser), located in a loop, decreases the stability of IPT1 and further showed how this variant propagates its destabilizing effects to the neighboring IPT2 and IPT3 (Fig. 2b, Fig. 4a, and Supplementary Fig. S5). Likely, the p.(Gly1314Val) variant also alters the stability of the loop and the chemical environment in the IPT5-IPT6 connection, since the

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

383 measured folding stability showed an 8.6 °C decrease of unfolding temperature between WT
384 IPT5-6 fragment and the p.(Gly1314Val) variant (Fig. 4). This is the first study showing strong
385 evidence to support how missense variants locally affect the structural folding and stability of
386 PKHD1L1 fragments *in vitro*. Given the high conservation rate of 81.8% in amino acid sequence
387 identity between the *Hs* and *Mm* PKHD1L1, and the 100% conservation of the mutated sites
388 across the species analyzed, we believe our findings using *Mm* PKHD1L1 protein fragments
389 can be directly applied to *Hs* PKHD1L1. Future functional studies involving full-length protein
390 purification would allow for better understanding of various effects such variants might have on
391 the stability of the entire PKHD1L1 extracellular domain, its protein expression and proper
392 localization, which might be linked to the progression and hearing loss severity. Furthermore,
393 studies focused on uncovering the influence of mutations on the structure of the complete
394 PKHD1L1 extracellular domain will help in understanding the role of PKHD1L1 in hearing
395 function and beyond, since the PKHD1L1 has been suggested as a tumor suppressor (Yang et
396 al. 2023) and a human cancer biomarker (Kafita et al. 2023; Wang et al. 2023; Zheng et al.
397 2019).

398 Additional syndromic involvement was excluded in our three patients. However, in addition to
399 hearing impairment, disruption of PKHD1L1 has also been associated with increased
400 susceptibility to pentylenetetrazol-induced seizures in mice indicating a possible role in
401 maintenance of neuronal excitability in the central nervous system (Yu et al. 2023). PKHD1L1 is
402 expressed in the hippocampus and cerebral cortex in adult WT mice. Knockdown of PHKD1L1
403 using PHKD1L1-shRNA or PHKD1L1-shRNA-AAV increased susceptibility of seizures as
404 indicated by increased epileptiform bursting activity in cultured hippocampal neurons and
405 pentylenetetrazol-induced seizures of mice following knockdown, suggesting a role for
406 PKHD1L1 in the maintenance of normal excitation-inhibition balance (Yu et al. 2023). Although
407 these findings have not been validated in *Pkhd1l1*-knockout mice, an open question remains as
408 to whether similar pleiotropic effects paralleling, for example, the various phenotypes caused by
409 pathogenic variants in *TBC1D24* may also occur as more *PKHD1L1* patients are discovered.
410 The gene-phenotype relationship of *TBC1D24* is extensive and includes both autosomal
411 dominant (DFNA65, OMIM 616044) and recessive (DFNB86, OMIM 614617) nonsyndromic
412 deafness, as well as familial infantile myoclonic epilepsy (FIME, OMIM 605021) with or without
413 deafness, early infantile epileptic encephalopathy (EIEE16, OMIM 615338), progressive
414 myoclonic epilepsy (PME, OMIM 310370), or deafness, onychodystrophy, osteodystrophy,
415 impaired intellectual development, and seizures syndrome (DOORS, OMIM 220500) (as
416 reviewed by Mucha et al. 1993; Rehman et al. 2017). Further reporting of variants in *PKHD1L1*

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

will be critical to discovering a potential parallel between *PKHD1L1* and the complex syndromes using the *TBC1D24* analogy.

Finally, we conclude with a cautionary note for diagnostic laboratories, as this work serves as an extended invitation to include *PKHD1L1* screening in their future patient diagnostic screening. Determination of pathogenicity of genes causing hearing loss relies on multiple lines of supporting evidence. However, in the case of large genes with large amounts of natural variation, like *PKHD1L1*, the allele frequency criteria, specifically the PM2_Supporting pathogenicity criteria, becomes a point of contention for classification. The PM2_Supporting criterion as proposed by the ACMG/AMP hearing loss expert panel specification have calculated a ≤0.00007 frequency for autosomal recessive disorders to trigger application (Oza et al. 2018). With MAFs of 0.0007204 in the non-European Finnish and 0.003107 in the South Asian populations (gnomAD v3.1.2), the c.3941G>T, p.(Gly1314Val) and c.7437C>A, p.(His2479Gln) variants exceeded the ACMG/AMP criteria with hearing loss expert panel specification , respectively, and suggests the potential for missed diagnoses if the proposed frequency is applied for variant filtering (Oza et al. 2018). This merits attention because the PM2 criterion is most commonly applied among the criteria for hearing loss variant classification (Kim et al. 2022). Given the large size of *PKHD1L1* and that more missense variants are observed than expected (Z = -0.72, o/e = 1.04 (1.02 – 1.08)) minor allele frequency threshold usage should be carefully assessed for this gene. This study is a call to clinical laboratories to include careful screening of *PKHD1L1* biallelic variants in patients with a hearing loss ranging from mild-moderate to severe.

Conclusion

Here we provide data to support *PKHD1L1* as a cause of human nonsyndromic autosomal recessive congenital, mild-moderate to severe hearing loss. We demonstrated that all reported missense variants are of residues highly conserved throughout evolution, suggesting that the native residues are key for protein folding and function, while variants in these sites affect the thermal-folding stability of PKHD1L1 fragments in solution. Inclusion as a hearing loss gene is supported by multiple families segregating plausible variants, *in vitro* functional data confirming the impact of variants, as well as previously published mouse and zebrafish models demonstrating hearing loss. Further research will be needed to determine the effect of age, noise, or trauma on progression of hearing loss.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7: 248-9. https://doi.org/10.1038/nmeth0410-248

Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37: 420-423. https://doi.org/10.1038/s41587-019-0036-z

Brownell WE (1990) Outer hair cell electromotility and otoacoustic emissions. Ear Hear 11: 82-92. https://doi.org/10.1097/00003446-199004000-00003

Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, Alföldi J, Watts NA, Vittal C, Gauthier LD, Poterba T, Wilson MW, Tarasova Y, Phu W, Yohannes MT, Koenig Z, Farjoun Y, Banks E, Donnelly S, Gabriel S, Gupta N, Ferriera S, Tolonen C, Novod S, Bergelson L, Roazen D, Ruano-Rubio V, Covarrubias M, Llanwarne C, Petrillo N, Wade G, Jeandet T, Munshi R, Tibbetts K, Consortium gP, O'Donnell-Luria A, Solomonson M, Seed C, Martin AR, Talkowski ME, Rehm HL, Daly MJ, Tiao G, Neale BM, MacArthur DG, Karczewski KJ (2022) A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. bioRxiv: 2022.03.20.485034. https://doi.org/10.1101/2022.03.20.485034

Choudhary D, Narui Y, Neel BL, Wimalasena LN, Klanseck CF, De-la-Torre P, Chen C, Araya-Secchi R, Tamilselvan E, Sotomayor M (2020) Structural determinants of protocadherin-15 mechanics and function in hearing and balance perception. Proc Natl Acad Sci U S A 117: 24837-24848. https://doi.org/10.1073/pnas.1920444117

Consortium G (2013) The Genotype-Tissue Expression (GTEx) project. Nat Genet 45: 580-5. https://doi.org/10.1038/ng.2653

De-la-Torre P, Choudhary D, Araya-Secchi R, Narui Y, Sotomayor M (2018) A Mechanically Weak Extracellular Membrane-Adjacent Domain Induces Dimerization of Protocadherin-15. Biophys J 115: 2368-2385. https://doi.org/10.1016/j.bpj.2018.11.010

Downie L, Halliday J, Burt R, Lunke S, Lynch E, Martyn M, Poulakis Z, Gaff C, Sung V, Wake M, Hunter MF, Saunders K, Rose E, Lewis S, Jarmolowicz A, Phelan D, Rehm HL, Melbourne Genomics Health A, Amor DJ (2020) Exome sequencing in infants with congenital hearing impairment: a population-based cohort study. Eur J Hum Genet 28: 587-596. https://doi.org/10.1038/s41431-019-0553-8

Fong KP, Barry C, Tran AN, Traxler EA, Wannemacher KM, Tang HY, Speicher KD, Blair IA, Speicher DW, Grosser T, Brass LF (2011) Deciphering the human platelet sheddome. Blood 117: e15-26. https://doi.org/10.1182/blood-2010-05-283838

Hiller K, Grote A, Scheer M, Munch R, Jahn D (2004) PrediSi: prediction of signal peptides and their cleavage positions. Nucleic Acids Res 32: W375-9. https://doi.org/10.1093/nar/gkh378

Hogan MC, Griffin MD, Rossetti S, Torres VE, Ward CJ, Harris PC (2003) PKHDL1, a homolog of the autosomal recessive polycystic kidney disease gene, encodes a receptor with inducible T lymphocyte expression. Hum Mol Genet 12: 685-98.

Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh CL, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W (2016) REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 99: 877-885. https://doi.org/10.1016/j.ajhg.2016.08.016

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Jaiganesh A, De-la-Torre P, Patel AA, Termine DJ, Velez-Cortes F, Chen C, Sotomayor M (2018) Zooming in on Cadherin-23: Structural Diversity and Potential Mechanisms of Inherited Deafness. Structure 26: 1210-1225 e4. https://doi.org/10.1016/j.str.2018.06.003

Kafita D, Nkhoma P, Dzobo K, Sinkala M (2023) Shedding Light on the Dark Genome: Insights into the Genetic, CRISPR-based, and Pharmacological Dependencies of Human Cancers and Disease Aggressiveness. bioRxiv: 2023.08.15.552589. https://doi.org/10.1101/2023.08.15.552589

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Genome Aggregation Database C, Neale BM, Daly MJ, MacArthur DG (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581: 434-443. https://doi.org/10.1038/s41586-020-2308-7

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647-9. https://doi.org/10.1093/bioinformatics/bts199

Kim SY, Kim BJ, Oh DY, Han JH, Yi N, Kim NJ, Park MK, Keum C, Seo GH, Choi BY (2022) Improving genetic diagnosis by disease-specific, ACMG/AMP variant interpretation guidelines for hearing loss. Sci Rep 12: 12457. https://doi.org/10.1038/s41598-022-16661-x

Krieger F, Moglich A, Kiefhaber T (2005) Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. J Am Chem Soc 127: 3346-52. https://doi.org/10.1021/ja042798i

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-8. https://doi.org/10.1093/bioinformatics/btm404

Letunic I, Khedkar S, Bork P (2021) SMART: recent updates, new developments and status in 2020. Nucleic Acids Res 49: D458-D460. https://doi.org/10.1093/nar/gkaa937

Lewis MA, Schulte J, Matthews L, Vaden KI, Steves CJ, Williams FMK, Schulte BA, Dubno JR, Steel KP (2023) Accurate phenotypic classification and exome sequencing allow identification of novel genes and variants associated with adult-onset hearing loss. medRxiv. https://doi.org/10.1101/2023.04.27.23289040

Makrogkikas S, Cheng RK, Lu H, Roy S (2023) A conserved function of Pkhd1l1, a mammalian hair cell stereociliary coat protein, in regulating hearing in zebrafish. J Neurogenet: 1-8. https://doi.org/10.1080/01677063.2023.2187792

Maynard DM, Heijnen HF, Horne MK, White JG, Gahl WA (2007) Proteomic analysis of platelet alpha-granules using mass spectrometry. J Thromb Haemost 5: 1945-55. https://doi.org/10.1111/j.1538-7836.2007.02690.x

Michalski N, Petit C (2015) Genetics of auditory mechano-electrical transduction. Pflugers Arch 467: 49-72. https://doi.org/10.1007/s00424-014-1552-9

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. Nat Methods 19: 679-682. https://doi.org/10.1038/s41592-022-01488-1

Mucha BE, Hennekam RCM, Sisodiya S, Campeau PM (1993) TBC1D24-Related Disorders. In: Adam MP, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Amemiya A (eds) GeneReviews((R)), Seattle (WA)

Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. Genome Res 11: 863-74. https://doi.org/10.1101/gr.176601

Niu M, McGrath M, Sammon D, Gardner S, Morgan RM, Di Maio A, Liu Y, Bubeck D, Hohenester E (2023) Structure of the transmembrane protein 2 (TMEM2) ectodomain and its apparent lack of hyaluronidase activity. Wellcome Open Res 8: 76. https://doi.org/10.12688/wellcomeopenres.18937.2

Oza AM, DiStefano MT, Hemphill SE, Cushman BJ, Grant AR, Siegert RK, Shen J, Chapin A, Boczek NJ, Schimmenti LA, Murry JB, Hasadsri L, Nara K, Kenna M, Booth KT, Azaiez H, Griffith A, Avraham KB, Kremer H, Rehm HL, Amr SS, Abou Tayoun AN, ClinGen Hearing Loss Clinical Domain Working G (2018) Expert specification of the ACMG/AMP variant interpretation guidelines for genetic hearing loss. Hum Mutat 39: 1593-1613. https://doi.org/10.1002/humu.23630

Perry J, Redfield S, Oza A, Rouse S, Stewart C, Khela H, Srinivasan T, Albano V, Shearer E, Kenna M (2023) Exome Sequencing Expands the Genetic Diagnostic Spectrum for Pediatric Hearing Loss. Laryngoscope 133: 2417-2424. https://doi.org/10.1002/lary.30507

Petit C, Richardson GP (2009) Linking genes underlying deafness to hair-bundle development and function. Nat Neurosci 12: 703-10. https://doi.org/10.1038/nn.2330

Rehman AU, Friedman TB, Griffith AJ (2017) Unresolved questions regarding human hereditary deafness. Oral Dis 23: 551-558. https://doi.org/10.1111/odi.12516

Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M (2019) CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res 47: D886-D894. https://doi.org/10.1093/nar/gky1016

Rockowitz S, LeCompte N, Carmack M, Quitadamo A, Wang L, Park M, Knight D, Sexton E, Smith L, Sheidley B, Field M, Holm IA, Brownstein CA, Agrawal PB, Kornetsky S, Poduri A, Snapper SB, Beggs AH, Yu TW, Williams DA, Sliz P (2020) Children's rare disease cohorts: an integrative research and clinical genomics initiative. NPJ Genom Med 5: 29. https://doi.org/10.1038/s41525-020-0137-0

Rouse SL, Florentine MM, Taketa E, Chan DK (2022) Racial and ethnic disparities in genetic testing for hearing loss: a systematic review and synthesis. Hum Genet 141: 485-494. https://doi.org/10.1007/s00439-021-02335-7

Sato S, Miyazaki M, Fukuda S, Mizutani Y, Mizukami Y, Higashiyama S, Inoue S (2023) Human TMEM2 is not a catalytic hyaluronidase, but a regulator of hyaluronan metabolism via HYBID (KIAA1199/CEMIP) and HAS2 expression. J Biol Chem 299: 104826. https://doi.org/10.1016/j.jbc.2023.104826

Schwarz JM, Cooper DN, Schuelke M, Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11: 361-2. https://doi.org/10.1038/nmeth.2890

Shearer AE, Smith RJ (2015) Massively Parallel Sequencing for Genetic Diagnosis of Hearing Loss: The New Standard of Care. Otolaryngol Head Neck Surg 153: 175-82. https://doi.org/10.1177/0194599815591156

Shihab HA, Gough J, Mort M, Cooper DN, Day IN, Gaunt TR (2014) Ranking non-synonymous single nucleotide polymorphisms based on disease concepts. Hum Genomics 8: 11. https://doi.org/10.1186/1479-7364-8-11

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Sloan-Heggen CM, Bierer AO, Shearer AE, Kolbe DL, Nishimura CJ, Frees KL, Ephraim SS, Shibata SB, Booth KT, Campbell CA, Ranum PT, Weaver AE, Black-Ziegelbein EA, Wang D, Azaiez H, Smith RJH (2016) Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss. Hum Genet 135: 441-450. https://doi.org/10.1007/s00439-016-1648-8

Smith RJ, Bale JF, Jr., White KR (2005) Sensorineural hearing loss in children. Lancet 365: 879-90. https://doi.org/10.1016/S0140-6736(05)71047-3

Vona B, Mazaheri N, Lin SJ, Dunbar LA, Maroofian R, Azaiez H, Booth KT, Vitry S, Rad A, Ruschendorf F, Varshney P, Fowler B, Beetz C, Alagramam KN, Murphy D, Shariati G, Sedaghat A, Houlden H, Petree C, VijayKumar S, Smith RJH, Haaf T, El-Amraoui A, Bowl MR, Varshney GK, Galehdari H (2021) A biallelic variant in CLRN2 causes non-syndromic hearing loss in humans. Hum Genet 140: 915-931. https://doi.org/10.1007/s00439-020-02254-z

Wang L, Chen Q, Liu T, Bai T, Zhang M, Hu Y, Li J, Chang F (2023) Role and mechanism of benzo[a]pyrene in the transformation of chronic obstructive pulmonary disease into lung adenocarcinoma. J Cancer Res Clin Oncol 149: 4741-4760. https://doi.org/10.1007/s00432-022-04353-y

Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C (2019) MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. Hum Mutat 40: 1030-1038. https://doi.org/10.1002/humu.23798

Wu X, Ivanchenko MV, Al Jandal H, Cicconet M, Indzhykulian AA, Corey DP (2019) PKHD1L1 is a coat protein of hair-cell stereocilia and is required for normal hearing. Nat Commun 10: 3801. https://doi.org/10.1038/s41467-019-11712-w

Yamamoto H, Tobisawa Y, Inubushi T, Irie F, Ohyama C, Yamaguchi Y (2017) A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. J Biol Chem 292: 7304-7313. https://doi.org/10.1074/jbc.M116.770149

Yang Y, Pang Q, Hua M, Huangfu Z, Yan R, Liu W, Zhang W, Shi X, Xu Y, Shi J (2023) Excavation of diagnostic biomarkers and construction of prognostic model for clear cell renal cell carcinoma based on urine proteomics. Front Oncol 13: 1170567. https://doi.org/10.3389/fonc.2023.1170567

Yoshida H, Nagaoka A, Kusaka-Kikushima A, Tobiishi M, Kawabata K, Sayo T, Sakai S, Sugiyama Y, Enomoto H, Okada Y, Inoue S (2013) KIAA1199, a deafness gene of unknown function, is a new hyaluronan binding protein involved in hyaluronan depolymerization. Proc Natl Acad Sci U S A 110: 5612-7. https://doi.org/10.1073/pnas.1215432110

Yu J, Wang G, Chen Z, Wan L, Zhou J, Cai J, Liu X, Wang Y (2023) Deficit of PKHD1L1 in the dentate gyrus increases seizure susceptibility in mice. Hum Mol Genet 32: 506-519. https://doi.org/10.1093/hmg/ddac220

Zheng C, Quan R, Xia EJ, Bhandari A, Zhang X (2019) Original tumour suppressor gene polycystic kidney and hepatic disease 1-like 1 is associated with thyroid cancer cell progression. Oncol Lett 18: 3227-3235. https://doi.org/10.3892/ol.2019.10632

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

643 # Statements and Declarations

644 ## Acknowledgments

645 We thank the families for their participation.

646 ## Funding

647 This work was supported by NIDCD K08 DC019716 to AES, the De Garay Family Fund to MAK
648 with support from the Boston Children's Rare Disease Cohort Initiative, and the German
649 Research Foundation (DFG) VO 2138/7-1 grant 469177153 to BV, and by NIH R01DC017166
650 (NIDCD) and R01DC020190 (NIDCD) to AAI. Funding for work in Pakistan was provided by the
651 University of the Punjab (SN).

652 ### Author contributions

653 SER and PD contributed equally as shared first authors and AAI, AES and BV contributed
654 equally as shared last authors. AES, PD, and AAI conceived the study. PD, TM, and AAI
655 designed methodology. PD and TM contributed software. MZ, PD, and AAI performed formal
656 analyses that were validated by MZ, MK, PD, TM, and AAI. MZ, HK, MK, SN, MAK, PD, TM,
657 AAI, and BV conducted the investigation process. HG, GS, MAK, AES, SER, and PD secured
658 patient and experimental resources. MZ, HG, GS, BV, SER, AES, PD, and AAI performed data
659 curation. BV, AES, PD, and AAI wrote the original draft. MZ, HK, MK, SN, HG, BV, SER, AES,
660 PD, and AAI contributed to reviewing and editing. BV, AES, SER, PD, and AAI prepared figures.
661 SN, AES, PD, and AAI provided supervision. BV, AES, and AAI oversaw project administration.
662 All authors reviewed and approved of the manuscript.

663 ## Data Availability

664 All data needed to evaluate the conclusions in the paper are present in the paper and/or the
665 Supplementary Materials. Additional data related to this paper may be requested from the
666 authors.

667 ## Compliance with ethical standards

668 ## Conflict of interest

669 Go Hun Seo is an employee of 3Billion, Inc. The other authors have no conflicts to declare.

670 ## Ethics approval

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

This study was approved by the institutional review boards of Boston Children's Hospital (IRB P-00031494), University Medical Center Göttingen (No. 3/2/16), and the School of Biological Sciences, University of Punjab, Lahore, Pakistan (IRB No. 00005281).

Consent to participate

Written informed consent was obtained from all participants.

Consent to publish

Consent to publish was obtained from all participants.

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

678 ## Figure Legends

679 **Fig. 1. Pedigrees & Audiograms.** Pedigree and audiometric information for three families with
680 biallelic *PKHD1L1* variations. (a) Pedigrees for families 1-3. Each proband with SNHL is
681 indicated with shading and arrow. (b) Pure tone audiometry for probands 1-3; x represents
682 results for the left ear and o represents the right. Audiometric evaluation performed at 10-15-
683 years-old, for probands 1 and 3, and 6-10-years-old, for probands 2, respectively.

684
685

686 **Fig. 2. PKHD1L1 protein domain prediction and evolutionary analysis.** (a) Schematic of a
687 hair-cell stereocilia bundle under force stimulation highlighting the stereocilia surface coat. (b)
688 Protein domain composition prediction from SMART using the *Hs* PKHD1L1 protein sequence
689 as in NCBI accession code NP_803875.2, including the signal peptide (See Supplementary
690 Table S1). Positions of each variant reported in this study is presented with a green arrowhead.
691 The red star represents a newly predicted TMEM2-like domain. Black and purple arrow-headed
692 lines represent the sequence fragments used for AlphaFold2 modelling of IPT1-2, IPT5-6, and
693 TMEM2-like domain, respectively. (c) Topological description of *Hs* PKHD1L1 protein sequence
694 as a reference. (d-e) Multiple protein sequence alignments comparing IPT1 and IPT6 domains
695 from 10 different species, respectively (see Supplementary TableS1 for details about the
696 selected species and Supplementary Fig. S3 for full PKHDL1 sequence alignment). IPT1 has a
697 sequence pairwise identity conservation of 82.3%, while IPT6 has a pairwise identity of 74.9%.
698 Missense variants are highlighted by green triangles. Blue circles represent cysteine residues
699 forming disulfide bonds showed in panels *f-i*. Each alignment was color-coded for sequence
700 similarity (35% threshold) using Jalview. White-colored residues show the lowest similarity and
701 dark blue report the highest (see Methods). Species were chosen based on sequence
702 availability and taxonomical diversity (Choudhary et al. 2020; De-la-Torre et al. 2018; Jaiganesh
703 et al. 2018). (f) Superposed AlphaFold2 models of both native *Hs* IPT1-2 (orange) and *Hs* IPT1-
704 2 p.(Gly120Ser) variant (mauve) are shown. (g) Higher magnification image of the mutated site.
705 S129 in lime and G129 in cyan. No apparent structural changes are predicted by AlphaFold2.
706 (h) Structural model of *Hs* IPT5-6 showing the native Gly1314 position. (i) Superposed *Hs* IPT5-
707 6 (orange) and *Hs* IPT5-6 p.(Gly1314Val) (mauve) structures showing a structural change
708 predicted by AlphaFold2 as a result of p.(Gly1314Val) substitution. β-strands and loops do not
709 overlap, with a dashed black arrow reporting the loop shift. (j) Higher magnification image of the
710 mutated site showing the conformation change of β-strands and loops. V1314 (lime) cause
711 steric hindrance in the area inducing an expanded conformation to the variant structure in
712 mauve. See dashed arrows.

713
714

715 **Fig. 3. PKHD1L1 structural modelling of the protein fragment containing the**
716 **p.(His2479Gln) variant.** Based on AlphaFold2 predictions, this fragment of PKHD1L1 shares a

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

common fold with the TMEM2 protein within the region carrying the p.(His2479Gln) variant. (a) Protein sequence alignment of a protein segment of *Hs* TMEM2 against 10 different species of PKHD1L1 proteins (See Supplementary Table S1 for details of the selected species, Supplementary Fig. S4 for sequence alignment of this specific fragment, and Supplementary Fig. S3 for full PKHDL1 sequence alignment). Residue numbering for TMEM2 as in PDB: 8C6I (Niu et al. 2023), while residue numbering for *Hs* PKHD1L1 as in NCBI accession code NP_803875.2 with the signal peptide included (Supplementary Table S1, 26 residues are suggested, see Methods). Green triangles point to the location of the human variant p.(His2479Gln), orange circles (*left*) indicate 100% amino acid sequence identity for this PKHD1L1 fragment between the *Hs*, *Pt*, and *Mm2* species (See supplementary Table S1 for details about the selected species). Green circles represent depicted residues in panels *b*-*e*. The alignment was color-coded for sequence similarity (35% threshold) using Jalview. White-colored residues show the lowest similarity and dark blue report the highest (see Methods). Species were chosen based on sequence availability and taxonomical diversity. (b) The simulated protein structure covering the protein fragment highlighted by purple arrow-headed line in Fig. 2b. Front view of the structure showing IPT14 linked to the PKHD1L1 TMEM2-like domain. The red star points to the linker connection. Residues at the mutation site are depicted as cyan sticks. (c) Side view from *panel a* showing a clear view of the stacked β-strand motifs. (d) Superposed structural protein alignment between WT *Hs* PKHD1L1 TMEM2-like domain model (orange) with the X-ray crystal structure of *Hs* TMEM2 protein (PDB: 8C6I, magenta). Residues at the native TMEM2-histidine finger site are depicted as green sticks. (e) Higher magnification image of the potential conserved histidine-finger site between PKHD1L1 (orange) and TMEM2 (magenta) protein fragments and the Ni^{2+} ion shown as lime sphere. (f) Displayed residues between both proteins surrounding the Ni^{2+}-ion site highlighted in *panel a* in green circles.

Fig. 4. Thermodynamic and folding stability evaluation of two missense variants using NanoDSF. (a) NanoDSF melting temperatures for WT *Mm* IPT1-3 (orange) and *Mm* IPT1-3 p.(Gly129Ser) variant (pink). Measurements show at least three T_m peaks (orange dotted line) for the WT IPT1-3, likely because the protein fragment includes multiple IPT domains that unfold sequentially. Measured T_m values are shifted to the left (pink dotted line) showing a decrease of the thermal folding stability. Temperatures are labeled for each T_m transition point. (b) Results for WT *Mm* IPT5-6 and *Mm* IPT5-6 p.(Gly1314Val) showing a reduced thermal stability (2 replicates, see Methods section). Traces correspond to the normalized first derivate of the fluorescence ratio showing the inflection point of the fluorescence ratio, which corresponds to the melting temperature of the sample. T_{onset} values and protein purification experiments are shown in Supplementary Fig. S5.

Table

Table 1. *PKHD1L1* variants identified.

Family	Chr8 Genomic location (g.)	Coding DNA (c.)	Protein (p.)	Zyg	AF gnomAD (v.3.1.2)	MAF gnomAD (v3.1.2)	MAF Pop gnomAD (v3.1.2))	AF gnomAD (v.2.1.1)	MAF gnomAD (v.2.1.1)	MAF Pop gnomAD (v2.1.1)	SIFT	PP-2	FATHMM	MT	REVEL	CADD
1	109382539G>A	385G>A	Gly129Ser	Het	0.000006576	0.00001471	European (non-Finnish)	-	-	-	0.165	0.995	-2.15	0.99	0.55	26.2
1	109439077G>T	3941G>T	Gly1314Val	Het	0.0003681	0.0007204	Ashkenazi Jewish	0.0002793	0.0005713	European (non-Finnish)	0.004	0.98	-2.56	1	0.86	26.6
2	109491899C>T	10141C>T	Arg3381Ter	Hom	0.00001979	0.0001936	East Asian	0.00004321	0.0002067	East Asian	-	-	-	1	-	39
3	109464269C>A	7437C>A	His2479Gln	Hom	0.00009864	0.003107	South Asian	0.0002259	0.001809	South Asian	0	0.88	-3.3	0.004	0.72	22.9

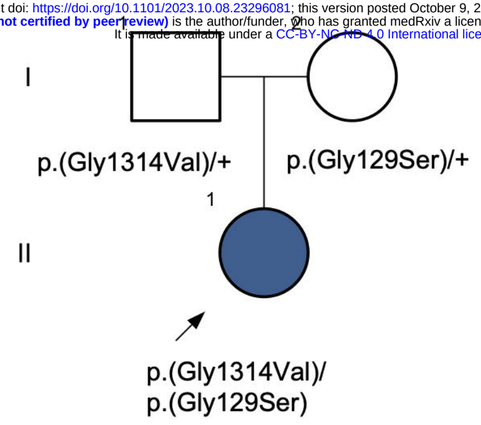
All variants are annotated according to NC_000008.11 (genomic) and NM_177531.6 (coding DNA). Residue position according to NP_803875.2 including the signal peptide (Supplementary Table S1). Abbreviations: AF, allele frequency; chr, chromosome; MAF, maximum allele frequency; MT, MutationTaster; Pop, population; PP2, PolyPhen-2; Zyg, zygosity

medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

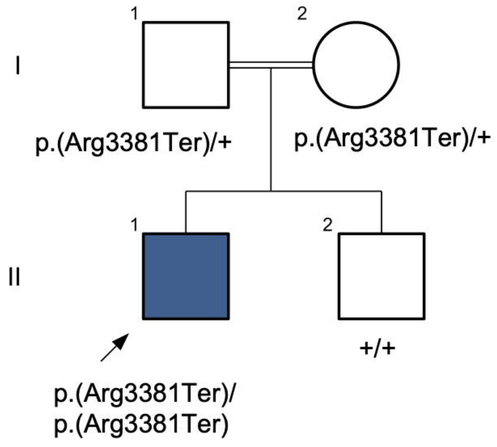
medRxiv preprint doi: https://doi.org/10.1101/2023.10.08.23296081; this version posted October 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

A

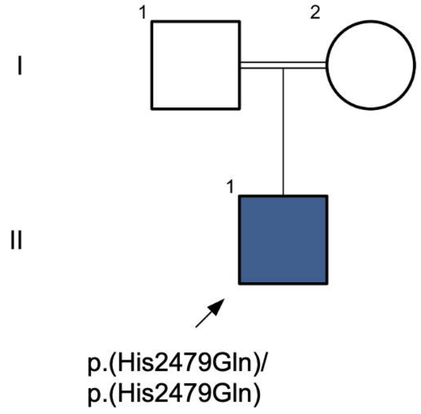
Family 1



Family 2

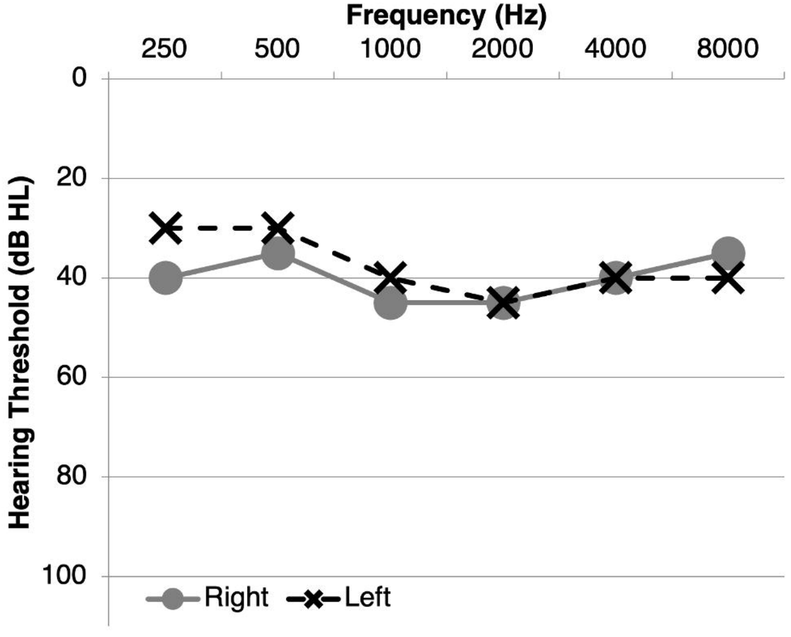


Family 3

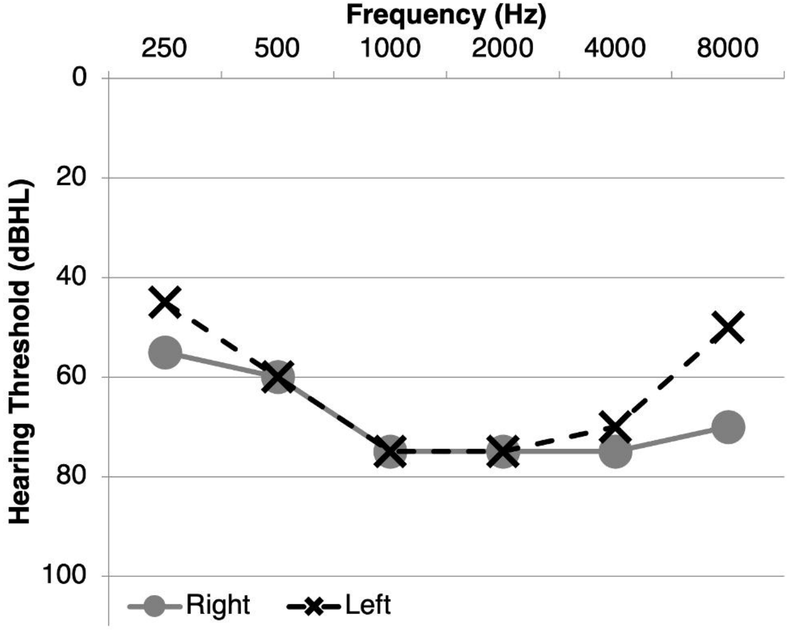


B

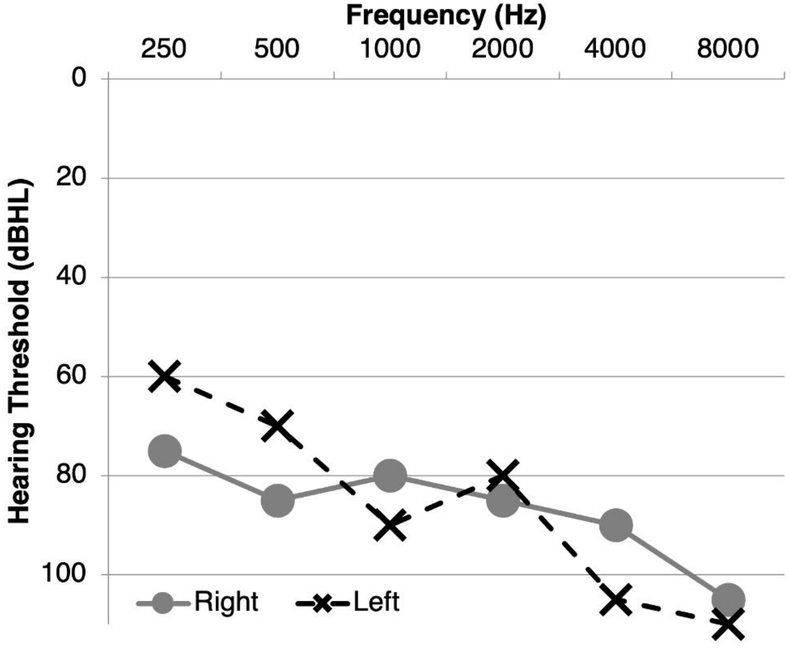
Family 1

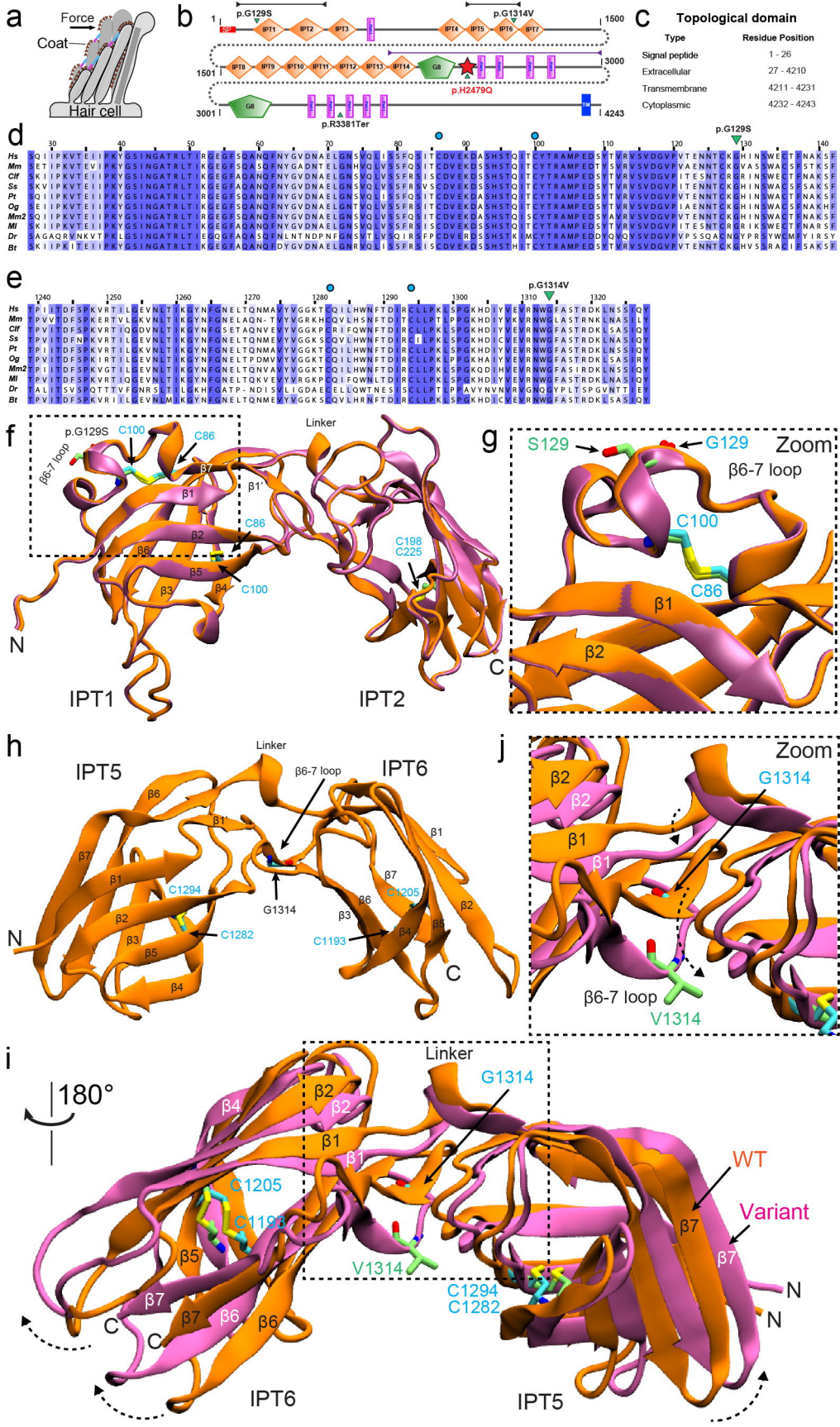


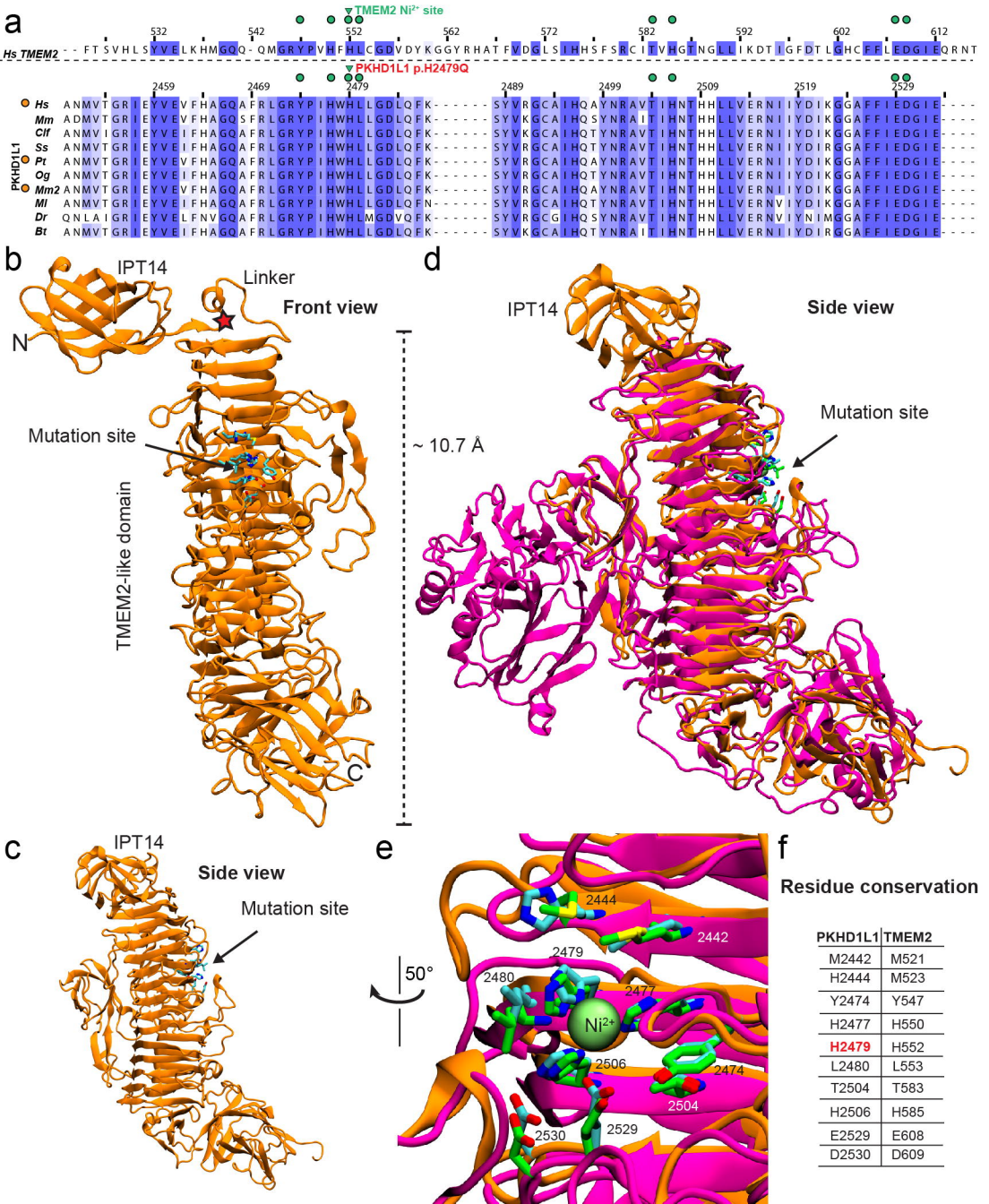
Family 2



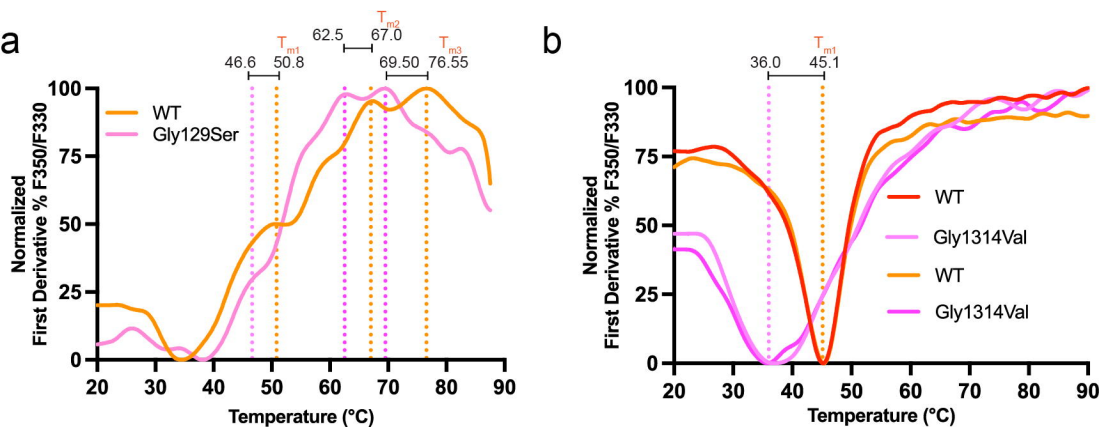
Family 3




a
Force
Coat
Hair cell
b
p.G129S
p.G1314V
SP
IPT1
IPT2
IPT3
IPT4
IPT5
IPT6
IPT7
IPT8
IPT9
IPT10
IPT11
IPT12
IPT13
IPT14
G8
TM
p.H2479Q
p.R3381Ter
1
1500
1501
3000
3001
4243
c
Topological domain
Type
Residue Position
Signal peptide
1 - 26
Extracellular
27 - 4210
Transmembrane
4211 - 4231
Cytoplasmic
4232 - 4243
d
p.G129S
Hs
Mm
Clf
Ss
Pt
Og
Mm2
Mf
Dr
Bt
e
p.G1314V
Hs
Mm
Clf
Ss
Pt
Og
Mm2
Mf
Dr
Bt
f
p.G129S
C100
C86
Linker
β6-7 loop
β7
β1
β1'
β2
C86
β6
β5
β3
β4
C100
C198
C225
N
C
IPT1
IPT2
g
S129
G129
Zoom
β6-7 loop
C100
C86
β1
β2
h
IPT5
Linker
β6-7 loop
IPT6
β6
β1'
β7
β1
C1294
G1314
β2
C1282
β3
β5
β4
C1205
β6
β3
C1193
β4
β2
β1
β5
N
C
j
Zoom
β2
β2
G1314
β1
β1
β6-7 loop
V1314
i
180°
Linker
G1314
β4
β2
β2
β1
β1
WT
Variant
C1205
C1193
β7
β7
β5
V1314
C1294
C1282
N
N
β7
β7
β6
β6
C
C
IPT6
IPT5



PKHD1L1	TMEM2
M2442	M521
H2444	M523
Y2474	Y547
H2477	H550
H2479	H552
L2480	L553
T2504	T583
H2506	H585
E2529	E608
D2530	D609


a
T_{m1}
T_{m2}
T_{m3}
46.6
50.8
62.5
67.0
69.50
76.55
WT
Gly129Ser
Normalized
First Derivative % F350/F330
100
75
50
25
0
20
30
40
50
60
70
80
90
Temperature (°C)
b
T_{m1}
36.0
45.1
WT
Gly1314Val
WT
Gly1314Val
Normalized
First Derivative % F350/F330
100
75
50
25
0
20
30
40
50
60
70
80
90
Temperature (°C)