Human Molecular Genetics, 2020, Vol. 29, No. 9 1520–1536

doi: 10.1093/hmg/ddaa075
Advance Access Publication Date: 27 April 2020
General Article

GENERAL ARTICLE

A rare genomic duplication in 2p14 underlies autosomal dominant hearing loss DFNA58

Karina Lezirovitz[1,2,*,†], Gleiciele A. Vieira-Silva[1,2], Ana C. Batissoco[1,2], Débora Levy[3], Joao P. Kitajima[4], Alix Trouillet[5], Ellen Ouyang[5], Navid Zebarjadi[5], Juliana Sampaio-Silva[1], Vinicius Pedroso-Campos[1], Larissa R. Nascimento[1,2], Cindy Y. Sonoda[1], Vinícius M. Borges[6], Laura G. Vasconcelos[2], Roberto M.O. Beck[2], Signe S. Grasel[2], Daniel J. Jagger[7], Nicolas Grillet[5], Ricardo F. Bento[1,2], Regina C. Mingroni-Netto[6] and Jeanne Oiticica[1,2]

[1]Otorhinolaryngology/LIM32, Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo, São Paulo 01246-000, Brazil, [2]Departamento de Otorrinolaringologia, Faculdade de Medicina FMUSP, Universidade de São Paulo, São Paulo 05403-000, Brazil, [3]Lipids, Oxidation, and Cell Biology Group, Head, Laboratory of Immunology (LIM19), Heart Institute (InCor), Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo, São Paulo 05403-900, Brazil, [4]Mendelics Genomic Analysis, São Paulo 04013-000, Brazil, [5]Department of Otolaryngology - Head and Neck Surgery, Stanford University, Stanford, CA 94305, USA, [6]Centro de Pesquisas sobre o Genoma Humano e Células-Tronco, Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo 05508-900, Brazil and [7]UCL Ear Institute, University College London, London WC1E 6BT, UK

*To whom correspondence should be addressed. Tel: +55-11-3061-7166; Fax: 3091-7553; E-mail: lezi.karina@gmail.com

Abstract

Here we define a ~200 Kb genomic duplication in 2p14 as the genetic signature that segregates with postlingual progressive sensorineural autosomal dominant hearing loss (HL) in 20 affected individuals from the DFNA58 family, first reported in 2009. The duplication includes two entire genes, *PLEK* and *CNRIP1*, and the first exon of *PPP3R1* (protein coding), in addition to four uncharacterized long non-coding (lnc) RNA genes and part of a novel protein-coding gene. Quantitative analysis of mRNA expression in blood samples revealed selective overexpression of *CNRIP1* and of two lncRNA genes (*LOC107985892* and *LOC102724389*) in all affected members tested, but not in unaffected ones. Qualitative analysis of mRNA expression identified also fusion transcripts involving parts of *PPP3R1*, *CNRIP1* and an intergenic region between *PLEK* and *CNRIP1*, in the blood of all carriers of the duplication, but were heterogeneous in nature. By *in situ* hybridization and

[†]Karina Lezirovitz, http://orcid.org/0000-0002-7495-9789
Received: April 2, 2020. **Revised:** April 2, 2020. **Accepted:** April 20, 2020

immunofluorescence, we showed that *Cnrip1, Plek* and *Ppp3r1* genes are all expressed in the adult mouse cochlea including the spiral ganglion neurons, suggesting changes in expression levels of these genes in the hearing organ could underlie the DFNA58 form of deafness. Our study highlights the value of studying rare genomic events leading to HL, such as copy number variations. Further studies will be required to determine which of these genes, either coding proteins or non-coding RNAs, is or are responsible for DFNA58 HL.

Introduction

Hearing loss (HL) is the most common sensory defect in humans and can occur at any age, from congenital to age-related HL, also known as presbycusis (1). HL prevalence grows with age: the HL onset is congenital in 2% of overall cases, and prevalence increases between ages 6 and 19 years to 11–12%, and during adulthood (between 20 and 59 years) to 20–32% of the population, constituting the largest affected group (2). The origin of HL may be majorly attributed to a deregulation of homeostatic or apoptotic pathways, in addition to several genetic and environmental risk factors that may affect its onset and progression (3,4). Sensorineural presbycusis is usually associated with age-dependent loss of hair cells and/or their afferent neurons, which detect the sound and transmit the information centrally, respectively (3). Monogenic forms of postlingual progressive HL are likely to share molecular pathways as well as pathophysiological mechanisms with presbycusis (5–7).

Despite the high prevalence of HL, our understanding of its molecular etiology is limited, as is the prospect of developing therapies based on this knowledge (8). Non-syndromic HL accounts for 70% of all cases (9). Hereditary non-syndromic HL is genetically heterogeneous and may show autosomal recessive, autosomal dominant, X-linked or mitochondrial inheritance (10). Genetic mapping of deafness-related genes in large human families and mouse models have led to the discovery of 120 genes associated with hereditary non-syndromic HL (11–13). During the last two decades, gene identification has not followed as fast as locus mapping, but it has speeded up recently, owing to the improvement and cost-efficiency of targeted genomic enrichment and massively parallel sequencing (14). Challenges remain not only in identifying candidate variants, but also in finding many affected individuals with variants in the same gene, either from a single large family or from different small ones, and in performing proper functional studies that give support to the pathogenic nature of those variants. In addition, given that the causal gene of at least 44 mapped loci have not been characterized (11), and also given that computational predictions estimate that only half of the human deafness genes are already known (15), it is evident that more crucial genes for auditory function are to be revealed.

Recent reports using targeted-sequencing found that copy number variations (CNVs) are frequent causes of HL (16–18). Many genome wide sequencing studies searching for genetic causes of deafness focus the analysis on variants of one or few nucleotides (SNVs and InDels), neglecting CNVs (19–21). A significant number of genes, thus, may escape identification if there is no assessment of their copy number. The proteins encoded by deafness-related genes participate in a variety of molecular/cellular pathways and structures, crucial for the development and/or maintenance of hearing. Some are unique to the cochlea, such as motor activity and cell adhesion in stereocilia, ionic homeostasis, tight junction formation, synaptic transmission or apoptotic pathways (10,3). Better understanding of these mechanisms paves the way for the development of innovative therapeutic strategies.

In this study, we characterized the genetic cause of the late onset sensorineural autosomal dominant form of HL DFNA58 (22) (OMIM 615654), showing that a rare genomic duplication of 200 Kb in 2p14 segregates with postlingual progressive HL in 20 hearing impaired family members. The duplication comprises three protein coding genes *CNRIP1*, *PLEK* and *PPP3R1* and a novel protein-coding gene in addition to four uncharacterized lncRNA genes. We report also the selective overexpression of *CNRIP1*, and two lncRNA genes, *LOC102724389* and *LOC107985892*, as well as heterogeneous fusion transcripts involving parts of *PPP3R1*, *CNRIP1* and an intergenic region between *PLEK* and *CNRIP1*, in blood of carriers of the duplication. In addition, we found that the three protein-coding genes are expressed in adult murine cochlea. Together, our data suggests that all protein-coding genes, at least two of the lncRNA genes and the fusion transcripts, could be related to the origin of HL.

Results

The primary functional impairment in patients with DFNA58 associated HL is of cochlear origin

Twenty-three family members affected by sensorineural postlingual HL were clinically investigated (Fig. 1, Supplementary Table S1). None of the hearing-impaired family members showed clinical signs of optic or neurological issues. Among the 23 affected subjects (3 phenocopies, with distinct clinical presentation of the HL), 20 had bilateral and progressive sensorineural HL, with ages of onset varying between childhood to 33 years (mean of 18.6 years, Fig. 2A). Tinnitus was reported by near 80% of these 20 family members (Supplementary Table S1). About half of the patients reported to have occasionally mild dizziness and imbalance, probably vestibular symptoms, but these usually were related to episodes of severe tinnitus or migraine. Vectoeletronystagmography, an exam to evaluate vestibular function by means of involuntary eye movements (nystagmus), yielded diagnosis of irritative peripheral vestibular syndrome in one affected family member (VI:2), even without complaints of vestibular symptoms.

Progression of the HL was assessed by the comparison of the 25 audiometric profiles of 17 subjects with DFNA58 HL divided in four different groups, according to their age at examination (Fig. 2B). Regarding four of these 17 subjects, audiograms obtained in different ages were available. Positive correlation between average audiometric thresholds (0.5, 1, 2 and 4 Hz) and age at examination was obtained using linear regression (Fig. 2C). ABR (auditory brainstem responses) records were obtained from 11 subjects with DFNA58 HL: eight showed normal results with no signs of retro-cochlear dysfunction (disorder occurring at the central or neural nerve), two had abnormal ABR recordings with poor morphology and only one showed absent responses. There was a positive correlation between average audiometric thresholds and ABR recording results ($r^2 = 0.4102$, $P = 0.0338$). Accordingly, normal ABR responses were observed in subjects with mild to moderate HL, abnormal responses in

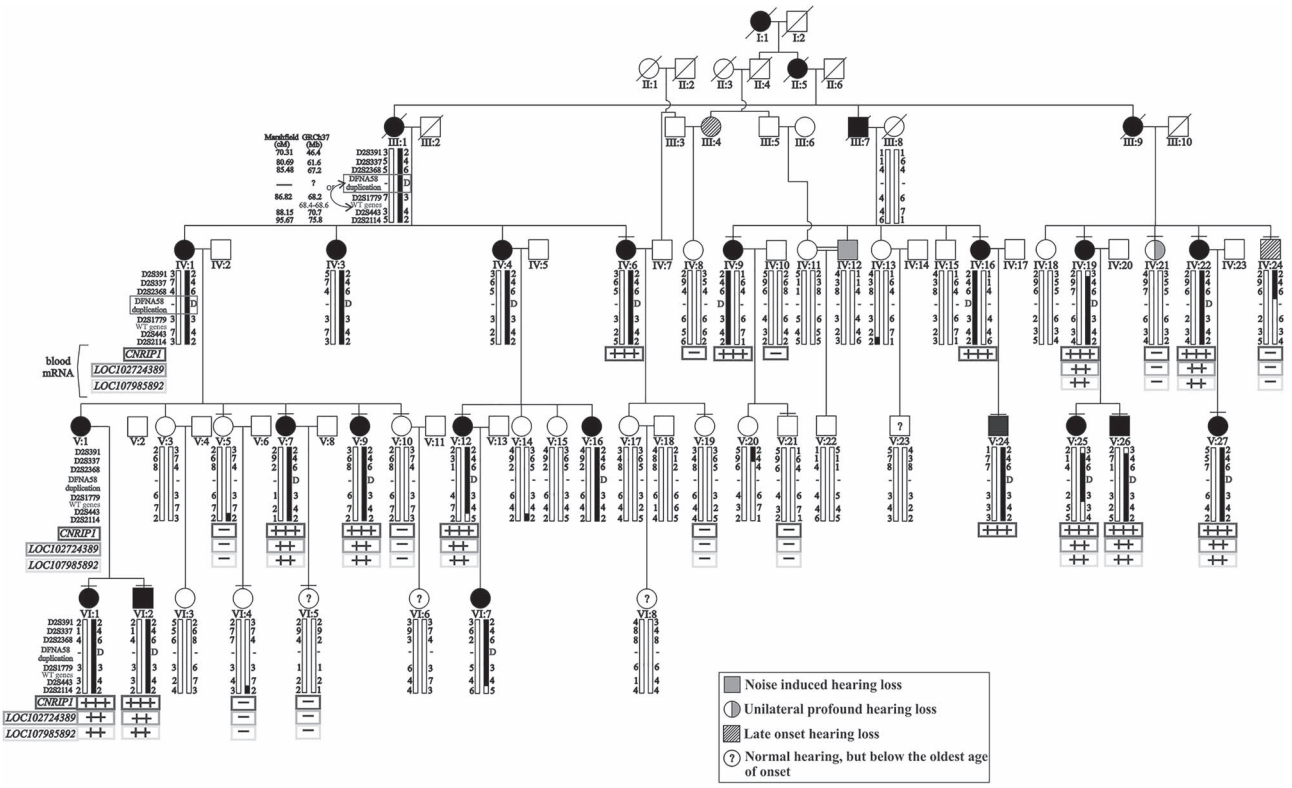

Figure 1. Pedigree showing the segregation of the DFNA58 duplication (− =non-duplicated; D = duplicated) and flanking microsatellite markers haplotype. Expression data from *CNRIP1*, *LOC102724389* and *LOC107985892* (mRNA) are also shown in the box below the haplotypes of each subject: +++ means RNA overexpression in RT-qPCR; −— means *RNA* normal expression in RT-qPCR.

moderate/severe HL, and absent ABR responses were observed in the subject with profound HL, who had the earliest referred age of onset during childhood (Supplementary Table S1). Distortion product otoacoustic emissions were absent in all affected subjects, except V:27, who had just started to report mild HL by age 33. On average, DFNA58 patients demonstrated relatively good speech recognition (>50%) suggesting that DFNA58 HL is not related to the auditory neuropathy spectrum. Overall, clinical data indicated that the primary functional impairment underlying the DFNA58 HL resides in the cochlea.

Three subjects exhibited different clinical presentations of their HL and were considered preliminarily as phenocopies: IV:12 had occupational noise induced HL at age 45, IV:21 experienced unilateral sudden HL after severe dizziness and IV:24 had mild HL starting after age 50, much older than the oldest reported age of onset (33 years old) from the subjects with bilateral postlingual progressive HL (Supplementary Table S1). Twenty-four unaffected family members were also included in the molecular study. Among them, one was a consanguineous spouse, three were unrelated spouses, and four who were below the highest age of onset of DFNA58 related HL.

A rare 200 kb duplication segregates with the DFNA58 HL

Since the initial DFNA58 locus mapping, we refined the candidate chromosomal region, from 30 cM to ∼3 cM (between markers D2S2368 and D2S443, GRCh38/hg38: chr2: 66986473-70 661 255; GRCh37/hg19: chr2: 67213605–70 788 387), by means of the analysis of recombinant microsatellite markers haplotypes of three subjects (Fig. 1, V:14, V:20 and IV:24). The individual V:20 inherited part of the DFNA58 haplotype between markers D2S391 and D2S337, and we obtained the clinical confirmation that she was unaffected. One of the phenocopies (IV:24) inherited the DFNA58 haplotype between markers D2S391 and D2S2368. Combining these recombinant haplotypes, the telomeric border of the candidate region was defined by marker D2S2368. In addition, the affected subject V:25 inherited the DFNA58 haplotype between markers D2S337 and D2S1779, defining D2S443 marker as the centromeric border of the candidate chromosomal region. The maximum multipoint LOD score (2p12-p21, Markers D2S391, D2S337, D2S2368, D2S1779, D2S443 and D2S2114, Fig. 1), recalculated with the additional affected and unaffected subjects, reached a highly significant value of 10.2 in the position of marker D2S1779 and in the candidate region between D2S1779 and marker D2S2368 (Fig. 3A).

To identify the genetic alteration associated with HL in the family, we performed next generation exome sequencing analyses in a sample from one affected subject. A total of 8.04 Gbp was sequenced with an average sampling depth of 114X after BWA alignment, 98.3% of all exome bases were sampled by 10 sequences or more. In total, 74 566 SNVs and InDels were genotyped with 69 303 (92%) of them being PASS variants. A total of 44 632 PASS variants (64%) were heterozygous and 24 671 (36%) were in homozygous state. In total, 60 663 PASS variants (87%) were SNVs and 8640 (13%) were InDels; 815 variants were in the candidate region (2p21-p12, that is, for b37, from chr2:41800000 until 2:83300000) with 629 of them in heterozygous state. Twenty-nine heterozygous variants inside the candidate region had zero frequency in the Broad gnomAD database (http://gnomad.broadinstitute.org/) (23) as well as zero frequency in 1000 genomes (http://www.internationalgenome.org/)

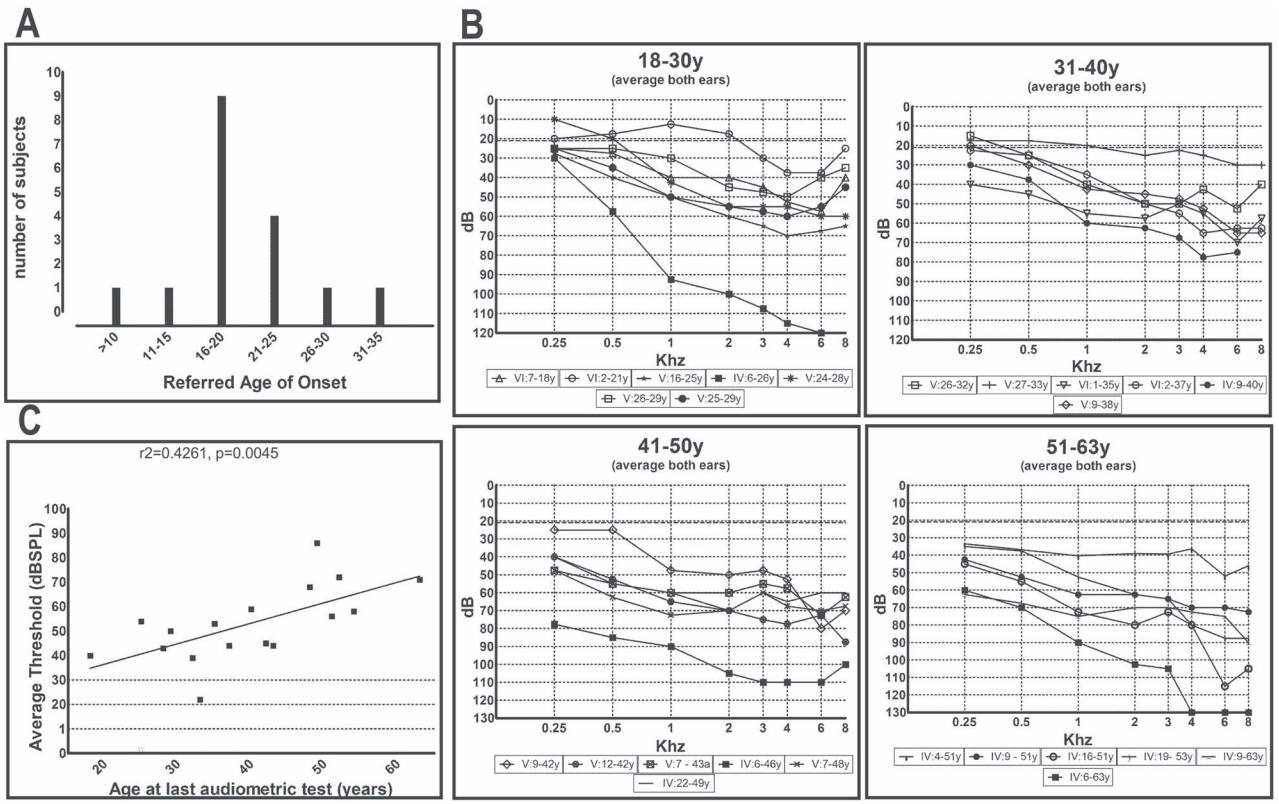

Figure 2. (**A**) Age distribution of onset of 17 patients with the DFNA58 HL. (**B**) 25 audiometric profiles of 17 affected patients divided in in four different age at exam groups. Subject IV:2 had a cholesteatoma in her right ear, in addition to the bilateral postlingual progressive sensorineural HL, thus explaining why her HL is not symmetrical, being profound in the right ear and moderate in the left ear. (**C**) Distribution of average thresholds of six affected individuals according to their age at examination showing fit to a linear regression, only when excluding the subject IV:6 with childhood onset.

(24). None of the variants were in genes with OMIM known phenotype, and none of them were in coding regions. A total of 65 CNVs were detected by the ExomeDepth software (25) with 45 of them considered as PASS events (score of 10 or larger). In the refined candidate chromosomal region (2p14-p13.3), only one duplication was detected (chr2:68479213-68 623 062; GRCh37, Fig. 3B) with score 43.5 (R ExomeDepth Bayes factor—the log10 of the likelihood ratio of data for the CNV call divided by the normal copy number). No other similar event (in size and location) was found in DGV, Decipher or ExAC CNV databases (26–28). In the Decipher databases (https://decipher.sanger.ac.uk/) (27), duplication events involving completely this region were much larger and were detected in three patients with developmental disorders, with no report of hearing deficits.

The duplication includes two entire protein-coding genes (*PLEK* and *CNRIP1*) and, partially, a third protein-coding gene, the *PPP3R1* gene (Figs 3 and 4D). In addition, other less characterized genes map inside the duplication segment: part (exons 1 and 2) of an uncharacterized gene encoding a novel protein (*AC017083.3*) and four predicted, not characterized, lncRNA (long non-coding RNA) genes: *AC017083.1*, *LOC107985892* (mRNAs *XR_001739527.1* and *XR_001739526.1* that differ from three bases in the beginning of exon 2), *LOC102724389* (mRNA *XR_940224.3*) and *LOC101927723/AC015969.1–201* (mRNA *XR_245020.3*).

The novel protein coding *AC017083.3* gene shares exons with *PPP3R1* and *WDR92*, such as its exons 3–7 overlap extensively with *PPP3R1's* exons 2–6 and its exons 8–15 overlap extensively with *WDR92's* exons 1–8 (UCSC Genome Browser—https://genome.ucsc.edu/) (Fig. 3B). The *AC017083.3* exon 2 (150 bp) is fully inside the single exon (979 bp) of the lncRNA *AC017083.1* gene, but they are transcribed in opposite directions.

LOC102724389 is an intergenic lncRNA gene distant ~13 Kb from *CNRIP1* and ~19 Kb from *PLEK* (29,30). *LOC107985892* may be considered antisense overlapping to both *CNRIP1* and *PLEK*, since its exon 1 is located in *PPP3R1* intron 1 and exon 2 located in *CNRIP1* intron 2, and is transcribed in opposite direction to both *PPP3R1* and *CNRIP1* (29,30). Part of the exonic sequence of the lncRNA *AC017083.1* gene overlaps with the exonic sequence of the novel protein-coding *AC017083.3* gene, but transcribed in opposite direction, it might be classified as an antisense overlapping lncRNA (29,30). *LOC101927723* may be an antisense overlapping lncRNA to *PLEK* gene since its first exon is completely contained inside *PLEK* intron 1 and its second exon has no overlaps with protein-coding genes (29,30).

Array comparative genome hybridization (CGH) 180 K Agilent oligo-platform confirmed the presence of the duplication in two affected subjects (duplicated segment chr2: 68476931-68 658 774; GRCh37). A customized MLPA panel of 10 probes (seven inside and three outside the duplication) demonstrated co-segregation of this duplication with DFNA58 HL in the 20 affected subjects (Figs 1 and 3B), but none of the three suspected phenocopy cases, none of the 21 unaffected family members inherited the duplication (four of them below the average age of onset) or the three unrelated spouses. This pattern of segregation also yields a highly significant two-point LOD score value of 10.7, indicating linkage between the DFNA58 HL and this duplication (Fig. 1).This value of LOD score was similar to the one obtained with multipoint LOD score

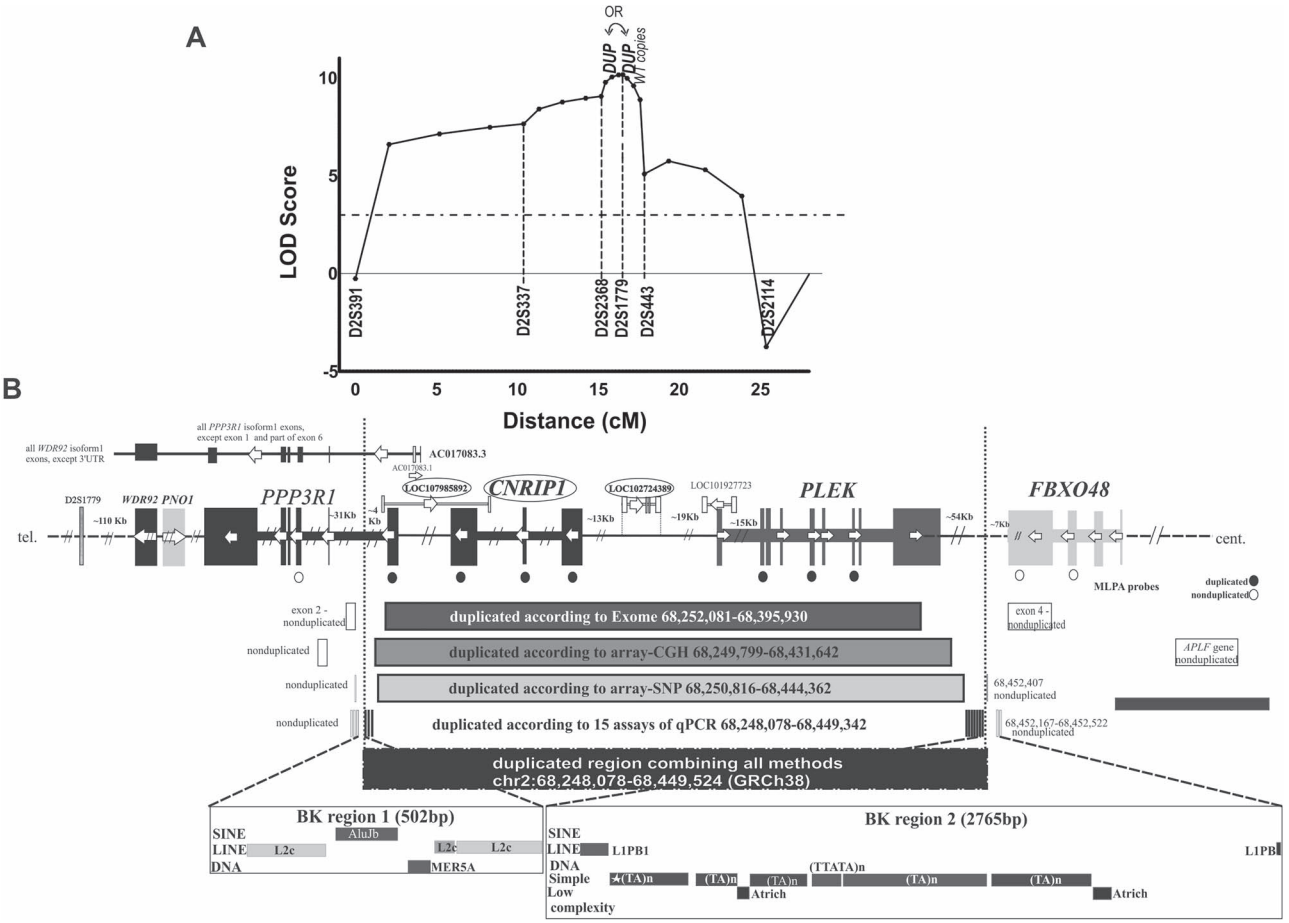

Figure 3. (**A**) Multipoint LOD scores calculated with Morgan software for markers in the 2p12-p21 chromosomal region using a penetrance estimate of 96%. (**B**) Sketch of the duplication showing the location of MLPA probes, as well as extension of the duplication according to the different methods and combining all [NC_000002.12:g.(68 247 572_68 248 077)_(68 449 525_68 452 166)dup]. The arrows indicate transcription orientation. The exons of each gene producing the longest mRNA and with RefSeq entry are represented, GRCh38 as the reference genome assembly. Both breakpoints (BK regions) locate within repetitive elements. Circled Genes are observed overexpression in the blood of duplication carriers are indicated by a circle.

calculations with microsatellite markers in the 2p14-p21 region (Fig. 3A). Curiously, the peak multipoint microsatellite LOD score maps between markers D2S2368 (~67.2 Mb, GRCh37/hg19) and D2S1779 (~68.3 Mb, GRCh37/hg19) instead of between D2S1779 and D2S443 (~70.8 Mb), where the *PPP3R1*, *PLEK* and *CNRIP1* are located (Fig. 3A). Screening of CNVs involving these three genes through the customized MLPA, in a collection of samples of 100 probands from familial cases with autosomal dominant sensorineural HL produced negative results.

DFNA58 duplication breakpoints map within or nearby repetitive elements

Illumina BeadChip 850 K® SNP-array analyses and 15 qPCR assays allowed refinement of the duplication's breakpoints (Fig. 3B), the first being located in the intron 1 of the *PPP3R1* gene within a ~0.5 Kb region (chr2:68474704–68 475 209: GRCh37; chr2: 68247572–68 248 077:GRCh38) and the second residing between *PLEK* and *FBXO48* genes, within a ~2.7 Kb region (chr2: 68676657–68 679 298:CRCh37; chr2:68 449 525–68 452 166:GRCh38). Exon 1 of *PPP3R1* is a coding exon in two of the three isoforms, but in isoform 1 it harbors only the ATG start codon. The duplication extension, as estimated combining all techniques, is between 201.4–204.6 Kb [NC_000002.12:g.(68 247 572_68 248 077)_(68 449 525_68 452 166)dup]. There are many repetitive elements (SINEs, LINEs, DNA repeats, simple repeats and low complexity repeats) within both breakpoint regions, which most likely mediated a non-homologous recombination event (Fig. 3B). The repetitive nature of the region makes it very difficult to define precisely the breakpoints.

CNRIP1, *LOC107985892* and *LOC102724389* mRNA are highly overexpressed in the blood of duplication carriers

To test if one or more of the duplicated protein-coding genes and long non-coding RNA genes would have their expression levels changed in the duplication carriers, we performed quantitative RT-PCR from their blood (Fig. 4).

Expression of the RNA of each of the three protein-coding genes had been previously detected in cells present in the blood (http://biogps.org/) (31). Indeed, we were able to detect their presence in non-carrier family members (Fig. 4A and Supplementary Figure S1). However, in all 14 hearing impaired carriers tested we found a selective mRNA upregulation only for *CNRIP1* (mean of all carriers = ~ 27.7X; LOD Score = 6.3; Fig. 4A) not detected in 14 non-carriers, eight unaffected members of the family and six unrelated controls.

Human *CNRIP1* encodes two isoforms, which share exons 1 and 2, but have a different exon 3. Human *CNRIP1* isoform 1 (Crip1a) is conserved among mammals and fish, but the presence of CRIP1b is limited to primates (32). Six different primer

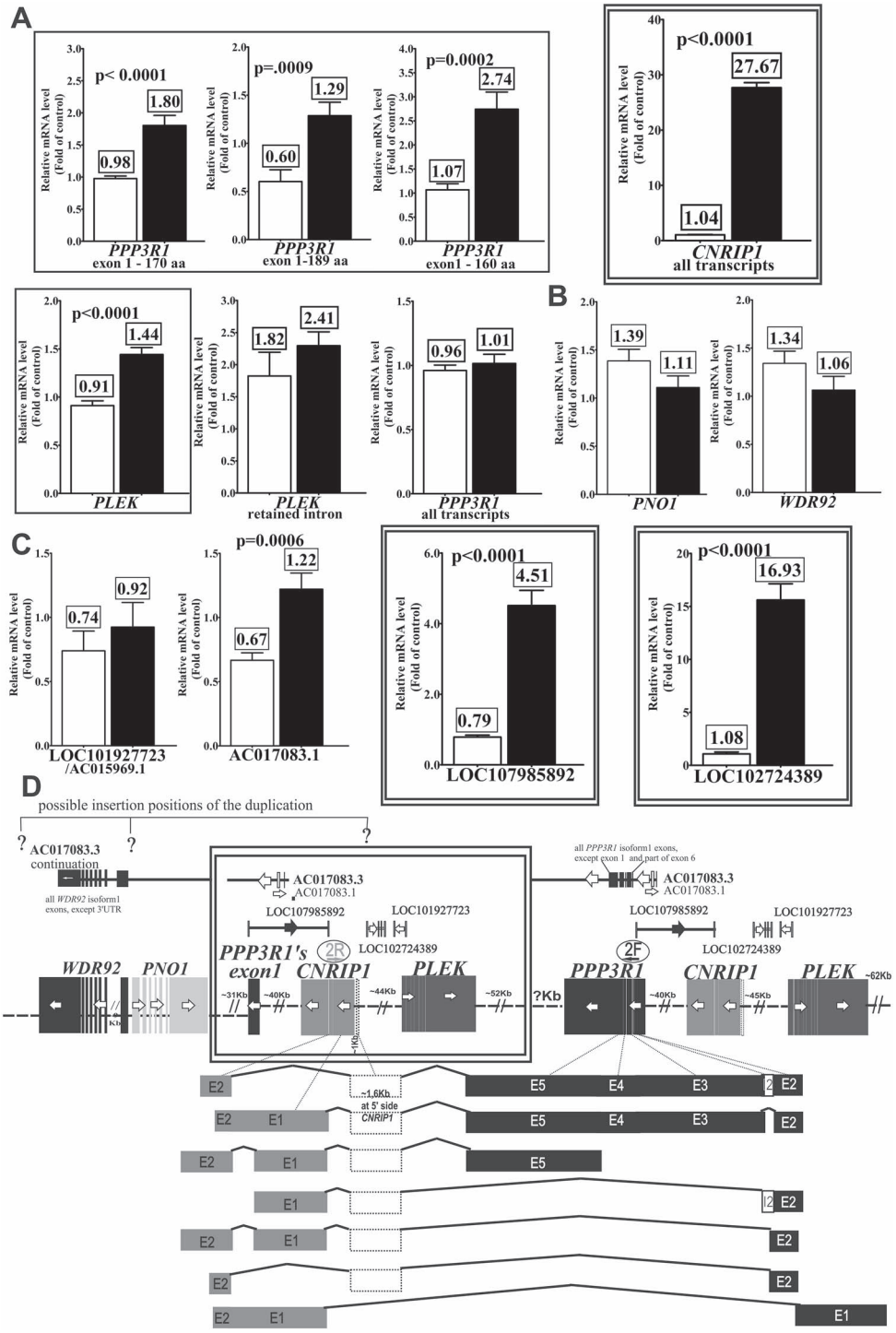

Figure 4. (**A–C**) RT-qPCR results showing mRNA blood levels in duplication carriers (mean of at least 10 subjects, indicated in black) and non-carriers (mean of at least eight subjects, indicated in white) of genes involved in the duplication or close to its probable insertion position: (A) Genes involved in the duplication: *PPP3R1* (exon 1), *CNRIP1* (all transcripts), *PLEK* (protein coding transcript and retained intron transcript) and *PPP3R1*(overall expression of all transcripts). (B) mRNA levels of *PNO1* and *WDR92*, two neighbor genes of the possible duplication insertion position. (C) mRNA levels of uncharacterized lncRNA genes contained inside the duplication segment: *LOC101927723*, *AC017083.1*, *LOC102724389* and *LOC107985892*. Graphs showing genes that are mildly overexpressed (1.2–1.8X) were indicated by a single line border: *PLEK* (protein coding transcript), *PPP3R1* (exon 1 of all three transcripts) and *AC017083.1* (lncRNA gene). Graphs showing genes that were highly overexpressed (4-27X) comparing the means of carriers X non-carriers were indicated by double line border: *CNRIP1*, *LOC102724389* and *LOC107985892*. (**D**) Possible insertion positions of the duplicated segment (?). The size of the duplicated segment is indicated by double line border. Depending on the insertion position, the new gene *AC017083.3* may or may not be interrupted. Examples of the low-abundant fusion *CNRIP1-PPP3R1* transcripts found in the duplication carriers are shown below.

pairs, spanning all exons of both human isoforms were tested for the *CNRIP1* cDNA quantitative PCR (Supplementary Table S2; results from the first pair that detects all transcripts are shown in Fig. 4C and Supplementary Figure S1) and all six primer pairs indicated significant overexpression of *CNRIP1*. Small but significant overexpression of the protein coding transcript of *PLEK* (but not the retained intron transcript, Ensembl ENSG00000115956), as well as of exon 1 of *PPP3R1*, in all three protein-coding transcripts, was revealed, comparing the average expression of carriers and non-carriers, but not all affected individuals showed overexpression (Fig. 4A and Supplementary Figure S1). No significant difference in *PPP3R1* overall expression (primers annealing to exons 2 and 3) was observed among the carriers and non-carriers averages (Fig. 4A, and Supplementary Figure S1).

Additionally, mRNA expression patterns of the four other uncharacterized lncRNA genes (*LOC107985892*, *LOC102724389*, *AC017083.1* and *LOC101927723*) and two protein coding genes located close to the most likely duplication insertion region (*WDR92* and *PNO1*) were assessed in blood (Fig. 4B and C). For two of the lncRNA genes, *LOC107985892* and *LOC102724389*, significant overexpression was observed, 4.51X and 16.93X, respectively, in the duplication carriers, but not in the non-carriers. As observed for *CNRIP1*, the upregulation was detected for the mean (mRNA levels) of all carriers, for both genes, (Fig. 4C) as well as for mRNA level of each carrier (Supplementary Figure S1). The lncRNA *AC017083.1* also showed significant overexpression of the mean of mRNA levels of carriers, but not in non-carriers, but in lower levels (1.42X) when compared to the other two lncRNA genes described above (Fig. 4C). However, this upregulation was not observed for every carrier (Supplementary Figure S1). Expression of *LOC101927723* did not differ among carriers and non-carriers (Fig. 4C, and Supplementary Figure S1).

We were unable to amplify *AC017083.3* mRNA from blood. Even though it is classified as a novel protein gene (ENSG00000273398), the transcript biotype is 'Nonsense mediated decay' in Ensembl ENST00000406334.3, which means this transcript might be detected as having nonsense codons and degraded in order to prevent translation of truncated or erroneous proteins.

Since a fusion transcript between the *PPP3R1* and *CNRIP1* genes, identified by RNA-seq technique, has been described both in lung adenocarcinoma tissue and tissue not involved in the cancer (33), we sought to test if we could detect selective fusion transcripts between *PPP3R1* and *CNRIP1* in the duplication carriers, not presented by the non-carriers. Indeed, low abundance transcripts involving fusion of sequences from *PPP3R1* and *CNRIP1* were amplified in all duplication carriers through RT-PCR, using a forward primer on *PPP3R1* exons 1 or 2 and a reverse primer on *CNRIP1* exon 2. These transcripts were not amplified from the cDNA of non-carriers. Sanger sequencing revealed that these transcripts were not identical in all duplication carriers, and some carriers amplified two or more different transcripts with different exon composition (Fig. 4D). Regardless of the size or composition of these fusion transcripts, they include parts of the exonic sequences of *CNRIP1* and *PPP3R1*. Besides, in most cases, *CNRIP1* and *PPP3R1* sequences were interspersed by an intergenic sequence at ~1.6 Kb from 5'of *CNRIP1* and ~44.4 Kb from 3′ of *PLEK* (Fig. 4D). The formation of these fusion gene transcripts and the orientation of the primers used for their amplification allowed us to conclude that the most likely position of the duplication is at 3′ downstream of the wild type location of the *PPP3R1* gene. Accordingly, the duplication is not inserted in either of the breakpoints defined by its borders, inside *PPP3R1* intron 1 or intergenic between *PLEK* and *FBXO48* (Fig. 4D). Indeed, haplotype segregation analysis as well as the multipoint LOD scores indicated a region closer to D2S1779 as the most likely position for the HL causative mutation. Given the most likely insertion position of the duplication, we also looked for fusion transcripts regarding *PPP3R1* exon 1 and coding sequences of *WDR92* and *PNO1*. No fusion transcript, present in carriers and absent in non-carriers, was detected. It is possible that the duplication is interrupting the *AC017083.3* gene, if its insertion position is located anywhere between *PPP3R1* and *WDR92*, but we were unable to test this hypothesis because we were not able to detect the *AC017083.3* transcript in blood.

Cnr1p1, *Ppp3r1* and *Plek* are expressed in the mouse cochlea

The auditory phenotype of DFNA58 is typical of sensorineural HL. Because of the high conservation between mouse and human deafness related genes, we investigated the expression pattern of the three protein-coding genes involved in the DFNA58 duplication in the mouse cochlea, predicting that the causal gene(s) would be expressed in this organ. Using *in situ* hybridization (ISH) on sagittal sections of postnatal day (P) 7 mice, we detected a robust expression of *Cnrip1* mRNA in spiral ganglion neurons (SGNs, white arrowhead) and a weaker one in the tympanic borders cells (Fig. 5). *Cnrip1* expression was already strong in SGNs at P4 (data not shown). *Plek* mRNA expression was below ISH detection level in the inner ear at P7 (Fig. 6), while it was detected in the dorsal root ganglion (data not shown). We detected *Ppp3r1* mRNA in the pillar cells (Fig. 7C, white dotted lines,) and at a weaker level in the SGN (white arrowhead) at P7. Notably, no mRNA expression of these genes was detected in the cochlear hair cells at this age, while *Loxhd1* was present (34) (Supplementary Figure S2). In early adulthood (P21–P35) all three genes (*Cnrip1*, *Ppp3r1* and *Plek*) showed expression in the organ of Corti and in the spiral ganglion neurons. In addition, *Plek* expression was also observed in the spiral limbus fibrocytes and *Ppp3r1* in the spiral ligament fibrocytes (Figs 5–7).

Next, we investigate the protein expression of these genes, whenever specific commercial antibodies for the mouse orthologue were available: By early adulthood (P21–P35), CNRIP1 expression was evident within SGN cell bodies and their neurites extending toward the organ of Corti and those extending toward the brainstem (Fig. 8A and B), and more weakly in the supporting cells of the organ of Corti. CNRIP1 immunofluorescence appeared relatively stronger in a subset of SGNs located around the edge of the ganglion, and subsequent co-labeling with an anti-peripherin antibody confirmed their identity as type II SGNs (Fig. 8B) afferent neurons that innervate outer hair cells (35). By early adulthood, the PLEK antibody strongly labeled fibrocytes lining the perilymphatic face of the bony spiral limbus, and cells within stria vascularis and the organ of Corti displayed diffuse fluorescence (Fig. 9A) (36). The cell bodies of SGNs were also distinctly immunolabeled (Fig. 9A and B). Regarding the PPP3R1 protein, the best antibody available (ThermoFisher PA5–29939) was predicted to detect two murine gene products, Pp3r1 and Pp3r2, preventing the investigation.

In summary, the expression of the three candidate protein-coding genes could be localized to key loci within the cochlea, as well as within the primary afferent pathway. Thus, all three genes could potentially play an important role in the inner ear.

Discussion

We obtained detailed clinical data from hearing-impaired members of the DFNA58 family that suggest that the primary

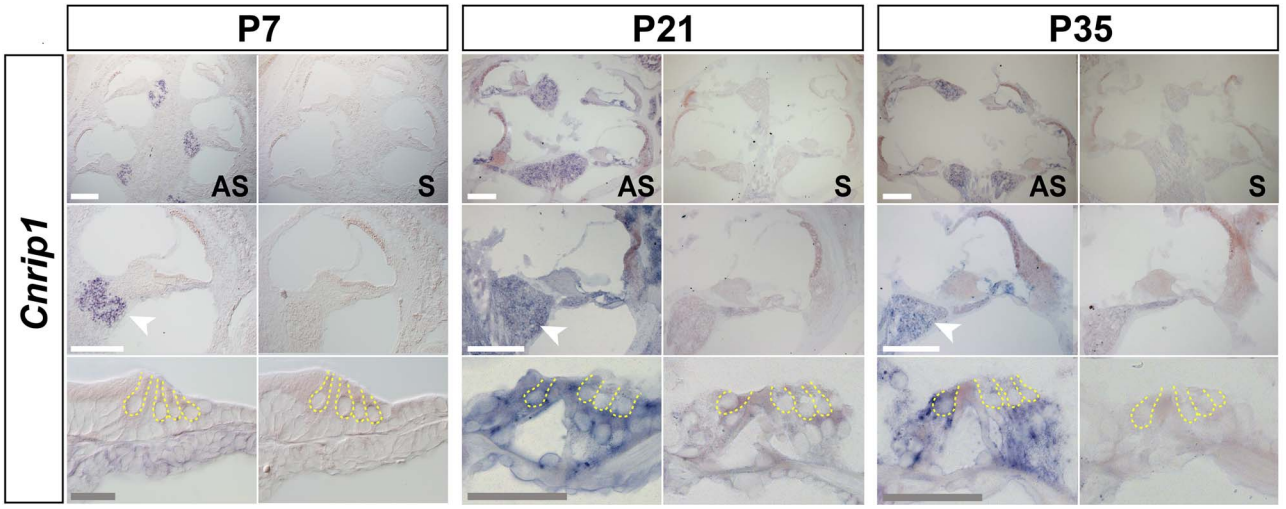

Figure 5. ISH in sagittal sections of P7, P21 and P35 mouse cochlea showing the expression of the candidate gene *Cnrip1*. AS, antisense probe; S, sense probe (negative control).White scale: 200 µm; Gray scale: 50 µm. White arrowhead: SGN; yellow dotted contour: inner and outer hair cells. The hair-cell specific marker *Loxhd1*, used as control, is in Supplementary Figure S2.

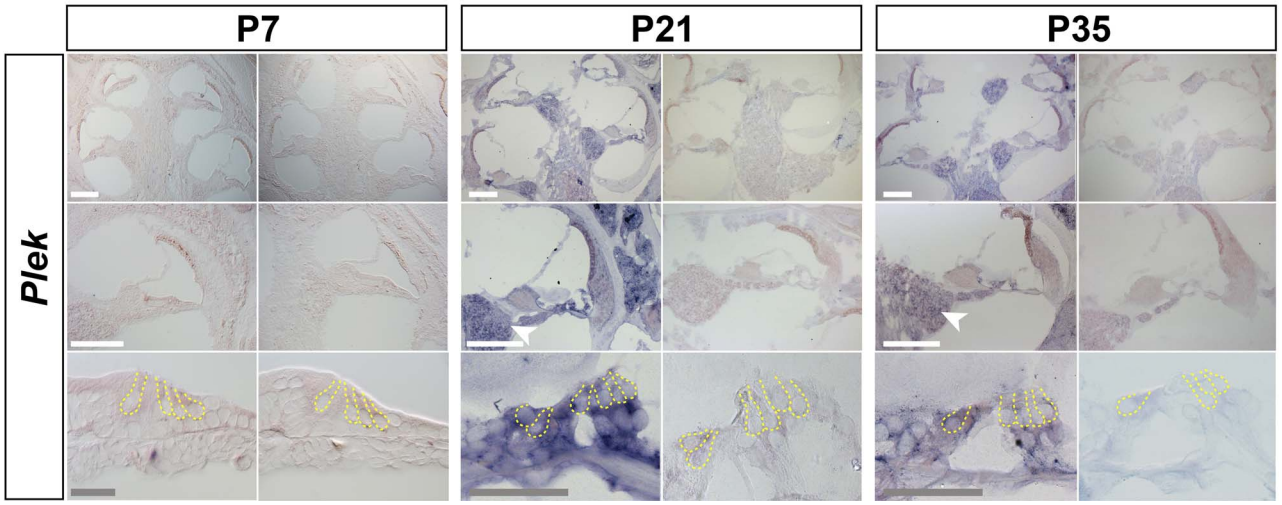

Figure 6. ISH in sagittal sections of P7, P21 and P35 mouse cochlea showing the expression of the candidate gene *Plek*.AS, antisense probe; S, sense probe (negative control). White scale: 200 µm; Gray scale: 50 µm. White arrowhead: SGN; yellow dotted contour: inner and outer hair cells. The hair-cell specific marker *Loxhd1*, used as control, is in Supplementary Figure S2.

functional impairment leading to HL is cochlear in origin. Relevant vestibular symptoms were not reported by affected individuals, but irritative peripheral vestibular syndrome was diagnosed after vectoeletronystagmography exam in one affected family member (VI:2), with no complaints of vestibular symptoms.

Through CNV analysis of data from exome sequencing, MLPA and qPCR, we revealed the genetic alteration responsible for DFNA58 HL as a ~200 kb duplication in 2p14. Linkage and haplotype analysis, as well as the sequence of low abundant aberrant fusion transcripts detected, suggest that the duplication is inserted in the same chromosomal region and orientation, more likely 3′ downstream the last exon of *PPP3R1* (Fig. 4D). The relevance of causative CNVs to non-syndromic HL has been pointed out in a study of 686 North-American patients, in which 18.7% were explained by pathogenic CNVs (16) and in a study of 100 Brazilian cases with 4% carrying causative CNVs (37). The duplication in DFNA58 patients includes two entire known protein-coding genes, *PLEK* and *CNRIP1*, and the first exon of another known protein-coding gene *PPP3R1*. Our results are consistent with bioinformatics prediction of deafness genes based on inner ear expression levels: using a classifier algorithm that utilizes the genome-wide transcriptional expression patterns of genes differentially expressed between mouse cochlear and vestibular sensory epithelia and between developmental ages, *PPP3R1*, *CNRIP1* and *PLEK* have been predicted as deafness genes with probabilities of 1.52, 1.33 and 1.31%, respectively, probabilities similar to many known deafness genes (*SLC22A4*–1.53%, *PRPS1*–1.41%, *DIAPH1*–1.18%, *HGF*-1.02%; *DFNA5*–0.95%, *DIAPH3*–0.89%, *NARS2*–0.75%) (15).

In all previously described cases of duplications involving 2p14 (published studies or from databases), those with or without overlap with our small duplication, the patients were evaluated during childhood, and these descriptions do not rule out the possibility of progressive HL manifesting during adolescence or adulthood. Two of the Decipher (27) patients with

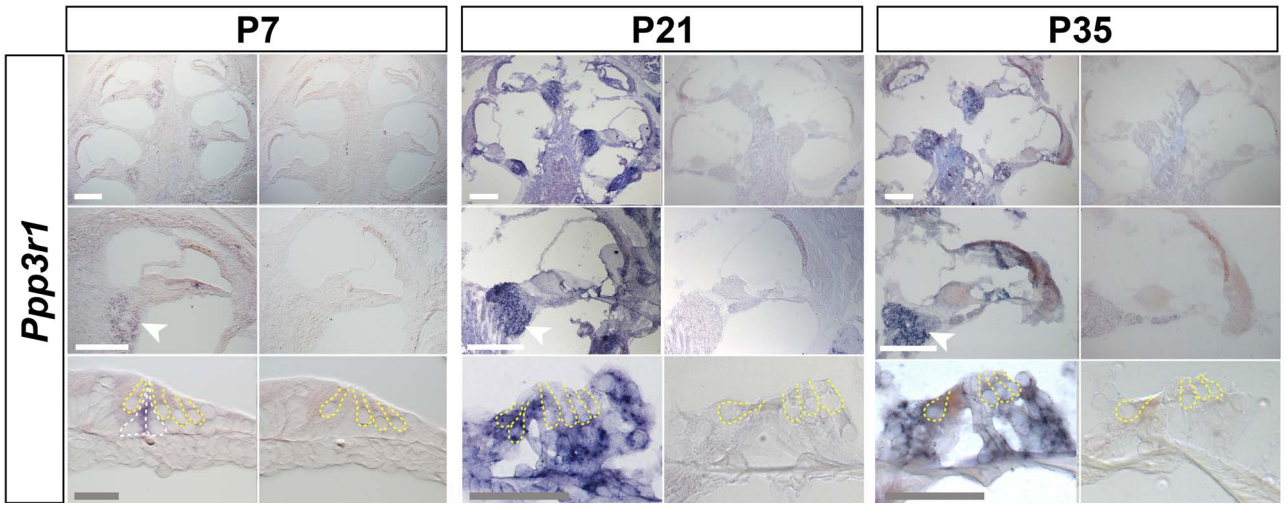

Figure 7. ISH in sagittal sections of P7, P21 and P35 mouse cochlea showing the expression of the candidate gene *Ppp3r1*. AS, antisense probe; S, sense probe (negative control). White scale: 200 µm; Gray scale: 50 µm. White arrowhead: SGN; white dotted contour: pillar cells; yellow dotted contour: inner and outer hair cells. The hair-cell specific marker *Loxhd1*, used as control, is in Supplementary Figure S2.

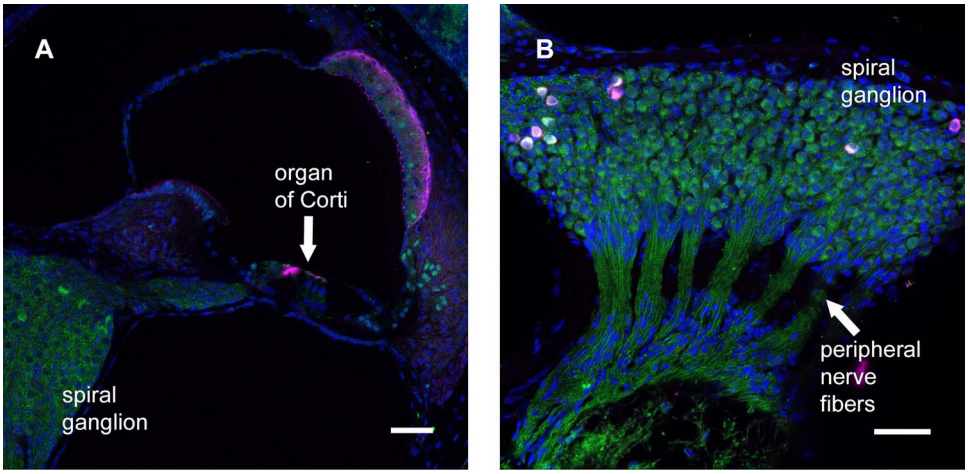

Figure 8. CNRIP1 immunofluorescence in sagittal sections of young adult murine cochleae. (**A**) Relatively intense CNRIP1 staining (green) of SGNs located at the periphery of the spiral ganglion at P21. Some nuclei were stained within the spiral ligament and stria vascularis. Actin stained using fluorescently tagged phalloidin (magenta). (**B**) At P35, neurons most strongly expressing CNRIP1 double labeled using anti-peripherin (magenta), a specific marker of type II SGNs. Anti-CNRIP1 immunofluorescence also detected in peripheral nerve fibers leaving the ganglion, toward the auditory brainstem. Scale bars = 50 µm.

larger duplications that included our duplicated segment were assessed clinically during the first year of life (Patient 262 515—duplicated region chr2:55 572 601–72 071 295:GRCh38; Patient 286 233—duplicated region chr2:10 000-88 750 930:GRCh38). As to the third one, who was ascertained at the age of 16, there was no phenotypic data (Patient 254 864—assessed below of the maximum age of onset of DFNA58 HL). Besides, patients with serious development disorders might have hearing deficit overlooked. Two Decipher subjects had a duplication that includes only *PLEK*, among the genes involved in the DFNA58 duplication: patient 285 363 (duplicated region chr2:68 328 045-68 431 525:GRCh38) who was last clinically assessed before the age of 1 year, presented hypoplasia of the corpus callosum, and patient 252 471 (duplicated region chr2: 68 370 141–68 802 061) with no phenotype information available except that the duplication was inherited from normal parents (age of parents was not reported). In three control samples from DGV (26), smaller duplication events covered only the *PLEK* gene. Three events (in 37 125 samples) of the *CNRIP1* duplication were present in ExAC (28) CNV databases, but they were also part of larger duplications. Although duplications involving *CNRIP1* were listed in ExAC (28), this fact does not exclude that duplication of this gene may cause teenager-to-young adult onset HL. In addition, CNV computation in the ExAC database also has limitations regarding the difficulty of making accurate CNV calls from read depth data, in particular from targeted short read sequencing, combined with the rarity of these events, which add noise to estimates of frequencies (see http://exac.broadinstitute.org/faq) (26). Some published studies have reported large duplications involving 2p14. In one study, the authors hypothesized that *CNRIP1* overdose, one of the 47 genes duplicated, could be responsible for the neurological and mental disorders due to its wide expression in the central nervous system (38). The segregation of a duplication involving *CNRIP1* in 20 adult patients without significant neurological or mental disorders argues against this hypothesis. The other few reports of duplications in 2p did not overlap with ours or did not indicate their precise physical positions (38–40). In another

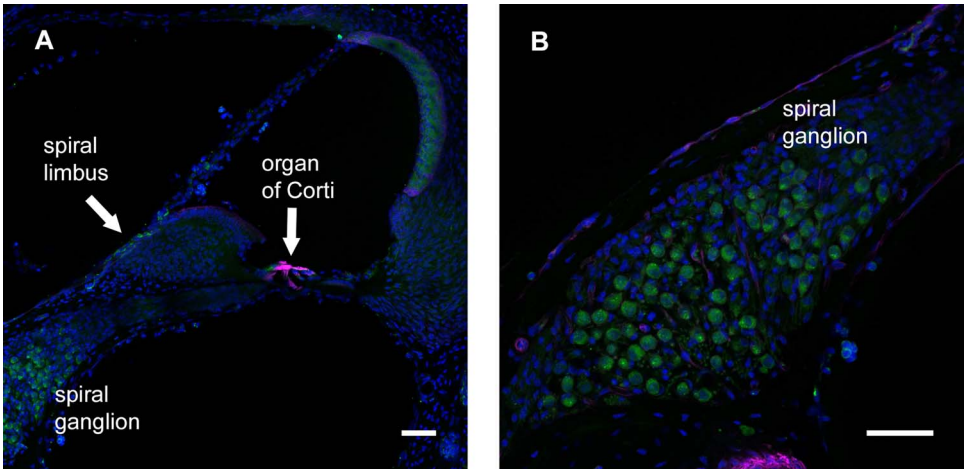

Figure 9. PLEK immunofluorescence in sagittal sections of young adult murine cochleae. (**A**) In the basal turn at P35, PLEK immunofluorescence (green) evident within SGN cell bodies and fibrocytes on the perilymphatic face of spiral limbus. Actin stained using fluorescently tagged phalloidin (magenta). (**B**) Detail of the spiral ganglion in the basal cochlear region. Scale bars = 50 µm.

recent investigation, hypercalcaemia was described in a patient with 2p13.2-p16.1 duplication and the authors proposed it was associated with *PPP3R1* duplication (41).

The gene expression analyses in patients' blood revealed significant overexpression of *CNRIP1* and of two of the four uncharacterized lncRNA genes inside the duplication, *LOC107985892* (mRNAsXR_001739527.1/XR_001739526.1) and *LOC102724389* (mRNA XR_940224.3). Besides, fusion *PPP3R1–CNRIP1* transcripts were also detected in all duplication carriers, but their sequence were not identical among carriers. Thus, it seems unlikely that they represent the main cause of HL, but a contribution to the phenotype could not be ruled out (Fig. 4D). Ideally, the expression patterns of the candidate genes should be tested in source material with disease-representative expression patterns, but human cochlear tissue is rarely available for biopsy in the literature, and was not available in our study. Likewise, with the emerging and promising strategy of using RNAseq for disease gene discovery, the main challenge is to have a disease representative tissue for this type of gene expression analysis. Since inner ear samples cannot be obtained from patients, we turned to blood samples. The relevance of blood as a proxy for studying disease patterns of gene expression can vary depending on the target tissues and the pathology. For example, in neuromuscular disorders, blood was not found to be representative of myotubes expression (42). Nonetheless, altered expression patterns analysis in blood have already been successful in pointing out candidate genes linked to the disease phenotype for rare diseases of varying pathophysiology, including neurological diseases, where blood was not previously assumed to be a representative tissue (43–45). Thus, we cannot exclude the possibility of a different gene expression pattern in the human cochlea from what was found in blood. However, whole-gene duplication leading to documented overexpression associated with HL has been reported before in the case of the DFNA51 locus, in which a 270 Kb duplication including the *TJP2* gene, as the sole protein-coding gene, segregated with autosomal dominant postlingual progressive HL in an Israeli family (44). The authors observed overexpression of the *TJP2* gene in lymphoblastoid cells obtained from duplication carriers' blood and proposed that the duplication maintains the *TJP2* gene expression elevated as age progresses, which leads to increased *TJP2* and GSK-3β mediated susceptibility to apoptosis of the inner ear cells (46). Although the observed overexpression of *CNRIP1* and of the two non-coding RNAs in DFNA58 patients' blood may not represent the expression patterns in the inner ear, a currently inaccessible tissue, it is tempting to speculate that a similar mechanism might explain DFNA58 related HL.

The localization of *CNRIP1* expression (RNA and protein) in the mouse cochlea, enriched in the SGNs, and preferentially in type II neurons, is in agreement with four transcriptome studies of the inner ear (47–50). First, in a cell type–specific transcriptome analysis of the mouse inner ear (47), *CNRIP1* was one of the 3000 genes differentially expressed, but *PPP3R1* and *PLEK* were not. The expression of *CNRIP1* in neuronal cells was about 10X higher than in sensory epithelia (47). Secondly, in a study of the adult human inner ear transcriptome, *CNRIP1* gene was found to be significantly expressed in the inner ear and included in the >3500 genes differentially expressed between cochlea and vestibule but the *PPP3R1* and the *PLEK* genes were not differentially expressed (48). Finally, in two studies focused on SGNs gene expression (49,50), *CNRIP1* and *PPP3R1* (but not *PLEK*) expression was detected, both in type I and type II SGNs, but type II SGN expression was higher compared to that in type I SGNs (49). The difference between the two neuronal types was more pronounced in adults, in both genes (https://lallemendlabcochlea.shinyapps.io/shinyapp-sgns_diversity) (49). There are known genes responsible for sensorineural HL that are mainly expressed in SGNs, such as *DIAPH1* (51), *GAB1* and *METTL13* (52).

The human *CNRIP1* (CB1 cannabinoid receptor-interacting protein 1, OMIM*618538) encodes two protein isoforms: CRIP1a (164 aa), largely expressed in the brain and detected throughout vertebrates, and CRIP1b (128 aa), unique to primates (32). CRIP1a is known to interact *in vivo* with the human and rat cannabinoid type 1 receptor (*CNR1*, *CB1R*, OMIM*114610) and mGlu8R (or *GRM8*, OMIM* 601116) (32,53). It has been described that CRIP1a may act as a presynaptic modulator of neurotransmitter release and also in mediation of long-term depression of synaptic responses (54–57); It regulates CB1 activity upon agonist or antagonist binding as well as CB1 internalization (58,59,32); suppresses CB1-mediated tonic inhibition of voltage-gated Ca^{2+} channels (32); enhances CB1R signaling and diminishes the severity of chemically induced acute epileptiform seizure upon induced overexpression *in vivo* in the mouse hippocampus (54); reduces the proliferative and migration abilities of colon cancer cells upon induced overexpression *in vitro* (60). *CNRIP1* downregulation (due to hyper-methylation of its promoter)

is a biomarker in some cancers, such as adenomas, gastric cancer, non-Hodgkin lymphoma and cholangiocarcinoma (61–66). Functional data from the literature above described, besides the expression data herein presented, indicate *CNRIP1* as a putative deafness gene. To date, some HL genes are also known to be associated with cancer and/or apoptotic pathways (*GAB1*-DFNB26 and *METTL13*-DFNM1, *HGF*-DFNB39, *MSRB3*-DFNB74, *MET*-DFNB97, *GSDME*-DFNA5, *EYA4*-DFNA10, *TJP2*-DFNA51, *SMAC/DIABLO*–DFNA64, *MCM2*-DFNA70, *CEACAM16*-DFNA4B/DFNB113).

The *PLEK* gene encodes the pleckstrin protein, a major Protein Kinase C substrate of platelets and leucocytes (67). Pleckstrin plays an important role in exocytosis (68), protein homodimerization activity, actin cytoskeleton reorganization and cell projection organization (69–73), cortical actin cytoskeleton organization. It is reported as an important intermediate in the secretion and activation pathways of pro-inflammatory cytokines TNF-α and IL-1β (74). This gene has been found to be up-regulated in periodontitis and in chronic inflammatory diseases, such as cardiovascular disease, Rheumatoid Arthritis and Ulcerative Colitis (75). The *PPP3R1* gene encodes for the 19-kD regulatory unit of calcineurin (calcineurin B), which is a calcium-dependent serine/threonine protein phosphatase that is stimulated by calmodulin, conferring calcium sensitivity. It plays a role in many pathways including regulation of synaptic activity and neuronal excitability (76), T-cell activation in immune system responses (77), Na/K ion transportation (78) and cell death (79). It was also related to neurodegenerative disease, cardiac hypertrophy (80), Alzheimer's disease (81) and with the efficacy of the drug tacrolimus therapy for idiopathic membranous nephropathy (82). Various *in vitro* and *in vivo* observations indicate that excess of calcium may cause cell death through a number of pathways that include calcineurin (83,84), which has also been demonstrated in the cochlea, where chemical inhibition of calcineurin was shown to avoid noise-induced HL (85). Functional assays, like those above described in the literature (*PLEK*: 71, 72; *PPP3R1*: 76, 78, 85) as well as expression patterns here observed in the murine cochlea, suggest *PLEK* and *PPP3R1* are also good deafness candidate genes.

To date, there is no knowledge about function or orthologues of the two lncRNA genes highly overexpressed in the blood of DFNA58 patients. The lncRNA genes belong to a heterogeneous class of non-coding transcripts, longer than 200 nucleotides, that may be related to diverse mechanisms of transcriptional regulation of gene expression (86), such as mediation of epigenetic modifications of DNA (87), organization of nuclear domains, regulation of expression of neighboring genes (88) and distant genomic sequences (89), regulation of proteins or RNA molecules (90), precursors for functional small RNAs (with or without function themselves). Some annotated lncRNAs actually encode for small proteins (91). The biological role of most of them remain enigmatic (92), but the biological role of intergenic lncRNA, either as positive regulators of gene expression of neighboring protein-coding genes or negative regulators, has already been demonstrated (91). Likely, the lncRNA genes within the duplication could act as expression regulators of *CNRIP1*, *PPP3R1*, and/or *PLEK* or, less likely, of the other neighbor genes.

In conclusion, we revealed the genetic defect responsible for DFNA58 related postlingual progressive HL (22) as a 200 Kb rare duplication in 2p14 [NC_000002.12:g.(68 247 572_68 248 077)_(68 449 525_68 452 166)dup]. Expression patterns of RNA and protein in the murine cochlea suggested all three protein coding genes (*CNRIP1*, *PPP3R1* and *PLEK*) involved in the duplication are potential candidates to explain the HL. The observed overexpression of *CNRIP1* and two lncRNAs in blood of duplication carriers, and the detected fusion transcripts, allows us to speculate that they contribute to the postlingual HL in these patients. It is possible that a complex mechanism involving the expression of more than one gene (protein-coding and/or lncRNA) underlies DFNA58 pathophysiology. Further studies, beyond the scope of the present work, are required to confirm protein expression, since the antibodies available at present cannot do definitively for all three proteins. Stem cell lines derived from duplication carriers and/or the generation of mouse Knockout/Knockin models are interesting tools to unravel the pathophysiology mechanisms of DFNA58 HL.

Materials and Methods

Subjects and clinical evaluation

This study was approved by the Ethics Committee for Analysis of Research Projects from Institute of Biosciences and School of Medicine (both of University of Sao Paulo) and the National Committee on Ethics in Research (CONEP). Written informed consent was obtained from all hearing impaired individuals or their legal guardians, their relatives and control individuals. Animal procedures were approved by the Animal Care and Use Committee (CEUA) from University of São Paulo School of Medicine. The four-generation Brazilian family described in Lezirovitz *et al.* (22), in which the locus DFNA58 was mapped was reassessed, with new samples collected, both from already enrolled members, as well from additional affected and unaffected subjects. The updated pedigree is shown in Figure 1.

Physical examinations and complete anamneses were performed in order to acknowledge any other clinical symptoms. Malformations of the inner ear were examined by high-resolution computed tomography and/or magnetic resonance imaging of the temporal bones in the subjects IV:6, IV:9 and VI:2. In general, pure tone audiometry was performed to test for air conduction (250–8000 Hz) and bone conduction (500–4000 Hz) in affected and unaffected subjects. The affected individuals from the DFNA58 family recently ascertained or reevaluated in the last 2 years, which were available, (IV:6, IV:9, IV:16, IV:19, IV:22, V:7, V:9, V:25, V:27, V:28 andVI:2) were also examined by otomicroscopy and tympanometry followed by click auditory brainstem responses (click ABR), bone conduction ABR and distortion-product otoacoustic emissions (DPOAE), obtained in a sound-treated room. The tests were run using the Navigator Pro SCOUT and AEP (Natus Bio-Logic Systems Corp., Mundelein, IL) software. Click stimuli (duration: 100 milliseconds) were presented with ER3A insert earphones at the maximum level of 90 dBHL (Hearing Level) at a rate of 21.1/s with rarefaction and condensation polarity. The responses were considered to be absent when rarefaction and condensation waves showed no responses at 90 dBHL (Hearing Level). Bone conduction was tested with a B71 bone oscillator at the maximum intensity of 55 dB HL using alternated click stimuli and contralateral masking of the same intensity. For DPOAE we applied the diagnostic 750 to 8000 Hz test protocol (Navigator Pro SCOUT, Natus Bio-Logic Systems Corp., Mundelein, IL). We considered positive responses when the signal-to noise ratio was 6 dB or more in at least five out of eight frequencies, according to the Expanded Boys Town reference data.

Preparation of DNA

A total of 47 blood samples were obtained for DNA analysis, 23 from affected individuals and 24 from unaffected (three of

them unrelated married-in). In addition, 100 probands from families with postlingual HL and presumptive autosomal dominant inheritance (autosomal dominant non-syndromic HL) were also included in the molecular study of the duplication. Genomic DNA was extracted with commercial kits from whole blood. DNA quantity and quality was assessed by Nanodrop spectrometer (Nanodrop Technologies, Wilmington, DE, USA) and agarose gel.

Identification of the causative mutation in DFNA58

Whole exome sequencing was performed in a sample from patient IV:16. The DNA quality and quantity were verified with a Qubit fluorometer (Life Technologies, Grand Island, NY, USA). Capture libraries were prepared according to the manufacturer's instructions using the SureSelectXT Human All Exon V4 (Agilent Technologies, Santa Clara, CA, USA). Sequencing was conducted with an Illumina HiSeq2000 analyzer (Illumina, San Diego, CA, USA) by Mendelics Genomic Analysis. Base call was performed using Illumina program bcl2fastq. FastQ alignment to the reference sequence of the human genome (b37 GRC/NCBI) was done with BWA and recalibration of aligned sequences with GATK (http://www.broadinstitute.org/gatk/) (93). Genotyping of SNPs and InDels was also achieved with GATK. CNVs detection was performed using ExomeDepth (http://cran.r-project.org/web/packages/ExomeDepth/index.html) (25) and Pindel (https://trac.nbic.nl/pindel/) (94). Variants were annotated with Annovar (http://annovar.openbioinformatics.org/) (95) and SnpEff (http://snpeff.sourceforge.net/) (96). Population frequency filtering of the variant set was performed using public databases (for example, 1000G and ExAC) (24,28), including Brazilian variant banks ABraOM: Brazilian genomic variants (97). VAAST (http://www.yandell-lab.org/software/vaast.html) (98) and PolyPhen (http://genetics.bwh.harvard.edu/pph2/) (99) were used to evaluate biological effect of each variant, defining a priority list of candidate variants.

Duplication confirmation and fine mapping of the borders

Array-CGH was performed in DNA samples from two patients (V:12 and IV:16) using an 180 K oligo-platform from Agilent Technologies, as described by the manufacturer. Data were processed with feature extraction software and subsequently analyzed with the Genomic Workbench software (Agilent Technologies, Santa Clara, CA, USA). Gains and losses of genomic sequences were determined using the aberration detection statistical algorithm ADM-2, with a sensitivity threshold of 6.7. For each sample, we used two reversed labeled hybridizations; alterations not detected in both dye-swap experiments were disregarded. The detected CNVs was compared to data from oligoarray studies documented in the Database of Genomic Variants (DGV; http://projects.tcag.ca/variation/) (26).

To narrow down the mapping of the borders, we used the Illumina BeadChip 850 K® array, analyzed with BlueFuse™ Multi Analysis software in one duplication carrier (IV:16). We also used 15 pairs of primers designed to cover both breakpoints by qPCR amplicons. The qPCR reactions were carried out in a Step One System (Thermo Scientific, Waltham, MA, USA) with 20 ng of DNA and 2x PowerUp™ SYBR™ Green Master Mix (cat. A25780, Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, EUA), according to the manufacturer's protocol. Melting curves were performed after each amplification in order to verify if only one fragment was being amplified. The $2^{-\Delta\Delta Ct}$ model was used for CNV estimation (100). Controls samples were used as reference for normalization and carriers were compared to non-carriers of the duplication to define the status (duplicated or not) of each qPCR amplicon.

Segregation of the duplication in the DFNA58 pedigree and screening of 100 pedigrees with autosomal dominant non-syndromic HL

We designed a customized kit of 10 MLPA probes (Multiplex Ligation-Probe Amplification), using RaW-Probe software (version 0.15β) according to the manufacturer's instructions (MRC-Holland, Amsterdam, The Netherlands), to test all family members (affected and unaffected). Samples were collected from 100 other probands of pedigrees with autosomal dominant non-syndromic HL to test for the presence of CNVs (deletions or duplications) involving the three duplicated known protein coding genes. Three pairs of probes mapped outside the candidate duplicated region, in *PPP3R1*'s exon 3 (NM_000945) and in *FBXO48*'s exons 3 and 4 (NM_001024680). Seven pairs of probes mapped inside the duplication: three pairs in *CNRIP1* (exons 1–2-3; NM_015463), three pairs in *PLEK* (exons 3–5-7; NM_002664) and one pair in *PPP3R1* (exon 1). The amplification products obtained according to the manufacture's protocol, together with Rox Size Standard, were submitted to capillary electrophoresis in the ABI 3730 DNA Analyzer (Applied Bioystems, Foster City, CA, USA A). The results were analyzed with the Coffalyser.net software (MRC-Holland).

Multipoint LOD scores (DFNA58 HL phenotype X microsatellite markers) and Two-point (DFNA58 HL phenotype X duplication and DFNA58 HL phenotype X *CNRIP1* mRNA overexpression) were calculated using Morgan 3.4 software (101), under an autosomal dominant model and assuming penetrance of 96%. Marker allele frequencies were calculated based on genotype data from the family, while the disease allele frequency and the duplication was set at 0.0001. The four subjects below age of onset were included in the analysis with unknown phenotype.

Analysis of mRNA levels in blood of affected and unaffected individuals (RT-qPCR)

For RNA analysis, samples from 24 family members were collected, 16 affected (14 duplication carriers and two non-carriers), and eight unaffected (Fig. 1). Mononuclear cells from peripheral blood were isolated by Ficoll Gradient Separation and their total RNA was extracted with the RNeasy Microarray Tissue Mini Kit (Qiagen, Hilden, Germany). Synthesis of cDNA was performed from 500 ng to 1 µg of RNA with SuperScript™ III or IV First-Strand Synthesis System, using both random hexamers and oligoDT (Invitrogen, Thermo Scientific, Waltham, MA, USA). Primers pairs were designed using Primer-Blast (102) to analyze the expression of the three known protein-coding genes involved in the duplication (*PPP3R1, CNRIP1* and *PLEK*), the four uncharacterized lncRNA genes (*LOC102724389*, *LOC107985892*, *LOC101927723* and *AC017083.1*), part of a novel protein-coding (*AC017083.3*) and of two neighbor genes of the most likely duplication insertion position (*WDR92* and *PNO1*). The PCR reactions were carried out in a Step One System, using HOT FIREPol® Evagreen® qPCR Supermix (Solis Biodyne, Tartu, Estonia) or PowerUp™ SYBR™ Green Master Mix (Thermo Scientific, Waltham, MA, USA), 0.05 to 0.5 µM final primer concentration and 2 µl of cDNA dilution (20 to 70 ng of RNA, depending on the level of gene expression). Different reference genes were tested in order to obtain consistent and robust results (*B2M*, *TBP*, *GUSB* and *HPRT1* for the human samples). The primer pairs used for quantitative real time

PCR are listed in Supplementary Table S2. For all experiments, reference gene and reference sample were run together with tested gene, carriers and non-carriers samples. A negative control of the cDNA synthesis (without reverse transcriptase) was used in RT-qPCR, especially when primer pairs annealed to the same exon, in order to control for possible genomic DNA contamination. Technical replicates (2X) were done for every sample tested for every gene in all experiment runs. The $2^{-\Delta\Delta Ct}$model was also used for gene expression quantitative analysis to obtain the mean RNA level, fold of control/reference (103). In case the replicates cycle threshold (Ct) values had a standard deviation higher than 0.2 or the calculated RNA level significantly differed from the other samples from the same category (carriers/non-carriers), they were repeated. The mean and standard deviations shown in Supplementary Figure S1 represent technical replicates either in the same experiment plate or in different ones, but with the exact same conditions. For the genes with altered expression, analysis were conducted with two different reference genes for at least some (3–5) of the 14 carrier samples in order to check for reproducibility. Genes were classified as normally expressed, mildly overexpressed and highly overexpressed based on the statistically difference (two-tailed unpaired t-test, 95% confidence interval) between the mean RNA level of carriers and mean RNA level of non-carriers, as well as the magnitude of RNA level overexpression. Accordingly, genes were classified as normally expressed when there was no statistically significant difference between carriers and non-carriers; genes were classified as mildly overexpressed, when there was statistically significant difference between carriers and non-carriers mean, and RNA level mean of carriers ranged from 1.2–1.8X (higher than reference); genes were classified as highly overexpressed when there was statistically significant difference between carriers and non-carriers mean, and carriers RNA level mean ranged from 4.5–27.6X (higher than reference). Interestingly, these genes, classified as mildly expressed genes, were not overexpressed in every carrier, but those classified as highly overexpressed were significantly overexpressed in every duplication carrier tested. No difference on the classification (normally expressed, mildly overexpressed and highly overexpressed) was observed comparing *B2M* or *TBP* as reference genes. The RT-qPCR results of genes *PLEK*, *CNRIP1* and *PPP3R1* shown in Figure 4 and Supplementary Figure S1 used *B2M* gene as reference and for the others, *TBP* gene was the reference.

RT-PCR and sanger sequencing

In order to verify if the duplication carriers had aberrant fusion transcripts with sequences of both *PPP3R1* and *CNRIP1*, RT-PCR was performed with primer forward on exon 1 or 2 of *PPP3R1* and primer reverse on exon 2 of *CNRIP1*. The amplified fragments were sequenced using BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and analyzed in the ABI PRISM® 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Similarly, primers spanning *PPP3R1* exon 1 were combined with primers from *WDR92* and *PNO1* genes for the detection of possible fusion transcripts. All primers are listed in Supplementary Table S2.

In situ hybridization

A 787 bp Plek (GenBank NM_019549.2) cDNA fragment was amplified using the primers ATCACTACAACCAAACAGCAGGACCACT and TGGAAGTGGCTGCCTGCAAGTAATA. A 940 bp Cnrip1 (GenBank NM_029861) cDNA fragment was amplified using the primers GCCTGCTGCCTCCATGCTGTCTCT and CAGGAACACCAGCATACAATAGCAAAAG. A 777 bp Ppp3r1 (GenBank NM_024459) cDNA fragment was amplified using the primers AGCAAGATGGGAAATGAGGCGAGTTACC and AACCCCTTCCCTTTCTCCACCATACTGA. Fragments were cloned in pGEM-T (Promega) to generate RNA probes. ISH was performed in cryosections of murine cochlea as reported in Grillet *et al.* (34). Samples older than P7 were decalcified with 0.17 M EDTA in 4% PFA 1% PBS for 24 h.

Immunofluorescence

Whole cochleae were harvested from C57BL/6 mice. Animal work conformed to United Kingdom legislation outlined in the Animals (Scientific Procedures) Act 1986. A small hole was made in the apex of the otic capsule using a hypodermic needle allowed direct perfusion of paraformaldehyde (4% in phosphate buffered saline, PBS) into the cochlea. Fixation was carried out for 40 min at room temperature. Fixed cochleae were washed several times in PBS, then decalcified in 4% EDTA in PBS for 48 h at 4°C. Decalcified otic capsules were mounted in 4% low-melting point agarose (Sigma A9045, St. Louis, MO, USA) and sectioned on a vibratome (1000 plus system, Intracel) at 150- to 200-µm intervals. For antibody labeling, sections were blocked and permeabilized (10% normal goat serum and 0.2% Triton X-100 in PBS) for 1 h at room temperature. Primary antibodies, rabbit anti-CRIP1a from Dr Ken Mackie's lab (58,103), mouse anti-peripherin (Santa Cruz sc-377 093), and goat anti-Pleckstrin (Abcam, ab115514), were diluted 1:100 in lysine blocking solution (0.1 M lysine and 0.2% Triton X-100 in PBS). The anti-CRIP1a affinity-purified antibody is directed against the entire hCRIP1a protein, which shares 96% identity with the mouse CRIP1a. Its specificity has been validated by immunofluorescence in cell line overexpressing tagged-hCRIP1a, but also by peptide competition on mouse retina sections by Hu *et al.* (103). The anti-Pleckstrin antibody is a commercially available antibody that targets a C-terminal epitope in the human protein, sharing 93% homology with the same region in mouse Pleckstrin. The antibody labels a single band at the predicted molecular weight in western blots of human peripheral blood mononuclear cells (www.abcam.com/pleckstrin-antibody-ab115514.html). Slices were incubated in primary antibodies for 3 h at room temperature. The primary antibodies were omitted in control experiments. Following several washes in PBS, the sections were incubated in Alexa Fluor-tagged anti-rabbit and anti-mouse secondary antibodies (diluted 1:400 in lysine blocking solution; Invitrogen, Carlsbad, California, USA) and/or fluorescently-tagged phalloidin (1:1000; Sigma P1951) at room temperature for 1 h. After a final set of washes in PBS, sections were mounted in Vectashield containing DAPI (Vector Laboratories, Burlingame, CA, USA). Images were acquired using a 20X air immersion objective (N.A. 0.75) on an LSM880 confocal microscope (Carl Zeiss Microscopy, Oberkochen, Germany). Data shown are representative of at least three animals at each developmental age.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

This work was supported by the FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO [CEPID 2013/08028-1 to R.C.M.N; 2014/13071-6 and 2018/03433-9 to K.L.], Stanford Otolaryngology Head and Neck Surgery startup package and the

National Institute on Deafness and other Communication Disorders (NIDCD) [grant RO1 DC016409-01A1 for A.T., N.Z. and N.G] and the Biotechnology and Biological Sciences Research Council [BB/M019322/1 and BB/R000549/1 to D.J.]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We specially thank all family members for their kind collaboration. We also thank Dr Ken Mackie, Linda and Jack Gill Chair of Neuroscience and Distinguished Professor from the Department of Psychological & Brain Sciences, Indiana University, for kindly providing a validated anti-CRIP1a antibody. Daniela Tiaki Uehara for helping in MLPA probe design, Dr Carla Rosenberg and Erika Freitas for performing the array-CGH analysis. We would like also to thanks the Stanford Initiative to Cure Hearing Loss contributors for their support, and in particular the Eberts and Oberndorf families.

Conflict of interest statement

None declared.

Data Availability Statement

These data have been submitted to the European Variation Archive (EVA) under the Project: PRJEB35159; Analyses: ERZ1106453.

Web Resources

ABraOM: Brazilian genomic variants—http://abraom.ib.usp.br/

Annovar (http://www.openbioinformatics.org/annovar/).

BIOGPS (http://biogps.org/).

Decipher (https://decipher.sanger.ac.uk/).

DGV; http://projects.tcag.ca/variation/

ExomeDepth (http://projects.tcag.ca/variation/).

GATK (http://www.broadinstitute.org/gatk).

Gene expression in spiral ganglion neuron subtypes—https://lallemendlabcochlea.shinyapps.io/shinyapp-sgns_diversity

UCSC Genome Browser—https://genome.ucsc.edu/

NIDCD Epidemiology and Statistics Program. Age at Which Hearing Loss Begins. Source: National Health Interview Survey, 2007. https://www.nidcd.nih.gov/health/statistics/age-which-hearing-loss-begins. Updated in November 2012.

Pindel (https://trac.nbic.nl/pindel/).

PolyPhen (http://genetics.bwh.harvard.edu/pph2/).

SnpEff (http://snpeff.sourceforge.net/).

VAAST (http://www.yandell-lab.org/software/vaast.html).

Van Camp, G., Smith, R.J.H. Hereditary Hearing Loss Homepage. https://hereditaryhearingloss.org. (5/2019).

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G. *et al*.. Ensembl 2018. *Nucleic Acids Res*., **46**, D754-D76.– http://grch37.ensembl.org/Homo_sapiens/

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M.R., Armean, I.M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C. (2019) Ensembl 2019. *Nucleic Acids Res*., **47**, D745-D751.

References

1. Hilgert, N., Smith, R.J. and Van Camp, G. (2009) Forty-six genes causing nonsyndromic hearing impairment: which ones should be analyzed in DNA diagnostics? *Mutat. Res*., **681**, 189–196.
2. NIDCD Epidemiology and Statistics Program. Age at which hearing loss begins. Source: National Health Interview Survey, 2007 https://www.nidcd.nih.gov/health/statistics/age-which-hearing-loss-begins (updated in November 2012).
3. Op de Beeck, K., Schacht, J. and Van Camp, G. (2011) Apoptosis in acquired and genetic hearing impairment: the programmed death of the hair cell. *Hear. Res*., **281**, 18–27.
4. Bowl, M.R. and Brown, S.D.M. (2018) Genetic landscape of auditory dysfunction. *Hum. Mol. Genet*., **27**, R130–R135.
5. Yan, D., Zhu, Y., Walsh, T., Xie, D., Yuan, H., Sirmaci, A., Fujikawa, T., Wong, A.C., Loh, T.L., Du, L. *et al*. (2013) Mutation of the ATP-gated P2X(2) receptor leads to progressive hearing loss and increased susceptibility to noise. *Proc. Natl. Acad. Sci. USA*., **110**, 2228–2233.
6. Bowl, M.R. and Dawson, S.J. (2015) The mouse as a model for age-related hearing loss - a mini-review. *Gerontology*, **61**, 149–157.
7. Vona, B., Nanda, I., Hofrichter, M.A., Shehata-Dieler, W. and Haaf, T. (2015) Non-syndromic hearing loss gene identification: a brief history and glimpse into the future. *Mol. Cell. Probes*, **29**, 260–270.
8. Ren, Y., Landegger, L.D. and Stankovic, K.M. (2019) Gene therapy for human Sensorineural hearing loss. *Front. Cell.Neurosci*., **13**, 323–335.
9. Bitner-Glindzics, M. (2002) Hereditary deafness and phenotyping in humans. *Br. Med. Bull*., **63**, 73–94.
10. Dror, A.A. and Avraham, K.B. (2010) Hearing impairment: a panoply of genes and functions. *Neuron*, **68**, 293–308.
11. Van Camp, G., Smith, R.J.H. (1/2020) hereditary hearing loss homepage https://hereditaryhearingloss.org.
12. Dror, A.A. and Avraham, K.B. (2009) Hearing loss: mechanisms revealed by genetics and cell biology. *Annu. Rev. Genet*., **43**, 411–437.
13. Richardson, G.P., de Monvel, J.B. and Petit, C. (2011) How the genetics of deafness illuminates auditory physiology. *Annu. Rev. Physiol*., **73**, 311–334.
14. Pandya, A. (2016) Genetic hearing loss: the journey of discovery to destination—how close are we to therapy? *Mol. Genet. Genomic Med*., **4**, 583–587.
15. Perl, K., Shamir, R., Avraham, K. and B. (2018) Computational analysis of mRNA expression profiling in the inner ear reveals candidate transcription factors associated with proliferation, differentiation, and deafness. *Hum. Genomics*, **12**, 30–49.
16. Shearer, A.E., Kolbe, D.L., Azaiez, H., Sloan, C.M., Frees, K.L., Weaver, A.E., Clark, E.T., Nishimura, C.J., Black-Ziegelbein, E.A. and Smith, R.J. (2014) Copy number variants are a common cause of non-syndromic hearing loss. *Genome Med*., **6**, 37–47.
17. Bademci, G., Foster, J., 2nd., Mahdieh, N., Bonyadi, M., Duman, D., Cengiz, F.B., Menendez, I., Diaz-Horta, O., Shirkavand, A., Zeinali, S. *et al*. (2016) Comprehensive analysis via exome sequencing uncovers genetic etiology in autosomal recessive nonsyndromic deafness in a large multiethnic cohort. *Genet. Med*., **18**, 364–371.
18. Cabanillas, R., Diñeiro, M., Cifuentes, G.A., Castillo, D., Pruneda, P.C., Álvarez, R., Sánchez-Durán, N., Capín, R., Plasencia, A., Viejo-Díaz, M. *et al*. (2018) Comprehensive genomic diagnosis of non-syndromic and syndromic hereditary hearing loss in Spanish patients. *BMC Med. Genet*., **11**, 58–74.
19. Wesdorp, M., de KoningGans, P.A.M., Schraders, M., Oostrik, J., Huynen, M.A., Venselaar, H., Beynon, A.J., van Gaalen, J., Piai, V., Voermans, N. *et al*. (2018) Heterozygous

missense variants of LMX1A lead to nonsyndromic hearing impairment and vestibular dysfunction. *Hum. Genet.*, **137**, 389–400.
20. Lv, Y., Gu, J., Qiu, H., Li, H., Zhang, Z., Yin, S., Mao, Y., Kong, L., Liang, B., Jiang, H. *et al.* (2019) Whole-exome sequencing identifies a donor splice-site variant in SMPX that causes rare X-linked congenital deafness. *Mol. Genet. Genomic. Med.*, **11**, e967–e975.
21. Morgan, A., Koboldt, D.C., Barrie, E.S., Crist, E.R., García, G.G., Mezzavilla, M., Faletra, F., Mosher, T.M., Wilson, R.K., Blanchet, C. *et al.* (2019) Mutations in PLS1, encoding fimbrin, cause autosomal dominant non-syndromic hearing loss. *Hum. Mutat.*, **40**, 2286–2295.
22. Lezirovitz, K., Braga, M.C., Thiele-Aguiar, R.S., Auricchio, M.T., Pearson, P.L., Otto, P.A. and Mingroni-Netto, R.C. (2009) A novel autosomal dominant deafness locus (DFNA58) maps to 2p12-p21. *Clin. Genet.*, **75**, 490–493.
23. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P. *et al.* (2019) Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*. doi: 10.1101/531210.
24. Clarke, L., Fairley, S., Zheng-Bradley, X., Streeter, I., Perry, E., Lowy, E., Tassé, A.M. and Flicek, P. (2017) The international genome sample resource (IGSR): a worldwide collection of genome variation incorporating the 1000 genomes project data. *Nucleic Acids Res.*, **45**, D854–D859.
25. Plagnol, V., Curtis, J., Epstein, M., Mok, K.Y., Stebbings, E., Grigoriadou, S., Wood, N.W., Hambleton, S., Burns, S.O., Thrasher, A.J. *et al.* (2012) A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics*, **28**, 2747–2754.
26. MacDonald, J.R., Ziman, R., Yuen, R.K., Feuk, L. and Scherer, S.W. (2014) The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.*, **42**, D986–D992.
27. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M. and Carter, N.P. (2009) DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am. J. Hum. Genet.*, **84**, 524–533.
28. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B. *et al.* (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, **536**, 285–291.
29. Ma, L., Bajic, V.B. and Zhang, Z. (2013) On the classification of long non-coding RNAs. *RNA Biol.*, **10**, 925–933.
30. Fernandes, J.C.R., Acuña, S.M., Aoki, J.I., Floeter-Winter, L.M. and Muxel, S.M. (2019) Long non-coding RNAs in the regulation of gene expression: physiology and disease. *Noncoding RNA.*, **5**, 1–25.
31. Wu, C., Jin, X., Tsueng, G., Afrasiabi, C. and Su, A.I. (2016) BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res.*, **44**, D313–D316.
32. Niehaus, J.L., Liu, Y., Wallis, K.T., Egertová, M., Bhartur, S.G., Mukhopadhyay, S., Shi, S., He, H., Selley, D.E., Howlett, A.C. *et al.* (2007) CB1 cannabinoid receptor activity is modulated by the cannabinoid receptor interacting protein CRIP 1a. *Mol. Pharmacol.*, **72**, 1557–1566.
33. Galvan, A., Frullanti, E., Anderlini, M., Manenti, G., Noci, S., Dugo, M., Ambrogi, F., De Cecco, L., Spinelli, R., Piazza, R. *et al.* (2013) Gene expression signature of non-involved lung tissue associated with survival in lung adenocarcinoma patients. *Carcinogenesis*, **34**, 2767–2773.
34. Grillet, N., Schwander, M., Hildebrand, M.S., Sczaniecka, A., Kolatkar, A., Velasco, J., Webster, J.A., Kahrizi, K., Najmabadi, H., Kimberling, W.J. *et al.* (2009) Mutations in LOXHD1, an evolutionarily conserved stereociliary protein, disrupt hair cell function in mice and cause progressive hearing loss in humans. *Am. J. Hum. Genet.*, **85**, 328–337.
35. Jagger, D.J. and Housley, G.D. (2003) Membrane properties of type II spiral ganglion neurones identified in a neonatal rat cochlear slice. *J. Physiol.*, **552**, 525–533.
36. Furness, D.N., Lawton, D.M., Mahendrasingam, S., Hodierne, L. and Jagger, D.J. (2009) Quantitative analysis of the expression of the glutamate-aspartate transporter and identification of functional glutamate uptake reveal a role for cochlear fibrocytes in glutamate homeostasis. *Neuroscience*, **162**, 1307–1321.
37. Rosenberg, C., Freitas, É.L., Uehara, D.T., Auricchio, M.T.B.M., Costa, S.S., Oiticica, J., Silva, A.G., Krepischi, A.C. and Mingroni-Netto, R.C. (2016) Genomic copy number alterations in non-syndromic hearing loss. *Clin. Genet.*, **89**, 473–477.
38. Kasnauskiene, J., Cimbalistiene, L., Utkus, A., Ciuladaite, Z., Preiksaitiene, E., Pečiulytė, A. and Kučinskas, V. (2012) Two new de novo interstitial duplications covering 2p14-p22.1: clinical and molecular analysis. *Cytogenet. Genome Res.*, **139**, 52–58.
39. Dobyns, W.B., Mirzaa, G., Christian, S.L., Petras, K., Roseberry, J., Clark, G.D., Curry, C.J., McDonald-McGinn, D., Medne, L., Zackai, E. *et al.* (2008) Consistent chromosome abnormalities identify novel polymicrogyria loci in 1p36.3, 2p16.1–p23.1, 4q21.21–q22.1, 6q26–q27, and 21q2. *Am. J. Med. Genet. A*, **146A**, 1637–1654.
40. Guilherme, R. (2009) Abnormal muscle development of the diaphragm in a fetus with 2p14–p16 duplication. *Am. J. Med. Genet. A*, **149A**, 2892–2897.
41. Lodefalk, M., Frykholm, C., Esbjörner, E. and Ljunggren, Ö. (2016) Hypercalcaemia in a patient with 2p13.2-p16.1 duplication. *Horm. Res. Paediatr.*, **85**, 213–218.
42. Gonorazky, H.D., Naumenko, S., Ramani, A.K., Nelakuditi, V., Mashouri, P., Wang, P., Kao, D., Ohri, K., Viththiyapaskaran, S. and Tarnopolsky, M.A. (2019) Expanding the boundaries of RNA sequencing as a diagnostic tool for rare Mendelian disease. *Am. J. Hum. Genet.*, **104**, 1007–1025.
43. Zeng, Y., Wang, G., Yang, E., Ji, G., Brinkmeyer-Langford, C.L. and Cai, J.J. (2015) Aberrant gene expression in humans. *PLoS Genet.*, **11**, e1004942–e1004961.
44. Zhao, J., Akinsanmi, I., Arafat, D., Cradick, T.J., Lee, C.M., Banskota, S., Marigortam, U.M., Bao, G. and Gibson, G. A burden of rare variants associated with extremes of gene expression in human peripheral blood. *Am. J. Hum. Genet.*, **98**, 299–309.
45. Frésard, L., Smail, C., Ferraro, N.M., Teran, N.A., Li, X., Smith, K.S., Bonner, D., Kernohan, K.D., Marwaha, S., Zappala, Z. *et al.* (2019) Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat. Med.*, **25**, 911–919.
46. Walsh, T., Pierce, S.B., Lenz, D.R., Brownstein, Z., Dagan-Rosenfeld, O., Shahin, H., Roeb, W., McCarthy, S., Nord, A.S., Gordon, C.R. *et al.* (2010) Genomic duplication and overexpression of TJP2/ZO-2 leads to altered expression of apoptosis genes in progressive nonsyndromic hearing loss DFNA51. *Am. J. Hum. Genet.*, **87**, 101–109.

47. Hertzano, R., Elkon, R., Kurima, K., Morrisson, A., Chan, S.L., Sallin, M., Biedlingmaier, A., Darling, D.S., Griffith, A.J., Eisenman, D.J. *et al.* (2011) Cell type-specific transcriptome analysis reveals a major role for Zeb1 and miR-200b in mouse inner ear morphogenesis. *PLoS Genet.*, **7**, e1002309–e1002323.
48. Schrauwen, I., Hasin-Brumshtein, Y., Corneveaux, J.J., Ohmen, J., White, C., Allen, A.N., Lusis, A.J., Van Camp, G., Huentelman, M.J. and Friedman, R.A. (2016) A comprehensive catalogue of the coding and non-coding transcripts of the human inner ear. *Hear. Res.*, **333**, 266–274.
49. Petitpré, C., Wu, H., Sharma, A., Tokarska, A., Fontanet, P., Wang, Y., Helmbacher, F., Yackle, K., Silberberg, G., Hadjab, S. *et al.* (2018) Neuronal heterogeneity and stereotyped connectivity in the auditory afferent system. *Nat. Commun.*, **9**, 3691–3703.
50. Sun, S., Babola, T., Pregernig, G., So, K.S., Nguyen, M., Su, S.M., Palermo, A.T., Bergles, D.E., Burns, J.C. and Müller, U. (2018) Hair cell Mechanotransduction regulates spontaneous activity and spiral ganglion subtype specification in the auditory system. *Cell*, **174**, 1247–1263.e15.
51. Neuhaus, C., Lang-Roth, R., Zimmermann, U., Heller, R., Eisenberger, T., Weikert, M., Markus, S., Knipper, M. and Bolz, H.J. (2017) Extension of the clinical and molecular phenotype of DIAPH1-associated autosomal dominant hearing loss (DFNA1). *Clin. Genet.*, **91**, 892–901.
52. Yousaf, R., Ahmed, Z.M., Giese, A.P., Morell, R.J., Lagziel, A., Dabdoub, A., Wilcox, E.R., Riazuddin, S., Friedman, T.B. and Riazuddin, S. (2018) Modifier variant of METTL13 suppresses human GAB1-associated profound deafness. *J. Clin. Invest.*, **128**, 1509–1522.
53. Mascia, F., Klotz, L., Lerch, J., Ahmed, M.H., Zhang, Y. and Enz, R. (2017) CRIP1a inhibits endocytosis of G-protein coupled receptors activated by endocannabinoids and glutamate by a common molecular mechanism. *J. Neurochem.*, **141**, 577–591.
54. Guggenhuber, S., Alpar, A., Chen, R., Schmitz, N., Wickert, M., Mattheus, T., Harasta, A.E., Purrio, M., Kaiser, N. and Elphick, M.R. (2016) Cannabinoid receptor-interacting protein Crip1a modulates CB1 receptor signaling in mouse hippocampus. *Brain Struct. Funct.*, **22**, 2061–2074.
55. Kreitzer, A.C. and Regehr, W.G. (2001) Cerebellar depolarization-induced suppression of inhibition is mediated by endogenous cannabinoids. *J. Neurosci.*, **21**, RC174–RC178.
56. Sjöström, P.J., Turrigiano, G.G. and Nelson, S.B. (2003) Neocortical LTD via coincident activation of presynaptic NMDA and cannabinoid receptors. *Neuron*, **39**, 641–654.
57. Wilson, R.I. and Nicoll, R.A. (2001) Endogenous cannabinoids mediate retrograde signaling at hippocampal synapses. *Nature*, **410**, 588–592.
58. Smith, T.H., Blume, L.C., Straiker, A., Cox, J.O., David, B.G., McVoy, J.R., Sayers, K.W., Poklis, J.L., Abdullah, R.A., Egertová, M. *et al.* (2015) Cannabinoid receptor-interacting protein 1a modulates CB1 receptor signaling and regulation. *Mol. Pharmacol.*, **87**, 747–765.
59. Blume, L.C., Eldeeb, K., Bass, C.E., Selley, D.E. and Howlett, A.C. (2015) Cannabinoid receptor interacting protein (CRIP1a) attenuates CB1R signaling in neuronal cells. *Cell. Signal.*, **27**, 716–726.
60. Zhang, T., Cui, G., Yao, Y.L., Wang, Q.C., Gu, H.G., Li, X.N., Zhang, H., Feng, W.M., Shi, Q.L. and Cui, W. (2017) Value of CNRIP1 promoter methylation in colorectal cancer screening and prognosis assessment and its influence on the activity of cancer cells. *Arch. Med. Sci.*, **13**, 1281–1294.
61. Lind, G.E., Danielsen, S.A., Ahlquis, T., Merok, M.A., Andresen, K., Skotheim, R.I., Hektoen, M., Rognum, T.O., Meling, G.I., Hoff, G. *et al.* (2011) Identification of an epigenetic biomarker panel with high sensitivity and specificity for colorectal cancer and adenomas. *Mol. Cancer*, **10**, 85–99.
62. Oster, B., Thorsen, K., Lamy, P., Wojdacz, T.K., Hansen, L.L., Birkenkamp-Demtröder, K., Sørensen, K.D., Laurberg, S., Orntoft, T.F. and Andersen, C.L. (2011) Identification and validation of highly frequent CpG island hypermethylation in colorectal adenomas and carcinomas. *Int. J. Cancer*, **129**, 2855–2266.
63. Huang, Q., Yang, Q., Mo, M., Ye, X., Zhang, J., Zhang, L., Chen, B., Li, J. and Cai, C. (2017) Screening of exon methylation biomarkers for colorectal cancer via LC-MS/MS strategy. *J. Mass Spectrom.*, **52**, 860–866.
64. Chong, Y., Mia-Jan, K., Ryu, H., Abdul-Ghafar, J., Munkhdelger, J., Lkhagvadorj, S., Jung, S.Y., Lee, M., Ji, S.Y., Choi, E. *et al.* (2014) DNA methylation status of a distinctively different subset of genes is associated with each histologic Lauren classification subtype in early gastric carcinogenesis. *Oncol. Rep.*, **31**, 2535–2544.
65. Bethge, N., Lothe, R.A., Honne, H., Andresen, K., Trøen, G., Eknæs, M., Liestøl, K., Holte, H., Delabie, J., Smeland, E.B. *et al.* (2014) Colorectal cancer DNA methylation marker panel validated with high performance in non-Hodgkin lymphoma. *Epigenetics*, **9**, 428–436.
66. Andresen, K., Boberg, K.M., Vedeld, H.M., Honne, H., Jebsen, P., Hektoen, M., Wadsworth, C.A., Clausen, O.P., Lundin, K.E., Paulsen, V. *et al.* (2015) Four DNA methylation biomarkers in biliary brush samples accurately identify the presence of cholangiocarcinoma. *Hepatology*, **61**, 1651–1659.
67. Gailani, D., Fisher, T.C., Mills, D.C. and Macfarlane, D.E. (1990) P47 phosphoprotein of blood platelets (pleckstrin) is a major target for phorbol ester-induced protein phosphorylation in intact platelets, granulocytes, lymphocytes, monocytes and cultured leukaemic cells: absence of P47 in non-haematopoietic cells. *Br. J. Haematol.*, **74**, 192–202.
68. Lian, L., Wang, Y., Flick, M., Choi, J., Scott, E.W., Degen, J., Lemmon, M.A. and Abrams, C.S. (2009) Loss of pleckstrin defines a novel pathway for PKC-mediated exocytosis. *Blood*, **113**, 3577–3584.
69. Ma, A.D., Brass, L.F. and Abrams, C.S. (1997) Pleckstrin associates with plasma membranes and induces the formation of membrane projections: requirements for phosphorylation and the NH2-terminal PH domain. *J. Cell Biol.*, **136**, 1071–1079.
70. Ma, A.D. and Abrams, C.S. (1999) Pleckstrin induces cytoskeletal reorganization via a Rac-dependent pathway. *J. Biol. Chem.*, **274**, 28730–28735.
71. Roll, R.L., Bauman, E.M., Bennett, J.S. and Abrams, C.S. (2000) Phosphorylated pleckstrin induces cell spreading via an integrin-dependent pathway. *J. Cell Biol.*, **150**, 1461–1466.
72. Baig, A., Bao, X. and Haslam, R.J. (2009) Proteomic identification of pleckstrin-associated proteins in platelets: possible interactions with actin. *Proteomics*, **9**, 4254–4258.
73. Hu, M.H., Bauman, E.M., Roll, R.L., Yeilding, N. and Abrams, C.S. (1999) Pleckstrin 2, a widely expressed paralog of pleckstrin involved in actin rearrangement. *J. Biol. Chem.*, **274**, 21515–21518.
74. Ding, Y., Kantarci, A., Badwey, J.A., Hasturk, H., Malabanan, A. and Van Dyke, T.E. (2007) Phosphorylation of pleckstrin

increases proinflammatory cytokine secretion by mononuclear phagocytes in diabetes mellitus. *J. Immunol.*, **179**, 647–654.
75. Lundmark, A., Davanian, H., Båge, T., Johannsen, G., Koro, C., Lundeberg, J. and Yucel-Lindberg, T. (2015) Transcriptome analysis reveals mucin 4 to be highly associated with periodontitis and identifies pleckstrin as a link to systemic diseases. *Sci. Rep.*, **5**, 18475–18488.
76. Reese, L.C. and Taglialatela, G. (2011) A role for Calcineurin in Alzheimer's disease. *Curr. Neuropharmacol.*, **9**, 685–692.
77. Kincaid, R.L. (1995) The role of calcineurin in immune system responses. *J. Allergy Clin. Immunol.*, **96**, 1170–1177.
78. Tumlin, J.A. (1997) Expression and function of calcineurin in the mammalian nephron: physiological roles, receptor signaling, and ion transport. *Am. J. Kidney Dis.*, **30**, 884–895.
79. Asai, A., Qiu, J., Narita, Y., Chi, S., Saito, N., Shinoura, N., Hamada, H., Kuchino, Y. and Kirino, T. (1999) High level calcineurin activity predisposes neuronal cells to apoptosis. *J. Biol. Chem.*, **274**, 34450–34458.
80. Rusnak, F. and Mertz, P. (2000) Calcineurin: form and function. *Physiol. Rev.*, **80**, 1483–1521.
81. Peterson, D., Munger, C., Crowley, J., Corcoran, C., Cruchaga, C., Goate, A.M., Norton, M.C., Green, R.C., Munger, R.G., Breitner, J.C. *et al.* (2014) Variants in PPP3R1 and MAPT are associated with more rapid functional decline in Alzheimer's disease: the Cache County dementia progression study. *Alzheimers Dement.*, **10**, 366–371.
82. Zhu, Y., Zhang, M., Wang, F., Lu, J., Chen, R., Xie, Q., Sun, J., Xue, J., Hao, C. and Lin, S. (2018) The calcineurin regulatory subunit polymorphism and the treatment efficacy of tacrolimus for idiopathic membranous nephropathy. *Int. Immunopharmacol.*, **65**, 422–428.
83. Jayaraman, T. and Marks, A.R. (2000) Calcineurin is downstream of the inositol 1,4,5-trisphosphate receptor in the apoptotic and cell growth pathways. *J. Biol. Chem.*, **275**, 6417–6420.
84. Orrenius, S., Zhivotovsky, B. and Nicotera, P. (2003) Regulation of cell death: the calcium-apoptotic link. *Nat. Rev. Cell. Biol.*, **4**, 552–565.
85. Minami, S.B., Yamashita, D., Schacht, J. and Miller, J.M. (2004) Calcineurin activation contributes to noise-induced hearing loss. *J. Neurosci. Res.*, **78**, 383–392.
86. Esteller, M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.*, **12**, 861–874.
87. Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L. *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, **464**, 1071–1076.
88. Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M. and Attardi, L.D. (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, **142**, 409–419.
89. Guttma, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R. and Bruhn, L. (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, **477**, 295–300.
90. Ulitsky, I. and Bartel, D.P. (2013) lincRNAs: genomics, evolution, and mechanisms. *Cell*, **154**, 26–46.
91. Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G. *et al.* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.*, **22**, 1775–1789.
92. Kopp, F1., Mendell, J.T. (2018) Functional classification and experimental dissection of long noncoding RNAs. *Cell*, **172**, 393–407.
93. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M. *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.*, **43**, 491–498.
94. Ye, K., Schulz, M.H., Long, Q., Apweiler, R. and Ning, Z. (2009) Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*, **25**, 2865–2871.
95. Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from next-generation sequencing data. *Nucleic Acids Res.*, **38**, e164.
96. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*, **6**, 80–92.
97. Naslavsky, M.S., Yamamoto, G.L., de Almeida, T.F., Ezquina, S.A.M., Sunaga, D.Y., Pho, N., Bozoklian, D., Sandberg, T.O.M., Brito, L.A., Lazar, M. *et al.* (2017) Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum. Mutat.*, **38**, 751–763.
98. Rope, A.F., Wang, K., Evjenth, R., Xing, J., Johnston, J.J., Swensen, J.J., Johnson, W.J., Moore, B., Huff, C.D., Bird, L.M. *et al.* (2011) Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal Acetyltransferase deficiency. *Am. J. Hum. Genet.*, **89**, 28–43.
99. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.
100. Livak, K.J. and Schmittgenm, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta DeltaC(T)) method. *Methods*, **25**, 402–408.
101. Tong, L. and Thompson, E. (2008) Multilocus lod scores in large pedigrees: combination of exact and approximate calculations. *Hum. Hered.*, **65**, 142–153.
102. Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. and Madden, T.L. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinf.*, **13**, 134–145.
103. Hu, S.S., Arnold, A., Hutchens, J.M., Radicke, J., Cravatt, B.F., Wager-Miller, J., Mackie, K. and Straiker, A. (2010) Architecture of cannabinoid signaling in mouse retina. *J. Comp. Neurol.*, **518**, 3848–3866.