npg

ORIGINAL ARTICLE

Exome sequencing identifies a novel *CEACAM16* mutation associated with autosomal dominant nonsyndromic hearing loss DFNA4B in a Chinese family

Honghan Wang[1,2,3,8], Xinwei Wang[1,3,8], Chufeng He[1,3], Haibo Li[4], Jie Qing[5], M'hamed Grati[5], Zhengmao Hu[4], Jiada Li[4], Yiqiao Hu[4], Kun Xia[4], Lingyun Mei[1,3], Xingwei Wang[1,3], Jianjun Yu[2], Hongsheng Chen[1,3], Lu Jiang[1,3], Yalan Liu[1,3], Meichao Men[1,6], Hailin Zhang[2], Liping Guan[7], Jingjing Xiao[7], Jianguo Zhang[7], Xuezhong Liu[1,5] and Yong Feng[1,3,4]

Autosomal dominant nonsyndromic hearing loss (ADNSHL/DFNA) is a highly genetically heterogeneous disorder. Hitherto only about 30 ADNSHL-causing genes have been identified and many unknown genes remain to be discovered. In this research, genome-wide linkage analysis mapped the disease locus to a 4.3 Mb region on chromosome 19q13 in SY-026, a five-generation nonconsanguineous Chinese family affected by late-onset and progressive ADNSHL. This linkage region showed partial overlap with the previously reported DFNA4. Simultaneously, probands were analyzed using exome capture followed by next-generation sequencing. Encouragingly, a heterozygous missense mutation, c.505G > A (p.G169R) in exon 3 of the *CEACAM16* gene (carcinoembryonic antigen-related cell adhesion molecule 16), was identified via this combined strategy. Sanger sequencing verified that the mutation co-segregated with hearing loss in the family and that it was not present in 200 unrelated control subjects with matched ancestry. This is the second report in the literature of a family with ADNSHL caused by *CEACAM16* mutation. Immunofluorescence staining and western blots also prove CEACAM16 to be a secreted protein. Furthermore, our studies in transfected HEK293T cells show that the secretion efficacy of the mutant CEACAM16 is much lower than that of the wild type, suggesting a deleterious effect of the sequence variant.

INTRODUCTION

Hearing loss (HL) is one of the most common sensorineural deficits in humans worldwide.[1] It can be due to genetic causes, environmental factors or a complex combination of the two. More than 50% of the profound early-onset deafness cases are caused by genetic factors in developed countries.[2,3] Genetic HL may occur as part of a multisystem disorder, or as disease restricted to the cochlea and vestibular system with no additional abnormalities. The latter is the main form accounting for about 70% of neonates failing newborn hearing screening.[4] Furthermore, autosomal dominant nonsyndromic HL (ADNSHL/DFNA) accounts for 20–25% of hereditary nonsyndromic sensorineural deafness cases and is most often postlingual in onset.[5]

Identification of the pathogenic mutations underlying ADNSHL is still difficult because of the incredible genetic heterogeneity regarding HL, and likelihood that ADNSHL is not entirely caused by monogenic mutations (also known as Mendelian disorder).[6] Hitherto, only 30 distinct deafness genes have been identified while 64 DFNA loci have been mapped in the past two decades (http://hereditaryhearingloss.org). The genetic heterogeneity is a major hurdle to the efficient discovery of disease-causing genes. Traditional positional candidate approach, which was once the most promising strategy for cloning genes underlying Mendelian disorder, has amassed an impressive list of pathogenic gene discoveries for ADNSHL. Still, several challenges remain before this approach becomes firmly entrenched, such as locus

[1]Department of Otolaryngology, Xiangya Hospital, Central South University, Changsha, Hunan, China; [2]Department of Head and Neck Surgery, Hunan Provincial Tumor Hospital and the Affiliated Tumor Hospital of Xiangya Medical School, Central South University, Changsha, Hunan, China; [3]Province Key Laboratory of Otolaryngology Critical Diseases, Changsha, Hunan, China; [4]State Key Laboratory of Medical Genetics of China, Changsha, Hunan, China; [5]Department of Otorhinolaryngology, University of Miami, Miller School of Medicine, Miami, FL, USA; [6]Health Management Center, Xiangya Hospital, Central South University, Changsha, Hunan, China and [7]BGI-Shenzhen, Shenzhen, Guangdong, China
[8]These authors contributed equally to this work.
Correspondence: Professor Y Feng, Department of Otolaryngology, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, 410008, Hunan, China.
E-mail: fengyong_hn@hotmail.com
or Professor X Liu, Department of Otorhinolaryngology, University of Miami, Miller School of Medicine, 1601 NW 12 Ave, Rm. 6023, Miami, 33136, FL, USA.
E-mail: xliu@med.miami.edu
Received 7 July 2014; revised 21 November 2014; accepted 25 November 2014; published online 15 January 2015

heterogeneity whereby hundreds of genes assemble in the candidate region and in several cases the availability of only small families where linkage mapping would not be informative.[7,8] Recently, the application of next-generation sequencing of targeted regions has been developing dramatically and especially exome sequencing technology has already provided a new powerful tool to identify causative genes.[9] In the present study, we identified efficiently a novel *CEACAM16* (carcinoembryonic antigen-related cell adhesion molecule 16) mutation in a Chinese family with ADNSHL by exome sequencing in parallel with linkage analysis. *In vitro*, functional analyses suggest a deleterious effect of a single nucleotide variant.

MATERIALS AND METHODS

Ethics statement
This research involving human participants was formally approved by the Medical Ethics Committee of Xiangya Hospital, Central South University, Changsha, China. Written informed consent was obtained from all subjects (including the control individuals) or their legal guardians.

Subjects and clinical assessment
We investigated a five-generation ADNSHL family SY-026 from Hunan province of mainland China. Thirty-four members, including 13 deaf (III:2, III:12, III:15, III:17, III:19, IV:13, IV:16, IV:22, IV:24, IV:26, IV:31, V:10 and V:13), 17 normal hearing (II:8, III:4, III:6, III:8, III:13, III:14, III:16, III:18, IV:14, IV:15, IV:17, IV:18, IV:20, IV:21, IV:29, IV:30 and IV:32) and 4 subjects younger than the onset age whose hearing status is ambiguous (V:11, V:12, V:14, V:16) were recruited (Figure 1a; Table 1). All participants were subjected to particular physical and otological examinations by two experienced otologists independently. Pure-tone audiometry was performed to define hearing threshold levels (dB HL) for both air and bone conduction at frequencies of 250, 500, 1000, 2000, 4000, 6000 and 8000 Hz. Previous audiological tests were collected if available. Pure-tone audiometry average thresholds of the air conduction, which were based on the frequencies at 500, 1000 and 2000 Hz in the better hearing ear, were used to determine the degree of HL. The list of the classification criteria was as follows: normal (<15 dB HL), slight (16–25 dB HL), mild (26–40 dB HL), moderate (41–55 dB HL), moderately severe (56–70 dB HL), severe (71–90 dB HL) and profound HL (>90 dB HL).[10] Two affected individuals underwent computed tomography scan of the temporal bone and vestibular testing. Vestibular function was assessed by videonystagmography using the System 2000 (Micromedical Technologies, Chatham, IL, USA). Videonystagmography protocol included saccade test, eye tracking test, optokinetic test, gaze test, spontaneous nystagmus test, positional test, Dix-Hallpike test, Roll test and caloric test. The detailed medical history was obtained by questionnaire to eliminate the interference of environmental factors. Genomic DNA was extracted from peripheral venous blood by standard phenol extraction protocols. Mutations in the *GJB2*, *GJB3*, *EYA4* and *KCNQ4* genes have been excluded in this family. After being informed, 200 Han Chinese control subjects (between the ages of 30 and 65), whose health physical examination items contained pure-tone audiometry and the result showed the hearing threshold was less than 15 dB HL, were also collected through health management center, Xiangya Hospital, Central South University. All control subjects had no family history of HL.

Genome-wide genotyping and linkage analysis
The purification of DNA samples from 22 subjects whose phenotypes were clear (including 12 affected and 10 unaffected family members) were genotyped using commercially available HumanLinkage-24 BeadChip Kit (Illumina, San Diego, CA, USA), which contains a total of 6000 high-density single nucleotide polymorphism (SNP) markers (Figure 1a). The affected member III:12 and other four participants (V:11, V:12, V:14 and V:16) were excluded from the linkage analysis studies because of positive medical history of meningitis and exposure to ototoxic drug or less than the onset of age. Two-point logarithm of odds scores were calculated by the Merlin programs (version 5.2) under an autosomal dominant inheritance model with a mutant allele frequency of 0.0001 and complete penetrance. The allele frequencies of markers and the recombination frequencies were assumed to be equal to each other.[11,12]

Exome sequencing and bioinformatics analysis
The genomic DNA samples were randomly sheared to obtain fragments with an average length between 250 and 300 bp using adaptive focused acoustics from Covaris Inc., Woburn, MA, USA. Then adapters were ligated to both ends of the fragments. Extracted DNA was then amplified by ligation-mediated PCR, purified and hybridized to the custom Nimblegen SeqCap EZ Library (Roche/NimbleGen, Madison, WI, USA) for enrichment. Both non-captured and captured ligation-mediated PCR products were subjected to quantitative PCR to estimate the magnitude of enrichment. Each captured library was then loaded on a Hiseq 2000 platform (Illumina Inc.). High-throughput sequencing was performed to meet the desired average sequencing depth. Raw image files were processed by Illumina basecalling Software 1.7 (Illumina) for base-calling with default parameters and the sequences of each individual were generated as 90 bp pair-end reads.

SOAPaligner (soap2.21) and Burrows-Wheeler Aligner tools were used to align clean reads to the human reference genome (hg 19 version) for SNP and insertion/deletion (Indel), respectively.[13,14] And then, the software SOAPsnp was used to assemble the consensus sequence and call genotypes in target regions.[15] We filtered candidate SNPs with the following criterion: SNP quality was ⩾20, the sequencing depth was between 4- and 1000-fold, the estimated copy number was ⩽2, and the distance between two SNPs was >5. Indels were detected through the alignment result with Genome Analysis Toolkit.[16] Functional effects of the variants were predicted by programs of Sorting Intolerant From Tolerant.[17]

Sanger sequencing
PCR was carried out using the standard protocol (Qiagen, Germantown, MD, USA). The PCR products were treated with the shrimp alkaline phosphatase and exonuclease-I to degrade deoxynucleotide triphosphates and unincorporated PCR primers. The purified amplicons were mixed with 10 picomoles of the forward and reverse PCR primers for bidirectional sequencing on an ABI-Prism 3100 DNA sequencer via dye termination chemistry (Applied Biosystems, Foster City, CA, USA). DNA sequence variants were identified using the SeqMan II program (DNA-STAR Inc., Madison, WI, USA).

Plasmid DNA construction
The pRc/CMV-hCEACAM16 was kindly provided by Tumor Immunology Laboratory, LIFE Center, University Hospital of Munich, Ludwig-Maximilians-University Munich, Germany. The h*CEACAM16* was sequenced and cloned using the primers: 5′-CGATA*GAATTC*GCCACCATGGCGCTGACTGGGTA CAG-3′ and 5′-GCATA*TCTAGA*CCCCAGGGGGGCCACCTGCA-3′. The coding sequence of *CEACAM16* was inserted into pRK5-Flag at *Eco*RI/*Xba*I site to add a Flag tag to the C-terminus of *CEACAM16*. The mutant type G169R of *CEACAM16* was generated by using QuikChange II Site-Directed Mutagenesis (GE Healthcare Inc., Buckinghamshire, UK). All constructions were verified by Sanger sequencing.

Plasmid transfection, amplification and extraction
The two kinds of plasmid pRK5-wild type (WT) and G169R mutant (MT) *CEACAM16*-Flag were transformed into One Shot TOP10 competent *E. coil* (Life Technologies Inc., Carlsbad, CA, USA) on LB culture plates with ampicillin. Then the monoclonal flora was picked up and mass cultured in LB medium with ampicillin. The amplified plasmids were isolated, purified from bacterial cultures by HiSpeed Plasmid Midi Kit (Qiagen).

Culture and immunofluorescence of COS7 cells
The COS7 cells were dissociated with trypsin/EDTA (Invitrogen Inc., Manassas, VA, USA) and plated onto coverslips with 2% fetal bovine serum/DMEM medium (Invitrogen) in six-well culture plates. Three micrograms of plasmid was diluted into 200 μl of jetPRIME buffer and mixed with 4 μl jetPRIME reagent (jetPrime DNA transfection reagent, appled by Polyplus Transfection Inc., New York, NY, USA). The mixture was incubated at room temperature for 15 min and added to the cells (Polyplus Transfection Inc.). Cells were
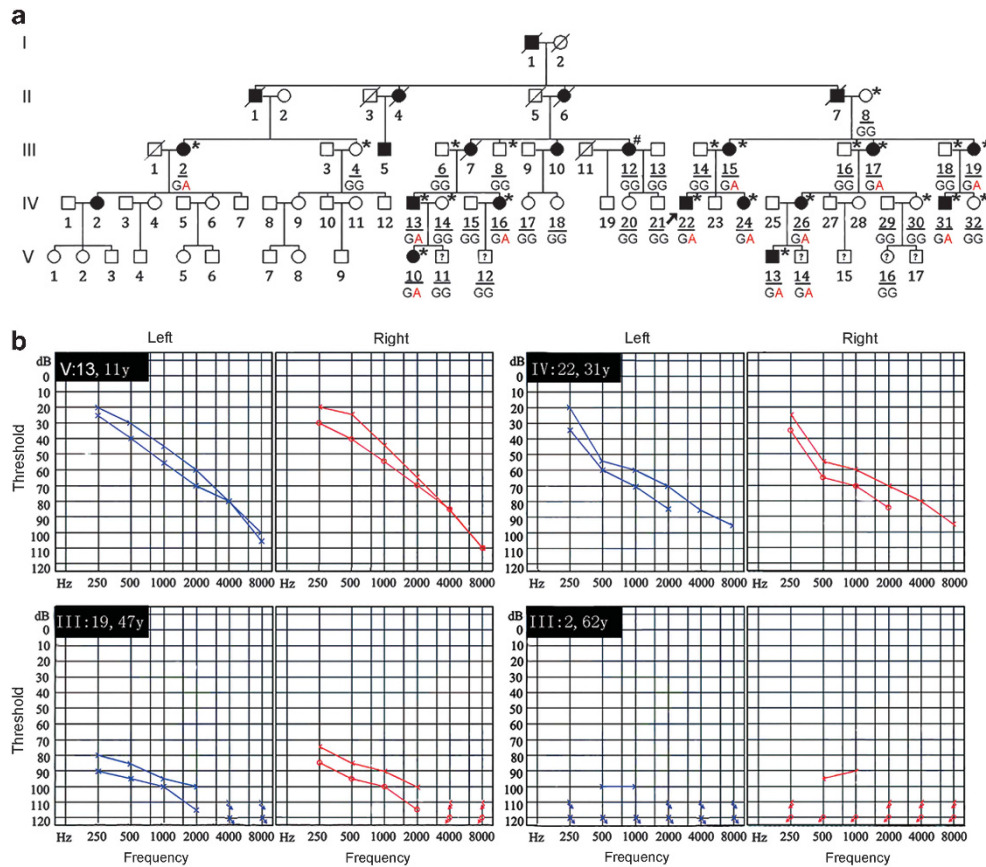
Figure 1 Pedigree of a large Chinese family (SY-026) with late-onset ADNSHL carrying the missense G169R mutation in *CEACAM16* and the audiograms of four affected subjects from the family. (**a**) Pedigree of the family shows an autosomal-dominant inheritance pattern. The circular and square symbols represent female and male, and the black and white ones indicate affected and unaffected individuals, respectively. The subjects younger than the onset age whose hearing status is ambiguous are marked by question mark and the deceased are differentiated by a slash. Arrow shows the proband (IV:22). The bars below each symbol indicate individuals involved in this study. Twenty-two family members included in the linkage analysis are designated by asterisk on each right shoulder of the symbol. The genotypes at the c.505G>A mutation site of *CEACAM16* are also presented for each enrolled individual. (Note: #The affected member of III:12 was excluded from linkage analysis because the hearing impairment was caused by environmental factors.) (**b**) Audiograms of four affected individuals from the family. By convention, the air conduction results are displayed on audiogram using blue '×' for left ear and red '○' for right ear. And likewise, bone conduction results are displayed on audiogram using blue 'greater than' angle brackets '>' for left ear and red 'less than' angle brackets '<' for right ear.

incubated for overnight, at 37 °C, and then washed by phosphate-buffered saline (PBS) solution gently.

COS7 cells were fixed with 4% paraformaldehyde, washed with PBS, permeabilized with 0.2% Triton/PBS, blocked with 10% goat serum in PBS for 1 h, incubated with rabbit anti-flag and phalloidin Alexa fluor-647 for 1 h, stained with Alexa Fluor 488, 568 for 30 min, and then washed with PBS. At last, Gold antifade reagent was added to mount onto coverslips. The cells were captured using confocal microscope. The image data were analyzed using MetaMorph software (Molecular Devices, LLC, Inc., Downingtown, PA, USA).

HEK293T cells cultures and western blotting

The HEK293T cells were grown at 37 °C under 5% CO_2 in DMEM which contained 10% fetal bovine serum and 100 U ml^{-1} of penicillin/streptomycin. One day before transfection, the HEK293T cells were seeded at a proximate 50% confluency in 10 cm culture plates. Cells were transfected with 6 µg plasmids for each. After 6–8 h, we replaced the cell culture medium with DMEM only with 100 U ml^{-1} of penicillin/streptomycin.

At 48 h after transfection, cells were lysed in 2×SDS sample buffer, containing 1mM phenylmethanesulfonyl fluoride and 1×protease inhibitor cocktail (Sigma-Aldrich Inc., St. Louis, MO, USA). Samples were incubated for 10 min, at 98 °C, and quantified using BCA Protein Assay Kit (Thermo Fisher

Scientific Inc., Waltham, MA, USA). Culture medium from *CEACAM16*-transfected cells were collected and centrifuged at 4000 r.p.m. for 10 min, 4 °C, and filtered through 0.22 µm filters, then centrifuged at 14000×*g* for 10 min, 4 °C. Proteins in supernatants were precipitated by 100% trichloroacetic acid, and washed by acetone for four times. The same amount of protein from cell abstract and supernatant (both were 20 µg) were separated on 8% SDS-PAGE and transferred onto a PVDF membrane. The membranes were blocked in Tris-buffered saline supplemented with 5% non-fat milk for 1 h at room temperature and then incubated with mouse monoclonal anti-Flag M2 antibody (1:1000 dilution, Sigma-Aldrich) or mouse anti-β-actin antibody (1:1000 dilution, Sigma-Aldrich) at room temperature for 2 h. After washing with Tris-buffered saline supplemented with 0.1% Tween 20 (Sigma-Aldrich), the membranes were incubated for 1 h at room temperature with a horseradish peroxidase-conjugated secondary anti-mouse IgG antibody (1:10000 dilution, Sigma-Aldrich). Detection was performed using the ECL plus Western blotting detection system (GE Healthcare) according to the manufacturer's instruction.

RESULTS

Clinical findings and the features of the family

The Chinese SY-026 family's HL was consistent with ADNSHL. The majority of the affected family members appeared to have developed a

Table 1 Clinical data and genotypic characteristics of all participants from SY-026 family

Subjects	Gender	Age at test (years)	Age of onset (years)	Severity[a]	Tinnitus	Vertigo	Environmental factors	Alleles[b]
III:2	F	62	25	Profound	+	−	−	G/A
III:12	F	50	3	Profound	+	−	Meningitis; aminoglycoside (+)	G/G
III:15	F	54	28	Profound	+	−	−	G/A
III:17	F	49	22	Profound	+	−	−	G/A
III:19	F	47	21	Profound	+	−	−	G/A
IV:13	M	35	10	Severe	+	−	−	G/A
IV:16	F	32	13	Severe	+	−	−	G/A
IV:22	M	31	15	Severe	+	−	−	G/A
IV:24	F	21	14	moderately severe	−	−	−	G/A
IV:26	F	30	16	Severe	+	−	−	G/A
IV:31	M	28	14	moderately severe	+	−	−	G/A
V:10	M	12	12	Moderate	-	−	−	G/A
V:13	M	11	11	Moderate	-	−	−	G/A
II:8	F	72	−	Normal	−	−	−	G/G
III:4	F	60	−	Normal	−	−	−	G/G
III:6	M	61	−	Normal	−	−	−	G/G
III:8	M	56	−	Normal	−	−	−	G/G
III:13	M	52	−	Normal	−	−	−	G/G
III:14	M	59	−	Normal	−	−	−	G/G
III:16	M	53	−	Normal	−	−	−	G/G
III:18	M	50	−	Normal	−	−	−	G/G
IV:14	F	35	−	Normal	−	−	−	G/G
IV:15	M	36	−	Normal	−	−	−	G/G
IV:17	F	22	−	Normal	−	−	−	G/G
IV:18	F	16	−	Normal	−	−	−	G/G
IV:20	F	12	−	Normal	−	−	−	G/G
IV:21	M	10	−	Normal	−	−	−	G/G
IV:29	M	25	−	Normal	−	−	−	G/G
IV:30	F	23	−	Normal	−	−	−	G/G
IV:32	F	23	−	Normal	−	−	−	G/G
V:11	M	9	−	intangibility	−	−	−	G/G
V:12	M	6	−	intangibility	intangibility	−	−	G/G
V:14	M	7	−	intangibility	intangibility	−	−	G/A
V:16	F	3	−	intangibility	intangibility	−	−	G/G

Abbreviations: F, female; M, male.
+, present; −, not present.
[a]According to criteria in published literature (Clark[10]).
[b]Alleles refer to CEACAM16 c.505, at position 45207410 (hg19).

high-frequency tinnitus at the onset of HL. The affected subjects had very similar phenotypes, characterized by bilateral, postlingual, sensorineural, progressive HL (except III:12). Audiological examination of affected members suggested that HL in this pedigree was late onset and that higher frequencies were initially decreased in the first or second decade and progressed to profound HL involving all frequencies. Specially, the III:12 had normal hearing at birth but was affected with sudden profound sensorineural HL at 3 years of age by aminoglycoside exposure to treat meningitis. There was no any apparent evidence for vestibular dysfunction or otological associated malformations by computed tomography and vestibular testing. The main clinical and audiological findings of affected subjects are shown in Table 1 and Figure 1b.

Linkage analysis combined with exome sequencing
To identify the disease locus, two-point genome-wide linkage analysis was performed on the family members that were indicated by asterisks (Figure 1a). The result identified a single locus on chromosome (chr) 19 with a maximum logarithm of odds score of 4.448 and no significant logarithm of odds scores (>3)[18] were obtained for any

other regions of the genome. The critical interval is flanked by genetic markers rs11671074 and rs1236093 (physical position: chr19: 44,804,119–49,104,308 and physical map: chr19q13.31–13.33), and spans approximately 4.3 Mb (Figure 2a and b). The logarithm of odds score obtained for each genetic marker is presented in Supplementary Table S1.

Exome sequencing was performed on three affected subjects (IV:24, IV:13 and III:2) and one unaffected subject (III:4). The mean depth of target regions ranged from 49- to 76-fold, the coverage of target region reached up to 99% and the average read length was about 90 bp. The detail information of the data production can be found in Supplementary Table S2. Following alignment and quality assessment of the data, >10 000 variants (SNPs and Indels) were identified in subjects. Synonymous variants were excluded and the other were then filtered against several control data sets, including dbSNP135, the 1000 Genomes project, exome data from eight HapMap individuals previously sequenced and YH database (YanHuang Project). The remaining variants were further selected by segregating from the unaffected member (control) and the three affected members. Finally, a solitary heterozygous missense variant, c.505G>A in exon 3 of
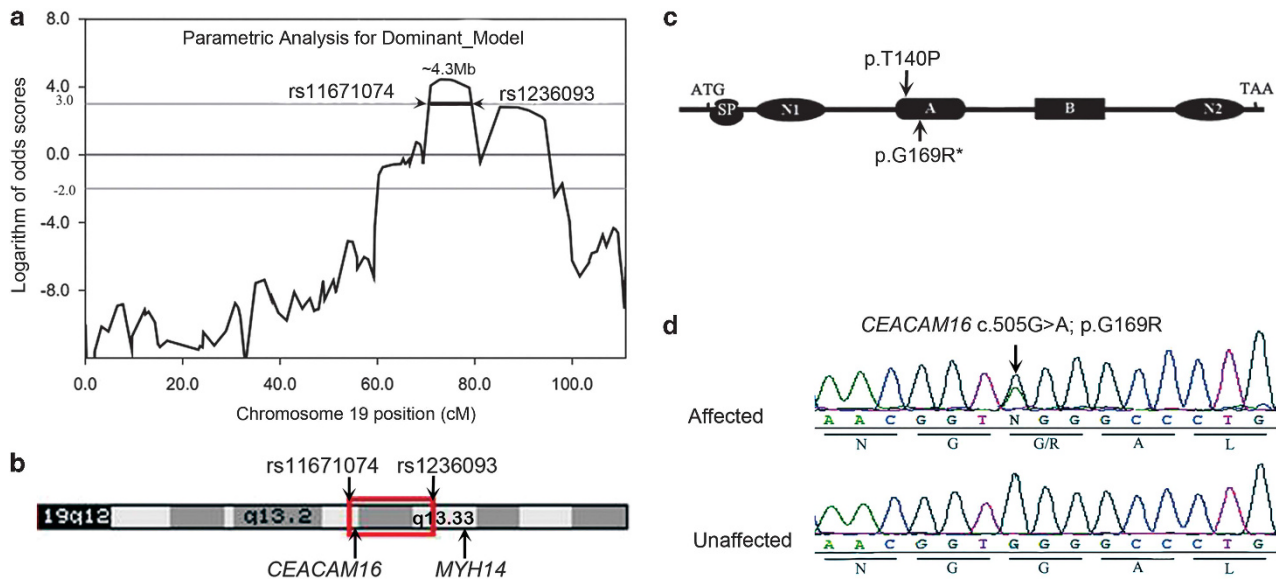
Figure 2 Combinational strategy of linkage analysis and exome sequencing identifies a novel *CEACAM16* mutation as causing ADNSHL. (**a**) SNP linkage analysis maps the disorder locus to chromosome 19 (two-point lod scores of more than 3). The critical interval is flanked by genetic markers rs11671074 and rs1236093 and spans approximately 4.3 Mb. (**b**) Genome bioinformatics analysis maps physical position to chr19: 44,804,119–49,104,308 and physical map to chr19q13.31–13.33 according to UCSC (Unverisity of California, Santa Cruz) Genome Bioinformatics (http://genome.ucsc.edu/cgi-bin/hgGateway). (**c**) Schematic structure of *CEACAM16* shows both mutations previously reported in American 1070 family and in the Chinese SY-026 family in this report occurred in the IgC-like domain of subtype A. (**d**) Identification of a novel heterozygous missense G169R mutation in the *CEACAM16* gene. Arrow shows the position of the mutation.

CEACAM16, was mapped to the linkage region on chr19q13.31–13.33 (Figure 2b). This variant was predicted to be damaging by Sorting Intolerant From Tolerant and co-segregated with HL in the whole family by Sanger sequencing (Figure 2d; Figure 1a). We further confirmed the absence of the mutation in 200 unrelated, ethnically and geographically matched controls. This missense mutation was also absent from the Human Gene Mutation Database (http://www.hgmd.org/) and has not been reported in any published literature, to our knowledge.

We transfected pRK5-WT and MT *CEACAM16*-flag plasmids into both COS7 cells and HEK293 cells. Immunocytochemical studies on transfected COS7 cells showed the expression quantity of CEACAM16 protein, both intracellular and extracellular. The cytoskeleton was counterstained for actin to observe the extracellular precipitation clearly. There was no distinct difference in the expression quantity between WT and MT CEACAM16 proteins (Figure 3a and b). However, the western blotting result in transfected HEK293T cells seemed to be different (Figure 3c). The cell medium used to culture those transfected HEK293T cells were collected separately and centrifuged to remove cell and possible exosomes. The extracted proteins were detected as: a monomeric band migrating at about 50–55 kDa,[19] an oligomeric ladder (dimer to hexamer, ~ 100–300 kDa) and bands of β-actin. Although the β-actin was seen in both WT and MT CEACAM16-transfected cell pellets, either no or very little oligomeric band could be found in the culture medium of MT CEACAM16-transfected cell.

DISCUSSION

In this study, we described the clinical features and genetic pattern of a five-generation Chinese family (SY-026) with clinical traits of ADNSHL. We identified a candidate region mapping to chr19q13.31–13.33 that partially overlaps with the previously reported DFNA4

locus.[20] Prior to this research, the DFNA4 locus was originally mapped to chromosome 19q13 in an American family (1070) in 1995.[20] Subsequently, the second family from Germany (Leipzig 9) and then several other families around the world were identified whose mutations could be mapped to DFNA4 locus.[21–25] Because the *MYH14* gene, encoding a non-muscle myosin heavy-chain, was the first gene known to correspond to DFNA4, we sequenced it in the probands of family SY-026 and no pathogenic mutation was revealed. Then, we performed exome sequencing analysis on four family members and successfully identified a c.505G > A (p.G169R) in exon 3 of the *CEACAM16* gene. The p.G169 is conserved in some species, including elephant and a number of primates (such as *Pan troglodytes*, *Gorilla gorilla*, *Pongo abelii*, *Nomascus leucogenys* and *Callithrix jacchus*).[26] This mutation co-segregated with the phenotype and was absent in 200 controls. At the time of the study, Zheng *et al.*[19] reported their identification of a heterozygous missense *CEACAM16* mutation in the family 1070 (that is, originally reported by Chen *et al.*[20] in 1995) by positional candidate gene approach. The clinical features of that family were fluctuating, progressive, sensorineural HL that began in the second decade and led to severe-to-profound HL by the age of 40 years. Likewise, the phenotypes in our family was very similar to those of the family 1070, except III:12 whose HL was definitely caused by meningitis and aminoglycoside exposure. By contrast, HL was non-fluctuating and several individuals manifested an earlier onset age (the youngest was at the age of 10 years) in our family. Moreover, the concomitant symptom of bilateral continuous high-frequency tinnitus was first reported in most of cases with DFNA4, to our knowledge.

The *CEACAM16* gene belongs to the CEACAMs family, a subgroup of the immunoglobulin (Ig) superfamily. CEACAMs comprise a cluster of both conserved and independently evolved proteins that are involved in various physiological and pathophysiological processes,
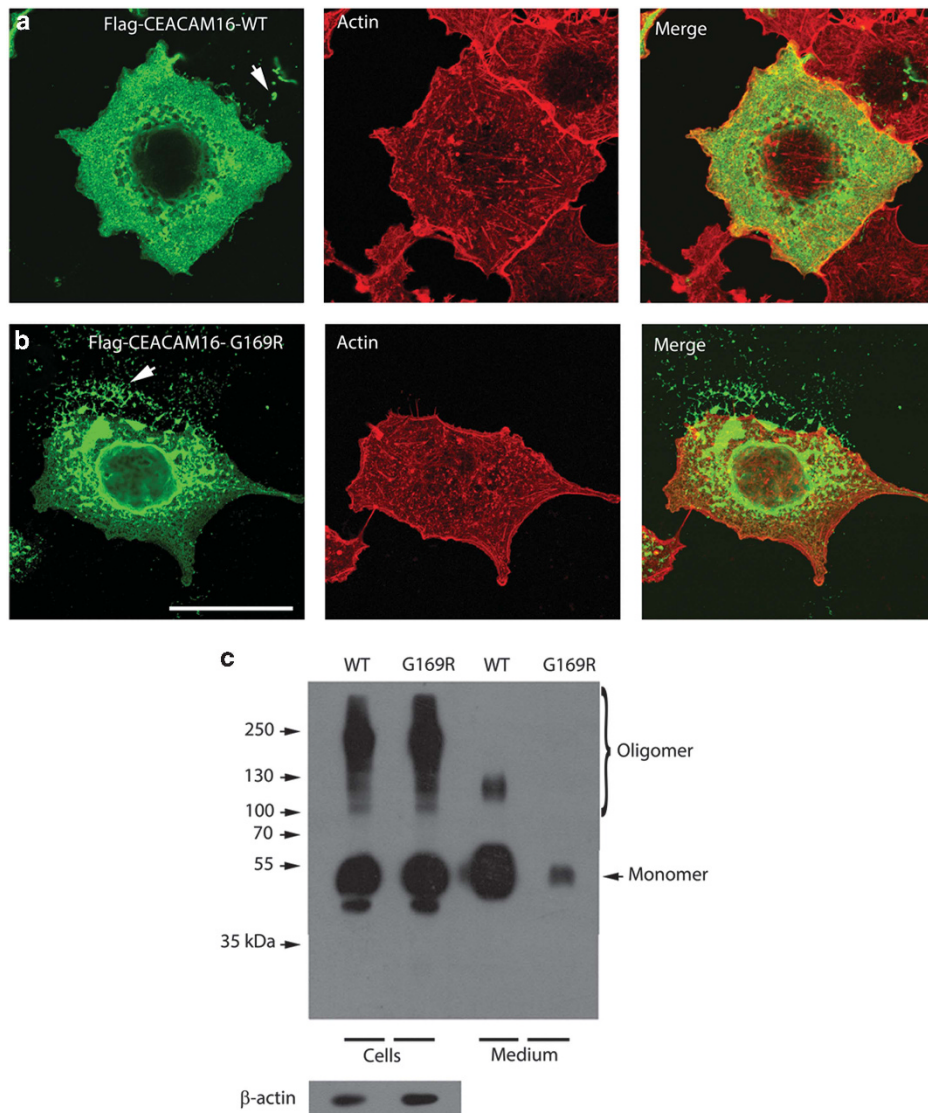
Figure 3 Effects of G169R mutation on CEACAM16 trafficking and secretion. (**a**, **b**) Confocal images of COS7 cells expressing Flag-tagged wild-type (WT, **a**) and G169R mutant (**b**) CEACAM16 (green), counterstained for actin (red). Arrows in **a** and **b** show secreted proteins. (**c**) Western blot of protein extracts of HEK293T cells expressing wild-type (WT) and G169R mutant Flag-CEACAM16 and of culture media secreted protein precipitates (upper panel). Total protein expression of WT and G169R CEACAM16 is very similar (Cells); total β-actin loads from both preparations are comparable (lower panel). However, the secreted (Medium) amount of mutant CEACAM16 is much lower than the WT CEACAM16. Scale bar: 10 μm.

including cell differentiation and proliferation, tissue architecture, immune responses, angiogenesis, tumorigenesis and metastasis, and bacteria mucosal colonization.[27–34] Mouse CEACAM16 was initially cloned and detected only in the cerebellum with a low level expression in 2005.[35] Mouse CEACAM16 and its human counterpart are well conserved. Analysis by bioinformatics shows CEACAM16 molecules consists of an Ig variable-like N domain at its NH2-terminus (N1), followed by Ig constant (IgC)-like domains of subtypes A (A) and B (B) and another Ig variable-like N domain at its COOH-terminal (N2) (Figure 2c). The protein encoded by *CEACAM16* is a secreted glycoprotein with a signal peptide at its NH2-terminal and lacks a glycosylphosphatidylinositol anchor or transmembrane domain present in other CEACAM members.[19,35,36] EST database analysis showed CEACAM16 expression in inner/organ of Corti. The causative gene of the previous American DFNA4B family was first focused on the *CEACAM16* gene in 2011, but its function is still unknown.[19] Tectorial

membrane is an extracellular connective tissue that covers the mechanically sensitive hair bundles of the exquisite sensory receptor cells in the organ of Corti and plays a pivotal role in transforming sound to mechanical stimulation. *In vivo* and *in vitro* studies of *CEACAM16* indicated that CEACAM16 protein was highly expressed at the tip of cochlear outer hair cells and interacts with α-tectorin, a tectorial membrane protein coded by the *TECTA* gene.[19,26] Previous studies have identified mutations in *TECTA* can lead to DFNA12 or DFNB21.[37–39] In 2014, a study showed that TECTB (coded by the *TECTB* gene) levels were reduced, a clearly defined striated-sheet matrix did not develop, and Hensen's stripe, a prominent feature in the basal two-thirds of the tectorial membrane, was absent in *CEACAM16*-null mutant mice. Furthermore, CEACAM16 was confirmed to interact with TECTB.[40] Therefore, *CEACAM*16 protein, as another mammal-specific tectorial membrane component, may play a key role in the maintenance of proper hearing function, just like

α-tectorin. We had observed that the surface membrane of transfected COS7 cells was intact using confocal microscope. Figure 3a and b shows that the borders of cells are defined by the red fluorescence dye (the actin cytoskeleton) and a part of the Flag-tagged wild-type (a) and G169R mutant (b) CEACAM16 (green) seems to be secreted from the cells, without background staining. This suggests that the CEACAM16 protein may be a secreted protein, which is consistent with the western blotting result (Figure 3c). Furthermore, our studies in transfected HEK293T cells show that the secretion efficacy of MT CEACAM16 is much lower than that of the WT, and only CEACAM16 monomers could be detected. The decreased secretion and the defection of polymerization suggest the effect of the G169R mutation of CEA-CAM16. In previous studies, the second Ig variable domains were shown to mediate cell-cell adhesion and enable CEACAM16 to possibly link α-tectorin to the tip of stereocilia of outer hair cells as a fifth member of non-collagenous proteins specific to the acellular extracellular matrix.[19,26] The p.G169R in CEACAM16 is found in this deafness family. The glycine is small, with two carbons and one nitrogen. It is neutral and plays a hinge-like role in protein structures. In the protein sequence, G169 is located in the middle, and probably part of a loop. It is part of a predicted IgC_2 type 1 domain in CEACAM16, which is mainly a beta sheet domain where glycines typically play a hinge-like role in the loops between the sheets. In hugely sharp contrast, arginine has five carbons and three nitrogens, it is positively charged and plays a role in protein folding by creating hydrogen bonds with other amino acids in the peptide chain.[41] So apart from a natural disturbance of having to physically fit this large amino acid in place of the glycine, this substitute amino acid will communicate with others around it, potentially causing folds that would certainly disrupt the precise beta sheet arrangement that is typical of Ig–like type 1 domains. Interestingly, IgC_2 has been shown to be associated with CEACAM1-mediated bacterial internalization.[42] A study suggested that the absence of the $IgC_2(A)$ domain could not affect the CEACAM16 protein to form the oligomer,[26] but we suggest that the G169R mutation at the $IgC_2(A)$ domain may change the spatial structure and thereby decrease the stability and polymerizing ability of the protein, and possibly speed up its degradation. This alteration in turn may lead to a significant impact on the cochlear amplification mechanism and subsequently cause hearing impairment. However, future *in vivo* research on cochleae in mice with mutated CEACM16 cochlea is indicated.

In summary, we have reported the clinical characteristics of a Chinese family with ADNSHL and the high-efficiency identification of a novel missense mutation in the *CEACAM16* gene. Our study demonstrates that CEACAM16 is a secreted protein and that the G169R mutation in this gene may lead to a defect in the CEACAM16 protein, which can results in disruption of the tectorial membrane amplification mechanism.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

1 Kalatzis, V. & Petit, C. The fundamental and medical impacts of recent progress in research on hereditary hearing loss. *Hum. Mol. Genet.* **7**, 1589–1597 (1998).
2 Marazita, M. L., Ploughman, L. M., Rawlings, B., Remington, E., Arnos, K. S. & Nance, W. E. Genetic epidemiological studies of early-onset deafness in the U.S. school-age population. *Am. J. Med. Genet.* **46**, 486–491 (1993).
3 Reardon, W. Genetics of deafness: clinical aspects. *Br. J. Hosp. Med.* **47**, 507–511 (1992).
4 Van Camp, G., Willems, P. J. & Smith, R. J. Nonsyndromic hearing impairment: unparalleled heterogeneity. *Am. J. Hum. Genet.* **60**, 758–764 (1997).
5 Petersen, M. B. Non-syndromic autosomal-dominant deafness. *Clin. Genet.* **62**, 1–13 (2002).
6 Hilgert, N., Smith, R. J. & Van Camp, G. Forty-six genes causing nonsyndromic hearing impairment: which ones should be analyzed in DNA diagnostics? *Mut. Res.* **681**, 189–196 (2009).
7 Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K. & Dent, K. M. et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat. Genet.* **42**, 30–35 (2010).
8 Zhu, M. & Zhao, S. Candidate gene identification approach: progress and challenges. *Int. J. Biol. Sci.* **3**, 420–427 (2007).
9 Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J. & Nickerson, D. A. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* **12**, 745–755 (2011).
10 Clark, J. G. Uses and abuses of hearing loss classification. *ASHA* **23**, 493–500 (1981).
11 Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30**, 97–101 (2002).
12 Lathrop, G. M., Lalouel, J. M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. *Proc. Natl Acad. Sci. USA* **81**, 3443–3446 (1984).
13 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
14 Li, R., Li, Y., Kristiansen, K. & Wang, J. SOAP: short oligonucleotide alignment program. *Bioinformatics* **24**, 713–714 (2008).
15 Li, R., Li, Y., Fang, X., Yang, H., Wang, J. & Kristiansen, K. *et al.* SNP detection for massively parallel whole-genome resequencing. *Genome Res.* **19**, 1124–1132 (2009).
16 DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R. & Hartl, C et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
17 Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A. & Bork, P. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
18 Morton, N. E. Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* **7**, 277–318 (1955).
19 Zheng, J., Miller, K. K., Yang, T., Hildebrand, M. S., Shearer, A. E. & DeLuca, A. P. *et al.* Carcinoembryonic antigen-related cell adhesion molecule 16 interacts with alpha-tectorin and is mutated in autosomal dominant hearing loss (DFNA4). *Proc. Natl Acad. Sci. USA* **108**, 4218–4223 (2011).
20 Chen, A. H., Ni, L., Fukushima, K., Marietta, J., O'Neill, M. & Coucke, P. et al. Linkage of a gene for dominant non-syndromic deafness to chromosome 19. *Hum. Mol. Genet.* **4**, 1073–1076 (1995).
21 Donaudy, F., Snoeckx, R., Pfister, M., Zenner, H. P., Blin, N. & Di Stazio, M. *et al.* Nonmuscle myosin heavy-chain gene MYH14 is expressed in cochlea and mutated in patients affected by autosomal dominant hearing impairment (DFNA4). *Am J. Hum. Genet.* **74**, 770–776 (2004).
22 Mirghomizadeh, F., Bardtke, B., Devoto, M., Pfister, M., Oeken, J. & Konig, E. *et al.* Second family with hearing impairment linked to 19q13 and refined DFNA4 localisation. *Eur. J. Hum. Genet.* **10**, 95–99 (2002).
23 Pusch, C. M., Meyer, B., Kupka, S., Smith, R. J., Lalwani, A. K. & Zenner, H. P. *et al.* Refinement of the DFNA4 locus to a 1.44 Mb region in 19q13.33. *J. Mol. Med. (Berl)* **82**, 398–402 (2004).
24 Yang, T., Pfister, M., Blin, N., Zenner, H. P., Pusch, C. M. & Smith, R. J. Genetic heterogeneity of deafness phenotypes linked to DFNA4. *Am. J. Med. Genet. A.* **139**, 9–12 (2005).
25 Zong, L., Lu, C., Zhao, Y., Li, Q., Han, D. & Yang, W. *et al.* Clue to a new deafness gene: a large Chinese nonsyndromic hearing loss family linked to DFNA4. *J. Genet. Genomics* **39**, 653–657 (2012).
26 Kammerer, R., Ruttiger, L., Riesenberg, R., Schauble, C., Krupar, R. & Kamp, A. *et al.* Loss of mammal-specific tectorial membrane component carcinoembryonic antigen cell adhesion molecule 16 (CEACAM16) leads to hearing impairment at low and high frequencies. *J. Biol. Chem.* **287**, 21584–21598 (2012).
27 Benchimol, S., Fuks, A., Jothy, S., Beauchemin, N., Shirota, K. & Stanners, C. P. Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. *Cell* **57**, 327–334 (1989).
28 Boulton, I. C. & Gray-Owen, S. D. Neisserial binding to CEACAM1 arrests the activation and proliferation of CD4+ T lymphocytes. *Nat. Immunol.* **3**, 229–236 (2002).

29 Eidelman, F. J., Fuks, A., DeMarte, L., Taheri, M. & Stanners, C. P. Human carcinoembryonic antigen, an intercellular adhesion molecule, blocks fusion and differentiation of rat myoblasts. *J. Cell Biol.* **123**, 467–475 (1993).

30 Ergun, S., Kilik, N., Ziegeler, G., Hansen, A., Nollau, P. & Gotze, J. *et al.* CEA-related cell adhesion molecule 1: a potent angiogenic factor and a major effector of vascular endothelial growth factor. *Mol. Cell* **5**, 311–320 (2000).

31 Muenzner, P., Bachmann, V., Zimmermann, W., Hentschel, J. & Hauck, C. R. Human-restricted bacterial pathogens block shedding of epithelial cells by stimulating integrin activation. *Science* **329**, 1197–1201 (2010).

32 Sarantis, H. & Gray-Owen, S. D. Defining the roles of human carcinoembryonic antigen-related cellular adhesion molecules during neutrophil responses to Neisseria gonorrhoeae. *Infect. Immun.* **80**, 345–358 (2012).

33 Schmitter, T., Agerer, F., Peterson, L., Munzner, P. & Hauck, C. R. Granulocyte CEACAM3 is a phagocytic receptor of the innate immune system that mediates recognition and elimination of human-specific pathogens. *J. Exp. Med.* **199**, 35–46 (2004).

34 Zimmer, R. & Thomas, P. Mutations in the carcinoembryonic antigen gene in colorectal cancer patients: implications on liver metastasis. *Cancer Res.* **61**, 2822–2826 (2001).

35 Zebhauser, R., Kammerer, R., Eisenried, A., McLellan, A., Moore, T. & Zimmermann, W. Identification of a novel group of evolutionarily conserved members within the rapidly diverging murine Cea family. *Genomics* **86**, 566–580 (2005).

36 Kuespert, K., Pils, S. & Hauck, C. R. CEACAMs: their role in physiology and pathophysiology. *Curr. Opin. Cell Biol.* **18**, 565–571 (2006).

37 Balciuniene, J., Dahl, N., Borg, E., Samuelsson, E., Koisti, M. J. & Pettersson, U. *et al.* Evidence for digenic inheritance of nonsyndromic hereditary hearing loss in a Swedish family. *Am. J. Hum. Genet.* **63**, 786–793 (1998).

38 Meyer, N. C., Alasti, F., Nishimura, C. J., Imanirad, P., Kahrizi, K. & Riazalhosseini, Y. *et al.* Identification of three novel TECTA mutations in Iranian families with autosomal recessive nonsyndromic hearing impairment at the DFNB21 locus. *Am. J. Med. Genet. A.* **143A**, 1623–1629 (2007).

39 Mustapha, M., Weil, D., Chardenoux, S., Elias, S., El-Zir, E. & Beckmann, J. S. *et al.* An alpha-tectorin gene defect causes a newly identified autosomal recessive form of sensorineural pre-lingual non-syndromic deafness, DFNB21. *Hum. Mol. Genet.* **8**, 409–412 (1999).

40 Cheatham, M. A., Goodyear, R. J., Homma, K., Legan, P. K., Korchagina, J. & Naskar, S. *et al.* Loss of the tectorial membrane protein CEACAM16 enhances spontaneous, stimulus-frequency, and transiently evoked otoacoustic emissions. *J. Neurosci.* **34**, 10325–10338 (2014).

41 Borders, C. L. Jr., Broadwater, J. A., Bekeny, P. A., Salmon, J. E., Lee, A. S. & Eldridge, A. M. *et al.* A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens. *Protein Sci.* **3**, 541–548 (1994).

42 Voges, M., Bachmann, V., Naujoks, J., Kopp, K. & Hauck, C. R. Extracellular IgC2 constant domains of CEACAMs mediate PI3K sensitivity during uptake of pathogens. *PloS One* **7**, e39908 (2012).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)