RESEARCH ARTICLE

Whole-exome sequencing of de novo genetic variants in a Chinese family with a sporadic case of congenital nonsyndromic hearing loss [version 1; peer review: 1 approved, 1 approved with reservations]

Sijing Hu1, Hao Zhang2, Yunqiang Liu2, Mohan Liu2, Jingjing Li1, Shunyao Liao1

1Diabetes Center & Institute of Organ Transplantation, Sichuan Provincial People’s Hospital, University of Electronic Science and Technology of China, Chengdu, China, 610072, China
2Department of Medical Genetics and Division of Morbid Genomics, State Key Laboratory of Biotherapy, West China Hospital, West China Medical School, Sichuan University, Chengdu, China, 610041, China

First published: 02 Feb 2021, 10:61
https://doi.org/10.12688/f1000research.27739.1
Latest published: 24 Aug 2021, 10:61
https://doi.org/10.12688/f1000research.27739.2

Abstract

Background: We examined the genetic variants of a Chinese family with a 22-month-old infant with sporadic non-syndromic sensorineural hearing loss (NSHL).

Methods: The whole-exome sequence data in the family, especially the de novo variants presented in the patient, were analyzed and the effect of the disease-causing genetic variants on the protein expression level and cellular localization were examined by cell-based functional assay.

Results: The infant had no known NSHL-causing variants, except two compound heterozygous variants in connexin26 gene GJB2; one was the c.79G>A, c.341A>G haplotype from the asymptomatic mother which was benign, and the other was a de novo pathogenic c.262G>C (p.A88P). In vitro, GJB2 with c.262G>C was weakly expressed and displayed a punctate distribution in the cytoplasm and cytomembrane, while wild type GJB2 was robustly expressed in the cytomembrane. We deduced that the de novo pathogenic GJB2 c.262G>C exacerbated loss-of-function in the context of leaky variants c.79G>A, c.341A>G in the patient. Interestingly, further analysis of exome sequences revealed that the occurrence of de novo pathogenic variants in the infant was frequent. Among the total ~47,000 variants, 143 were de novo in the patient, whereas among all 74 variants predicted to be pathogenic/likely pathogenic, 21 were heterozygous and two were homozygous de novo. The occurrence rate of de novo deleterious variants was much higher (31.1%, 23/74) than that in total (0.34%, 143/47,000). It is notable that most genes with de novo deleterious variants were environment-sensitive, such as GJB2, MNK1, MNK2, MUC4, RAD21 and DNA copy number variations.

Open Peer Review

Invited Reviewers
1
2
version 2
(revision)
24 Aug 2021
version 1
02 Feb 2021

1. Kun Zhang, Chengdu Medical College, Chengdu, China
2. Yuande Tan, PENN State University, Hershey College of Medicine, Hershey, USA

Any reports and responses or comments on the article can be found at the end of the article.
Conclusions: The full picture of genetic variants in the exome might help us to interpret the NSHL-causing variants. More research is needed into the causes of de novo deleterious variants and gene-environment interactions in congenital NSHL.

Keywords
Whole-exome sequencing, de novo pathogenic variant, nonsyndromic hearing loss, compound heterozygosity, genetic and environmental interaction
Introduction

Hearing loss is one of the most common birth defects. Variants in the Gap Junction Protein Beta 2 gene (GJB2, HGNC: 4284), which encodes a beta-2 gap junction protein (connexin 26; Cx26), have been shown to be the leading genetic cause of non-syndromic sensorineural hearing loss (NSHL) (OMIM: 121011). GJB2-related autosomal recessive deafness explains approximately 50% of congenital autosomal recessive deafness, and GJB2-related autosomal dominant deafness is extremely rare.

GJB2 constitutes cell-to-cell channels and facilitates the intercellular exchange of ions and molecules. The amino acid alanine at position 88 (p.A88) of GJB2, which is located in the second transmembrane domain of Cx26, is highly conserved in vertebrates. To date, five studies have reported five nucleotide changes in the p.A88 coding region that resulted in distinct clinical abnormalities and different inheritance patterns. Frei et al. first reported the heterozygous c.262G>T (p.A88S) variant in a male Austrian patient with NSHL. As the proband’s mother was an asymptomatic carrier, the authors inferred that the missense variant could be connected to deafness but not in a simple, monogenetic disease model. Gravian et al. found that the c.262G>C (p.A88P) variant in compound heterozygosity with the nonpathogenic variant p.V27I in an Argentina child with profound deafness, implicating the destructive potential of the c.262G>T variant. Other researchers have reported p.A88 coding variants at the 263rd nucleotide: one was the c.263C>G (p.A88G) variant in a Tunisian girl with autosomal recessive NSHL, where her consanguineous parents were healthy carriers; another was the c.263C>A (p.A88E) variant in a Chinese patient with sporadic NSHL where the variant was in compound heterozygosity with the disease-causing c.235delC; and another was the c.263C>T (p.A88V) variant in a Japanese girl with severe keratitis-ichthyosis-deafness syndrome and septic complications, with unaffected parents. To date, by directly sequencing the GJB2 genetic region, studies have demonstrated that variants in GJB2 p.A88 have been associated with hearing loss in children. However, descriptions of the penetrance of the variants have been inconsistent.

On the other hand, the GJB2 c.79G>A (p.V27I, rs2274084) in cis with c.341A>G (p.E114G, rs2274083) forming a haplotype of p.[V27I; E114G] occurs frequently in East Asian populations. P.V27I is located in the first transmembrane domain and p.E114G is located in the intracellular loop of Cx26. Both are classified as benign polymorphisms. However, several clinical studies have found the p.[V27I; E114G] haplotype to be a risk factor for hearing impairment, and functional assays in vitro have demonstrated that the channel activities of VG (p.E114G variant only) and IG (both p.V27I; p.E114G variants) were reduced. However, as both genotypes were detected in both patients and controls, the exact pathogenic role of these variants in NSHL remains controversial.

Whole-exome sequencing enables a comprehensive and precise genetic investigation of congenital disorders and allows us to search highly heterogeneous genetic causes. This study aimed to explore possible molecular abnormalities in a Chinese non-consanguineous family with a 22-month old daughter suffering from NSHL. We carried out whole-exome sequencing, assessed the cytological/clinical characteristics of the genetic variants, and evaluated the possible cause of de novo pathogenic variants in the patient’s exome.

Methods

Patient details

The family included in this study is of Han Chinese heritage and resides in Chengdu City of Southwest China. The proband was a 22-month-old girl with NSHL who had previously been born in our hospital by spontaneous delivery at full term. Both of her parents were healthy during pregnancy. The baby failed the newborn hearing examination but no prenatal or postnatal risk factors for hearing loss were identified. Similarly, no family history of hearing abnormalities was reported. When the parents brought the 22-month-old child back to the hospital in October 2018, physical, biochemical, and otoscopic examinations were carried out. A CT scan of the temporal bones and MR analysis of the child’s head were also done to search for any organic brain lesions, and pure tone audiometry was performed in the girl and her parents.

Written informed consent was obtained from both parents for them and their daughter to participate in the study. The work was approved by the Research Ethics Committee of Sichuan Provincial People’s Hospital, School of Medicine, University of Electronic Science and Technology of China.

Whole-exome and mitochondrial DNA sequencing

Blood genomic DNA and mitochondrial DNA were extracted from all family members according to standard procedures (Abcam, Cambridge, UK) and stored in -20°C. The DNA concentration and quality were examined using a NanoDrop 2000 (Thermo, USA).
The whole-exome sequencing, the entire mitochondrial DNA and genetic variations analysis are described in our previous work. The fragmented genomic DNA was enriched using a NimbleGen probe capture array SeqCap EZ Exome Kit v3.0 (Roche NimbleGen, Inc. Madison, WI). The kit using the SeqCap advanced design algorithm coupled with 2.1 million long oligonucleotide probes to achieve superior target enrichment performance, and detect genetic variants with ~98% sensitivity and 99% specificity. The enriched DNA fragments passed the qPCR test, and the size distribution and concentration were examined using the Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA). The samples were sequenced on an Illumina NovaSeq 6000 (Illumina, San Diego, CA), and two parallel reactions were performed. Raw image files were processed by the BclToFastq (Illumina) for base calling to generate the raw data. The low-quality variations were filtered out using the quality score =20 (Q20). The sequencing reads were aligned to the NCBI human reference genome (hg19) using Burrows-Wheeler Aligner (version 0.6.2). SAMtools and Pindel were used to analyze single nucleotide polymorphisms (SNPs) and insertion/deletion of the sequence. The coding variants and CNVs were filtered out in the dbSNP135, Exome Variant Server, 1000 Genomes, and in-house database with more than 100,000 Chinese exomes (Joy Oriental Co. Beijing, China). The variants and CNVs were also searched in the Human Gene Mutation Database (HGMD), ClinVar, and the Online Mendelian Inheritance in Man database (OMIM).

The entire mitochondrial DNA was enriched by long-range PCR followed by massively parallel sequencing.

The related primers are listed in Table 1.

Sanger sequencing
Sanger sequencing was used to verify the variations of the candidate genes in the family members. The primers for amplifying the targeted region of candidate genes are also shown in Table 1.

Variant functional assay
The wild-type GJB2 cDNA and GJB2 cDNA with the c.262G>C variant were amplified with the primers shown in Table 1. The HA-tagged wild-type and mutant coding sequences were inserted into the pcDNA3.1(+) vector (Invitrogen, Carlsbad, CA) using the Mut Express® II Fast Mutagenesis kit V2 (Vazyme, Nanjing, China). Human H1299 cells (ATCC, Manassas, VA) were transfected to express the vectors using the jetPRIME Transfection Kit (Polyplus, Illkirch, France) according to the manufacturer’s instructions. After 48 hours of transfection, the cells were collected for immunoblotting and immunohistochemical analysis.

Protein analysis
The online Clustal Omega and Conseq software programs were used to align the amino acid sequences in a variety of species. The Polyphen-2, SIFT and MutationTaster programs were used to predict the variants as “damaging” or “possibly damaging”. The clinical interpretation of genetic variants by the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) guidelines was followed to classify the variants into “benign”, “likely benign”, “uncertain significance”, “likely pathogenic”, and “pathogenic”.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primers for amplifying genomic DNA</td>
<td></td>
</tr>
<tr>
<td>GJB2-F</td>
<td>5’-AGCAACGCCCAGAGTAGAAG-3’</td>
</tr>
<tr>
<td>GJB2-R</td>
<td>5’-AAGATGACCAGGAAGAAGATGCT-3’</td>
</tr>
<tr>
<td>Primers for HA-tagged protein expression vector construction</td>
<td></td>
</tr>
<tr>
<td>WT-F</td>
<td>5’-CTTGTTACCGAGCTCGGATCCATGGATTGGGGCACGCTG-3’</td>
</tr>
<tr>
<td>WT-R</td>
<td>5’-TGCTTGGATATCTGCAGAATTCAACTGGGCAATGCGTTAAACTG-3’</td>
</tr>
<tr>
<td>Mut-F</td>
<td>5’-CCGCTCCTAGTGGCCATGCACGTGG-3’</td>
</tr>
<tr>
<td>Mut-R</td>
<td>5’-GTGGCGTGAGCAGAAGATCAGCTGCA-3’</td>
</tr>
<tr>
<td>Mitochondrial genome DNA sequencing</td>
<td></td>
</tr>
<tr>
<td>mt16426F</td>
<td>5’-CCGCACAAGAGTGCTAATCTCTCCTC-3’</td>
</tr>
<tr>
<td>mt16425R</td>
<td>5’-GATATTGACTTACCGAGAGAGGATGTTG-3’</td>
</tr>
</tbody>
</table>
Immunoblotting and western blot analysis

For immunoblotting analysis, the cells were lysed with RIPA buffer containing protease inhibitor cocktail (Roche Diagnostics GmgH, Mannheim, Germany). The lysate was centrifuged, collected, and boiled in SDS loading buffer. Then, the proteins were separated on 10% SDS-polyacrylamide gels. After the proteins were transferred onto polyvinylidene difluoride membranes (Millipore, USA), the membranes were blocked and incubated with rabbit polyclonal anti-HA antibody (Dilution: 1:1000, Cat No.: 51064-2-AP, Proteintech, Chicago, USA) and the secondary antibodies (Dilution: 1:10000, Cat No.: BA1055, Boster Wuhan, China), and the protein bands were visualized using an HRP chemiluminescent substrate kit (Millipore) and a ChemiDoc XRS+ System (Bio-Rad Company, Berkeley, CA).

For immunohistochemical analysis, the cells were fixed with 4% paraformaldehyde in 0.1 M PBS for 20 min and then rinsed three times in PBS. Then, the coverslips were immersed in cold methanol for 15 min at -20°C. The primary antisera and dilutions were as follows: rabbit anti-HA antibody at 1:100 (Proteintech) for WT/MUT GJB2. After incubation with primary antiserum at 4°C overnight, the cells were rinsed in PBS three times before adding Alexa Fluor 488- and/or Alexa Fluor 594-conjugated secondary antibodies (Dilution: 1:500, Cat. No.: A-11008, Invitrogen). ER was stained with ER-Tracker Red at 1:2000 dilutions (Beyotime, Shanghai, China) for 10 min at room temperature. Preimmune rabbit serum was used as the primary antibody for the negative controls. The images were visualized using a Zeiss Axio Imager Z2 microscope (Carl Zeiss, Jena, Germany).

Results

Clinical characteristics of the patient’s family

The pedigree of the family is shown in Figure 1A. The kinship connection between the proband and parents is confirmed by the exome sequence data\(^2\). The proband had normal physical, biochemical and otoscopic evaluations. No abnormality was found in her cranium by MR examination or in her cochlear, vestibular, and semicircular canals by CT scan. Pure tone audiometry indicated that her left and right hearing thresholds were 78 dB and 87 dB, respectively, with severe hearing loss in both ears (Figure 1B). Since there was no family history of HL and the child’s parents had normal hearing, the affected infant is considered to be a sporadic case of NSHL.

Analyses of variants detected in GJB2 gene in the patient’s exome

The mitochondrial sequencing showed no NSHL-causing variants or large deletions. The exome sequences revealed no known NSHL-causing variants in the family except that the proband had a \textit{de novo} heterogeneous variant c.262G>C in the GJB2 gene (MAF unknown, Table 2 and Figure 2), whereas her parents were wild type. The c.262G>C variant led to a missense variant of p.A88P, which was graded to be “damaging” with a SIFT score of 0.00 and a Polyphen-2 score of 1.00. To examine the effect of the c.262G>C variant on the protein expression level and cellular localization, we transiently transfected the GJB2 c.262G>C mutant into H1299 cells and found that the mutant was expressed weakly and...
<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene annotation</th>
<th>Genetic Location</th>
<th>rs</th>
<th>Proband</th>
<th>Father</th>
<th>Mather</th>
<th>de novo variant</th>
<th>Amino acid variation</th>
<th>Type</th>
<th>MAF (in-house database of Joy Oriental Co.)</th>
<th>Clinical significance defined by ACMG</th>
</tr>
</thead>
<tbody>
<tr>
<td>GJB2</td>
<td>Gap junction protein beta 2</td>
<td>13q11.1-q12 (20763459)</td>
<td>rs1267269489</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.262(G&gt; C)</td>
<td>Missense</td>
<td>p.A88P</td>
<td>0.007 Pathogenic</td>
<td>Pathogenic</td>
</tr>
<tr>
<td>LOC100509263</td>
<td></td>
<td>chr9 (35134638-35183484)</td>
<td></td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.250(G&gt;T)</td>
<td>Missense</td>
<td>p.R84Fs*18</td>
<td>0.007 Pathogenic</td>
<td>Pathogenic</td>
</tr>
<tr>
<td>MUC4</td>
<td>Mucin 4, cell surface associated</td>
<td>3q29 (195506650)</td>
<td>rs771640527</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.11801(C&gt;T)</td>
<td>Missense</td>
<td>p.A3934V</td>
<td>0.016 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>MUC4</td>
<td>Mucin 4, cell surface associated</td>
<td>3q29 (19550667-19550667.76)</td>
<td>rs1267269489</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.11777(C&gt;T)</td>
<td>Missense</td>
<td>p.A3934V</td>
<td>0.016 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>PHGR1</td>
<td>Proline, histidine and glycine rich 1</td>
<td>15q15.1 (40648396-40648398)</td>
<td>rs755350345</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.141(G&gt;T)</td>
<td>Missense</td>
<td>p.A47G</td>
<td>0.000123 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>AGAP4</td>
<td>ArFGAP with GTPase domain, ankyrin repeat and PH domain 4</td>
<td>10q11.12 (20632037)</td>
<td>rs87947523</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1318(G&gt;T)</td>
<td>Missense</td>
<td>p.K440E</td>
<td>0.000259 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>MAPKBP2</td>
<td>Mitogen-activated protein kinase 8 interacting protein 2</td>
<td>1p21.33 (51048346)</td>
<td>rs1267269489</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1816(G&gt;A)</td>
<td>Missense</td>
<td>p.E606K</td>
<td>0.0000456 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>ADAMTS14</td>
<td>Thrombospondin repeat-containing protein 1</td>
<td>1p21.33 (15029350)</td>
<td>rs1267269489</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1750(G&gt;T)</td>
<td>Missense</td>
<td>p.K440E</td>
<td>0.000259 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>MAPKBP1</td>
<td>Mitogen-activated protein kinase 8 interacting protein 1</td>
<td>1p21.33 (51048346)</td>
<td>rs1267269489</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1750(G&gt;T)</td>
<td>Missense</td>
<td>p.K440E</td>
<td>0.000259 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>FRG1</td>
<td>Fasicoscalopulumeral muscular dystrophy region gene-1</td>
<td>4q35 (190880984)</td>
<td>rs62345304</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.538(G&gt;T)</td>
<td>Missense</td>
<td>p.A176G</td>
<td>0.012 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td>ZNF880</td>
<td>Zinc finger protein LOC40072</td>
<td>1q21.33 (52888874-52888875)</td>
<td>rs87947523</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1242(G&gt;T)</td>
<td>Missense</td>
<td>p.A414G</td>
<td>0.006 Pathogenic</td>
<td></td>
</tr>
<tr>
<td>RAD21</td>
<td>protein involved in DNA double-strand break repair, sister chromatid cohesion 1</td>
<td>1q21.33 (52888874-52888875)</td>
<td>rs1267269489</td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1162(G&gt;T)</td>
<td>Missense</td>
<td>p.A176G</td>
<td>0.012 Likely pathogenic</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Heterozygote</td>
<td>Wild type</td>
<td>Wild type</td>
<td>c.1162(G&gt;T)</td>
<td>Missense</td>
<td>p.A176G</td>
<td>0.012 Likely pathogenic</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. All de novo variants which were predicted to be pathogenic or likely pathogenic in the proband’s whole exome.**
<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene annotation</th>
<th>Genetic Location</th>
<th>rs</th>
<th>Proband</th>
<th>Father</th>
<th>Mather</th>
<th>de novo variant</th>
<th>Amino acid variation</th>
<th>Type</th>
<th>MAF (in-house database of Joy Oriental Co.)</th>
<th>Clinical significance defined by ACMG</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPN1MW2</td>
<td>Opsin 1, medium wave sensitive 2</td>
<td>chrX</td>
<td>(153485284-153498649)</td>
<td>Loss of heterozygosity</td>
<td>Wild type</td>
<td>Wild type</td>
<td>loss1(EXON:1-6)(all)</td>
<td>Exon deletion</td>
<td>0.047</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>ARHGEF5</td>
<td>Rho guanine nucleotide exchange factor 5</td>
<td>chr7</td>
<td>(144059762-144072768)</td>
<td>Loss of heterozygosity</td>
<td>Wild type</td>
<td>Wild type</td>
<td>loss1(EXON:2-12)</td>
<td>Exon deletion</td>
<td>0.05</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>chr3</td>
<td></td>
<td>chr3</td>
<td>(15621416-15631119)</td>
<td>Loss of heterozygosity</td>
<td>Wild type</td>
<td>Wild type</td>
<td>loss1</td>
<td>Large CNV deletion</td>
<td>0.0064</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>chr10</td>
<td></td>
<td>chr10</td>
<td>(49218794-49237210)</td>
<td>Loss of heterozygosity</td>
<td>Wild type</td>
<td>Wild type</td>
<td>loss1</td>
<td>Large CNV deletion</td>
<td>0.025</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>chr19</td>
<td></td>
<td>chr19</td>
<td>(7981506-7985434)</td>
<td>Loss of heterozygosity</td>
<td>Wild type</td>
<td>Wild type</td>
<td>loss1</td>
<td>Large CNV deletion</td>
<td>0.0053</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>chr22</td>
<td></td>
<td>chr22</td>
<td>(50320902-50354549)</td>
<td>Loss of heterozygosity</td>
<td>Wild type</td>
<td>Wild type</td>
<td>loss1</td>
<td>Large CNV deletion</td>
<td>0.019</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>chr16</td>
<td></td>
<td>chr16</td>
<td>(2827007-28726011)</td>
<td>Single repeat</td>
<td>Wild type</td>
<td>Wild type</td>
<td>gain1</td>
<td>Large CNV duplication</td>
<td>0.039</td>
<td>Pathogenic</td>
<td></td>
</tr>
<tr>
<td>LOC81691</td>
<td></td>
<td>chr16</td>
<td>(2083679-20839859)</td>
<td>Single repeat</td>
<td>Wild type</td>
<td>Wild type</td>
<td>gain1(EXON:9-11)</td>
<td>Exon duplication</td>
<td>0.0056</td>
<td>Pathogenic</td>
<td></td>
</tr>
</tbody>
</table>
displayed a punctate distribution in the cytoplasm and cytomembrane. In contrast, wild-type GJB2 was expressed robustly and was distributed mainly in the cytomembrane (Figures 3A & B). This result confirmed that the GJB2 p.A88P mutant may fail to locate into the cell membrane and subsequently reduce the formation of gap junctions in quantity.

Because heterozygous c.262G>C missense variants were previously found in both patients and healthy carriers in the clinic2,4, we rechecked the exome sequences of the family to search for any other possible genetic causes of NSHL. We failed to find any other NSHL-causing variants, but noticed that the mother was a heterozygote of GJB2 c.79G>A (p.V27I), c.341A>G (p.E114G), the father was wild type, and the affected infant was a heterozygote of c.79G>A (p.V27I) and c.341A>G (p.E114G). As mentioned before, although no significant loss of function has been detected when VG and IG gap junctions coexist with the VE and IE types, the VG and IG types have displayed a moderate deficit in biochemical coupling and reduced channel activity \textit{in vitro}11. Hence, we deduced that the \textit{de novo} p.A88P mutants in the infant dislocated from the cell membrane, exacerbating GJB2 loss-of-function in the context of the p.[V27I; E114G], whereas the wild-type p.A88 in her mother could compensate for the loss, thus the infant’s \textit{compound} heterozygosity at p.A88P and p.[V27I; E114G] was affected while her mother is an asymptomatic carrier of p.[V27I; E114G]. This result indicates that the multiple genetic variants in \textit{GJB2} could influence protein function additively.

Analyses of \textit{de novo} variants detected in the patient’s exome

As we noticed that the c.262G pathogenic variant was \textit{de novo}, we examined the \textit{de novo} variants in the patient’s exome. It showed that there were approximately 47,000 variants, of which 143 variants were \textit{de novo} (0.34%, 143/47,000). Among these 47,000 variants, 74 (0.016%, 74/47,000) were predicted to be pathogenic or likely pathogenic. Remarkably, 23 \textit{de novo} variants were predicted to be pathogenic or likely pathogenic, including 21 heterozygous and two homozygous variants. The \textit{de novo} adverse variants accounted for approximately one-third (23/74) of all pathogenic or likely pathogenic variants (Table 2). Compared with the frequency of \textit{de novo} variants in total being only 0.34%, the frequency of \textit{de novo} adverse variants in all \textit{de novo} variants reached above 16% (23/143) which is surprisingly high. The 23 \textit{de novo} adverse variants were distributed in 19 different genetic areas, including 12 known genetic regions, two unclassified gene zones (LOC100509263 and LOC81691) and five other chromosome domains without defined roles (Table 2). Except for the \textit{de novo} GJB2 c.262G>C variant, the other adverse variants have not yet been reported to be NSHL-causing.

Notably, eight copy number variations (CNVs) - nearly half of the infant’s 17 total adverse CNVs - were \textit{de novo}. The \textit{de novo} pathogenic CNVs accounted for 42% (8/19) of the total \textit{de novo} adverse mutated genes (Table 2); however, the minor allele frequencies of these CNVs were below 0.05 in the Human Gene Mutation Database and our in-house database (Joy Oriental Co., Table 2). Of these CNVs, six lacked the relevant information about their function, except the exon deletion in Rho
Guanine Nucleotide Exchange Factor 5 (ARHGEF5, exon 2-12, 13,006 bp) and Opsin 1, Medium Wave Sensitive 2 (OPN1MW2, exon 1-6, 13,365 bp). The ARHGEF5 and OPN1MW2 gene products are crucial proteins that transduce external environmental cues into cellular signals across the cell membrane. Indeed, most CNVs do not encode important genes related to development and are thought to be subjected to adaptation to different environments\(^\text{14}\). Here, these recurrent de novo pathogenic CNVs in the patient remind us about the environmental influence on genetic components.

In total, four missense, four frameshift, three noncoding, two splice-site, one in-frame deletion and one stop gain variant which were predicted to be pathogenic/likely pathogenic, were de novo (Table 2). Interestingly two heterozygous pathogenic variants in mitogen-activated protein kinase 8 interacting kinases 1 and 2 (MNK1 and MNK2) were de novo: one MNK1 noncoding variant c.679 C>T in chromosome 11 (exon 12, MAF=0.000046) and the other MNK2 missense variant c.1816 G>A in chromosome 22 (exon 6, MAF unknown). Both MNK1 and MNK2 are serine/threonine kinases from the Ca\(^{2+}\)/calmodulin-dependent kinase family and take part in initiating mRNA translation in response to MAPK signaling, accordingly playing important roles concerning environmental stress and cytokines\(^\text{15}\). Also, four de novo pathogenic variants, including one homozygous missense variant c.11801C>T (MAF=0.016), accumulated in the cell surface-associated Mucin 4 gene (MUC4). Mucins are integral membrane glycoproteins on the cell surface. As the major constituents of mucus, mucins protect epithelial cells from outward stimuli. Additionally, two de novo heterozygous pathogenic variants c.1162-4(IVS9) insG (MAF unknown) and c.1162-5(IVS9) A>T (MAF=0.000008) were detected in the Rad1-like checkpoint DNA exonuclease gene (RAD21). The RAD21 gene encodes the major cohesion subunit, known as the component of a heterotrimeric cell cycle checkpoint complex, regulating the segregation of sister chromatids in cell cycle progression and connecting inducible gene expression in response to diverse stimuli\(^\text{16}\). It is assumed that the proband’s genes with de novo pathogenic variants, including the disease-causing GJB2 c.262G>C

Figure 3. Expression of the p.A88P Cx26 mutants in cells. The immunofluorescence staining showed the wild-type GJB2 was expressed robustly and distributed mainly in the cytomembrane, while the p.A88P mutant was expressed weakly and displayed a punctate distribution in the cytoplasm and cytomembrane (A, arrow); the ER tracker red demonstrated the cytoplasm (A). The immunoblotting confirmed the wild-type GJB2 largely localized in the cytomembrane, while the p.A88P mutants co-localized in the cytoplasm and cytomembrane (B).
of congenital nonsyndromic hearing loss. https://doi.org/10.17605/OSF.IO/DS7TW2

Open Science Framework: Whole-exome sequencing of de novo genetic variants in a Chinese family with a sporadic case of NSHL. 

Our study showed that the c.262G>T variant was expressed weakly and failed to regularly locate in the cell membrane, consequently reducing the formation of cell gap junctions. The GJB2 p.[V27I; E114G] variant may also additively impair GJB2 function, as the channel activities of homozygous p.[V27I; E114G] CX26 gap junctions has been previously shown to be reduced. Thus, it seems reasonable that carriers of the simple heterozygous c.262G>C could be asymptomatic, while carriers of c.262G>C in compound heterozygosity along with any other deafness-related variants such as c.235delC, p.V27I, etc. could experience HL, like in our study and a previous NSHL case. Hence, the pathogenic effects of these GJB2 variants could be additive.

It should be noted that the penetrance of the GJB2 c.262G>T seemed to be undetermined in the two previous NSHL cases: the heterozygous p.A88S in the Austrian patient with NSHL and his asymptomatic mother carrier; and the heterozygous p.A88V in the Japanese girl with severe keratitis-ichthyosis-deafness syndrome and her healthy parents. Both studies reported no other GJB2 variants except the heterozygous c.262G. In fact, only the candidate GJB2 genetic region was sequenced in their studies, so any other genetic disease-causing variants in the patients' genome are still unknown. Therefore, for an accurate variant interpretation and improved clinical care, we propose that more comprehensive details about the related variants, such as variant domain, effect, and reciprocal interaction, should be investigated.

In the current study, it is intriguing that there was a very high frequency of de novo adverse variants in the proband's exome and that most de novo variants are in the genetic regions characterized as environment-sensitive. Several notable results were found. First, the gene in which the de novo NSHL-causing GJB2 c.262G>C is located is immediately responsive to the surrounding changes. The gene product GJB2 is essential for gap channels, which allows the exchange of small substances including nutrients, metabolites, ions and second messengers, and regulates signaling pathways in intracellular communication. Second, variants in MNK1 and MNK2 - two downstream MAPK signaling effectors located in different chromosomes - were also de novo. Both MNKs are involved in guiding cellular responses to a diverse array of stimuli, such as mitogens, osmotic stress, heat shock and proinflammation. Third, the de novo pathogenic variants were aggregated in the MUC4 gene region. Mucins are integral membrane glycoproteins on the cell surface, covering epithelial surfaces such as those in the trachea, colon and cervix, and exert anti-adhesive effects on cell-cell and cell-extracellular matrix interactions. Fourth, there were two de novo pathogenic variants in the RAD21 gene. RAD21 participates in repairing DNA double-strand breaks and chromatin cohesion and can be affected by various agents, including ionizing radiation, topoisomerase inhibitors, cycloheximide, proteasome inhibitors, cytokines agents and inflammatory stimuli. Finally, there was a very high incidence of de novo pathogenic CNVs which have quite low MAFs. CNVs are often enriched in genes related to sensory perception of the external environment (e.g., smell, sight, and taste), neurodevelopmental processes, and response to chemical stimuli, immunity and other processes. Therefore, we wonder whether there might have been any direct external stimuli to trigger the fetal adaptive responses for the occurrence of such a considerable amount of de novo pathogenic variants, which thus led to the disease.

It should be noted that the sporadic congenital NSHL in the family that we studied here was rare and limited; more research in similar birth defect cases is needed to confirm the role of environmental factors in transformation of genetic variants in the fetus/offspring.

In summary, by whole-exome sequencing, we examined overall genetic variants, especially the compound heterozygous GJB2 variants and the high frequency of de novo pathogenic variants in a Chinese family with a rare sporadic case of NSHL. Though the reported case here is limited, we think the detailed full picture of genetic variants could improve our interpretation of the HL-associated genetic variants. In order to further advance our understanding of disease biology in birth defects, further research on environmental causes for de novo pathogenic variants may be needed.

Data availability

Underlying data

This project contains the following underlying data:

- The immunoblotting of the p.A88P Cx26 mutants in cells. (mut*.tif)
- The immunoblotting of the wild type Cx26 in cells. (WT*.tif)
- The data comparison of the exome DNA sequences of the family members. (*.xlsx)

NCBI Gene: Exome sequencing of a Chinese family with a sporadic congenital NSHL. Accession number PRJNA688744.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Ethical approval
The data were de-identified as sufficiently as possible and data sharing was approved by the Research Ethics Committee of Sichuan Provincial People’s Hospital, School of Medicine, UESTC (approval number 2019-065).

Acknowledgments
We sincerely thank the study participants, without their permission, this work would not be possible.

References

Open Peer Review

Current Peer Review Status: ✔️ ?

Version 1

Reviewer Report 15 April 2021

https://doi.org/10.5256/f1000research.30672.r82018

© 2021 Tan Y. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Yuande Tan
Institute of Personalized Medicine, PENN State University, Hershey College of Medicine, Hershey, PA, USA

This paper displays a variant profile of a Chinese family with a case of congenital nonsyndromic hearing loss using the whole-exome sequencing data. Authors attempted to utilize heterozygote of a de nova variant (c.262G>C) in gene GJB2 to explain case of nonsyndromic hearing loss occurring in an offspring individual. However, I have several major concerns to be addressed:

1. The results show among the total~47,000 variants, 143 were de novo occurring in the patient. According to definition of de nova from Wikipedia, a de nova variant is a variant occurring in offspring individuals, not in parents, that is, de nova variants results from mutation occurring in sperm or germ cells. However, 143/47000(0.304%) is too high. Since there is one offspring individual in this study, so, it is very difficult to determine whether these de nova variants are due to noise or true mutation. In addition, 47,000 variants detected in the patient are also too many and most of these so-called variants may be noisy or result from sequencing error. Authors should discuss this issue.

2. If it is known that nonsyndromic hearing loss is due to mutation in gene GJB2, it is unnecessary to do the whole exome sequencing, that is to say, the whole variant profiles of parents and daughter (proband) do not provide useful information for interpreting the nonsyndromic hearing loss. The de nova variants listed in Table 2 also do not make sense because they are not used to explain occurrence of the nonsyndromic hearing loss.

3. How heterozygous c.262G>C missense variants works for occurrence of the nonsyndromic hearing loss is not clear. Since the farther and mother have homozygote c.262GG and have no nonsyndromic hearing loss and heterozygous c.262G>C variant was also found in healthy individuals, the proband with heterozygous c.262G>C variant at this position in gene GJB2 should be normal hearing. However, the proband had heterozygous variants c.79G>A and c.341A>G in gene GJB2 derived from her mother but her mother is normal hearing. Therefore, there may be interaction (dominant epistasis or receive epistasis) between c.262G>C and c.79G>A or c.341A>G. Author should discuss this possible heredity mechanism.
Minor points:
1. On page 3, simple, monogenetic should be simple and monogenetic.
2. On page 3, “another was the c.263C>A (p.A88E) variant” should be “another case was the c.263C>A (p.A88E) variant”.
3. On page 3, “c.235delC; and another was the c.263C>T (p.A88V) variant” should be “c.235delC and the other variant was the c.263C>T (p.A88V)”.
4. On page 3, “several clinical studies have found the p.[V27I; E114G] haplotype to be a risk factor” should be “several clinical studies have found that the p.[V27I; E114G] haplotype is a risk factor”.
5. On page 3, “Written informed consent was obtained” should be “An informed consent was obtained”.
6. On page 4, what is “size distribution and concentration”?
7. On page 5, “the affected infant is considered to be a sporadic case of NSHL”. I don't think it is a sporadic case because you just found one case and you cannot determine it is a sporadic case or a pedigree case or a family case.
8. On page 5, what is “large deletions”?
9. On page 5, “The exome sequences revealed no known NSHL-causing variants” should be “The exome sequences revealed unknown NSHL-causing variants”.
10. On page 5, “whereas her parents were wild type” should be “whereas her parents were normal type”. I have never seen “wild type” in human.
11. Legend Figure 1 is unclear. It does not state what is Figure 1A.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Partly
Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Author Response 21 May 2021**

**Sonia Liao**, Sichuan Provincial People’s Hospital, University of Electronic Science and Technology of China, Chengdu, China

1. The results show among the total ~47,000 variants, 143 were de novo occurring in the patient. According to definition of de nova from Wikipedia, a de nova variant is a variant occurring in offspring individuals, not in parents, that is, de nova variants results from mutation occurring in sperm or germ cells. However, 143/47000 (0.304%) is too high. Since there is one offspring individual in this study, so, it is very difficult to determine whether these de nova variants are due to noise or true mutation. In addition, 47,000 variants detected in the patient are also too many and most of these so-called variants may be noisy or result from sequencing error. Authors should discuss this issue.

   Our answer: Thanks. We discussed that “sequencing errors remain one of the main obstacles in the identification of causative genetic variants and/or mutations” in the new version, please see the 4th paragraph of the discussion section.

   We may need to mention here, the genetic counseling center in our hospital, one of the earliest clinic genetic research groups in China, offers regular exome sequencing to identify and study disease-related genetic mutations for patients. For the quality of our exome sequencing in the current work, please refer to the chapter of Whole-exome and mitochondrial DNA sequencing in the materials and methods. And also, all sequence information is open at NCBI Gene: Exome sequencing of a Chinese family with a sporadic congenital NSHL. Accession number PRJNA688744.

2. If it is known that nonsyndromic hearing loss is due to mutation in gene GJB2, it is unnecessary to do the whole exome sequencing, that is to say, the whole variant profiles of parents and daughter (proband) do not provide useful information for interpreting the nonsyndromic hearing loss. The de nova variants listed in Table 2 also do not make sense because they are not used to explain occurrence of the nonsyndromic hearing loss.

   Our answer: Thanks. As we mentioned in the 1st paragraph of the introduction: “GJB 2 have been shown to be the leading genetic cause of NSHL (OMIM: 121011). GJB2-related autosomal recessive deafness can explain approximately 50% of congenital autosomal recessive deafness.” In fact, pathogenic variants were found in 49 genes according to GeneReviews of NSHL (https://www.ncbi.nlm.nih.gov/books/NBK1272/). In the current case of an NSHL newborn, we think that it is necessary to carry out the whole-exome sequencing to search for the genetic causes of the disease.

   In the new version, we added that “the pathogenic variants of non-syndromic sensorineural hearing loss (NSHL) (OMIM: 121011) were found in 49 genes (https://www.ncbi.nlm.nih.gov/books/NBK1272/)” in the 1st paragraph of the introduction,
in order to emphasize that GJB2 variants are not the only genetic causes of NSHL, and also, we pointed out that “We carried out whole-exome sequencing, assessed the cytological/clinical characteristics of the genetic variants, specifically in the GJB2 genetic variants, and evaluated the possible cause of de novo pathogenic variants in the patient’s exome.” in the last paragraph of the introduction.

3. How heterozygous c.262G>C missense variants works for occurrence of the nonsyndromic hearing loss is not clear. Since the farther and mother have homozygote c.262GG and have no nonsyndromic hearing loss and heterozygous c.262G>C variant was also found in healthy individuals, the proband with heterozygous c.262G>C variant at this position in gene GJB2 should be normal hearing. However, the proband had heterozygous variants c.79G>A and c.341A>G in gene GJB2 derived from her mother but her mother is normal hearing. Therefore, there may be interaction (dominant epistasis or receive epistasis) between c.262G>C and c.79G>A or c.341A>G. Author should discuss this possible heredity mechanism.

Our answer: Thanks. According to the definition of National Human Genome Research Institute, epistasis refers to a circumstance where the expression of one gene is affected by the expression of one or more independently inherited genes (https://www.genome.gov/genetics-glossary/Epistasis). In the current study, both c.262G>C and the haplotype of c.79G>A and c.341A>G are in the exon2 of GJB2 gene, and encode amino acids which were tightly connected in this protein (for details please see the figure2). We do not think that it should be necessary to discuss about the independent heredity mechanism of these variants in the same gene. Additionally, in our study, we noticed the insufficient GJB2 function because of these variants, while we have not seen any report about the epistasis in this exact situation.

Minor points:
1. On page 3, simple, monogenetic should be simple and monogenetic.
   Our answer: According to the suggestion, we modified the text.

2. On page 3, “another was the c.263C>A (p.A88E) variant” should be “another case was the c.263C>A (p.A88E) variant”.
   Our answer: The “case” was added in the context of the whole sentence.

3. On page 3, “c.235delC5 ; and another was the c.263C>T (p.A88V) variant” should be “c.235delC5 and the other variant was the c.263C>T (p.A88V)”.
   Our answer: Thanks. This sentence might be confusing due to the typesetting errors? “c.235delC5” was not correct here. Here, we listed that researchers have reported 3 p.A88 coding variants at the 263rd nucleotide: (1) c.263C>G (p.A88G) [reference4], (2) c.263C>A (p.A88E) in compound heterozygosity with c.235delC [reference5], (3) c.263C>T (p.A88V) [reference5]. Since there were totally 3 reported cases, the sentence we presented them like: “there are three pens, one is red, another is black, and another is green”

4. On page 3, “several clinical studies have found the p.[V27I; E114G] haplotype to be a risk factor” should be “several clinical studies have found that the p.[V27I; E114G] haplotype is a risk factor”.

5. On page 3, “our data” should be “our findings”.
   Our answer: The word “data” was replaced with “findings”.
5. On page 3, “Written informed consent was obtained” should be “An informed consent was obtained”.
   Our answer: Thanks. We found that “written informed consent” without any prefix (“a” or “the”) was used in many research articles, and listed below 2 of these examples from the Pubmed for your reference.
   ▪ “Consent for publication Written informed consent was obtained from the patient for the publication of this report and any accompanying images.”
   ▪ “Informed consent Written informed consent was obtained from the patient for publication of this case report.”
   And also, there are the same sentence examples at the websites of English sources like:
   https://www.lawinsider.com/dictionary/informed-written-consent and
   https://ludwig.guru/s/written+informed+consent+was+obtained+from+all+subjects.

6. On page 4, what is “size distribution and concentration”?
   Our answer: Here the “size distribution and concentration” belong to the qPCR of the enriched DNA fragments; we changed it into “the size distribution and concentration of these DNA fragments were examined...” Please see the revised sentence in the new version.

7. On page 5, “the affected infant is considered to be a sporadic case of NSHL”. I don’t think it is a sporadic case because you just found one case and you cannot determine it is a sporadic case or a pedigree case or a family case.
   Our answer: Commonly for patients with birth defect, we investigated their family history for any inheritance diseases; if we got no report of such case in their family, we would consider that the new case could be isolated and sporadic. We think that, to an individual family with birth defect, this kind of considerations and the searching for genetic causes of the disease could be useful, especially when these families want to prevent and intervene more disease happening. We believe that current case report about the de novo pathogenic variants could provide new and useful information that lead to further and vital research.

On page 5, what is “large deletions”?
   Our answer: According to the National Human Genome Research Institute: “deletion can be small, involving a single missing DNA base pair, or large, involving a piece of a chromosome”. Large deletions in genomic DNA have been reported to associate with many diseases. For an example:
In our study, because large deletions in mitochondrial DNA have been reported to associate with hearing loss [some related references were listed below], we stated our sequencing result as that “the mitochondrial sequencing showed no NSHL-causing variants or large deletions.”


8. On page 5, “The exome sequences revealed no known NSHL-causing variants” should be “The exome sequences revealed unknown NSHL-causing variants”. Our answer: Here we want to say that “The exome sequences did not revealed any known NSHL-causing variants”, NOT about any unknown NSHL-causing variants.

9. On page 5, “whereas her parents were wild type” should be “whereas her parents were normal type”. I have never seen “wild type” in human. Our answer: Thanks. We searched “wild type in human” in the Pubmed and listed some results here for your reference:


10. Legend Figure 1 is unclear. It does not state what is Figure 1A. Our answer: Thanks. We made it clear now that figure1A shows the pedigree of the family in the new version.

**Competing Interests:** None.
Department of Genetics, School of Bioscience and Technology, Chengdu Medical College, Chengdu, China

The authors report an interesting observation describing a patient with non-syndromic hearing loss caused by a de novo heterozygous variant of GJB2 c.262G>C. The authors identified the variant by whole exome sequencing and provided biological evidence of the variant by in vitro analysis.

Here, several concerns were raised:

The number of de novo variants (23) seemed very high. De novo genetic variants are recently suggested to be important in human disease, one investigation of severe childhood cardiomyopathy identified a de novo variant in 46% of children with a pathogenic variant (Vasilescu et al., Genetics basis of severe childhood onset cardiomyopathies. Journal of the American College of Cardiology 2018), yet the overall frequency of de novo variants is largely unknown. Should the authors discuss more about the issue? It should be noted that one case here is limited.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Medical Genetics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 21 May 2021

**Sonia Liao**, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China
Thanks a lot. We presented the suggested reference in the new version, please see the 4th paragraph of the discussion section.

**Competing Interests:** No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)