Whole exome sequencing identified mutations causing hearing loss in five consanguineous Pakistan families

CURRENT STATUS: UNDER REVIEW

**BMC Medical Genetics  **

Yingjie Zhou  
Second Hospital of Hebei Medical University

Muhammad Tariq  
National Institute for Biotechnology and Genetic Engineering

Sijie He  
hesijie@genomics.cn  
BGI  
*Corresponding Author*  
*ORCiD: 0000-0002-4418-1785*

Uzma Abdullah  
NIBGE

Jianguo Zhang  
BGI

Shahid Mahmood Baig  
NIBGE

**DOI:** 10.21203/rs.2.19325/v1

**SUBJECT AREAS**  
*Medical Genetics*

**KEYWORDS**  
*hearing loss, whole exome sequencing, consanguineous pedigrees, clinical detection*
Abstract

**Background:** Hearing loss is the most common sensory defect that affects over 6% of the population worldwide. About 50%-60% of hearing loss patients are attributed to genetic causes. Currently more than 100 genes have been reported to cause non-syndromic hearing loss. It's possible and efficient to screen all potential disease-causing genes for hereditary hearing loss by whole exome sequencing (WES).

**Methods:** We collected 5 consanguineous pedigrees with hearing loss from Pakistan and applied WES on selected patients for each pedigree, followed by bioinformatics analysis and Sanger validation to identify the causing genes for them.

**Results:** Variants in 7 genes were identified and validated in these pedigrees. We identified single candidate for 3 pedigrees, which were *GIPC3* (c.937T>C), *LOXHD1* (c.2935G>A) and *TMPRSS3* (c.941T>C). And the remaining 2 pedigrees each contained two candidates, which were *TECTA* (c.4045G>A) and *MYO15A* (c.3310G>T and c.1705G>C) for one pedigree and *DFNB59* (c.494G>A) and *TRIOBP* (c.1952C>T) for the other pedigree. The candidates were validated in all available samples by Sanger sequencing.

**Conclusion:** The candidate variants in hearing loss genes were validated to be co-segregated in the pedigrees, which may indicate the reasons for such patients. We also suggested that WES may be suitable strategy for hearing loss gene screening in clinical detection.

**Background**

Hearing loss is the most common sensory defect that affects ~1/500 newborns [1] and 466 million people worldwide (https://www.who.int/pbd/deafness/estimates/en/).
About 50%~60% of hearing loss patients are attributed to genetic causes [1, 2]. Hereditary hearing loss is a genetically heterogeneous disorder [3] that can be divided into syndromic hearing loss and non-syndromic hearing loss, among which non-syndromic hearing loss is the predominant type with a proportion of ~80% [4]. Currently more than 100 genes have been reported to cause non-syndromic hearing loss (https://hereditaryhearingloss.org/), and the total number of genes related to hearing loss is expected to be several hundreds.

There are mature gene panels for hearing loss detection, and the genes involved range from 4 to more than 100. However, except for several genes, such as GJB2 [5-7] or SLC26A4 [8-10], most causing genes contribute a little fraction for the disorder. So, we may not obtain a satisfied result by gene panel screening for many cases in clinical detection. As the whole exome sequencing (WES) technology rapidly developed and its cost becomes cheaper, it’s possible and efficient to screen all potential disease-causing genes for hereditary hearing loss by WES [11, 12].

Recessive inheritance hearing loss is worth studying because such patients usually have normal parents, which makes the disorder seem to be “sudden strike”, and this situation is more difficult to prevent. Consanguineous pedigree is suitable natural model to study recessive disorders [13]. In Pakistan, there are numerous consanguineous pedigrees because of their custom, which may bring more opportunity to study and recognize such disorders [14, 15].

In this study, we collected 5 consanguineous pedigrees with hearing loss from Pakistan and applied WES to identify the causing genes for them. We identified several variants in hearing loss genes that co-segregated in the pedigrees, which may indicate the reasons for such patients.
Methods

Participants and clinical diagnosis

In the present study we collected 5 consanguineous pedigrees containing 22 patients of hearing loss from rural areas in Pakistan. All the patients showed different degrees of hearing loss. The pedigree showed clear autosomal recessive pattern of inheritance (Fig1). The study was approved by the ethical committee of National institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan and all participants provided written informed consent.

DNA extraction and Whole exome sequencing

According to the manufacturer's instructions, genomic DNA was isolated from peripheral blood leukocytes of all participants using DNA QIAamp mini kit (Qiagen, Hilden, Germany). One patient from each pedigree was selected and performed WES. Exons were captured using BGI-Exome kit V4 and sequenced by BGI-seq 500 with 100bp paired-end.

Bioinformatics analysis

Low-quality reads were removed by SOAPnuke [16], then the reads were mapped to the human genome reference (UCSCGRCh37/hg19) by Burrows–Wheeler Aligner (BWA-MEM, version 0.7.10) [17]. Variants were called using the Genome Analysis Tool Kit (GATK, version 3.3) [18]. Variant Effect Predictor (VEP) [19] was used to annotate and classify all the variants. After that all the variants were filtered based on their frequency in public databases (e.g., 1000 genome project, exome sequencing project and ExAC) and our in-house databases, and the variants with MAF<0.005 were retained. Then homozygous variants and compound heterozygous variants were selected because all the pedigrees showed clear autosomal recessive
pattern of inheritance. Finally, we applied several variants prediction tools including SIFT [20] (http://provean.jcvi.org/), PolyPhen2 [21] (http://genetics.bwh.harvard.edu/pph2/), MutationTaster [22] (http://www.mutationtaster.org/) and CADD [23] to predict the functional impact of candidate variants [24].

**Sanger validation**

DNA from all available samples in the five pedigrees was Sanger sequenced to validate the variants and confirm their co-segregation in pedigree. Forward and reverse primers were designed by Primer3. After PCR amplification, the purified product was sequenced on ABI 3730XL DNA Analyzer.

**Results**

**Clinical features**

All the patients showed different degrees of hearing loss. In the first pedigree (HL1), all patients showed severe deafness, and one patient (III1) was selected to perform WES. In the second pedigree (HL2), all patients showed congenital profound deafness and mute, and one patient (IV1) was selected to perform WES. In the third pedigree (HL3), all patients showed moderate deafness, their hearing loss started after seizures, and one patient (V1) was selected to perform WES. In the fourth pedigree (HL4), all patients showed congenital profound deafness, and one patient (IV1) was selected to perform WES. In the fifth pedigree (HL5), all patients showed moderate deafness, and one patient (III5) was selected to perform WES.

**Genetic analysis**

WES was applied on the selected patients. An average depth of target region was 146X with a coverage of 99.85%, and the coverage of target region that sequenced
at least 10 times (depth >= 10 X) was 98.20% (Table 1). For each individual, more
than ten thousand variants that may influence protein were identified. After the
frequency filtration (MAF<0.005), about 15~32 exon variants were retained. Further
inheritance model filtration kept 1~6 candidates left (Table 1).

We identified a stop-codon lost homozygous variant GIPC3: c.937T>C from the
patient in HL1 pedigree, and the variant prediction tools provided a benign
prediction. For the patient from HL2 pedigree, we identified 4 variants in 2 genes.
However, one gene was reported to cause autosomal dominant hearing loss so we
first analyzed the other one. And then a homozygous variant LOXHD1: c.2935G>A
was regarded as candidate for this pedigree. The variant prediction tools provided a
damaging prediction. For the patient from HL3 pedigree, 6 variants in 5 genes were
identified at first, further analysis indicated that only 2 genes may cause autosomal
recessive deafness. So, a homozygous variant TECTA: c.4045G>A and two
compound heterozygous variants c.3310G>T and c.1705G>C in MYO15A were
regarded as candidates. The variant prediction tools provided benign prediction for
the first two variants and damaging prediction for the last variant. For the patient
from HL4 pedigree, 3 variants in 3 genes were identified, and two of them may
cause autosomal recessive deafness. The two homozygous variants were DFN59:
c.494G>A and TRIOBP:c.1952C>T. The variant prediction tools provided damaging
prediction for the first variant and benign prediction for the last variant. For the
patient from HL5 pedigree, only one homozygous candidate was identified,
TMPRSS3: c.941T>C. And the variant prediction tools provided a damaging
prediction. All these variants were not reported to cause hearing loss before. The
detailed information was listed in Table 2.

In summary, we identified one most likely causing variant for HL1, HL2 and HL5
pedigrees and two most likely causing candidates for HL3 and HL4 pedigrees.

**Sanger validation**

To validate co-segregation in pedigree, we applied Sanger sequencing on all available samples. In total, 6 samples (II1~2 and III1~4) were sequenced for HL1 pedigree, 6 samples (III1~2 and IV1~4) were sequenced for HL2 pedigree, 4 samples (IV1 and V1~3) were sequenced for HL3 pedigree, and 21 samples (I1, II1~11 and III1~9) were sequenced for HL5 pedigree. For HL4 pedigree, the samples collected at the first time were degraded and we failed to collect additional samples. All the variants selected to perform Sanger sequencing were co-segregated in pedigree (Fig 2).

**Discussion**

In this study, we identified several variants in genes reported to cause hearing loss that co-segregated in the pedigree. For HL1 pedigree, the variant in *GIPC3* gene may cause the disorder. *GIPC3* encodes a 312-residue protein that contains three predicted low complexity regions and a central conserved PDZ domain [25]. The PDZ domain of *GIPC3* is required for the survival of spiral ganglion and hair cells in the mouse ears. This gene was reported to cause autosomal recessive deafness 15 and non-syndromic genetic deafness. Currently 11 pathogenic variants were reported in this gene from clinvar.

For HL2 pedigree, the variant in *LOXHD1* gene was the likely causative one. *LOXHD1* encodes a highly conserved stereociliary protein consisting of 15 polycystin-1/lipoxygenase/alpha-toxin (PLAT) domains, which facilitates proteins interacting with the plasma membrane [26]. *Loxhd1* in mice plays a crucial role in maintaining normal function of cochlear hair cells [27]. It was reported to cause disorders
including autosomal recessive deafness 77. Currently 28 pathogenic variants were reported in this gene from clinvar.

For HL3 pedigree, a homozygous variant in TECTA and two compound heterozygous variants in MYO15A were the likely candidates. TECTA encodes a protein containing 2,155 amino acids that is one of the major non-collagenous glycoproteins of the tectorial membrane, a non-cellular matrix overlying the cochlear neuroepithelium that lies over stereocilia of the hair cells and is critical for the mechanical amplification and transmission of sound [28, 29]. This gene was reported to cause autosomal recessive deafness 21, and 40 pathogenic variants were reported in this gene from clinvar. The protein encoded by MYO15A is a member of the unconventional myosin super-family and plays an indispensable role in the graded elongation of stereocilia and actin organization in hair cells of the inner ear, which are essential for normal hearing function [30]. MYO15A was reported to cause autosomal recessive deafness 3, and 112 pathogenic variants were reported in this gene from clinvar. Considering the transmission of variants in the pedigree (this is a consanguineous pedigree), we thought the homozygous variant in TECTA was more likely to cause the disorder in this pedigree.

For HL4 pedigree, homozygous variants were detected in both DFNB59 and TRIOBP genes. DFNB59 encodes a protein contains 352 amino acids which plays a crucial role in auditory nerve signaling transmission [31]. It was reported to cause autosomal recessive deafness 59, and 9 pathogenic variants were reported in this gene from clinvar. TRIOBP encodes a protein contains 652 amino acids which plays a role in regulation of adherens junctions as well as reorganization of the actin cytoskeleton [32]. Actually, little is known about the exact function of TRIOBP, and multiple roles of this gene raised the issue why pathogenic variants in this gene do
not lead to other pathologies than isolated hearing loss. This gene was reported to cause autosomal recessive deafness 28, and 26 pathogenic variants were reported in this gene from clinvar. Further study is needed to prioritize the causing gene of this pedigree.

For HL5 pedigree, the variant in TMPRSS3 may cause the disorder. The protein encoded by this gene contains a serine protease domain, a transmembrane domain, an LDL receptor-like domain, and a scavenger receptor cysteine-rich domain. It plays an important role in activating the ENaC sodium channel, which is regulated by serine protease activity [33], and maintains a low Na\(^+\) concentration in the endolymph of the inner ear [34]. TMPRSS3 was reported to cause autosomal recessive deafness 8, and 23 pathogenic variants were reported in this gene from clinvar.

We calculated the density of reported pathogenic variants in these genes and they were 11.7/kb, 4.2/kb, 6.2/kb, 10.6/kb, 8.5/kb, 3.7/kb and 16.8/kb for GIPC3, LOXHD1, TECTA, MYO15A, DFNB59, TRIOBP and TMPRSS3, respectively. The density may indicate the degree of understanding or focusing for different genes. And the genes with low density such as LOXHD1 and TRIOBP may have potential research value.

The majority of causing genes we identified for these pedigrees were not common hearing loss genes. If we applied common hearing loss gene panel to screen these patients, we would obtain negative results and the causing gene/variant for the patients would be missed. So, WES may be better strategy than panel sequencing for hearing loss screening even in clinical detection.

**Conclusion**

In conclusion, we applied WES on five consanguineous pedigrees (one patient per
pedigree) with hearing loss from Pakistan, followed by Sanger sequencing for all available samples among the pedigrees to identify the causing genes for them. Several variants in hearing loss genes were validated to be co-segregated in the pedigrees, which may indicate the reasons for such patients. Moreover, we suggested that WES may be suitable strategy for hearing loss screening in clinical detection.

Declarations

**Ethics approval and consent to participate**

This study was approved by the ethical committee of National institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan and all participants provided written informed consent.

**Consent for publication**

Not applicable.

**Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. The data that support the findings of this study have been deposited in the CNSA (https://db.cngb.org/cnsa/) of CNGBdb with accession code CNP0000508-sub010963.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

YZ analyzed and interpreted the WES data and write the manuscript. MT analyzed and interpreted the patient data and modified the manuscript. SH analyzed and interpreted the WES data. UA performed Sanger sequencing. JZ designed the project
and guided the analysis of WES data. SB designed the project, guided the analysis of patient data and helped the manuscript. All authors read and approved the final manuscript.

**Acknowledgments**

The authors thanked all the participating patients and the families in this study for their cooperation.

**References**

1. Morton CC, Nance WE. Newborn hearing screening--a silent revolution. The New England journal of medicine. 2006 May 18;354(20):2151-64.

2. Hu S, Sun F, Zhang J, Tang Y, Qiu J, Wang Z, et al. Genetic Etiology Study of Ten Chinese Families with Nonsyndromic Hearing Loss. Neural plasticity. 2018;2018:4920980.

3. Dror AA, Avraham KB. Hearing impairment: a panoply of genes and functions. Neuron. 2010 Oct 21;68(2):293-308.

4. Morton NE. Genetic epidemiology of hearing impairment. Annals of the New York Academy of Sciences. 1991;630:16-31.

5. Bakhchane A, Bousfiha A, Charoute H, Salime S, Detsouli M, Snoussi K, et al. Update of the spectrum of GJB2 gene mutations in 152 Moroccan families with autosomal recessive nonsyndromic hearing loss. European journal of medical genetics. 2016 Jun;59(6-7):325-9.

6. Tlili A, Al Mutery A, Kamal Eddine Ahmad Mohamed W, Mahfood M, Hadj Kacem H. Prevalence of GJB2 Mutations in Affected Individuals from United Arab Emirates with Autosomal Recessive Nonsyndromic Hearing Loss. Genetic testing and molecular biomarkers. 2017 Nov;21(11):686-91.
7. Dalamon V, Lotersztein V, Beheran A, Lipovsek M, Diamante F, Pallares N, et al. GJB2 and GJB6 genes: molecular study and identification of novel GJB2 mutations in the hearing-impaired Argentinean population. Audiology & neuro-otology. 2010;15(3):194-202.

8. Park HJ, Shaukat S, Liu XZ, Hahn SH, Naz S, Ghosh M, et al. Origins and frequencies of SLC26A4 (PDS) mutations in east and south Asians: global implications for the epidemiology of deafness. Journal of medical genetics. 2003 Apr;40(4):242-8.

9. Tsukada K, Nishio SY, Hattori M, Usami S. Ethnic-specific spectrum of GJB2 and SLC26A4 mutations: their origin and a literature review. The Annals of otology, rhinology, and laryngology. 2015 May;124 Suppl 1:61S-76S.

10. Albert S, Blons H, Jonard L, Feldmann D, Chauvin P, Loundon N, et al. SLC26A4 gene is frequently involved in nonsyndromic hearing impairment with enlarged vestibular aqueduct in Caucasian populations. European journal of human genetics : EJHG. 2006 Jun;14(6):773-9.

11. Zazo Seco C, Wesdorp M, Feenstra I, Pfundt R, Hehir-Kwa JY, Lelieveld SH, et al. The diagnostic yield of whole-exome sequencing targeting a gene panel for hearing impairment in The Netherlands. European journal of human genetics : EJHG. 2017 Feb;25(3):308-14.

12. Bademci G, Foster J, 2nd, Mahdieh N, Bonyadi M, Duman D, Cengiz FB, et al. Comprehensive analysis via exome sequencing uncovers genetic etiology in autosomal recessive nonsyndromic deafness in a large multiethnic cohort. Genetics in medicine : official journal of the American College of Medical Genetics. 2016 Apr;18(4):364-71.

13. Bittles A. Consanguinity and its relevance to clinical genetics. Clinical
14. Hamamy H, Antonarakis SE, Cavalli-Sforza LL, Temtamy S, Romeo G, Kate LP, et al. Consanguineous marriages, pearls and perils: Geneva International Consanguinity Workshop Report. Genetics in medicine : official journal of the American College of Medical Genetics. 2011 Sep;13(9):841-7.

15. Li L, Chen Y, Jiao X, Jin C, Jiang D, Tanwar M, et al. Homozygosity Mapping and Genetic Analysis of Autosomal Recessive Retinal Dystrophies in 144 Consanguineous Pakistani Families. Investigative ophthalmology & visual science. 2017 Apr 1;58(4):2218-38.

16. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. GigaScience. 2018 Jan 1;7(1):1-6.

17. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England). 2009 Jul 15;25(14):1754-60.

18. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010 Sep;20(9):1297-303.

19. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome biology. 2016 Jun 6;17(1):122.

20. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols. 2009;4(7):1073-81.

21. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature
methods. 2010 Apr;7(4):248-9.

22. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nature methods. 2014 Apr;11(4):361-2.

23. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics. 2014 Mar;46(3):310-5.

24. Li MX, Kwan JS, Bao SY, Yang W, Ho SL, Song YQ, et al. Predicting mendelian disease-causing non-synonymous single nucleotide variants in exome sequencing studies. PLoS genetics. 2013;9(1):e1003143.

25. Reed BC, Cefalu C, Bellaire BH, Cardelli JA, Louis T, Salamon J, et al. GLUT1CBP(TIP2/GIPC1) interactions with GLUT1 and myosin VI: evidence supporting an adapter function for GLUT1CBP. Molecular biology of the cell. 2005 Sep;16(9):4183-201.

26. Bateman A, Sandford R. The PLAT domain: a new piece in the PKD1 puzzle. Current biology : CB. 1999 Aug 26;9(16):R588-90.

27. Grillet N, Schwander M, Hildebrand MS, Sczaniecka A, Kolatkar A, Velasco J, et al. Mutations in LOXHD1, an evolutionarily conserved stereociliary protein, disrupt hair cell function in mice and cause progressive hearing loss in humans. American journal of human genetics. 2009 Sep;85(3):328-37.

28. Verhoeven K, Van Laer L, Kirschhofer K, Legan PK, Hughes DC, Schatteman I, et al. Mutations in the human alpha-tectorin gene cause autosomal dominant non-syndromic hearing impairment. Nature genetics. 1998 May;19(1):60-2.

29. Balciuniene J, Dahl N, Jalonen P, Verhoeven K, Van Camp G, Borg E, et al. Alpha-tectorin involvement in hearing disabilities: one gene--two phenotypes.
30. Redowicz MJ. Myosins and deafness. Journal of muscle research and cell motility. 1999 Apr;20(3):241-8.

31. Delmaghani S, del Castillo FJ, Michel V, Leibovici M, Aghaie A, Ron U, et al. Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy. Nature genetics. 2006 Jul;38(7):770-8.

32. Shahin H, Walsh T, Sobe T, Abu Sa'ed J, Abu Rayan A, Lynch ED, et al. Mutations in a novel isoform of TRIOBP that encodes a filamentous-actin binding protein are responsible for DFNB28 recessive nonsyndromic hearing loss. American journal of human genetics. 2006 Jan;78(1):144-52.

33. Vallet V, Chraibi A, Gaeggeler HP, Horisberger JD, Rossier BC. An epithelial serine protease activates the amiloride-sensitive sodium channel. Nature. 1997 Oct 9;389(6651):607-10.

34. Guipponi M, Vuagniaux G, Wattenhofer M, Shibuya K, Vazquez M, Dougherty L, et al. The transmembrane serine protease (TMPRSS3) mutated in deafness DFNB8/10 activates the epithelial sodium channel (ENaC) in vitro. Human molecular genetics. 2002 Nov 1;11(23):2829-36.

Tables

Table 1. Sequencing and variants data
## Table 2. Detailed information for candidate variants

| Pedigree | Variant | Gene   | Strand | DNA change | AA change          | Type          |
|----------|---------|--------|--------|------------|--------------------|---------------|
| HL1      | 19-3590186-T-C | GIPC3   | +      | c.937T>C   | stopcodon lost     | homozygous   |
| HL2      | 18-44063569-C-T | LOXHD1  | -      | c.2935G>A  | p.Glu979Lys        | homozygous   |
| HL3      | 11-121016765-G-A | TECTA   | +      | c.4045G>A  | p.Ala1349Thr       | homozygous   |
| HL3      | 17-18025424-G-T | MYO15A  | -      | c.3310G>T  | p.Gly1104Cys       | compound     |
| HL3      | 17-18069800-G-C | MYO15A  | -      | c.1705G>C  | p.Glu569Gln        | compound     |
| HL4      | 2-179320814-G-A | DFNB59  | -      | c.494G>A   | p.Ser165Asn        | homozygous   |
| HL4      | 22-38120515-C-T | TRIOBP  | +      | c.1952C>T  | p.Ser651Phe        | homozygous   |
| HL5      | 21-43795850-A-G | TMPRSS3 | -      | c.941T>C   | p.Leu314Pro        | homozygous   |

**Figures**
Figure 1

Pedigree figures of the five consanguineous families. The squares and the circles...
Figure 2

Sanger results of the candidate variants. The 6 variants are shown. Each variant is...