

ARTICLE

Novel form of X-linked nonsyndromic hearing loss with cochlear malformation caused by a mutation in the type IV collagen gene *COL4A6*

Simone Rost^{*1}, Elisa Bach¹, Cordula Neuner¹, Indrajit Nanda¹, Sandra Dysek², Reginald E Bittner², Alexander Keller³, Oliver Bartsch⁴, Robert Mlynski⁵, Thomas Haaf^{*1}, Clemens R Müller¹ and Erdmute Kunstmann¹

Hereditary hearing loss is the most common human sensorineural disorder. Genetic causes are highly heterogeneous, with mutations detected in >40 genes associated with nonsyndromic hearing loss, to date. Whereas autosomal recessive and autosomal dominant inheritance is prevalent, X-linked forms of nonsyndromic hearing impairment are extremely rare. Here, we present a Hungarian three-generation family with X-linked nonsyndromic congenital hearing loss and the underlying genetic defect. Next-generation sequencing and subsequent segregation analysis detected a missense mutation (c.1771G>A, p.Gly591Ser) in the type IV collagen gene *COL4A6* in all affected family members. Bioinformatic analysis and expression studies support this substitution as being causative. *COL4A6* encodes the alpha-6 chain of type IV collagen of basal membranes, which forms a heterotrimer with two alpha-5 chains encoded by *COL4A5*. Whereas mutations in *COL4A5* and contiguous X-chromosomal deletions involving *COL4A5* and *COL4A6* are associated with X-linked Alport syndrome, a nephropathy associated with deafness and cataract, mutations in *COL4A6* alone have not been related to any hereditary disease so far. Moreover, our index patient and other affected family members show normal renal and ocular function, which is not consistent with Alport syndrome, but with a nonsyndromic type of hearing loss. *In situ* hybridization and immunostaining demonstrated expression of the *COL4A6* homologs in the otic vesicle of the zebrafish and in the murine inner ear, supporting its role in normal ear development and function. In conclusion, our results suggest *COL4A6* as being the fourth gene associated with X-linked nonsyndromic hearing loss.

European Journal of Human Genetics advance online publication, 29 May 2013; doi:10.1038/ejhg.2013.108

Keywords: next-generation sequencing; X-linked nonsyndromic hearing loss; type IV collagen; *COL4A6*

INTRODUCTION

Childhood hearing loss is one of the most common and heterogeneous diseases worldwide. In addition to various environmental causes, for example, intrauterine, peri- or postnatal infections, numerous genetic defects leading to hearing impairment are known. Hereditary hearing loss is classified into two main forms: syndromic and nonsyndromic. Syndromic hearing loss includes autosomal dominant (for example, Waardenburg syndrome, MIM 193500, branchio-oto-renal syndrome, MIM 113650 and Stickler syndrome, MIM 108300), autosomal recessive (for example, Usher syndrome, MIM 276900, Pendred syndrome, MIM 274600 and Jervell and Lange-Nielsen syndrome, MIM 220400) and X-linked forms (for example, Alport syndrome, MIM 301050).¹ Each of these syndromes is caused by mutations in more than one gene.^{1,2} Syndromic hearing loss can also occur in a number of mitochondrial diseases.³ Similarly, nonsyndromic hearing loss comprises autosomal dominant (DNFA), autosomal recessive (DFNB) and X-linked (DFNX or DFN) subtypes, as well as some mitochondrial forms (Hereditary Hearing Loss Homepage, <http://hereditaryhearingloss.org/>). Mutations in the

GJB2 gene (MIM 121011) encoding connexin 26 (CX26)⁴ are responsible for most cases of autosomal recessive nonsyndromic hearing loss, accounting for >30% of affected children in Western Europe.⁵ More than 40 different genes have been identified for autosomal recessive deafness (Hereditary Hearing Loss Homepage), whereas X-linked nonsyndromic hearing loss is much rarer, accounting for 1–3% of hereditary deafness.¹ Only three genes have been associated with X-linked nonsyndromic hearing loss until now: *PRPS1* (MIM 311850; DFNX1, formerly DFN2) on Xq22 that encodes phosphoribosyl pyrophosphate synthetase 1,⁶ *POU3F4* (MIM 300039; DFNX2, formerly DFN3) on Xq21, encoding a member of a transcription factor family that contains a POU domain⁷ and the recently identified gene for the small muscle protein, X-linked (*SMPX*; MIM 300226) on Xp22, accounting for DFNX4 (formerly DFN6).^{8,9} An additional locus has been mapped to the short arm of the X chromosome (DFNX3, formerly DFN4) overlapping with the *DMD* locus on Xp21.2.^{10,11}

Here, we present data on a Hungarian family in which only male subjects suffer from severe congenital hearing loss, whereas female

¹Department of Human Genetics, University of Würzburg, Würzburg, Germany; ²Department of Neuromuscular Research, Center for Anatomy and Cell Biology, Medical University of Vienna, Vienna, Austria; ³DNA Analytics Core Facility and Department of Animal Ecology and Tropical Biology, University of Würzburg, Würzburg, Germany; ⁴Institute of Human Genetics, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany; ⁵Department of Oto-Rhino-Laryngology, Plastic, Aesthetic and Reconstructive Head and Neck Surgery, Comprehensive Hearing Center, University of Würzburg, Würzburg, Germany
*Correspondence: Dr S Rost or Professor Dr T Haaf, Department of Human Genetics, University of Würzburg, Biocenter, Am Hubland, Würzburg 97074, Germany.
Tel: +49 931 31 84095 or +49 931 31 88738; Fax: +49 931 31 84069; E-mail: simone.rost@biozentrum.uni-wuerzburg.de or thomas.haaf@uni-wuerzburg.de
Received 30 October 2012; revised 8 April 2013; accepted 19 April 2013

subjects are not or only mildly to moderately affected. Based on the highly suggestive X-linked recessive inheritance and after ruling out the known X-linked genes *PRPS1* and *POU3F4*, next-generation sequencing (NGS) of the X-chromosomal exome was performed on three family members in order to identify the underlying genetic defect. Segregation, bioinformatic and expression analyses identified a missense mutation in the *COL4A6* gene underlying the hearing impairment in this family.

MATERIALS AND METHODS

Karyotyping and SNP array

To verify the presence of structural changes, conventional cytogenetic analysis was performed on chromosomes prepared from 72-h PHA-stimulated peripheral blood lymphocytes of the index patient using the standard procedure. Metaphase chromosomes were then evaluated by Giemsa–Trypsin–Giemsa (GTG) banding at about the 500 band level, according to the International System for Human Cytogenetic Nomenclature (ISCN, 2009).

In order to ascertain genomic alterations, a genome-wide SNP array was performed using the HumanCytoSNP-12 BeadChip (Illumina Inc, San Diego, CA, USA) encompassing 300 000 SNPs dispersed across the genome, with an average distance of 10 kb between markers. In brief, 200 ng of genomic DNA from the index patient, mother and cousin was hybridized to the BeadChip according to the manufacturer's instruction. After hybridization, the BeadChip was scanned with the Illumina BeadArray Reader, and the data were analysed by examining signal intensity (log *R* ratio) and allelic composition (BAF) with GenomeStudio v2010.1 and cnvPartition v3.1.6 software (Illumina Inc). A minimum of a five probe cutoff value was used to define a copy number change.

Marker and MLPA analysis

CA-repeat markers located in or adjacent to the dystrophin gene (*DMD*) on Xp21.2 (5'DYS-I, 5'-7n4, STR-45, STR-49, 3'-19n8, DXS992; according to the Leiden Muscular Dystrophy pages, http://www.dmd.nl/ca_dmd.html) were examined in all available family members using fluorescence-labeled primers and standard PCR conditions. Fragment analyses were performed on an ABI 3130xl capillary sequencer (Applied Biosystems, Carlsbad, CA, USA) according to the standard protocols. Data were evaluated using GeneMapper v4.0 (Applied Biosystems), and haplotypes were constructed by Cyrillic v2.1 (CyrillicSoftware, Oxfordshire, UK).

DNA samples of the index patient and his mother were analysed by multiplex ligation-dependent probe amplification (MLPA) for quantitative genomic changes in the *POU3F4* gene using the P163-C1 probe mix (MRC Holland, Amsterdam, The Netherlands), according to the manufacturer's instructions. MLPA PCR products were separated on an ABI 3130xl automatic sequencer, pre-analysed by GeneMapper v4.0 and finally evaluated by the Excel-based software Coffalyser v5.2 (MRC Holland).

DNA sequencing analysis

Sanger sequencing was performed for the coding exons of the X-linked genes *PRPS1* and *POU3F4* in the index patient on an ABI 3130xl capillary sequencer by using standard conditions. Sequence data were analysed by Gensearch v3.8

(PhenoSystems, Lillois, Belgium) and automatically aligned to reference sequences from GenBank (NCBI, <http://www.ncbi.nlm.nih.gov/genbank/>).

The whole exome of the X chromosome was sequenced in two affected male subjects (Figure 1, III.3 and III.5) and an obligatory carrier female subject (II.2) by NGS on a HiSeq2000 (Illumina Inc) after target enrichment with the SureSelect Human Chromosome X Exome Kit (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions. Sequence reads were mapped using the Burrows–Wheeler Aligner (BWA)¹² after trimming and sorting of the raw data. Processed data were aligned to the X chromosome sequence of the human reference sequence hg19 (GRCh37) and analyzed using the software GensearchNGS (PhenoSystems), which enables variant detection based on VarScan.¹³ Detected variants of the three examined patients were automatically compared in a table generated by the software (data not shown) and additional filtering of the variants was performed considering quality (Phred quality score >20), coverage (>20×), variant frequency (>0.8 for male patients, >0.3 and <0.7 for the female carrier), intragenic location (elimination of extragenic variants) and biochemical effect of the variant (severity > synonymous variants).

Candidate variants were verified by Sanger sequencing using standard protocols, and their segregation was studied within the family.

DNA samples of 93 healthy Hungarian women, as well as 91 male and 65 female blood donors from Central Europe were screened for the presence of the c.1771 variant in the *COL4A6* gene, either by Sanger sequencing or by pyrosequencing on a PyroMark Q96 MD system (Qiagen GmbH, Hilden, Germany) according to the manufacturer's conditions.

The *COL4A6* gene was sequenced in 94 unrelated male and two female patients suffering from hearing loss, following target enrichment of all 46 exons (including the two alternative exons: 1A of NM_001847.2 and 1B of NM_033641.2) and parts of the 5' and 3'-UTRs with the Access Array System (Fluidigm Europe BV, Amsterdam, The Netherlands), according to the instructions provided by Fluidigm. Amplicon sequencing was performed on a Genome Sequencer 454 FLX System (Roche Applied Science, Mannheim, Germany) after amplification by emulsion PCR using reagents and conditions provided by Roche Applied Science. Sequence data were processed and analyzed using the software GensearchNGS, as described above.

Bioinformatic analysis

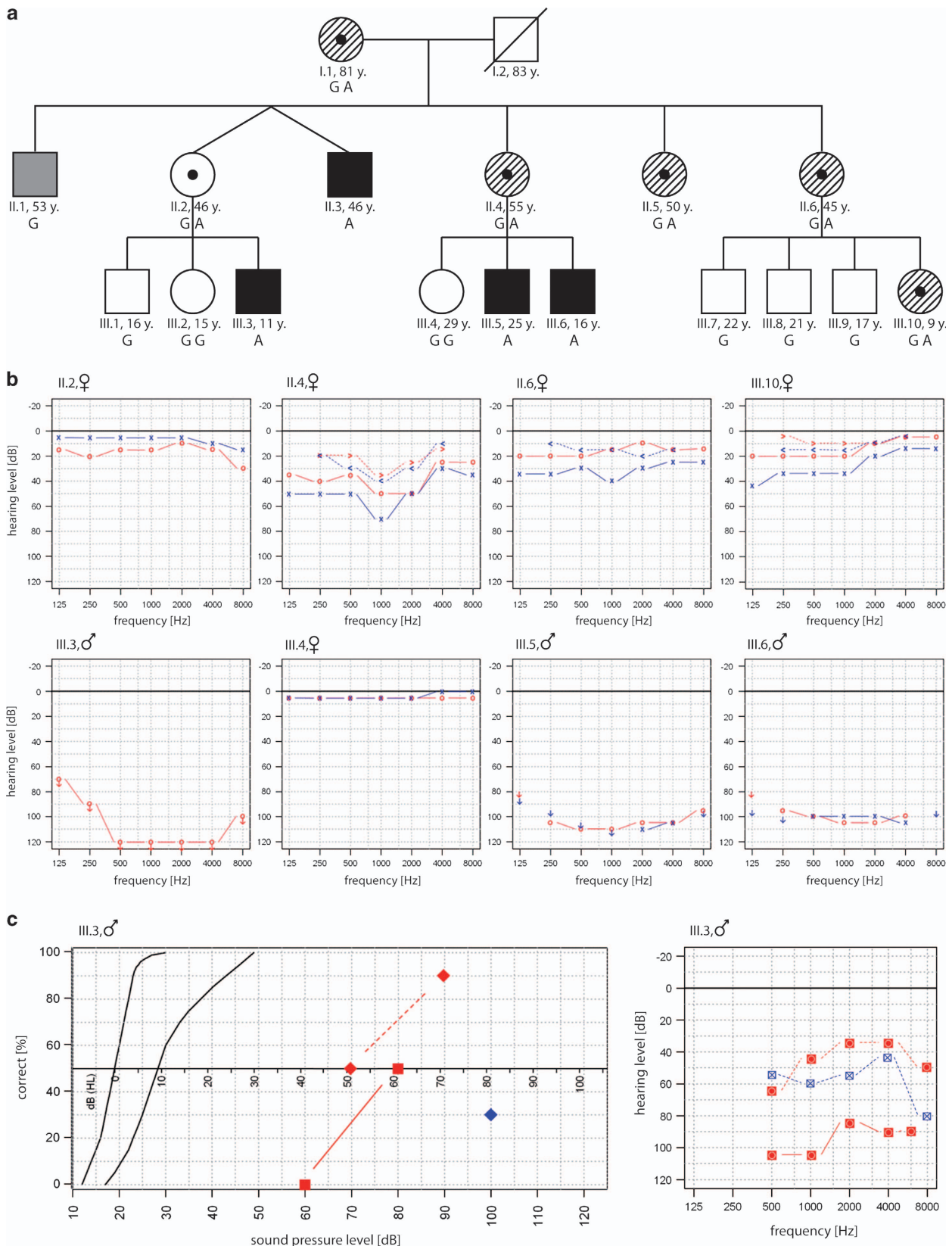
Prediction of the effect of amino acid substitutions on protein function was done by the bioinformatic tools SIFT (<http://sift.jcvi.org/>), PolyPhen-2 (Polymorphism Phenotyping v2, <http://genetics.bwh.harvard.edu/pph2/index.shtml>) and Mutation Taster (<http://www.mutationtaster.org/>) that are included in Alamut version 2.2 (Interactive Biosoftware, Rouen, France). Splice-site prediction was also performed using Alamut, which comprises five different prediction tools: SplicesiteFinder-like (Alamut-specific tool), Max-EntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html), NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html), GeneSplicer (<http://www.cbcb.umd.edu/software/genesplicer/>) and Human Splicing Finder (<http://www.umd.be/HSF/>). The alignment tool MultAlin (<http://multalin.toulouse.inra.fr/multalin/>)¹⁴ was used for multiple sequence alignments of *COL4A6* protein sequences from different species (mammals, fish and reptiles) and of the six alpha chains ($\alpha 1$ – $\alpha 6$) of human type IV collagens.

Figure 1 (a) Pedigree of the family. Black squares indicate affected male subjects with severe hereditary hearing loss, the gray square (II.1) indicates a male with acquired deafness, circles with black dots show mutation carriers who are healthy to date (unshaded) or who are affected mildly to moderately (shaded). The detected nucleotides at position c.1771 of the *COL4A6* gene are given for every family member: G or GG means hemi- or homozygous for the wild-type sequence, GA means c.1771G>A heterozygous and A denotes the mutation in hemizygous form. (b) Representative pure tone audiometry of eight family members. The female subjects II.2 and III.4 show regular thresholds at their current age. The females II.6 (at 41 years of age) and III.10 (at 7 years of age) show normal thresholds of their right ears, and suffer from low-grade conductive hearing loss of maximum 25 dB at 1 kHz (II.6) and a maximum of 20 dB between 0.5 and 1 kHz (III.10) on their left side. Female II.4 (at present) expresses bilateral moderate mixed hearing loss of maximum 35 dB (right) and 40 dB (left) at 1 kHz, and an additional conductive component of maximum 25 dB (right) at 2 kHz and 30 dB (left) at 250 Hz and 1 kHz. Thresholds of patients III.3 (at 8 years), III.5 (at 24 years) and III.6 (at 12 years) at their latest follow-up, showing only residual hearing in single frequencies (red—right side, blue—left side, circles and crosses—air conduction, arrow heads—bone conduction, vertical arrows—frequency not heard). (c) Behavior audiometry under best aided conditions (bilateral hearing aids) of the index patient III.3 at the age of 3 years prior to cochlea implantation of the left ear (dashed lines), and at the age of 8 years before cochlear implantation of the right ear. On the left: speech audiometry 'Mainzer Kindertest' (solid diamonds) and 'Freiburger monosyllables' (solid boxes), on the right: functional gain using hearing aids measured with warble thresholds (red—right ear, blue—left ear).

Bioinformatic analysis for protein stability and structure prediction was performed using the software The BuScr 1.06 (<http://structbio.biochem.dal.ca/jrailey/THBuScr.html>), SCWRL 4.0 (<http://dunbrack.fccc.edu/Software.php>) and UCSF Chimera 1.5 (<http://www.cgl.ucsf.edu/chimera/>).

COL4A6 mRNA analysis

Fibroblasts from skin biopsies of the index patient and his mother were grown in Dulbecco's Modified Eagle Medium (DMEM; Gibco, Life Technologies GmbH, Darmstadt, Germany) supplemented with 15% fetal bovine serum



(FBS; Gibco). After harvesting the cultured cells with trypsin/EDTA solution (Gibco), total mRNA was isolated using RNeasy Mini Kit (Qiagen). Complementary DNA (cDNA) was synthesized by reverse transcriptase (Superscript II, Invitrogen, Life Technologies GmbH, Darmstadt, Germany). Amplification of exons 22–24 of *COL4A6* cDNA was done using specific primers located in the adjacent exons (21 forward: 5'-AAAGGAGCCAGAGGAGATCG-3', 25 reverse: 5'-GAGGCCTCGAGACCCTTTAG-3'). Obtained fragments were sequenced on an ABI 3130xl capillary sequencer using standard conditions.

Expression studies

We performed an immunohistochemical staining of the mouse inner ear (C57Bl/10 male mouse, 12 weeks of age) with a *COL4A6*-specific antibody (rabbit polyclonal antibody M137, Santa Cruz Biotechnology Inc, Santa Cruz, CA, USA), following published procedures.¹⁵ Antibody specificity was checked by omitting the first antibody (data not shown).

Whole-mount *in situ* hybridization according to established protocols¹⁶ was done to assess the *col4a6* expression pattern in wild-type developing embryos of the zebrafish. A 620-bp fragment corresponding to the coding region (nts 1487–2106) adjacent to the 3'-UTR of the *col4a6* mRNA (XM_001921595.2) was isolated from cDNA of whole embryos (26–28 h post fertilization, hpf) and cloned into pCRII TOPO vector (Invitrogen). After linearization of the plasmid, both sense and antisense digoxigenin-labelled probes were generated through *in vitro* transcription using Sp6 and T7 RNA polymerase. The control sense probe did not reveal a detectable signal. The riboprobe sequence specificity was checked by BLAST analysis against the entire zebrafish genome, which showed exclusive full-length homology to the *col4a6* on chromosome 7.

RESULTS

Clinical data

The family examined in the present study consists of 17 available members from three generations with nine male subjects, five of which suffering from bilateral severe sensorineural hearing loss, whereas three female subjects are healthy and five females suffer from mild to moderate hearing impairment. The pedigree of the non-consanguineous family is shown in Figure 1a. The index patient (III.3) did not pass the standardized hearing screening during pediatric examination at the age of 3 months. Brain stem audiometry revealed bilateral severe hearing loss with unaided hearing thresholds of 80 dB right side and 80–100 dB left side (data not shown). Lacking any obvious postnatal diseases or treatments with potential impact on hearing, the hearing loss was considered congenital and the patient received bilateral hearing aids. He developed speech in two to three word sentences at the age of 3 years, but articulation was insufficient. At the age of 3 years, magnetic resonance imaging (MRI) of the cranium and high-resolution computed tomography (HRCT) of the temporal bone showed bilateral malformation of the cochlea, with incomplete separation of the cochlea from the internal auditory canal (Figure 2a). No further malformation of the temporal bone, labyrinth and cerebellopontine angle was detected. The patient received his first cochlear implant (CI) on the left side at the age of 4 years. Losing his usable residual hearing at the right side by the age of 8 (Figure 1c), he underwent cochlear implantation of the second, right side for bilateral rehabilitation. Intraoperatively, a gusher phenomenon (profuse escape of cerebrospinal fluid after opening of the cochlea) was observed in both ears. Although the index patient is fully focused on spoken language communication, his hearing perception is reduced in situations of background noise. His performance in speech perception is well within the range of patients with inner ear malformation.¹⁷ No further clinical manifestations were obvious. In particular, renal laboratory parameters, abdominal ultrasound and ophthalmological examination were normal.

Three further male family members (II.3, III.5 and III.6) also suffer from severe hearing loss due to malformation of the cochlea, as evidenced by HRCT and MRI imaging (Figures 1b, 2b and c). Language

skills of these affected males are severely restricted, indicating prelingual onset of hearing loss and lack of appropriate treatment with hearing aids and CIs. Patient II.1 showed normal development in early childhood until he suffered from encephalitis at the age of 1 year, and presumably developed deaf-mutism and severe mental retardation as a consequence of this infection. Computed tomography of his cochlea showed no malformation but signs of a post-inflammation status, supporting the assumption that hearing loss is not inherited but acquired.

Four female family members (I.1, II.4, II.5 and II.6) developed mild to moderate hearing impairment during the third and fourth decade of their lives, respectively. Female II.4 (aged 55 years) suffers from mixed hearing loss and uses conventional hearing aids since ~20 years (Figure 1b). Female II.6 (aged 45 years) uses a hearing aid since 10 years due to a conductive hearing loss on the left side (Figure 1b). Reliable data for the reason of hearing loss in patients I.1 and II.5 are missing. Female III.10 shows normal thresholds of her right ear and suffers from low-grade conductive hearing loss on the left side (Figure 1b). The 46-year-old mother (II.2) of the index patient shows no hearing impairment to date, and computed tomography of her cochlea indicated no pathological findings (Figure 2d).

As only male family members are affected by severe sensorineural hearing loss already in early childhood, nonsyndromic X-chromosomal recessive hearing loss was most likely in this family.

Informed consent for genetic analyses was given by all ascertained family members.

Mapping and sequencing data

The pedigree of the family fits best to X-chromosomal inheritance of the hearing loss. Owing to the malformation of the cochlea, *POU3F4* was the most promising candidate gene. Sanger sequencing and MLPA analysis of *POU3F4* including the promoter region revealed no mutation. No distinct chromosomal rearrangements were detected by standard cytogenetic analysis using GTG banding, ruling out a translocation disrupting the *POU3F4* gene. Large copy number variations (CNVs) were excluded by SNP array analysis.

Further, the X-linked gene *PRPS1* was excluded in the index patient by Sanger sequencing, and linkage to the *DFNX3* locus on Xp21.2 was ruled out by analysis of markers located in or adjacent to the *DMD* gene (data not shown).

Sequencing of the X-chromosomal exome of the index patient (III.3), his mother (II.2) and his cousin (III.5) revealed 86 million passed filter reads (100 bp single reads), and an average coverage between 87 and 103 for the complete X-chromosomal exome. After alignment against the X-chromosomal reference sequence of hg19, we obtained 8–10 million reads per patient. About 10 000 exonic variants were scanned for each patient after quality filtering (Phred > 20 and coverage > 20 ×). Owing to additional filtering of variant frequencies and mutation type, as well as comparison of the detected variants in the female carrier and both male patients, we obtained a list of 31 shared variants in 23 different genes. About half of these variants could be excluded because they were annotated SNPs. Of the 15 remaining variants (in eight genes), six could not be confirmed by Sanger sequencing, and therefore represent NGS technical artefacts or can be explained by mis-aligned NGS reads due to existing pseudogenes. Segregation analyses of the remaining nine variants (in four genes: *NHS*, MIM 300457; *SLC6A8*, MIM 300036; *BCAP31*, MIM 300398 and *COL4A6*, MIM 303631) revealed that only a missense mutation in exon 23 of the gene encoding the alpha-6 chain of collagen type IV (*COL4A6*, NM_001847.2) perfectly co-segregated with the disease in the family (Figure 1a). The transition c.1771G>A resulting in the amino acid substitution p.Gly591Ser was

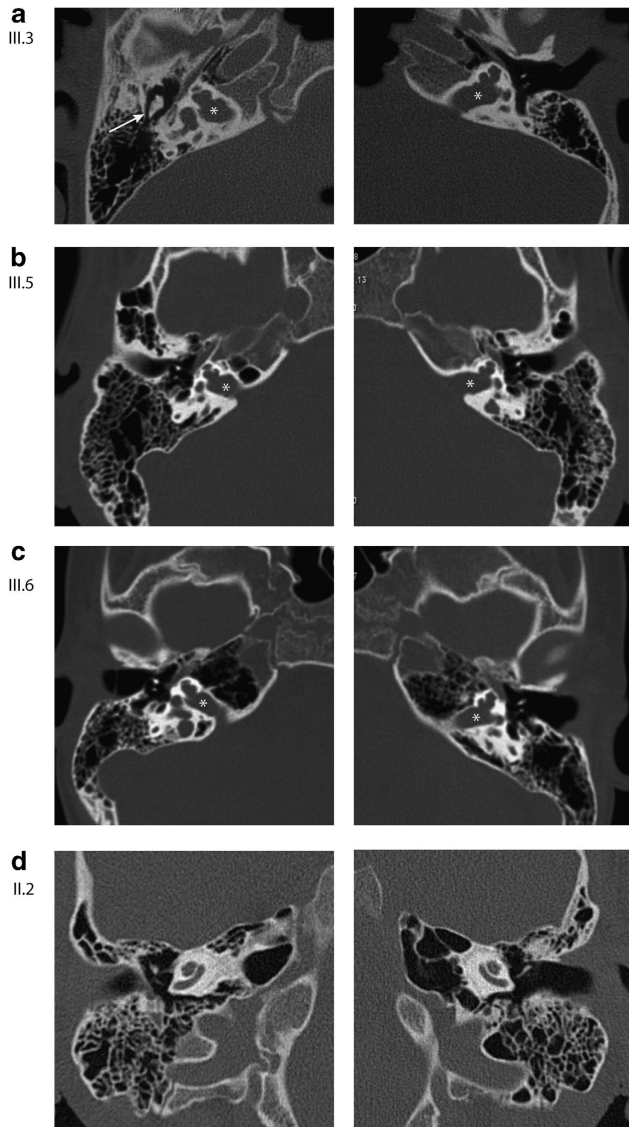


Figure 2 High-resolution imaging of the temporal bone of the index patient (III.3) at the age of 3 years, of his affected cousins (III.5 and III.6) and his healthy mother (II.2). Both sides are shown in each case. (a–c) Axial HRCT scans showing the isolated malformed cochlea, with incomplete partition of the cochlea and incomplete separation from the internal auditory canal (asterisks). The vestibular labyrinth is normal. The mastoid cavity and the middle ear ossicles appear regular (white arrow indicating malleo-incudial complex). (d) Representative images from a HRCT scan of the temporal bone in the mother of the index patient demonstrating a regular cochlea.

present in all affected male patients in a hemizygous form, except for patient II.1 who had suffered from an infantile encephalitis. All three obligatory female carriers and the slightly affected female family members carried the mutation in a heterozygous form. The substitution was absent from all unaffected male family members. Linkage analysis of the detected *COL4A6* variant revealed a LOD score of 2.36.

None out of 407 examined X chromosomes of male and female anonymous blood donors from Hungary and Central Europe showed the c.1771G>A transition in *COL4A6*, thus excluding a common polymorphism. Moreover, the mutation has not been described in any public SNP database.

Amplicon sequencing of *COL4A6* in 96 patients suffering from hearing loss (in whom mutations in *GJB2* had been excluded) could

not detect the candidate mutation c.1771G>A or any other causal mutation for hearing loss in the *COL4A6* gene.

Bioinformatic data

Bioinformatic tools that predict the effect of amino acid substitutions on protein function classified the missense mutation c.1771G>A, p.Gly591Ser as pathogenic (SIFT: deleterious, score: 0.00; PolyPhen: probably damaging, score: 1.00; Mutation Taster: disease causing, *P*-value: 0.983). According to Alamut, both nucleotide c.1771G and amino acid glycine-591 are highly conserved. Multiple sequence alignments of *COL4A6* protein sequences from different species (Figure 3a) and of the six alpha chains of human type IV collagens using MultAlin also showed high evolutionary conservation for glycine-591, which is substituted by serine in the affected members of the Hungarian family.

Five different splice-site prediction tools that are included in Alamut predicted minor to moderate effects on mRNA splicing in the mutated sequence, compared with the wild-type sequence ranging from –3.9 to –36.5% (mean of all five tools: –18.9%) for predicted change in acceptor splice-site usage (Table 1).

Bioinformatic analysis for protein stability and structure prediction using the described software showed striking differences in the collagen triple-helix stability and structure of the region containing the p.Gly591Ser variant compared with the wild type (Figure 3b).

COL4A6 mRNA and expression data

As the identified *COL4A6* missense mutation affects the first nucleotide of exon 23 and splice-site prediction software indicated a moderate effect on splicing probability caused by the mutation, we studied the *COL4A6* mRNA from fibroblasts of the index patient and his mother. Except for the c.1771G>A substitution, no difference in size, sequence or quantity compared with the *COL4A6* cDNA from control fibroblasts was detected (data not shown).

Immunohistochemical staining of the mouse inner ear showed that apart from multiple structures where a weak and somewhat diffuse staining was observed, most pronounced immune signals of *COL4A6* were visible at membranous and osseous structures at the *stria vascularis* of the spiral ligament (Figure 4a) and at higher magnification, in a subgroup of ganglia cells of the *ganglion spirale* (Figure 4b).

Whole-mount *in situ* hybridization was performed to assess the *col4a6* expression pattern in wild-type developing zebrafish embryos (Figure 4c). Around the 15–16 somite stage, the *col4a6* transcript was clearly visible in the dorsal midbrain and notochord, with a weak expression in the otic vesicle. At the 18 somite stage, expression in the otic vesicle became more prominent, whereas expression in the dorsal midbrain region persisted (not shown). Furthermore, faint expression was seen in the hindbrain and pharyngeal arch. During the early pharyngeal period (28 hpf), in addition to strong expression in the otic vesicle, high expression was noted in the mid hindbrain boundary and pharyngeal arches (not shown). At the same time, the expression in the notochord and dorsal midbrain diminished and a low level of expression appeared in the neural tube. After 42 hpf, expression in the ear and pharyngeal arches remained strong, but additional expression occurred in the swim bladder (Figure 4c) along with weak expression in the lateral line. A similar pattern of expression was also observed at 96 hpf, which did not differ remarkably from the pattern described in 42 hpf.

DISCUSSION

X-chromosomal exome sequencing revealed a missense mutation in *COL4A6* encoding the alpha-6 chain of type IV collagen, which co-segregated in all the affected members of a Hungarian family

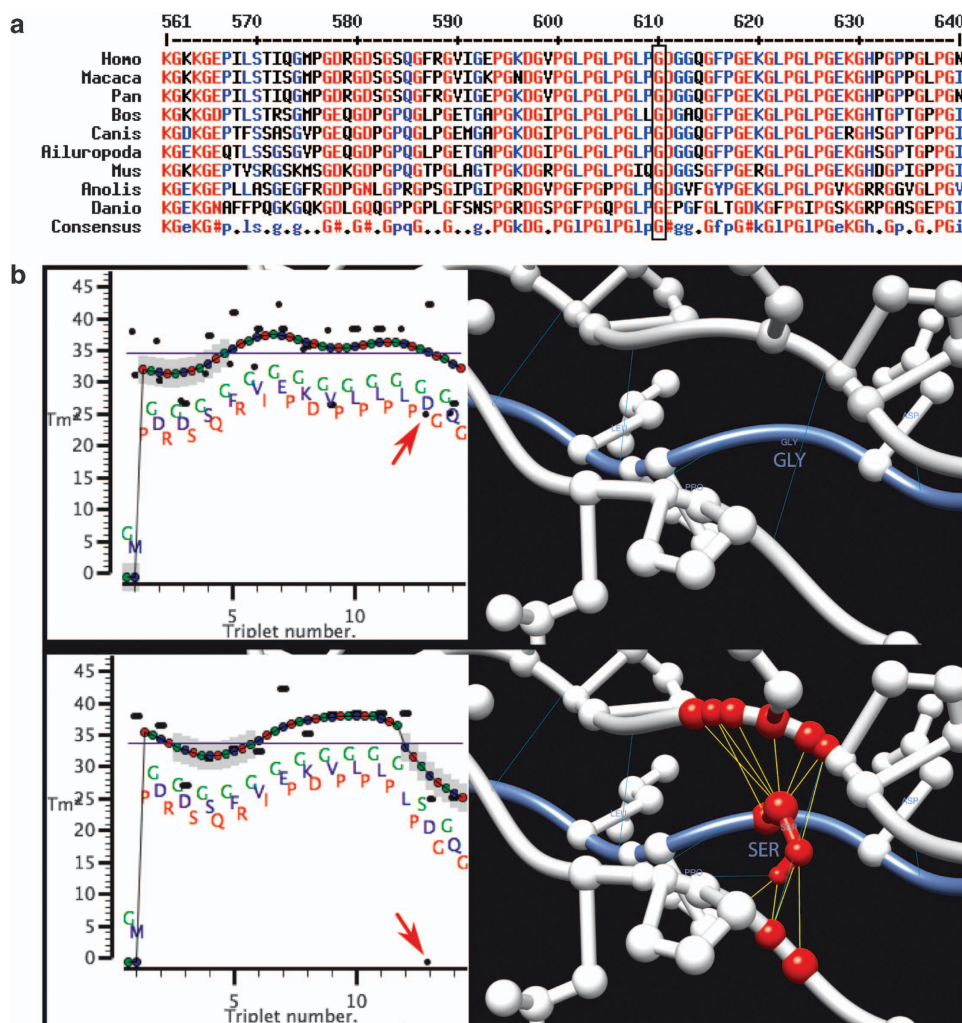


Figure 3 (a) Multiple sequence alignment of 80 amino acids encompassing glycine-591 of the human COL4A6 protein and orthologous proteins from other species using the software tool Multalign (<http://multalin.toulouse.inra.fr/multalin/>). Glycine-591 is perfectly conserved (black vertical bar; note that amino acid 591 is shifted by 19 positions to 610 in the alignment due to insertions in COL4A6 of the other species; red letters: high consensus, blue: low consensus, black: neutral). (b) Structural comparison of wild-type COL4A6 (upper part) with the Gly591Ser variant (lower part). Left subfigures illustrate estimated melting temperatures in degrees Celsius along the COL4A6 amino acid sequence. The respective mutation site is highlighted with an arrow. The Gly591Ser mutation seriously affects the melting temperature of the natural model. In the three-dimensional models (right subfigures), the COL4A6 chain is displayed in blue, whereas the white ribbons are the corresponding COL4A5 chains in the heterotrimeric structure. Atoms and bonds are colored in red, in case they clash or have unfavorable contacts with atoms of the other helices (van der Waals overlap $> 0.6 \text{ \AA}$). The pairings of clashing atoms are indicated by yellow lines. In the wild-type COL4A6, at position 591, Gly has no side chain at all and yields a clash-free conformation of the three chains. By contrast, the mutation to Ser produces an atypical large side-chain structure at the same position. In the model of the triple-helical collagen structure, this side chain has several hard clashes that reduce the stability of the interchain conformation and likely triggers disarrangement of the quaternary structure.

suffering from congenital nonsyndromic hearing loss. Bioinformatic analyses, expression studies and clinical data collectively provided evidence for the p.Gly591Ser variant underlying the hearing impairment in the three-generation family and for the type IV collagen having an important role in cochlea development. Although the missense mutation affects the first nucleotide of exon 23, aberrant splicing could not be demonstrated by *COL4A6* mRNA analysis from fibroblasts of the index patient. As splice-site prediction tools indicated low to moderate effects on splicing for the detected mutation, minor mis-splicing is considered possible, eventually resulting in an unstable product that escapes detection by PCR and subsequent sequencing. In addition, we cannot exclude abnormal splicing in the target tissue (inner ear), which was not accessible to analysis.

All members of the collagen superfamily display the characteristic triple-helical structure of three α chains, which is stabilized by the

presence of glycine as every third residue and a high content of proline and hydroxyproline.¹⁸ Thus, mutation of a glycine residue within the triple-helix structure, such as p.Gly591Ser in the α -1 chain, can be expected to exert a negative effect on both helical partners (two α -1 chains), destabilizing the whole collagen molecule that could be shown by bioinformatic collagen stability and structure prediction (Figure 3b). This is furthermore strongly supported by the fact that other genetic diseases of collagens, for example, osteogenesis imperfecta (*COL1A1*, MIM 120150; *COL1A2*, MIM 120160) or Ehlers–Danlos syndrome (*COL1A1*; *COL1A2*; *COL3A1*, MIM 120180; *COL5A1*, MIM 120215; *COL5A2*, MIM 120190) are predominantly caused by mutations of glycine residues in the corresponding collagen genes.¹⁹

The great diversity in the collagen superfamily is not only due to the existence of 28 different family members but also due to the

presence of several molecular isoforms of the triple helices formed either by homo- or heterotrimers.¹⁸ Type IV collagens are network-forming collagens that are made up of two molecular isoforms, for example, two alpha-5 chains combined with one alpha-6 chain [$\alpha 5(\text{IV})$]₂ $\alpha 6(\text{IV})$.²⁰ The genes *COL4A5* (MIM 303630) and *COL4A6* coding for the alpha-5 and alpha-6 chains of type IV collagens are located close to each other in a head-to-head orientation on chromosome Xq22.²¹ The transcription of both genes is controlled by a common bidirectional promoter element, but separate promoters for *COL4A5* and *COL4A6* allow also for tissue-specific gene

Table 1 Splice-site prediction (pred.) of five different software tools that are integrated in Alamut (Interactive Biosoftware), showing differences in acceptor splice-site usage of intron 22/exon 23 for the mutated sequence (Gly591Ser) compared with the wild-type *COL4A6* gene.

Prediction tool	Scale	Wild-type 3'	Gly591Ser 3'	Percentage of predicted change (%)
		splice site pred.	splice site pred.	
SpliceSiteFinder-like	0–100	80.6	76.7	–4.8
MaxEntScan	0–12	6.6	4.8	–27.3
NNSPLICE	0–1	0.9	0.7	–22.2
GeneSplicer	0–15	5.2	3.3	–36.5
Human Splicing Finder	0–100	80.2	77.1	–3.9

expression.²² All type IV collagens are components of basement membranes: $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ are expressed ubiquitously and the remaining alpha chains are only present in specific tissues. *COL4A5* is expressed in basement membranes of the kidney, skin, trachea, eye and neuromuscular junction, *COL4A6* is also present in these tissues except for the renal glomerular basement membrane.²² Consistently, the index patient of the Hungarian family showed normal renal function. *COL4A1–A5* could be localized in different structures of the murine cochlea by immunohistochemistry, however, *COL4A6* was not included in this investigation.¹⁵ In the present study, distinct expression of *COL4A6* in the *stria vascularis* and the *ganglion spirale* of the mouse inner ear were shown by immunohistochemical staining (Figures 4a and b). The current zebrafish genome (build Zv9) shows a single *COL4A6* ortholog on chromosome 7 (DRE7). Zebrafish *col4a6* is organized head-to-head with another type IV collagen gene, *col4a5*, on DRE7, demonstrating synteny conservation with human Xq22. Whole-mount *in situ* hybridization in wild-type developing zebrafish embryos showed dynamic expression of zebrafish *col4a6* in the ear, which is characterized by relatively low transcript levels in the otic vesicle during early developmental stages, followed by a more pronounced expression at later stages (42–96 hpf; Figure 4c). These data provide further evidence for the involvement of *COL4A6* in the hearing process.

Mutations in the human *COL4A3* and *COL4A4* genes are associated with autosomal recessive Alport syndrome (<20% of cases), whereas the more frequent X-linked form of Alport syndrome (>80%) is caused by mutations in *COL4A5*.²³ In Alport

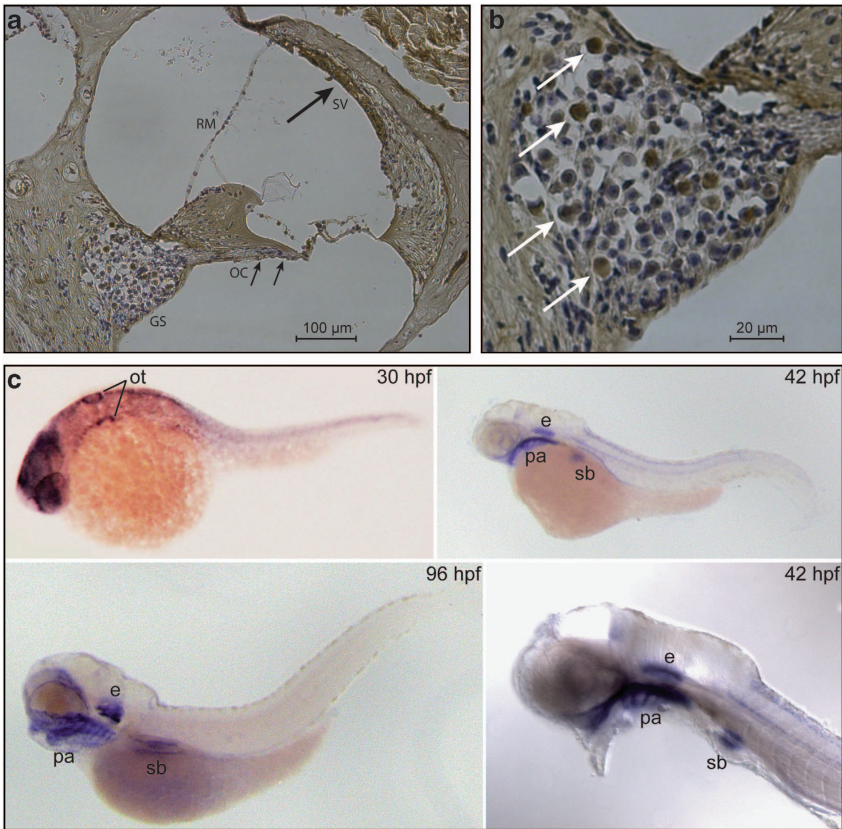


Figure 4 COL4A6 expression in the mouse inner ear and in the zebrafish embryo. (a) Thin sections of a demineralized murine inner ear were incubated with rabbit polyclonal Col4A6 antibody M137. COL4A6 is strongly expressed at membranous and osseous structures at the *stria vascularis* of the spiral ligament (black arrow), which were more intensely stained as compared with the basilar membrane underlying the organ of Corti (small black arrows). GS: *ganglion spirale*, OC: organ of Corti, RM: Reissner's membrane, SV: *stria vascularis*. (b) At higher magnification, a very distinct and pronounced reactivity was seen in a subgroup of ganglia cells of the *ganglion spirale* (white arrows). (c) Whole-mount *in situ* hybridization on zebrafish embryos of different developmental stages (e: ear, hpf: hours post fertilization, ot: otic vesicle, pa: pharyngeal arch, sb: swim bladder).

syndrome, renal basement membranes are affected, resulting in glomerulonephropathy and renal failure frequently accompanied by variable hearing impairment and ocular anomalies (MIM 301050 and 203780).²³ A subtype of X-linked Alport syndrome associated with diffuse leiomyomatosis is caused by a large deletion involving the two contiguous genes *COL4A5* and *COL4A6* (MIM 308940). This subtype is frequently combined with cataract and deafness.^{24,25} Mutations in *COL4A6* alone have not been associated with Alport syndrome or any other hereditary disorder so far. As renal problems have not been reported in any of the affected members of the Hungarian family in the present study, there is no evidence for Alport syndrome. Our results rather suggest that missense mutations in the *COL4A6* gene cause severe nonsyndromic hearing impairment in male patients. The female carriers in the family show variable phenotypes ranging from normal to mild or moderate hearing capabilities. Female II.2 showed normal hearing at the age of 46 (Figure 1b). In contrast, female subjects I.1, II.4 and II.6 developed hearing impairment in their third or fourth decade of life, requiring hearing aids. This variability in severity and onset in these female carriers might be explained by skewed X-inactivation, resulting in a higher or lower expression ratio of the mutated and wild-type *COL4A6* protein encoded by the two X-chromosomal gene copies. As there are no CT images available of the three affected carriers, no statement about the presence of cochlear malformations in female carriers can be made except for the healthy mother of the index patient who is a proven carrier but shows no malformation of her cochlea. The hearing impairment in females II.5 (left side only) and III.10 (Figure 1b) could be of mixed origin, as chronic otitis of the middle ear has been reported (more details not available).

Interestingly, the *COL4A6* gene is located only ~500 kb apart from the *DFNX1/PRPS1* locus. Mutations in the coding regions of *PRPS1* had been excluded in our family by Sanger sequencing. Most likely, the two genes cannot be discriminated by linkage analysis in pedigrees and, therefore, may not have been recognized before as separate loci. Microdeletions upstream of *POU3F4*, which is located about 24.6 Mb apart from *COL4A6* have been described repeatedly as causative for the most common form of X-linked nonsyndromic deafness associated with inner ear malformations.^{26,27} Both MLPA analysis comprising six probes in the critical upstream region and SNP array analysis did not provide any indication for CNVs in the *POU3F4* region. Additional Sanger sequencing of the coding region of *POU3F4* excluded this locus to be linked to hearing loss in the index patient.

In conclusion, our study provides evidence to suggest *COL4A6* as the fourth gene for X-linked hearing loss. In all affected male subjects of the Hungarian family, the p.Gly591Ser mutation segregates with severe hearing impairment and cochlear malformation, whereas further known X-linked deafness loci have been excluded. The alpha-6 chain of collagen type IV is part of the basement membrane of the inner ear and may have an important role in cochlea development. We propose that the p.Gly591Ser mutation is responsible for the observed malformation of the cochlea in the affected male subjects.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are deeply grateful to all the family members for their participation in this study, and to Dr Jörg Schröder, Professor Wafaa Shehata-Dieler and Professor Annerose Keilmann for clinical support. We thank Professor Tiemo Grimm for

calculation of the LOD scores, Dr Eberhard Schneider for inspecting radiographs of patients with hearing loss, and Dr Wolfram Kress for constructive discussions. We appreciate the donation of control samples by Dr Veronika Karcagi and Dr Ulrich Zechner. We also thank Professor Manfred Scharlt for his support with the whole-mount *in situ* experiment in zebrafish embryos. Additionally, we would like to thank Simone Günther, Birgit Halliger-Keller, Jens Gräf and Mohammed Uddin for technical assistance. This work was supported by the German Research Foundation (HA 1374/7–2).

- Bayazit YA, Yilmaz M: An overview of hereditary hearing loss. *ORL J Otorhinolaryngol Relat Spec* 2006; **68**: 57–63.
- Yan D, Liu XZ: Genetics and pathological mechanisms of Usher syndrome. *J Hum Genet* 2010; **55**: 327–335.
- Kokotas H, Petersen MB, Willems PJ: Mitochondrial deafness. *Clin Genet* 2007; **71**: 379–391.
- Kelsell DP, Dunlop J, Stevens HP et al: Connexin 26 mutations in hereditary nonsyndromic sensorineural deafness. *Nature* 1997; **387**: 80–83.
- Tekin M, Arnos KS, Pandya A: Advances in hereditary deafness. *Lancet* 2001; **358**: 1082–1090.
- Liu X, Han D, Li J et al: Loss-of-function mutations in the *PRPS1* gene cause a type of nonsyndromic X-linked sensorineural deafness, DFN2. *Am J Hum Genet* 2010; **86**: 65–71.
- de Kok YJ, van der Maarel SM, Bitner-Glindzicz M et al: Association between X-linked mixed deafness and mutations in the POU domain gene *POU3F4*. *Science* 1995; **267**: 685–688.
- Huebner AK, Gandia M, Frommolt P et al: Nonsense mutations in *SMPX*, encoding a protein responsive to physical force, result in X-chromosomal hearing loss. *Am J Hum Genet* 2011; **88**: 621–627.
- Schraders M, Haas SA, Weegerink NJ et al: Next-generation sequencing identifies mutations of *SMPX*, which encodes the small muscle protein, X-linked, as a cause of progressive hearing impairment. *Am J Hum Genet* 2011; **88**: 628–634.
- Lalwani AK, Brister JR, Fex J et al: A new nonsyndromic X-linked sensorineural hearing impairment linked to Xp21.2. *Am J Hum Genet* 1994; **55**: 685–694.
- Pfister MH, Apaydin F, Turan O et al: A second family with nonsyndromic sensorineural hearing loss linked to Xp21.2: refinement of the DFN4 locus within DMD. *Genomics* 1998; **53**: 377–382.
- Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
- Koboldt DC, Chen K, Wylie T et al: VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 2009; **25**: 2283–2285.
- Corpet F: Multiple sequence alignment with hierarchical clustering. *Nucl Acids Res* 1988; **16**: 10881–10890.
- Cosgrove D, Samuelson G, Pinnt J: Immunohistochemical localization of basement membrane collagens and associated proteins in the murine cochlea. *Hear Res* 1996; **97**: 54–65.
- Winkler C, Schafer M, Duschl J, Scharlt M, Volff JN: Functional divergence of two zebrafish midline growth factors following fish-specific gene duplication. *Genome Res* 2003; **13**: 1067–1081.
- Buchman CA, Copeland BJ, Yu KK, Brown CJ, Carrasco VN, Pillsbury HC: Cochlear implantation in children with congenital inner ear malformations. *Laryngoscope* 2004; **114**: 309–316.
- Ricard-Blum S: The collagen family. *Cold Spring Harb Perspect Biol* 2011; **3**: a004978.
- Dalgleish R: The human collagen mutation database 1998. *Nucleic Acids Res* 1998; **26**: 253–255.
- Borza DB, Bondar O, Ninomiya Y et al: The NC1 domain of collagen IV encodes a novel network composed of the alpha 1, alpha 2, alpha 5, and alpha 6 chains in smooth muscle basement membranes. *J Biol Chem* 2001; **276**: 28532–28540.
- Sugimoto M, Ohashi T, Ninomiya Y: The genes *COL4A5* and *COL4A6*, coding for basement membrane collagen chains alpha 5(IV) and alpha 6(IV), are located head-to-head in close proximity on human chromosome Xq22 and *COL4A6* is transcribed from two alternative promoters. *Proc Natl Acad Sci USA* 1994; **91**: 11679–11683.
- Segal Y, Zhuang L, Rondeau E, Sraer JD, Zhou J: Regulation of the paired type IV collagen genes *COL4A5* and *COL4A6*. Role of the proximal promoter region. *J Biol Chem* 2001; **276**: 11791–11797.
- Kashtan CE: Alport syndrome: an inherited disorder of renal, ocular, and cochlear basement membranes. *Medicine* 1999; **78**: 338–360.
- Cochat P, Guibaud P, Garcia-Torres R, Roussel B, Guarner V, Larbre F: Diffuse leiomyomatosis in Alport syndrome. *J Pediatr* 1988; **113**: 339–343.
- Anker MC, Arnemann J, Neumann K, Ahrens P, Schmidt H, König R: Alport syndrome with diffuse leiomyomatosis. *Am J Med Genet* 2003; **119A**: 381–385.
- de Kok YJ, Vossenaar ER, Cremers CW et al: Identification of a hot spot for microdeletions in patients with X-linked deafness type 3 (DFN3) 900 kb proximal to the DFN3 gene *POU3F4*. *Hum Mol Genet* 1996; **5**: 1229–1235.
- Song MH, Lee HK, Choi JY, Kim S, Bok J, Kim UK: Clinical evaluation of DFN3 patients with deletions in the *POU3F4* locus and detection of carrier female using MLPA. *Clin Genet* 2010; **78**: 524–532.