Biallelic mutation of *CLRN2* causes non-syndromic hearing loss in humans

Barbara Vona[1,2], Neda Mazaheri[3*], Sheng-Jia Lin[4*], Lucy A Dunbar[5*], Reza Maroofian[6], Hela Azaiez[7], Kevin T. Booth[7,8], Sandrine Vitry[9], Aboulfazl Rad[2], Pratishtha Varshney[4], Ben Fowler[10], Kumar N. Alagramam[11,12,13], David Murphy[6], Gholamreza Shariati[14,15], Alireza Sedaghat[16], Henry Houlden[6], Shruthi VijayKumar[4], Richard J. H. Smith[7], Thomas Haaf[1], Aziz El-Amraoui[9], Michael R. Bowl[5,17**], Gaurav K. Varshney[4**], Hamid Galehdari[3**]

*These authors contributed equally

**These authors contributed equally

[1]Institute of Human Genetics, Julius Maximilians University Würzburg, Würzburg, Germany

[2]Department of Otolaryngology—Head and Neck Surgery, Tübingen Hearing Research Centre, Eberhard Karls University Tübingen, Tübingen, Germany

[3]Department of Genetics, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran

[4]Genes & Human Disease Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, United States

[5]Mammalian Genetics Unit, MRC Harwell Institute, Harwell Campus, OX11 0RD, UK

[6]Department of Neuromuscular Disorders, UCL Queen Square Institute of Neurology, London, WC1N 3BG, UK

[7]Molecular Otolaryngology and Renal Research Laboratories, Department of Otolaryngology and Interdisciplinary Graduate Program in Molecular Medicine, Carver College of Medicine, University of Iowa, Iowa City, IA, USA

[8]Harvard Medical School, Department of Neurobiology, Boston Massachusetts

[9]Unit Progressive Sensory Disorders, Pathophysiology and Therapy Institut Pasteur, Institut de l'Audition, INSERM-UMRS1120, Sorbonne Université, 63 rue de Charenton, 75012 Paris, France

[10]Imaging & Histology Core, Oklahoma Medical Research Foundation, Oklahoma City, OK, United States

[11]Department of Otolaryngology, University Hospitals Cleveland Medical Center, School of Medicine, Case Western Reserve University, 11100 Euclid Avenue, Cleveland, Ohio, 44106, USA

[12]Department of Neurosciences, Case Western Reserve University, 11100 Euclid Avenue, Cleveland, Ohio, 44106, USA

[13]Department of Genetics and Genomic Sciences, Case Western Reserve University, Cleveland, Ohio 44106, USA

[14]Department of Medical Genetics, Faculty of Medicine, Ahvaz Jundishapur, University of Medical Sciences, Ahvaz, Iran

[15]Narges Medical Genetics and Prenatal Diagnostics Laboratory, East Mihan Ave., Kianpars, Ahvaz, Iran

[16]Diabetes Research Center, Health Research Institute, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

[17]UCL Ear Institute, University College London, 332 Gray's Inn Road, WC1X 8EE, London, UK

Correspondence:

Barbara Vona : Email: barbara.vona@uni-wuerzburg.de, Tel: +49 (0)931-31-84244

Current Email: barbara.vona@uni-tuebingen.de, Tel: +49 (0)7071-29-88154

Michael R. Bowl: Email: m.bowl@har.mrc.ac.uk, Tel: +44 (0)1235 841161

Disclosure

The authors declare no conflict of interest.

Abstract

Deafness, the most frequent sensory deficit in humans, is extremely heterogenous with hundreds of genes probably involved. Clinical and genetic analyses of an extended consanguineous family with pre-lingual, moderate-to-profound autosomal recessive sensorineural hearing loss, allowed us to identify *CLRN2,* encoding a tetraspan protein as a new deafness gene. Homozygosity mapping followed by exome sequencing identified a 15.2 Mb locus on chromosome 4p15.32p15.1 containing a missense pathogenic variant in *CLRN2* (c.494C>A, NM_001079827.2) segregating with the disease. Using *in vitro* RNA splicing analysis, we show that the *CLRN2* c.494C>A mutation leads to two events: 1) the substitution of a highly conserved threonine (uncharged amino acid) to lysine (charged amino acid) at position 165, p.(Thr165Lys), and 2) aberrant splicing, with the retention of intron 2 resulting in a stop codon after 26 additional amino acids, p.(Gly146Lysfs*26). Expression studies and phenotyping of newly produced zebrafish and mouse models deficient for clarin 2 further confirm that clarin 2, expressed in the inner ear hair cells, is essential for normal organization and maintenance of the auditory hair bundles, and for hearing function. Together, our findings identify *CLRN2* as a new deafness gene, which will impact future diagnosis and treatment for deaf patients.

Keywords: autosomal recessive hearing loss (DFNB), clarin 2 (*CLRN2*); deafness; homozygosity mapping; non-syndromic hearing loss

Introduction

The mammalian inner ear is an exquisite and highly complex organ, made up of the vestibule, the organ responsible for balance, and the cochlea, the sensory organ for hearing. The auditory sensory cells of the inner ear are called the inner and outer hair cells that are responsible for transduction of sound wave-induced mechanical energy into neuronal signals (Gillespie and Müller, 2009; Hudspeth, 1997). The functional mechanoelectrical transduction machinery involves intact formation and maintenance of a highly specialized and organized structure, the hair bundle. The hair bundle contains a few dozen F-actin-filled stereocilia, arranged in a highly interconnected and highly organized staircase-like pattern, which is critical for function (Kazmierczak et al., 2007). Knowledge of the mechanisms of formation, maintenance, and function of the transduction complex is limited (Cunningham and Müller, 2019). In this regard, identification of novel genes that encode protein products essential for hearing is likely to improve our understanding of the physical, morphological and molecular properties of hair cells and associated mechanistic processes.

Hereditary hearing loss is one of the most common and genetically heterogeneous disorders in humans (Wright et al., 2018). Sensorineural hearing loss has an incidence of one to two per 1000 at birth (Morton and Nance, 2006). It displays extraordinary phenotypic, genetic and allelic heterogeneity, with up to 1,000 different genes potentially involved (Ingham et al., 2019). So far, about 120 genes and more than 6,000 disease causing variants (Azaiez et al., 2018) have been identified as responsible for non-syndromic hearing loss in humans (see http://hereditaryhearingloss.org/ and http://deafnessvariationdatabase.org/), and many more are yet to be discovered.

Genetic factors predominate the etiological spectrum and most of hereditary hearing loss appears to follow an autosomal recessive inheritance pattern (Smith et al., 2005). Approximately 80% of the currently known autosomal recessive genes have been identified by studying extended consanguineous families (Hofrichter et al., 2018). There are many forms of hearing loss that are clinically indistinguishable but caused by distinct genetic entities that are presently unknown. Identification of additional genes essential for auditory function, through the study of families exhibiting hereditary hearing loss, will not only help increase our understanding of the biology of hearing, but will also identify new molecular targets for therapeutic intervention.

Through the study of an extended consanguineous Iranian family, we have identified a *CLRN2* coding lesion as the cause of hearing loss in family members that are homozygous for the allele. We have established that clarin 2 likely plays a critical role in mechanotransducing stereocilia of the hair bundle in zebrafish and mouse. *CLRN2* belongs to the clarin (CLRN) family of proteins that are comprised of three orthologues named clarin 1, 2, and 3 that encode four-transmembrane domain proteins. Pathogenic variants in *CLRN1* (clarin 1) cause either non-syndromic retinitis pigmentosa (RP) (Khan et al., 2011) or Usher syndrome type 3A (USH3A), that is characterized by progressive hearing loss, RP and variable vestibular dysfunction (Adato et al., 2002; Joensuu et al., 2001; Ness et al., 2003; Plantinga et al., 2005). This study establishes clarin 2 as essential for inner ear function in zebrafish, mice and humans, with a loss-of-function allele leading to autosomal recessive non-syndromic sensorineural hearing loss (ARNSHL).

Materials and Methods

Patient clinical and audiometry data

A three generation Iranian family of Lurs ethnicity was ascertained as part of a large ethnically diverse Iranian population rare disease study. After obtaining written informed consent from all participants with approval by the Faculty of Medicine ethics commissions at the University of Würzburg (46/15) and Shahid Chamran University of Ahvaz (#EE/97.24.3 17654), pure-tone audiograms and medical information were collected from participating members. Clinical examination excluded additional syndromic features.

Individuals IV-1, IV-6, and V-1 (Figure 1) underwent complete ear, nose and throat examination, including binocular ear microscopy and external ear inspection. Routine pure-tone audiometry was performed according to current standards that measured hearing thresholds at frequencies 0.25, 0.5, 1, 2, 4, 6 and 8 kHz. Both air- and bone-conduction thresholds were determined. Severity of hearing loss was defined as previously described (Mazzoli et al., 2003). Individuals IV-1 and IV-6 underwent additional tympanometry and speech recognition threshold testing. Audiometry testing for individuals IV-1, IV-6, and V-1 was performed at ages 29 years, 44 years, and 20 years, respectively.

Genotyping, homozygosity mapping, copy number variation and exome sequencing data analyses

Due to parental consanguinity and suspected autosomal recessive mode of inheritance, we assumed that the causative variant would be homozygous and identical

by descent in affected individuals in the fourth generation of the family. Blood samples from 14 family members were obtained and genomic DNA was isolated from whole blood using standard procedures. DNA from affected (IV-1, IV-6, and IV-8) and unaffected (IV-2, IV-3, IV-4, and IV-5) individuals were genotyped using the Infinium Global Screening Array-24 v1.0 BeadChip (Illumina, SanDiego, CA, USA) according to manufacturer's protocols. Homozygosity mapping was performed using Homozygosity Mapper to identify common homozygous intervals among the affected individuals (Seelow et al., 2009). Runs of homozygosity with a maximum threshold of 0.99 were checked after the exome-wide analysis was completed. Copy number variation calling was performed using GenomeStudio v.2011.1 and cnvPartition 3.2.0 (Illumina).

For exome sequencing, DNA samples from two affected individuals (IV-1 and IV-6) were used. The data from individual IV-6 was analyzed exome-wide and data from individual IV-1 was used for determination of allele sharing. Exome capture using genomic DNA was performed using the SureSelect Target Enrichment v6 (Agilent) kit following manufacturer's recommendations. The libraries were sequenced with a HiSeq4000 (Illumina). Data analysis was performed using the Burrows-Wheeler Alignment (BWA) tool for read mapping to the human reference genome GRCh37 (hg19), Picard for duplicate removal, GATK for local re-alignment, base recalibration, variant calling, and variant annotation, and SnpEff for variant annotation. Variant filtering was based on: coverage > 10X, Phred quality score ≥ 30, and minor allele frequency (MAF) ≤ 0.005 as reported in 1000 Genomes Project and EVS6500. Variants were filtered based on coding effect (non-synonymous, indels, and splice site variants), and artifact-prone genes (*HLA*s, *MAGE*s, *MUC*s, *NBPF*s, *OR*s, *PRAME*s) were excluded.

Visualization was performed using the Integrative Genomics Viewer. Analysis of homozygous and compound heterozygous variants between the two sequenced affected individuals (IV-6 and IV-1) followed. We analyzed missense variants by using a combination of criteria that scored conservation using GERP++ and PhyloP, and deleterious or pathogenic scores in Combined Annotation Dependent Depletion (CADD) (Kircher et al., 2014), LRT (Chun and Fay, 2009), MutationTaster (Schwarz et al., 2014), PolyPhen-2 (Adzhubei et al., 2010), and SIFT (Ng and Henikoff, 2001). Missense variants were excluded when three out of five *in silico* pathogenicity prediction tools yielded a benign score. Manual MAF analysis used gnomAD (Lek et al., 2016), GME (Scott et al., 2016), and Iranome (Fattahi et al., 2019). Potential effects on splicing were assessed using ESEfinder (Cartegni et al., 2003) and RESCUE-ESE (Fairbrother et al., 2004).

Segregation, sequence and *in vitro* splicing analyses of the *CLRN2* c.494C>A pathogenic variant

To confirm segregation of the *CLRN2* c.494C>A; p.(Thr165Lys) (NM_001079827.2) homozygous variant, Sanger sequencing was completed in all 14 family members using the following primers (*CLRN2* Ex3 F: 5'-AAATGCCACCTCTTACAGAGTTGC-3' and *CLRN2* Ex3 R: 5'-ACCGTGGCCTCTTCGATTTTGGTC-3') and standard PCR and sequencing parameters.

To document residue conservation, CLRN1 (UniProt: P58418) and CLRN2 (UniProt: A0PK11) were aligned and visualized in Jalview (Waterhouse et al., 2009) with an overview of the pathogenic and likely pathogenic missense and nonsense

188 *CLRN1* variants retrieved from the Deafness Variation Database v 8.2 (Azaiez et al.,
189 2018).

190 In addition, secondary protein structure prediction of human CLRN2
191 (NP_001073296.1) that included the wild-type (WT) and mutated amino acid residues
192 was performed using I-TASSER (Yang et al., 2015).

193 To assess the splicing effect of the c.494C>A variant, *in vitro* splicing assays,
194 also called mini-genes, were carried out as described (Booth et al., 2018a; Booth et al.,
195 2018b). Concurrently, the mini-gene splice assay experiment was conducted in a
196 double-blind manner as previously described (Lekszas et al., 2020). A detailed
197 description of this method can be found in the Supplementary Methods.

198 **CRISPR/Cas9-mediated inactivation of *clrn2* in zebrafish**

199 Zebrafish (*Danio rerio*) were raised and maintained in an AALAC accredited
200 facility at the Oklahoma Medical Research Foundation (OMRF) under standard
201 conditions. Zebrafish embryos/larvae were maintained in embryo medium with
202 0.00002% methylene blue and raised at 28˚C. All animal experiments were performed
203 as per protocol (17-01) and approved by the Institutional Animal Care Committee of
204 OMRF (IACUC). All zebrafish handling, embryo care, and microinjections were
205 performed as previously described (Westerfield, 2000). WT zebrafish strain NHGRI-1
206 was used for all experiments (LaFave et al., 2014). The zebrafish embryonic staging
207 was determined by morphological features according to Kimmel et al (Kimmel et al.,
208 1995).

209 To produce zebrafish *clrn2* crispants, the sgRNA target sequences were selected
210 from the UCSC genome browser tracks generated by the Burgess lab. Five

independent targets were chosen and sgRNAs were synthesized by *in vitro* transcription as described earlier (Varshney et al., 2016). sgRNAs and Cas9 protein complex were used to generate indels. A 6 µL mixture containing 2 µL of 40 µM Spy Cas9 NLS protein (New England Biolabs, MA, USA), 200 ng each of five sgRNAs (in 2 µL) and 2 µL of 1 M potassium chloride was injected into one-cell-stage WT embryos. Injection volumes were calibrated to 1.4 nL per injection. Insertion/deletion (indel) variants were detected by amplifying the target region by PCR and Sanger sequencing as described earlier (Varshney et al., 2016). The sequencing data were analyzed by Inference of CRISPR Edits (ICE) v2 CRISPR analysis tool. The sgRNA target sequences and PCR primer sequences are listed in Table S1.

Zebrafish RNA extraction and real-time quantitative PCR (RT-qPCR)

Total RNA at different developmental stages, adult tissues, and CRISPR/Cas9 injected larvae were extracted using the TRIzol Reagent (Thermo Fisher Scientific, CA, USA) and purified by RNA clean and concentrator-5 kit (Zymo Research, CA, USA) according to the manufacturer's instructions. RNA concentration was measured by DeNovix DS-11 spectrophotometer (DeNovix Inc. USA). The cDNA was synthesized by iScript RT Supermix (Bio-Rad, USA), and was used as a template for performing the RT-qPCR with SYBR Green Supermix (Thermo Fisher Scientific, CA, USA) and the Light Cycler® 96 System (Roche, CA, USA). All RT-qPCR reactions were carried out using three biological and technical replicates. The housekeeping gene *18S* was used as a reference gene.

All RT-qPCR primer pairs were designed across exon-exon junctions using NCBI Primer-BLAST program. The sequences are listed in Table S1. The PCR cycling

conditions were used as per the manufacturer instructions, and the amplification specificity was assessed by dissociation curve analysis at the end of the PCR cycles. The cycle threshold values (Ct) data were imported into Microsoft Excel for the relative gene expression analysis. Quantification was based on 2^(-ΔΔCT) method (Livak and Schmittgen, 2001), and using 18 hours post fertilization (hpf) for *clarin 2* temporal expression, muscle for *clarin 2* in different tissue expression and the corresponding age-matched control for *clarin 2* CRISPR injected F_0 larvae as normalization control.

Distribution of *clrn2* and phalloidin staining in zebrafish

To determine *clrn2* expression, we used in situ hybridization on larvae and inner ear-containing cryosections. The full-length coding sequence of zebrafish *clarin 2* (NM_001114690.1) was PCR amplified from WT zebrafish cDNA using primer pairs with restriction enzymes *Bam*HI and *Xho*I restriction sites cloned into pCS2+ vector (a kind gift from Dr. Dave Turner, University of Michigan). After restriction digestion, the resulting clones were sequenced and used as templates for riboprobe synthesis. The digoxigenin-UTP-labeled riboprobes were synthesized according to the manufacturer's instructions (Millipore Sigma, MO, USA). Briefly, the *clarin 2* and the *pvalb9* plasmids (Horizon Discovery) were linearized by *Bam*HI and *Not*I restriction enzymes, respectively. The linearized plasmid was purified and used as template for in vitro transcription using T7 RNA polymerase to synthesize anti-sense probes. The sense probe was made using *Xba*I linearized *clarin 2* plasmid and SP6 RNA polymerase.

The whole-mount in situ hybridization (WISH) on 3 and 5 dpf zebrafish embryos/larvae was performed following the procedures as described by Thisse et al.

with minor modifications that can be found in the Supplementary Methods (Thisse and Thisse, 2008).

For preparation of cryo-section samples after WISH, the 5 dpf larvae were soaked in 25%, 30% (V/V) sucrose/PBS and optimum cutting temperature (OCT) each for at least two days, and embedded in OCT then Cryotome sectioned at a 10-micrometer thickness.

For phalloidin staining of the zebrafish inner ear, 5 dpf larvae were euthanized with tricaine and fixed in 4% (V/V) paraformaldehyde (PFA) at 5 dpf, fixed embryos were washed by PBSTx (1% PBS, 0.2% triton X-100) and incubated in 2% triton X-100 in PBS at room temperature for overnight with agitation until the otoliths were completely dissolved. The larvae were sequentially washed in PBSTx and incubated with Alexa Fluor 647 Phalloidin (1:100) (Thermo Fisher Scientific, CA, USA) in PBSTw (1% PBS, 0.1% Tween-20) at room temperature for 4 hours. The samples were washed in PBSTx after staining and mounted laterally in 75% glycerol on slides. Images were acquired with a Zeiss LSM-710 Confocal microscope.

Production and phenotyping of clarin 2 deficient mutant in mice

The *Clrn2*^{del629} mutant line was generated on a C57BL/6N background by the Molecular and Cellular Biology group at the Mary Lyon Centre (MLC), MRC Harwell Institute, using CRISPR/Cas9 genome editing (Dunbar et al., 2019). The mice were housed and maintained under specific pathogen-free conditions in individually ventilated cages, with environmental conditions as outlined in the Home Office Code of Practice. Animals were housed with littermates until weaned, and then housed with mice of the same sex and of similar ages, which was often their littermates. Both male and female

279 animals were used for all experiments. Animal procedures at the MRC Harwell Institute
280 were licenced by the Home Office under the Animals (Scientific Procedures) Act 1986,
281 UK and additionally approved by the Institutional Animal Welfare and Ethical Review
282 Body (AWERB). The *Clrn1*$^{-/-}$ mice (*Clrn1*$^{tm1.2Ugpa}$, MGI: 6099052) used for comparative
283 scanning electron microscopy analyses were previously described (Dulon et al., 2018).
284 Auditory Brainstem Response (ABR) tests were performed using a click stimulus
285 and frequency-specific tone-burst stimuli (at 8, 16 and 32 kHz) to screen mice for
286 auditory phenotypes and investigate auditory function (Hardisty-Hughes et al., 2010).
287 Distortion Product Oto-Acoustic Emission (DPOAE) tests were performed using
288 frequency-specific tone-burst stimuli from 8 to 32 kHz with the TDT RZ6 System 3
289 hardware and BioSig RZ software (Tucker Davis Technology, Alachua, FL, USA).
290 Detailed descriptions of ABR and DPOAE testing can be found in the Supplementary
291 Methods.
292 Mice were euthanized by cervical dislocation and inner ears were removed and
293 fixed in 2.5% glutaraldehyde (TAAB Laboratories Equipment Ltd.) in 0.1 M phosphate
294 buffer (Sigma-Aldrich) overnight at 4ºC. Following decalcification in 4.3% EDTA,
295 cochleae were sub-dissected to expose the sensory epithelium then 'OTO processed'
296 with alternating incubations in 1% osmium tetroxide (TAAB Laboratories Equipment
297 Ltd.) in 0.1 M sodium cacodylate (Sigma-Aldrich) and 1% thiocarbohydrazide (Sigma-
298 Aldrich) in ddH_2O. Ears were dehydrated through a graded ethanol (Fisher Scientific)
299 series (25% to 100%) at 4ºC and stored in 100% acetone (VWR Chemicals) until critical
300 point drying with liquid CO_2 using an Emitech K850 (EM Technologies Ltd). Ears were
301 mounted onto stubs using silver paint (Agar Scientific), sputter coated with palladium

using a Quorum Q150R S sputter coater (Quorum Technologies) and visualised with a JSM-6010LV Scanning Electron Microscope (JEOL). Micrographs were pseudo-coloured in Adobe Photoshop.

Results

Identification of *CLRN2* as a novel deafness gene in a consanguineous Iranian family exhibiting autosomal recessive non-syndromic sensorineural hearing loss

A three generation Iranian family of Lurs ethnicity was ascertained as part of a large ethnically diverse Iranian population rare disease study (Fig. 1A). Three individuals that included the proband (IV-6), his sibling (IV-1), and a cousin (IV-8), born form consanguineous marriages, have reported moderate-to-profound bilateral non-syndromic sensorineural hearing loss (Fig. 1B). The age of onset for these three individuals was between 2 and 3 years of age. Pure-tone air- and bone-conduction audiometry thresholds (Fig. 1B) show evidence of intrafamilial variability. Individual IV-1 has a down sloping audiogram, with bilateral moderate-to-profound deafness. Individual IV-6 presented a moderate-to-severe hearing loss with slightly better hearing at higher frequencies. Both individuals showed normal (type A) tympanograms bilaterally. Speech recognition thresholds for individual IV-1 were 80 dB and 75 dB at 84% and 88% for right and left ears, respectively, and a most comfortable level of 95 dB. Speech recognition thresholds for individual IV-6 were 75 dB and 80 dB, each at 84%, for right and left ears, respectively. Patients have normal neuromotor, speech and language development, and did not show signs of impaired balance. No other abnormalities, including potential vision deficit, were present in the affected individuals, who were last evaluated at the age of 29 (IV-1), 44 (IV-6), and 25 (IV-8) years. For comparison, pure-tone audiometry was also recorded from a family member (V-1), with no reported history of hearing problem.

To identify the underlying genetic lesion, we applied homozygosity mapping in the extended family to identify a 15.2 Mb locus on chromosome 4p15.32p15.1 (GRCh37/hg19, chr4:17,298,445-32,495,165), defined by the SNPs rs7692897 and rs17081424 (Fig. 1C, Fig. S1A, Table S2). This locus contains 30 genes, none of which are presently associated with deafness in humans (Table S3). This approach also revealed four much smaller homozygous intervals on chromosomes 2p21 (137.3 kb), 3p22.2 (262.5 kb), 13q13.1 (90.7 kb), and 17q21.31 (292.6 kb) (Fig. S1A, Table S2) that do not contain known deafness-associated genes (Table S3). Pathogenic copy number variations were excluded. Next, we undertook exome sequencing of affected individual IV-6 (arrow, Fig. 1A). This generated 56,387,543 mappable reads, with 75.5% on-target reads. The mean depth was 57.3-fold, with 97.3% of regions with a 10-fold read depth. Analysis of the exome data of individual IV-6 excluded any candidate pathogenic variants in known deafness-associated genes (Doll et al., 2020) prompting an exome-wide analysis followed by filtering and re-analysis of variants in homozygous intervals (Table S4). Further, close inspection of the exome sequencing data revealed complete sequencing coverage of genes in the homozygous intervals (Table S5). Variant filtering detected a homozygous missense variant in *CLRN2* c.494C>A, (p.(Thr165Lys)) (NM_001079827.2) in the 15.2 Mb homozygous interval on chromosome 4 (Fig. S1A, S1B). This variant was shared with individual IV-1 and segregated in the extended family comprising a total of 14 individuals (Fig. 1A, D). Only individuals homozygous for the *CLRN2* c.494C>A variant exhibit hearing loss confirming the recessive nature of the allele (Fig. 1A).

The *CLRN2* c.494C>A leads to a pathogenic missense substitution and aberrant splicing

The c.494C>A variant on chromosome 4p15.32 is unanimously predicted to be deleterious and disease causing by *in silico* tools (Table S6). The c.494C>A variant in *CLRN2* replaces a polar uncharged amino acid (threonine) with a positively charged amino acid (lysine) in clarin 2, (p.(Thr165Lys)) (Creixell et al., 2012). This variant, as well as homozygous loss-of-function alleles are absent in population frequency databases. This suggests *CLRN2* is intolerant to biallelic loss-of-function. Our in-house collection of 16,041 additional exomes of simplex patients with predominantly Iranian, Pakistani and Egyptian ethnicities, including 462 exomes of probands with autosomal recessive hearing loss, identified only one hearing impaired individual who carried the c.494C>A variant in *CLRN2* (allele frequency $3.12x10^{-5}$). An additional multiethnic cohort (Iranian, European American and Indian) that consists of 380 probands with autosomal recessive hearing loss ranging from moderate-to-profound was also screened for variants in *CLRN2*, but none were identified.

The c.494C>A variant involves the exchange of a novel polar threonine (Thr) residue to a basic lysine (Lys) amino acid that affects a highly conserved amino acid in the alpha-helix of the PMP-22/EMP/MP20/Claudin superfamily domain (Fig. 2A-C). Among clarin proteins, clarin 2 and clarin 1 show 34.9% identity with 81 identical and 91 similar amino acids (using UniProt (UniProt Consortium, 2018), Fig. 2B). The outcome of *CLRN1* pathogenic or likely pathogenic missense variants, as well as nonsense variants (queried from the Deafness Variation Database v8.2 (Azaiez et al., 2018)) are marked in red (Fig. 2B) along with the clarin 2 p.(Thr165Lys) amino acid substitution

(Fig. 2B, asterisk). Interestingly, nine out of the 19 clarin 1 amino acid mutated residues are identical in clarin 2. Three clarin 1 amino acid substitutions (p.(Leu163Pro), p.(Leu167Trp), and p.(Ile181Asn), NP_001182723.1) align in close proximity to the clarin 2 p.(Thr165Lys). Furthermore, clarin 1 p.Leu163Pro (Fields et al., 2002) and p.Ile181Asn (García-García et al., 2012), that are both reported in USH3A, are p.Leu150 and p.Ile168 in clarin 2. Most importantly, the threonine residue at position 165 (Thr165) CLRN2 is conserved across species and the corresponding amino acid in clarin 1 is a serine residue (Fig. 2A,B), a scenario often associated with conserved phosphorylation site residue, here by serine/threonine protein kinases (Creixell et al., 2012).

In addition to causing an amino acid missense substitution, computational analysis also predicts that the c.494C>A variant will create an exonic splicing enhancer (ESE) motif, modifying the ESE hexameric sequence landscape of exon 3, which could interfere with the normal processing of *CLRN2* mRNA (Figs. S2A, S3B; ESEfinder and RESCUE-ESE, Human Splicing Finder) (Cartegni et al., 2003; Desmet et al., 2009; Fairbrother et al., 2004). To investigate the effect of the c.494C>A variant on *CLRN2* splicing, we used mini-gene assays using two different exon-trapping vectors and three different cell lines, Cos-7, ARPE-19, and HEK 293T. The mini-gene contained the 3' end of intron 2, all of exon 3 (with and without the *CLRN2* variant), and ~50 bp of the 3' UTR (Fig. 3A) and was transfected into COS-7 and ARPE-19 cells. As a negative *CLRN2* control, we used the rs117875715 SNP, a common polymorphism, with a global minor allele frequency of ~1.25% and >100 homozygous alleles reported in gnomAD (Lek et al., 2016) (http://gnomad.broadinstitute.org/variant/4-17528480-G-A) that is 20

nucleotides away from c.494C>A. Given its frequency, rs117875715 is predicted to be benign for hearing loss. Of note, this polymorphism is absent in the proband and family members reported here. Since exon 3 is the last exon of *CLRN2*, we designed our PCR primers to exclude the human poly-A signal and used the poly-A signal native to the pET01 vector. As expected for WT *CLRN2* (c.494C), we detected the splicing of the 5' native pET01 exon only to exon 3 of *CLRN2* (Fig. 3A, B). The same normal splicing was obtained in all cell types transfected with *CLRN2* containing the control (rs117875715) variant (Fig. 3B). However, the c.494C>A variant yielded two bands; one ~650 bp band matching the expected normally spliced exon, and a second abnormal band that was approximately ~1,360 bp (Fig. 3B). Sequencing of these amplicons validated normal splicing including the c.494A variant and also revealed a retained intron 2 in the aberrantly spliced transcript (Fig. S3C). The retention of intron 2 results in a new reading frame that introduces a stop codon 26 amino acids after the native exon 2 splice site (p.(Gly146Lysfs*26)) (Fig. 3C). These results were replicated using the pSPL3b vector and HEK 293T cells (Fig. S3A-C), confirming the c.494C>A induced normal and aberrant splicing, independent of the cell type context. Following TA-cloning of cDNA amplicons from the homozygous individual (from Fig. S3B), 23 of 26 amplicons (88.5%) showed normal splicing, and 3 of 26 amplicons (11.5%) showed a retained intron.

The mini-gene splicing assays and sequence analyses clearly show that the c.494C>A affects a highly conserved and key residue in clarin 2 sequence, while also creating aberrant mRNA splicing *in vivo* likely leading to a truncated protein. Altogether, this further confirms that variants in *CLRN2* can lead to sensorineural hearing loss.

Clrn2, a hair cell expressed gene key to hearing also in zebrafish and mice

To further study the role of clarin 2 in the inner ear, we investigated its expression and analyzed potential impact of *Clrn2* loss-of-function in two other species, zebrafish and mice.

clrn2 in zebrafish

Taking advantage of larva transparency, we used zebrafish as a model to investigate the *clarin 2* expression during early embryonic development. The RT-qPCR at different developmental stages revealed that *clrn2* mRNA was first detected at 18 hpf (Fig. 4A), a stage when the otic placode begins to form the otic vesicle in zebrafish (this stage is similar to mouse embryonic day 9 (E9), a stage of otic placode formation) (Kopecky et al., 2012; Whitfield et al., 2002). *clrn2* mRNA expression increased (2-fold at 72 and 96 hpf compared to 18 hpf) and was maintained at later stages, up to 120 hpf (Fig. 4A). Comparative analyses of *clrn2* mRNA expression in different adult tissues of zebrafish revealed a significant enrichment in utricle, saccule and lagena of the inner ear (Fig. 4B). Our data are in agreement with RNA expression data from the Genotype-Tissue Expression (GTEx) project, wherein *CLRN2* mRNA in humans is enriched in the nervous system, testis, kidney, salivary gland, and lung. *CLRN1* has a similar expression profile in humans.

To determine *clrn2* cellular expression, we used WISH in the inner ear of 3- and 5-dpf embryos (Fig. 4C, D). Unlike the *clrn2* sense probe, the anti-sense c*lrn2* revealed strong expression in the otic vesicle, similar to the expression of anti-sense *pvalb9*, used as a marker of hair cells (Fig. 4C). Histological examination of 5 dpf embryos

440 further confirmed that *clrn2* expression is more specifically, restricted to hair cells, and is
441 not expressed in the supporting cells of the inner ear (Fig. 4D).

442 To elucidate the function of *clrn2* in zebrafish, we used CRISPR/Cas9 to
443 generate loss-of-function alleles. To maximize the knockout efficiency, we used five
444 sgRNAs targeting the first and second exon of *clrn2* gene (Fig. S4). Injected embryos
445 (crispants) were sequenced and, as expected, a mix of alleles in the form of deletions
446 ranging from 4 bp to 73 bp, as well as insertions spanning +1 to +11 bp were observed.
447 The majority of the variants were frameshift that would most likely create a premature
448 stop codon in the protein (Fig. S4). The RT-qPCR analyses on injected embryos
449 showed that c*lrn2* crispants have a significantly reduced amount of *clrn2* mRNA (Fig.
450 4E), suggesting nonsense mediated decay, leading to disrupted clarin 2 protein
451 function.

452 Considering the expression in hair cells (Fig. 4D), we investigated the
453 mechanosensory structures of the hair cell bundle, which are important for hearing and
454 balance function in zebrafish. Interestingly, fluorescent phalloidin staining of the hair
455 bundles of the inner ear in *clrn2* crispants (n=10) showed disrupted hair bundle structure
456 compared to the WT controls; the hair bundles are splayed, thin and split in *clrn2*
457 crispants (arrowheads in Fig. 4E). This defective phenotype, suggesting a critical role in
458 hair bundle structures, is similar to the hair bundles in zebrafish *clrn1* knockouts (Gopal
459 et al., 2015), the *orbiter* mutants (defective in *protocadherin 15* (*pcdh15*), a gene
460 associated with human Usher syndrome 1F) (Seiler et al., 2005) and *ush1c* morphants
461 and *ush1c* mutants (Phillips et al., 2011).

462 ***Clrn2 in mice***

To further assess the requirement of clarin 2 for auditory function in mammals, and assess further its role in auditory hair bundles, we extended our analyses to mouse. Consistent with expression data in zebrafish (Fig. 4A, C, D), single cell RNA-seq data available to visualize on the gEAR portal (*umgear.org*) show that in the mouse cochlear epithelium at postnatal day 1 (P1) and P7, *Clrn2* transcripts are almost exclusively detectable only in inner and outer hair cell populations (Kolla et al., 2020) (see also Fig. S5). We utilized a CRISPR/Cas9-engineered *Clrn2* loss-of-function mouse mutant, in which exon 2 has been deleted (*Clrn2*^{del629}) (Fig. 5A), and measured ABRs in P21 (± 1 day) mice in response to click and tone-burst stimuli.

Analysis of ABR thresholds, which is the lowest sound stimulus required to elicit measurable activity in the auditory nerve, showed that homozygous (*Clrn2*$^{del629/del629}$) mice display very elevated thresholds (>80 decibel sound pressure level (dB SPL)) at all frequencies tested: 8, 16 and 32 kHz (Fig. 5B). Whereas, *Clrn2*$^{del629/+}$ mice exhibit thresholds comparable with those of WT (*Clrn2*$^{+/+}$) littermates (<40 dB SPL), demonstrating the absence of a heterozygous auditory phenotype (Fig. 5B).

To further assess cochlear function, DPOAEs were measured in P28 (± 1 day) *Clrn2*$^{del629/del629}$ mice. Compared to their *Clrn2*$^{+/+}$ and *Clrn2*$^{del629/+}$ littermates, *Clrn2*$^{del629/del629}$ mice have reduced DPOAEs (Fig. 5B) suggesting impaired outer hair cell (OHC) function.

To investigate stereocilia bundle morphology in *Clrn2*$^{del629/del629}$ mice, we used scanning electron microscopy to examine the cochlear sensory epithelia. At P28 (± 1 day), the inner and outer hair cell stereocilia bundles of *Clrn2* mutant mice display the expected U- and V- shape, respectively, which contrasts with the grossly misshapen

OHC bundles found in *Clrn1* mutant mice (Fig. 5C). However, while the patterning of the bundles appears normal in *Clrn2*$^{del629/del629}$ mice the heights of their middle and short row stereocilia are visibly more variable compared with those of *Clrn2*$^{+/+}$ littermates, and many of the short row 'mechanotransducing' stereocilia are missing (Fig. 5C).

Together, our findings establish that clarin 2 is key to hearing function in zebrafish and mouse, and likely has a key role in the mechanotransducing stereocilia of the hair bundle.

Discussion

We identify *CLRN2* as a novel deafness gene in human and zebrafish and describe a new deafness-causing allele in mice. Genetic study using homozygosity mapping and exome sequencing of an extended Iranian family with multiple consanguineous marriages identified a pathogenic variant, c.494C>A in exon 2 of *CLRN2* segregating with pre-lingual ARNSHL. The c.494C>A variant results in a missense and splicing defect in clarin 2. By producing mutant zebrafish and mice lacking clarin 2, we demonstrated the key role the protein plays to ensure normal structural and functional integrity of the hair bundle, the sound- and motion- receptive structure of inner ear hair cells.

The clarin gene family also includes the *CLRN1* gene. Pathogenic variants in *CLRN1* have been linked to variable clinical outcomes, ranging from non-syndromic RP (Khan et al., 2011) to USH3A characterized by variable and progressive post-lingual hearing loss, RP, and variable vestibular responses (Plantinga et al., 2005). Several

cases of later onset HL and/or RP, as late as the sixth decade of life, have been reported for USH3A patients (Ness et al., 2003). Clinical examination of affected individuals in this family, at the age of 25 (IV-8), 29 (IV-1), and 44 (IV-6) years of age, excluded the presence of additional syndromic features showing that homozygosity for the c.494C>A variant causes non-syndromic hearing loss, ranging from moderate-severe (IV-6) to profound (IV-1) deafness (Fig. 1A, B).

The c.494C>A variant affects an amino acid that is highly conserved among PMP-22/EMP/EP20/Claudin superfamily proteins (Fig. 2A-C). In addition, the c.494 cytosine is highly conserved and the exchange to adenine is predicted to create an ESE site that likely impacts splicing efficiency in humans (Fig. S2A, B) but not zebrafish (Fig. S2C). We confirmed the effect on splicing using mini-gene assays. We showed that the c.494C>A variant acts in two ways: 1) as a missense variant (p.Thr165Lys) producing a mutant full length protein and 2) as a splice variant leading to intron retention (Fig. 3B, and Fig. S3B, C) expected to cause a premature stop codon 26 amino acids into intron 2 (p.Gly146Lysfs*26) (Fig. 3C).

Variants that disrupt splicing machinery signals can impact accurate recognition and removal of intronic sequences from pre-mRNA (Fairbrother et al., 2004) and are recognized as significant contributors to human genetic diseases (Xiong et al., 2015). ESE sequences are *cis*-acting elements primarily recognized by the SR family proteins that function by recruiting core splicing machinery components to splice sites or by acting antagonistically against nearby silencing elements (Fairbrother et al., 2004; Graveley, 2000; Kan and Green, 1999). ESEs are often associated with introns that

contain weak splicing signals, but they can also reside in exons and impact the splicing process.

Two potential mechanisms could synergistically contribute to the disruptive effect of the missense variant. *First*, the replacement of threonine with lysine, an amino acid with a positively charged 'bulky' side chain (lysine), may affect protein folding (Creixell et al., 2012) and transport to the plasma membrane. Membrane proteins sort to the plasma membrane via the conventional secretory pathway associated with ER-to-Golgi complex (Viotti, 2016). Misfolded membrane proteins are typically retained in the endoplasmic reticulum (ER) and degraded by the ER-associated degradation pathway (Kincaid and Cooper, 2007; Sano and Reed, 2013). It is possible that a small fraction of the misfolded clarin 2 p.(Thr165Lys) could reach the plasma membrane via the unconventional secretory pathway, similar to that reported for clarin 1 p.(Asn48Lys) (p.(N48K)) (Gopal et al., 2019). The unconventional secretory pathway is induced by the ER–associated misfolded or unfolded protein response (Kinseth et al., 2007; Schröder and Kaufman, 2005). However, the mutant clarin 2 reaching the surface may be functionally inactive. *Second*, evolutionarily conserved threonine residues are also conserved protein phosphorylation sites. Phosphorylation adds a negative charge to the side chain of the amino acid and it serves as an important post-translational mechanism for regulation of protein function (Pearlman et al., 2011). Loss of threonine at position 165 would potentially prevent functional activation of clarin 2. However, additional experiments are essential to test these hypotheses and unravel the true pathogenic mechanism associated with the p.(Thr165Lys) missense variant.

552 Our *in situ* hybridization in zebrafish (Fig. 4C, D) and *in silico* analyses (Fig. S5)
553 in mouse clearly support predominant expression of *Clrn2* in the sensory hair cells. To
554 examine further the key role of clarin 2 in the inner ear, we generated zebrafish and
555 mice lacking a functional protein. ABR measurements in *Clrn2*$^{del629/del629}$ mice revealed
556 an early-onset hearing loss with elevated hearing thresholds compared with their
557 *Clrn2*$^{+/+}$ littermate controls (mean click threshold 87 dB SPL ± 7 s.d. and 24 dB SPL ± 6
558 s.d., respectively). These data are consistent with early-onset hearing loss observed in
559 another loss-of-function *Clrn2* mutant (*Clrn2*^{clarinet}), which harbors an early truncating
560 nonsense variant (p.Trp4*) (Dunbar et al., 2019). Interestingly, unlike observations in
561 humans with *CLRN1* pathogenic variants, neither *Clrn1* (Dulon et al., 2018; Geller et al.,
562 2009; Geng et al., 2009) nor *Clrn2* (Dunbar et al., 2019) knockout mice exhibit a retinal
563 phenotype. In the inner ear, the reduced DPOAEs in both *Clrn2*$^{del629/del629}$ and in
564 *Clrn2*$^{clarinet/clarinet}$ mice (Dunbar et al., 2019), indicates impairment of OHCs function. This,
565 together with the severe-to-profound hearing loss already exhibited at P21 in *Clrn2*
566 mutant mice points to gene defects likely affecting both inner hair cells (IHCs) and
567 OHCs. This is further supported by scanning electron microscopy data showing loss of
568 shortest row stereocilia in both the cochlear IHCs and OHCs (Fig. 5C). Phalloidin
569 staining of *clrn2* crispants also confirms hair bundle abnormalities in zebrafish.

570 In regard to the observed progressivity of the hearing impairment in *clarinet* mice
571 (Dunbar et al., 2019), the earliest reported clinical diagnosis of hearing loss of the
572 *CLRN2* affected individuals in the family we present is between 2 and 3 years of age.
573 Newborn hearing screening was not routinely performed when the affected individuals
574 were born, so we cannot confirm hearing was normal at birth. In light of absent serial

audiograms, we cannot report if the hearing loss experienced in these patients is progressive, as is observed in the mouse model (Dunbar et al., 2019).

In conclusion, we demonstrate the c.494C>A variant affects exon 3 splicing efficiency. We showed, for the first time, that *CLRN2* is a deafness-causing gene in humans. A variant causes hearing loss in humans, replicated by animal studies. Additional reports of families segregating *CLRN2* biallelic variants will be crucial to refine and dissect the clinical course and characteristics of hearing loss due to this gene. Together, our studies in zebrafish and mice establish that hearing loss is probably due to defective protein in the hair cells, where the presence of clarin 2 is essential for normal organization and maintenance of the mechanosensitive hair bundles.

Acknowledgements

We would like to extend our gratitude to the family for their participation. We thank Dr. Caroline Lekszas, Dr. Daniel Liedtke, and Dr. Indrajit Nanda from the Institute of Human Genetics at the University of Würzburg for their technical expertise. This work was supported by Intramural Funding (*f*ortüne) at the University of Tübingen (2545-1-0 to B.V.), the Ministry of Science, Research and Art Baden-Württemberg (to B.V.), the Medical Research Council (MC_UP_1503/2 to M.R.B), ANR light4deaf (ANR-15-RHUS-0001), HearInNoise (ANR-17-CE16-0017), LHW-stiftung to A.E.), and a grant from NIH/COBRE GM103636 (Project 3); the Presbyterian Health Foundation (PHF) Grant to GKV. This study was funded in part by NIDCDs R01s DC002842 and DC012049 to RJS and T32 GM007748 to KTB. LAD is a Medical Research Council DPhil student (1774724).

References

Adato, A., S. Vreugde, T. Joensuu, N. Avidan, R. Hamalainen, O. Belenkiy, T. Olender, B. Bonne-Tamir, E. Ben-Asher, C. Espinos, J.M. Millán, A.E. Lehesjoki, J.G. Flannery, K.B. Avraham, S. Pietrokovski, E.M. Sankila, J.S. Beckmann, and D. Lancet. 2002. USH3A transcripts encode clarin-1, a four-transmembrane-domain protein with a possible role in sensory synapses. *Eur J Hum Genet* 10:339-350.

Adzhubei, I.A., S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, and S.R. Sunyaev. 2010. A method and server for predicting damaging missense mutations. *Nat Methods* 7:248-249.

Azaiez, H., K.T. Booth, S.S. Ephraim, B. Crone, E.A. Black-Ziegelbein, R.J. Marini, A.E. Shearer, C.M. Sloan-Heggen, D. Kolbe, T. Casavant, M.J. Schnieders, C. Nishimura, T. Braun, and R.J.H. Smith. 2018. Genomic Landscape and Mutational Signatures of Deafness-Associated Genes. *Am J Hum Genet* 103:484-497.

Booth, K.T., H. Azaiez, K. Kahrizi, D. Wang, Y. Zhang, K. Frees, C. Nishimura, H. Najmabadi, and R.J. Smith. 2018a. Exonic mutations and exon skipping: Lessons learned from DFNA5. *Hum Mutat* 39:433-440.

Booth, K.T., K. Kahrizi, H. Najmabadi, H. Azaiez, and R.J. Smith. 2018b. Old gene, new phenotype: splice-altering variants in CEACAM16 cause recessive non-syndromic hearing impairment. *J Med Genet* 55:555-560.

Cartegni, L., J. Wang, Z. Zhu, M.Q. Zhang, and A.R. Krainer. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31:3568-3571.

618 Chun, S., and J.C. Fay. 2009. Identification of deleterious mutations within three human
619 genomes. *Genome Res* 19:1553-1561.

620 Creixell, P., E.M. Schoof, C.S. Tan, and R. Linding. 2012. Mutational properties of
621 amino acid residues: implications for evolvability of phosphorylatable residues. *Philos*
622 *Trans R Soc Lond B Biol Sci* 367:2584-2593.

623 Cunningham, C.L., and U. Müller. 2019. Molecular Structure of the Hair Cell
624 Mechanoelectrical Transduction Complex. *Cold Spring Harb Perspect Med* 9:pii:
625 a033167.

626 Desmet, F.O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres, and C.
627 Beroud. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing
628 signals. *Nucleic Acids Res* 37:e67.

629 Doll, J., M.A.H. Hofrichter, P. Bahena, A. Heihoff, D. Segebarth, T. Müller, M. Dittrich, T.
630 Haaf, and B. Vona. 2020. A novel missense variant in MYO3A is associated with
631 autosomal dominant high-frequency hearing loss in a German family. *Mol Genet*
632 *Genomic Med* e1343.

633 Dulon, D., S. Papal, P. Patni, M. Cortese, P.F. Vincent, M. Tertrais, A. Emptoz, A. Tlili,
634 Y. Bouleau, V. Michel, S. Delmaghani, A. Aghaie, E. Pepermans, O. Alegria-Prevot, O.
635 Akil, L. Lustig, P. Avan, S. Safieddine, C. Petit, and A. El-Amraoui. 2018. Clarin-1 gene
636 transfer rescues auditory synaptopathy in model of Usher syndrome. *J Clin Invest*
637 128:3382-3401.

638 Dunbar, L.A., P. Patni, C. Aguilar, P. Mburu, L. Corns, H.R. Wells, S. Delmaghani, A.
639 Parker, S. Johnson, D. Williams, C.T. Esapa, M.M. Simon, L. Chessum, S. Newton, J.
640 Dorning, P. Jeyarajan, S. Morse, A. Lelli, G.F. Codner, T. Peineau, S.R. Gopal, K.N.
641 Alagramam, R. Hertzano, D. Dulon, S. Wells, F.M. Williams, C. Petit, S.J. Dawson, S.D.
642 Brown, W.A.-O. Marcotti, A.A.-O. El-Amraoui, and M.A.-O.X. Bowl. 2019. Clarin-2 is
643 essential for hearing by maintaining stereocilia integrity and function. *EMBO Mol Med*
644 11:e10288.

645 Fairbrother, W.G., G.W. Yeo, R. Yeh, P. Goldstein, M. Mawson, P.A. Sharp, and C.B.
646 Burge. 2004. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate
647 exons. *Nucleic Acids Res* 32:W187-190.

648 Fattahi, Z., M. Beheshtian, M. Mohseni, H. Poustchi, E. Sellars, H. Nezhadi, A. Amini, S.
649 Arzhangi, K. Jalalvand, P. Jamali, Z. Mohammadi, B. Davarnia, P. Nikuei, M. Oladnabi,
650 A. Mohammadzadeh, E. Zohrehvand, A. Nejatizadeh, M. Shekari, M. Bagherzadeh, E.
651 Shamsi-Gooshki, S. Borno, B. Timmermann, A. Haghdoost, R. Najafipour, H.R.K.
652 Khorshid, K. Kahrizi, R. Malekzadeh, M.R. Akbari, and H. Najmabadi. 2019. Iranome: A
653 catalogue of genomic variations in the Iranian population. *Hum Mutat*

654 Fields, R.R., G. Zhou, D. Huang, J.R. Davis, C. Möller, S.G. Jacobson, W.J. Kimberling,
655 and J. Sumegi. 2002. Usher syndrome type III: revised genomic structure of the USH3
656 gene and identification of novel mutations. *Am J Hum Genet* 71:607-617.

657 García-García, G., M.J. Aparisi, R. Rodrigo, M.D. Sequedo, C. Espinós, J. Rosell, J.L.
658 Olea, M.P. Mendívil, M.A. Ramos-Arroyo, C. Ayuso, T. Jaijo, E. Aller, and J.M. Millán.

2012. Two novel disease-causing mutations in the CLRN1 gene in patients with Usher syndrome type 3. *Mol Vis* 18:3070-3078.

Geller, S.F., K.I. Guerin, M. Visel, A. Pham, E.S. Lee, A.A. Dror, K.B. Avraham, T. Hayashi, C.A. Ray, T.A. Reh, O. Bermingham-McDonogh, W.J. Triffo, S.W. Bao, J. Isosomppi, H. Västinsalo, E.M. Sankila, and J.G. Flannery. 2009. CLRN1 Is Nonessential in the Mouse Retina but Is Required for Cochlear Hair Cell Development. *Plos Genet* 5:

Geng, R.S., S.F. Geller, T. Hayashi, C.A. Ray, T.A. Reh, O. Bermingham-McDonogh, S.M. Jones, C.G. Wright, S. Melki, Y. Imanishi, K. Palczewski, K.N. Alagramam, and J.G. Flannery. 2009. Usher syndrome IIIA gene clarin-1 is essential for hair cell function and associated neural activation dagger. *Hum Mol Genet* 18:2748-2760.

Gillespie, P.G., and U. Müller. 2009. Mechanotransduction by hair cells: models, molecules, and mechanisms. *Cell* 139:33-44.

Gopal, S.R., D.H.C. Chen, S.W. Chou, J.J. Zang, S.C.F. Neuhauss, R. Stepanyan, B.M. McDermott, and K.N. Alagramam. 2015. Zebrafish Models for the Mechanosensory Hair Cell Dysfunction in Usher Syndrome 3 Reveal That Clarin-1 Is an Essential Hair Bundle Protein. *J Neurosci* 35:10188-10201.

Gopal, S.R., Y.T. Lee, R. Stepanyan, B.M. McDermott, Jr., and K.N. Alagramam. 2019. Unconventional secretory pathway activation restores hair cell mechanotransduction in an USH3A model. *Proc Natl Acad Sci U S A* 116:11000-11009.

679 Graveley, B.R. 2000. Sorting out the complexity of SR protein functions. *RNA* 6:1197-
680 1211.

681 Hardisty-Hughes, R.E., A. Parker, and S.D. Brown. 2010. A hearing and vestibular
682 phenotyping pipeline to identify mouse mutants with hearing impairment. *Nat Protoc*
683 5:177-190.

684 Hofrichter, M.A.H., M. Mojarad, J. Doll, C. Grimm, A. Eslahi, N.S. Hosseini, M. Rajati, T.
685 Muller, M. Dittrich, R. Maroofian, T. Haaf, and B. Vona. 2018. The conserved p.Arg108
686 residue in S1PR2 (DFNB68) is fundamental for proper hearing: evidence from a
687 consanguineous Iranian family. *Bmc Med Genet* 19:81.

688 Hudspeth, A.J. 1997. How hearing happens. *Neuron* 19:947-950.

689 Ingham, N.J., S.A. Pearson, V.E. Vancollie, V. Rook, M.A. Lewis, J. Chen, A. Buniello,
690 E. Martelletti, L. Preite, C.C. Lam, F.D. Weiss, Z. Powis, P. Suwannarat, C.J. Lelliott,
691 S.J. Dawson, J.K. White, and K.P. Steel. 2019. Mouse screen reveals multiple new
692 genes underlying mouse and human hearing loss. *PLoS Biol* 17:e3000194.

693 Joensuu, T., R. Hamalainen, B. Yuan, C. Johnson, S. Tegelberg, P. Gasparini, L.
694 Zelante, U. Pirvola, L. Pakarinen, A.E. Lehesjoki, A. de la Chapelle, and E.M. Sankila.
695 2001. Mutations in a novel gene with transmembrane domains underlie Usher
696 syndrome type 3. *Am J Hum Genet* 69:673-684.

697 Kan, J.L.C., and M.R. Green. 1999. Pre-mRNA splicing of IgM exons M1 and M2 is
698 directed by a juxtaposed splicing enhancer and inhibitor. *Gene Dev* 13:462-471.

699 Kazmierczak, P., H. Sakaguchi, J. Tokita, E.M. Wilson-Kubalek, R.A. Milligan, U. Muller,
700 and B. Kachar. 2007. Cadherin 23 and protocadherin 15 interact to form tip-link
701 filaments in sensory hair cells. *Nature* 449:87-91.

702 Khan, M.I., F.F. Kersten, M. Azam, R.W. Collin, A. Hussain, S.T. Shah, J.E. Keunen, H.
703 Kremer, F.P. Cremers, R. Qamar, and A.I. den Hollander. 2011. CLRN1 mutations
704 cause nonsyndromic retinitis pigmentosa. *Ophthalmology* 118:1444-1448.

705 Kimmel, C.B., W.W. Ballard, S.R. Kimmel, B. Ullmann, and T.F. Schilling. 1995. Stages
706 of embryonic development of the zebrafish. *Dev Dyn* 203:253-310.

707 Kincaid, M.M., and A.A. Cooper. 2007. Misfolded proteins traffic from the endoplasmic
708 reticulum (ER) due to ER export signals. *Mol Biol Cell* 18:455-463.

709 Kinseth, M.A., C. Anjard, D. Fuller, G. Guizzunti, W.F. Loomis, and V. Malhotra. 2007.
710 The golgi-associated protein GRASP is required for unconventional protein secretion
711 during development. *Cell* 130:524-534.

712 Kircher, M., D.M. Witten, P. Jain, B.J. O'Roak, G.M. Cooper, and J. Shendure. 2014. A
713 general framework for estimating the relative pathogenicity of human genetic variants.
714 *Nat Genet* 46:310-315.

715 Kolla, L., M.C. Kelly, Z.F. Mann, A. Anaya-Rocha, K. Ellis, A. Lemons, A.T. Palermo,
716 K.S. So, J.C. Mays, J. Orvis, J.C. Burns, R. Hertzano, E.C. Driver, and M.W. Kelley.
717 2020. Characterization of the development of the mouse cochlear epithelium at the
718 single cell level. *Nat Commun* 11:

719 Kopecky, B., S. Johnson, H. Schmitz, P. Santi, and B. Fritzsch. 2012. Scanning thin-
720 sheet laser imaging microscopy elucidates details on mouse ear development. *Dev Dyn*
721 241:465-480.

722 LaFave, M.C., G.K. Varshney, M. Vemulapalli, J.C. Mullikin, and S.M. Burgess. 2014. A
723 defined zebrafish line for high-throughput genetics and genomics: NHGRI-1. *Genetics*
724 198:167-170.

725 Lek, M., K.J. Karczewski, E.V. Minikel, K.E. Samocha, E. Banks, T. Fennell, A.H.
726 O'Donnell-Luria, J.S. Ware, A.J. Hill, B.B. Cummings, T. Tukiainen, D.P. Birnbaum, J.A.
727 Kosmicki, L.E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout,
728 D.N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J.
729 Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M.I. Kurki, A.L. Moonshine, P. Natarajan,
730 L. Orozco, G.M. Peloso, R. Poplin, M.A. Rivas, V. Ruano-Rubio, S.A. Rose, D.M.
731 Ruderfer, K. Shakir, P.D. Stenson, C. Stevens, B.P. Thomas, G. Tiao, M.T. Tusie-Luna,
732 B. Weisburd, H.H. Won, D. Yu, D.M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S.
733 Donnelly, R. Elosua, J.C. Florez, S.B. Gabriel, G. Getz, S.J. Glatt, C.M. Hultman, S.
734 Kathiresan, M. Laakso, S. McCarroll, M.I. McCarthy, D. McGovern, R. McPherson, B.M.
735 Neale, A. Palotie, S.M. Purcell, D. Saleheen, J.M. Scharf, P. Sklar, P.F. Sullivan, J.
736 Tuomilehto, M.T. Tsuang, H.C. Watkins, J.G. Wilson, M.J. Daly, D.G. MacArthur, and C.
737 Exome Aggregation. 2016. Analysis of protein-coding genetic variation in 60,706
738 humans. *Nature* 536:285-291.

739 Lekszas, C., O. Foresti, I. Raote, D. Liedtke, E.M. König, I. Nanda, B. Vona, P. De
740 Coster, R. Cauwels, V. Malhotra, and T. Haaf. 2020. Biallelic TANGO1 mutations cause
741 a novel syndromal disease due to hampered cellular collagen secretion. *Elife* 9:e51319.

742 Livak, K.J., and T.D. Schmittgen. 2001. Analysis of relative gene expression data using
743 real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25:402-408.

744 Mazzoli, M., G. Van Camp, V. Newton, N. Giarbini, F. Declau, and A. Parving. 2003.
745 Recommendations for the description of genetic and audiological data for families with
746 nonsyndromic hereditary hearing impairment. *Audiol Med* 1:148-150.

747 Morton, C.C., and W.E. Nance. 2006. Newborn hearing screening--a silent revolution. *N*
748 *Engl J Med* 354:2151-2164.

749 Ness, S.L., T. Ben-Yosef, A. Bar-Lev, A.C. Madeo, C.C. Brewer, K.B. Avraham, R.
750 Kornreich, R.J. Desnick, J.P. Willner, T.B. Friedman, and A.J. Griffith. 2003. Genetic
751 homogeneity and phenotypic variability among Ashkenazi Jews with Usher syndrome
752 type III. *J Med Genet* 40:767-772.

753 Ng, P.C., and S. Henikoff. 2001. Predicting deleterious amino acid substitutions.
754 *Genome Res* 11:863-874.

755 Pearlman, S.M., Z. Serber, and J.E. Ferrell, Jr. 2011. A mechanism for the evolution of
756 phosphorylation sites. *Cell* 147:934-946.

757 Phillips, J.B., B. Blanco-Sanchez, J.J. Lentz, A. Tallafuss, K. Khanobdee, S. Sampath,
758 Z.G. Jacobs, P.F. Han, M. Mishra, T.A. Titus, D.S. Williams, B.J. Keats, P.

759 Washbourne, and M. Westerfield. 2011. Harmonin (Ush1c) is required in zebrafish
760 Müller glial cells for photoreceptor synaptic development and function. *Dis Model Mech*
761 4:786-800.

762 Plantinga, R.F., L. Kleemola, P.L. Huygen, T. Joensuu, E.M. Sankila, R.J. Pennings,
763 and C.W. Cremers. 2005. Serial audiometry and speech recognition findings in Finnish
764 Usher syndrome type III patients. *Audiol Neurootol* 10:79-89.

765 Sano, R., and J.C. Reed. 2013. ER stress-induced cell death mechanisms. *Biochim*
766 *Biophys Acta* 1833:3460-3470.

767 Schröder, M., and R.J. Kaufman. 2005. The mammalian unfolded protein response.
768 *Annu Rev Biochem* 74:739-789.

769 Schwarz, J.M., D.N. Cooper, M. Schuelke, and D. Seelow. 2014. MutationTaster2:
770 mutation prediction for the deep-sequencing age. *Nat Methods* 11:361-362.

771 Scott, E.M., A. Halees, Y. Itan, E.G. Spencer, Y. He, M.A. Azab, S.B. Gabriel, A.
772 Belkadi, B. Boisson, L. Abel, A.G. Clark, C. Greater Middle East Variome, F.S.
773 Alkuraya, J.L. Casanova, and J.G. Gleeson. 2016. Characterization of Greater Middle
774 Eastern genetic variation for enhanced disease gene discovery. *Nat Genet* 48:1071-
775 1076.

776 Seelow, D., M. Schuelke, F. Hildebrandt, and P. Nürnberg. 2009. HomozygosityMapper-
777 an interactive approach to homozygosity mapping. *Nucleic Acids Res* 37:W593-W599.

778 Seiler, C., K.C. Finger-Baier, O. Rinner, Y.V. Makhankov, H. Schwarz, S.C. Neuhauss,
779 and T. Nicolson. 2005. Duplicated genes with split functions: independent roles of
780 protocadherin15 orthologues in zebrafish hearing and vision. *Development* 132:615-
781 623.

782 Smith, R.J., J.F. Bale, Jr., and K.R. White. 2005. Sensorineural hearing loss in children.
783 *Lancet* 365:879-890.

784 Thisse, C., and B. Thisse. 2008. High-resolution in situ hybridization to whole-mount
785 zebrafish embryos. *Nat Protoc* 3:59-69.

786 UniProt Consortium, T. 2018. UniProt: the universal protein knowledgebase. *Nucleic*
787 *Acids Res* 46:2699.

788 Varshney, G.K., B. Carrington, W. Pei, K. Bishop, Z. Chen, C. Fan, L. Xu, M. Jones,
789 M.C. LaFave, J. Ledin, R. Sood, and S.M. Burgess. 2016. A high-throughput functional
790 genomics workflow based on CRISPR/Cas9-mediated targeted mutagenesis in
791 zebrafish. *Nat Protoc* 11:2357-2375.

792 Viotti, C. 2016. ER to Golgi-Dependent Protein Secretion: The Conventional Pathway.
793 *Methods Mol Biol* 1459:3-29.

794 Waterhouse, A.M., J.B. Procter, D.M. Martin, M. Clamp, and G.J. Barton. 2009. Jalview
795 Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*
796 25:1189-1191.

797 Westerfield, M. 2000. The zebrafish book. A guide for the laboratory use of zebrafish
798 (Danio rerio). Univ. of Oregon Press, Eugene, Oregon, USA.

799 Whitfield, T.T., B.B. Riley, M.Y. Chiang, and B. Phillips. 2002. Development of the
800 zebrafish inner ear. *Dev Dyn* 223:427-458.

801 Wright, C.F., D.R. FitzPatrick, and H.V. Firth. 2018. Paediatric genomics: diagnosing
802 rare disease in children. *Nat Rev Genet* 19:253-268.

803 Xiong, H.Y., B. Alipanahi, L.J. Lee, H. Bretschneider, D. Merico, R.K. Yuen, Y. Hua, S.
804 Gueroussov, H.S. Najafabadi, T.R. Hughes, Q. Morris, Y. Barash, A.R. Krainer, N. Jojic,
805 S.W. Scherer, B.J. Blencowe, and B.J. Frey. 2015. RNA splicing. The human splicing
806 code reveals new insights into the genetic determinants of disease. *Science*
807 347:125806.

808 Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. 2015. The I-TASSER Suite:
809 protein structure and function prediction. *Nat Methods* 12:7-8.
810

Figure legends

Fig. 1. Pedigree, audiological data, genetic data, and locus mapping

(A) The consanguineous family of Iranian origin with hearing loss and segregation of the *CLRN2* c.494C>A variant.

(B) Pure-tone audiograms from affected individuals IV-1 (red) and IV-6 (blue), as well as an unaffected heterozygous individual V-1 (green). Air conduction thresholds in dB HL for the right and left ears are represented by circles and crosses, respectively. Bone-conduction thresholds are represented by < and > for right and left ears, respectively, and confirm a sensorineural hearing loss in the affected individuals.

(C) Homozygosity mapping reveals a 15.2 Mb locus on chromosome 4 containing *CLRN2*.

(D) Sequence electropherograms showing the homozygous, heterozygous and WT images of the *CLRN2* c.494C>A; pThr165Lys pathogenic variants.

Fig. 2. Conservation of the p.Thr165 residue, and clarin 1/clarin 2 alignment.

(A) Overview of clarin 2 protein and modular structure of the PMP-22/EMP/EP20/Claudin superfamily, with amino acid residue coordinates and position of the p.(Thr165Lys) substitution shown (upper panel). An alignment of the amino acid sequences from the segment of clarin 2 (represented by dashed lines) from vertebrate species shows the Thr165 position (asterisk) is well conserved among vertebrates.

(B) Alignment of clarin 2 (UniProtKB: A0PK11, upper alignment) and clarin 1 (UniProtKB: P58418, lower alignment) amino acid residues. Transmembrane domains

are marked in grey, conservation is shown in yellow, and consensus sequences are shown below for the 232 amino acid proteins. Missense and nonsense variants in clarin 1 (Deafness Variation Database v8.2) and clarin 2 (present study, asterisk) are marked in red.

(C) The predicted secondary structure of human clarin 2 (NP_001073296.1) wild-type (Thr165) and mutated (Thr165Lys) protein. H represents alpha-helix, S represents beta-strand and C represents coil.

Fig. 3. Analysis of the *CLRN2* c.494C>A variant on splicing

(A) Schematic illustration of the mini-gene splice construct design. Genomic representation of *CLRN2,* including the position of the missense variant c.494C>A (arrow) on exon 3 with 3' UTR (green), and the 5' UTR, as well as exons 1 and 2 (grey) (upper panel). Regions captured by mini-gene PCR primers are represented in green. Schematic illustration of the mini-gene splice construct including exon 3 and its flanking sequence (green) cloned into multiple cloning sites (*Sal*I and *Sac*II sites) of pET01 backbone vector (lower panel). Blue boxes represent native exons of the pET01 vector.

(B) RT-PCR of transcripts from post-mini-gene transfected COS-7 cells. Amplicons derived from the transcripts of WT (*CLRN2*), a benign *CLRN2* polymorphism (rs117875715, chr4(GRCh37):g.17,528,480G>A), the *CLRN2* c.494C>A variant and a negative control, were visualized on a 1.5% agarose gel. The SNP, rs117875715, was used to test and validate the designed WT and mutant mini-gene assay. The ~650 bp amplicon was associated with the WT and validation control rs117875715. The

amplicon derived from the *CLRN2* c.494C>A transcripts showed two bands: a 650 bp band and a larger ~1360 bp band, indicating retention of intron separating the donor site of the 5' exon and the acceptor site of *CLRN2* exon 3.

(C) Retention of intron in *CLRN2* c.494C>A mini-gene results in a stop codon (TGA) after *CLRN2* exon 2.

Fig. 4. Clarin 2 is required for the inner ear function in zebrafish

(A) RT-qPCR of *clrn2* mRNA expression from 1 to 120 hpf of WT embryos/larvae. *clarin 2* mRNA expression can be detected starting from 18 hpf and then increased throughout development. The bar graphs showed the mean values ± SEM after normalization to the housekeeping gene *18S* level and then compared to 18 hpf.

(B) RT-qPCR of c*lrn2* mRNA expression in different adult tissues. The bar graphs displayed mean values ± SEM after normalization to the housekeeping gene *18S* level and then compared to muscle.

(C, D) Whole-mount in situ hybridization (WISH) using antisense *clrn2* probe reveals the inner ear expression of *clrn2* mRNA (relative dark purple color, black arrowhead) at 3 (C) and 5 (D) dpf embryos. Sense *clrn2* probe was used as negative control and relative light purple color is considered as background. *clrn2* mRNA was consistently expressed in hair cells within inner ear macula (C, D) with lined and arrayed structure. A known hair cell marker *pvalb9* was used as an indicator for hair cells in the inner ear of 3 dpf embryos (C). Cryosection was performed after *clrn2* WISH at 5 dpf to confirm the small

patch of signal on the macula is hair cells rather than supporting cells (D, black arrow middle panel). Scale bar = 100 μm, except middle panel in D (20 μm).

(E) RT-qPCR of *clrn2* mRNA expression level was decreased 70% in *clrn2* crispants compared to uninjected larvae, indicating *clrn2* was successfully knocked out (E, upper panels). Phalloidin staining on *clrn2* crispants show that the hair cells in the inner ear macula display splayed, thin and split structures (red arrowhead in lower panels). Scale bar = 10 μm.

Figure 5. Clarin 2 is required for hearing function in mouse

(A) The genomic structure of mouse *Clrn2* (ENSMUST00000053250), and domains of the encoded tetraspan-like glycoprotein (232 amino acids). The positions of the transmembrane (TM) domains (dark green) and the structures of the WT *Clrn2 and* $Clrn2^{del629}$ alleles are indicated.

(B) ABR threshold measurements at P21 (± 1 day) show that $Clrn2^{del629/del629}$ mice (red) exhibit a severe-to-profound hearing loss affecting all frequencies tested, with thresholds at 80 dB SPL and beyond. Age-matched $Clrn2^{+/+}$ (black) and $Clrn2^{del629/+}$ (grey) controls display thresholds within the expected range (15-40 dB SPL). Data shown are mean ± SD. ***p<0.001, one-way ANOVA. **(C)** Averaged DPOAE responses at P28 (± 1 day), showing significantly reduced responses in $Clrn2^{del629/del629}$ mice. Data shown are mean ± SD. *p<0.02, **p<0.01, one-way ANOVA.

(C) Pseudo-colored scanning electron micrographs illustrate the three full rows, tallest (red), middle (blue) and short (yellow), of P28 (± 1 day) stereocilia in IHC and OHC hair bundles. Unlike the fragmented hair bundle in *Clrn1*$^{-/-}$ mice, lack of clarin-2 does not affect the shape of IHC or OHC hair bundles. However, all the short row stereocilia have completely or partially regressed in the absence of either clarin protein.

Scale bar = 1 µm.