Mutat Res Fund Mol Mech Mutagen 800–802 (2017) 29–36

Contents lists available at ScienceDirect

Mutat Res Fund Mol Mech Mutagen

journal homepage: www.elsevier.com/locate/mut

# Targeted sequencing identifies novel variants involved in autosomal recessive hereditary hearing loss in Qatari families

Moza K. Alkowari[a], Diego Vozzi[b], Shruti Bhagat[c], Navaneethakrishnan Krishnamoorthy[a,d], Anna Morgan[b,e], Yousra Hayder[f], Barathy Logendra[c], Nehal Najjar[f], Ilaria Gandin[b], Paolo Gasparini[b,e], Ramin Badii[c], Giorgia Girotto[b,e,*], Khalid Abdulhadi[f]

[a] Division of Experimental Genetics, Sidra Medical and Research Centre, Doha, Qatar
[b] Medical Genetics, Institute for Maternal and Child Health – IRCCS "Burlo Garofolo", Trieste, Italy
[c] Molecular Genetics Laboratory, HMC, Doha, Qatar
[d] Heart Science Centre, National Heart and Lung Institute, Imperial College London, United Kingdom
[e] Medical Sciences, Chirurgical and Health Department, University of Trieste, Trieste, Italy
[f] Audiology and Balance Unit, HMC, Doha, Qatar

## ABSTRACT

Hereditary hearing loss is characterized by a very high genetic heterogeneity. In the Qatari population the role of *GJB2*, the worldwide HHL major player, seems to be quite limited compared to Caucasian populations. In this study we analysed 18 Qatari families affected by non-syndromic hearing loss using a targeted sequencing approach that allowed us to analyse 81 genes simultaneously. Thanks to this approach, 50% of these families (9 out of 18) resulted positive for the presence of likely causative alleles in 6 different genes: *CDH23, MYO6, GJB6, OTOF, TMC1* and *OTOA*. In particular, 4 novel alleles were detected while the remaining ones were already described to be associated to HHL in other ethnic groups. Molecular modelling has been used to further investigate the role of novel alleles identified in *CDH23* and *TMC1* genes demonstrating their crucial role in Ca2+ binding and therefore possible functional role in proteins. Present study showed that an accurate molecular diagnosis based on next generation sequencing technologies might largely improve molecular diagnostics outcome leading to benefits for both genetic counseling and definition of recurrence risk.

## 1. Introduction

Hearing loss (HL) is a remarkably complex and heterogeneous disease presenting with various phenotypes as a result of both genetic and environmental factors. Approximately 70% of cases affected by hereditary hearing loss (HHL) can be classified as non-syndromic hearing loss (NSHL) (i.e. with the absence of abnormalities in other organs), whereas the remaining 30% are classified as syndromic [1]. The vast majority of NSHL cases (70–80%) are transmitted as autosomal recessive traits (ARNSHL) with 64 known genes. The remaining 15–20% show an autosomal dominant pattern of inheritance (ADNSHL, 36 known genes) while a small proportion is X-linked (4 known genes) or mitochondrial (8 known mutations) (Hereditary Hearing Loss Homepage; http://hereditaryhearingloss.org/).

Neonates belonging to inbred populations, such as the Qatari one, might have a higher risk of developing ARNSHL [2] and other autosomal recessive diseases as well [3]. Through the Qatar's national program aimed to hearing loss early detection, the 5.2% of newborns have been diagnosed as affected by hearing impairment and 60.5% of them are born from consanguineous mating [4]. Previous studies have shown that mutations in *GJB2* are the leading cause of HHL in the Caucasian population [5] however within the Qatari population *GJB2* seems to play a less relevant role [6]. In this light, there is the strong need to investigate several Qatari families to look for new causative alleles/genes such as *BDP1* gene, recently identified in a consanguineous family [7].

Given the high genetic heterogeneity of HHL, high throughput approaches such as next generation sequencing (NGS) are the most appropriate and reliable tools to dissect the complexity of HHL [8,9].

In this light, here we report the data, obtained from a series of Qatari families affected by HHL and analysed by targeted re-sequencing for the presence of possible causative mutations in 81 different HHL genes.

* Corresponding author at: Medical Sciences, Chirurgical and Health Department, University of Trieste/Medical Genetics, Institute for Maternal and Child Health – IRCCS "Burlo Garofolo", Trieste, Italy.
*E-mail address:* giorgia.girotto@burlo.trieste.it (G. Girotto).

http://dx.doi.org/10.1016/j.mrfmmm.2017.05.001
Received 16 February 2017; Received in revised form 11 April 2017; Accepted 3 May 2017
Available online 04 May 2017
0027-5107/ 

## 2. Materials and methods

### 2.1. Ethical statement

Eighteen families were recruited for this study. A written informed consent was obtained from all participants; in case of minors, their next of kin provided written informed consent. The study was approved either by the Institutional Review Board of Hamad Medical Corporation, Doha, Qatar and the Institutional Review Board of IRCCS Burlo Garofolo, Trieste, Italy.

### 2.2. Patients: clinical evaluation and sample collection

Subjects affected by sensorineural NSHL were recruited at the ENT Department, Hamad Medical Corporation (HMC), Doha, Qatar. A total of 18 families (80 individuals including 27 probands) were enrolled in the study. Probands' age was ranging from 1 to 13 years, 14 of them were males and 13 were females. An accurate anamnestic history (clinical and instrumental) was collected for all of them. A complete medical evaluation was performed on each family to exclude hearing loss due to infections, trauma, or other non-genetic causes. The presence of a syndromic deafness was also excluded. All participants underwent pure tone audiometric testing (PTA) or auditory brainstem response (ABR) (depending on the probands' age) in order to characterise the severity of HL according to the following guidelines [1]:

- Slight: 16-25 dB (dB)
- Mild: 26–40 dB
- Moderate: 41–55 dB
- Moderately severe: 56–70 dB
- Severe: 71–90 dB
- Profound: 91 dB or more

All probands were completely negative for mutations in both *GJB2* and *MTRNR1* genes. Blood samples were taken from 77 participants, while saliva samples from the remaining ones. Genomic DNA was extracted using Promega Maxwell® 16 Blood DNA purification kit and Norgen Biotek's Saliva DNA Collection Preservation and Isolation Kit, from whole blood and saliva respectively. Quality and quantity of DNA was checked using NanoDrop 1000 spectrophotometer and Qubit 2.0 Fluorometer respectively. At least 3 individuals (both affected and unaffected) from each family were sequenced and included in the genetic co-segregation analysis.

### 2.3. Hereditary hearing loss panel

A panel of 81 genes was designed using the Ion AmpliSeq™ Designer v1.2 (Life Technologies, CA, USA). The genes were selected on the basis of the most updated literature survey and the most comprehensive HHL databases such as Deafness Variation Database (http://deafnessvariationdatabase.org/) and Hereditary Hearing Loss Database (http://hereditaryhearingloss.org/). The sequencing panel includes 25 known ARNSHL genes, 18 known ADNSHL genes, 6 genes for both ARNSHL and ADNSHL and 1 gene known to cause X-linked NSHL. The remaining 31 genes are known to be involved in hearing function or to be expressed in mouse inner ear. A complete list of genes is provided in Supplementary Table S1. Targeted regions include coding regions (CCDS) plus 50 bp flanking exons of each gene. The overall coverage is approximately 95% accounting for 317 kb.

### 2.4. Targeted Re-Sequencing (TRS) and data analysis

Sequencing was performed using the Ion Torrent™ platform (Life Technologies, USA) and libraries prepared as described in Supplementary S1. Among all the identified genetic variants, we filtered out those with quality score (QUAL) < 30, synonymous nucleotide substitutions and variants with minor allele frequencies (MAF) > 0.03. In particular, the following databases were used: NCBI dbSNP build142 (http://www.ncbi.nlm.nih.gov/SNP/), 1000 Genomes Project (http://www.1000genomes.org/), NHLBI Exome Sequencing Project (ESP) Exome Variant Server (http://evs.gs.washington.edu/EVS/), ExAC browser (http://exac.broadinstitute.org/). The pathogenicity of known variants was evaluated using ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/), Deafness Variation Database (http://deafnessvariationdatabase.org/) and The Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php). For novel variants several *in silico* tools, such as PolyPhen-2 [10], SIFT [11], MutationTaster [12], PhyloP [13], GERP$^{++}$ [14] and CADD score [15] were used, in order to consider both the effects on protein structure and the conservation of the affected residues. A schematic representation of the analytical workflow is reported in Supplementary Fig. S1. The Coverage Analysis Report, a collection of summary statistics and graphical representations of the reads coverage available for the Ion Torrent Suite™ (Life Technologies, USA), was used for the identification of any inconsistencies, including coverage falls due to the presence of homozygous deletions. On a patient-by-patient base, the candidate variants were discussed in the context of phenotypic data at weekly interdisciplinary meetings by a team of experts including clinicians, scientists, geneticists, genetic counselors, and bioinformaticians. Finally, Sanger sequencing analysis to confirm the mutations and to check the segregation was carried out.

### 2.5. Molecular modelling and structural analysis

The most frequent mutations detected in our families were further investigated at protein level. Briefly, the structural consequences of the missense mutations in *CDH23* and *TMC1* were analysed by molecular modelling (CDH23 UniProt: Q9H251 and TMC1 UniProt: Q8TDI8 respectively). In both cases, for CDH23 (cadherin domains 20–21) and TMC1 (residues 402–652) there were no 3D structures available. Thus, for CDH23, the crystal structure of mouse N-cadherin ectodomain protein (PDB ID: 3Q2W) [16] sharing 42% similarity with human CDH23 was used, while, for TMC1, the crystal structure of a calcium (Ca 2 + ) channel protein (PDB ID: 4WIS) from the fungus (*Nectria haematococca*) [17], that shares 26% similarity with human TMC1, was considered. The quality of the models was evaluated as previously described [18], showing that structures were biologically reliable. The constructed wild type structures were used for building mutational models (p.P2205L in CDH23 and p.R445H & p.L603H in TMC1) in the discovery studio (Accelrys Inc., San Diego, CA, USA) as previously described [19] and to identify functional (Ca2+ binding) sites and the key residues involved. Pymol was used to visualize the proteins and to prepare model representations (www.pymol.org).

## 3. Results

On average 95.48% of the targeted region was covered by sequencing data at least 20-folds, while 270-folds mean-depth total coverage was reached. An overall amount of 763 genetic variations was identified (i.e. 170 synonymous, 178 non-synonymous, 43 frameshift/3 non-frameshift INDELs and the remaining ones intronic and within UTRs). After the filtering process, six point mutations in 5 HHL genes (*CDH23, MYO6, GJB6, OTOF, TMC1*) and a large deletion in *OTOA* were identified thus characterizing 50% of families investigated (9 out of 18) (Table 1).

As expected, being consanguineous families, the possible causative variants were found at the homozygous state while healthy relatives were WT or heterozygous. Clinical and instrumental details of patients carrying these alleles and belonging to the 9 families characterized are reported in Supplementary Table S2.