Genetic contribution of \textit{GJB2} gene to hearing impairment in Pakistan

Hamna Tariq\textsuperscript{1}\textsuperscript{*}, Kalsoom Zaigham\textsuperscript{1}\textsuperscript{#}, Samra Kousar\textsuperscript{1}, Aysha Azhar\textsuperscript{2}

Abstract

Background: Hearing impairment (HI) is defined as inability to hear and is an extremely heterogeneous genetic disorder. HI is divided into syndromic (if associated with clinical manifestation in addition to hearing impairment) and non-syndromic forms. So far one hundred and seventeen loci/genes have been mapped for non-syndromic HI and mutations in DFNB1 locus (\textit{GJB2} gene) are the most prevalent cause among them. This study was intended to find the relative contribution of the DFNB1 locus/ \\textit{GJB2} gene for hearing loss in Pakistan and Azad Kashmir.

Methods: Twenty-one families were collected from different rural and urban regions of Pakistan and Azad Kashmir. The contribution of \textit{GJB2} gene was initially studied by linkage analysis using short tandem repeats (STR) microsatellite markers. Sanger sequencing was employed to identify the causative variants in coding region of the gene.

Results: Phenotype of four families were found linked with \textit{GJB2} gene and all affected individuals of these families segregating same mutation c.231G\textgreater{}A (p.Trp77\textsuperscript{*}) which was confirmed after Sanger sequencing.

Conclusion: The genetic causes of hearing impairment were studied in twenty one families segregating autosomal recessive pattern of inheritance with different ethnicities. We further established the founder effect for the one recurrent mutation in \textit{GJB2} gene in Pakistani and Kashmiri hearing impaired families for the very first time.
Introduction

Hearing impairment (HI) is a common sensory incapability that has worldwide incidence of 1 in 650 infants, with frequency of 2.7/1000 in minors and 3.5/1000 in adolescents making it the most frequently occurring inherited sensory impairment [1,2].

HI has multifactorial etiologies i.e. genetic and environmental aspects. Approximately 50-60% of hearing impairment cases are attributed to genetic causes. Clinically hearing impairment is categorized as either non-syndromic, where it expresses itself as a solo entity, or syndromic with additional physiological or metabolic malformation. Non-syndromic hearing impairment accounts for 70% of the cases and is highly heterogeneous [3]. Autosomal recessive pattern is most common pattern of inheritance in these cases. To date, 117 DFNB loci have been mapped to show linkage with non-syndromic hearing impairment (https://hereditaryhearingloss.org/recessive-loci/).

Pakistani population is a rich source to study the genetic disorders presents itself as a perfect genetic model for the study of hereditary diseases owing to its high rate of consanguineous marriages and defined lineage system [4]. Additionally, it is observed that the rate of hearing loss in Pakistan is approximately 1.6 per 1000 live birth which is higher than the world average of 1 per 1000 live births [5].

GJB2 gene variations have been recurrently reported in hearing impaired individuals, few of them are frequent in some populations and others are extremely rare. The mutation spectrum diverges considerably among different populations, and presents ethnic biases, for example mutations like 35delG is common among whites with a carrier rate of 2%-4% [6], 235delC has carrier rate of 1%-2% in the Japanese population [7], while 167delT is reported in the Ashkenazi Jews with carrier rate of 7.5% [8], and V37I in Taiwan (carrier rate of 11.6%) [9]. Regardless of this variation, the overall frequency of all the mutations reported in GJB2 gene is considerably high in various populations around the world and thus make this gene a clinically useful for genetic testing.

Current study was carried out to identify Pakistani families suffering from hearing impairment and to genetically analyze such families, for the most prevalent hearing impairment locus and their genetic alteration in population. The haplotypes obtained after linkage study of four families with DFNB1 locus harboring GJB2 gene with different ethnicities including two Punjabi families from Upper Punjab province i.e., Chakwal and Mandi Bahauddin, one family from South Punjab city of Rahim Yar Khan and a Kashmiri family from Azad Kashmir suggested that they share a common ancestor.

Methods

The approval for this research study was taken by Institutional Review Board (IRB) at the Centre of Excellence in Molecular Biology (CEMB), Lahore, Pakistan. Informed written consents and clinical information were taken from all the participants of the study. 10 ml of blood sample was drawn from all participating members in 50 ml Sterilin® polypropylene tubes with 200 µl of 0.5M EDTA (anticoagulant) for subsequent DNA extraction.

Genomic DNA extraction

The genomic DNA was extracted from the collected blood samples. The non-organic method of DNA extraction from white blood cells was optimized and followed [10]. The quality and quantity of extracted genomic DNA was estimated by gel electrophoresis and on Spectrophotometer (Bio-Rad, Hercules, CA).

Exclusion analysis, PCR amplification and genotyping

Exclusion study was performed to screen the most common hearing impairment locus DFNB1 in the world and to find its prevalence in Pakistani population. Linkage analysis was done with the help of highly polymorphic fluorescent STR markers designed to cover DFNB1 locus. Mixtures of PCR amplified products and size standards were prepared and run in 3730 Genetic Analyzer (Applied Biosystems Inc.). Alleles were assigned through GeneMapper® software (Applied Biosystems, version 4.0) was used.

Sanger Sequencing

Sanger Sequencing was carried out for one coding exon of GJB2 gene by using Big Dye (Big dye Terminator v3.1 Cycle Sequencing Kit- Applied Biosystems®). The sequencing products were run through 3730 Genetic Analyzer (Applied Biosystems Inc.) before further analysis. The data was analyzed with Chromas software version 2.6.6 to confirm segregation of alleles.

Results

In this study we aimed to find the prevalence of most common hearing impairment locus DFNB1 harboring GJB2 gene. A cohort of 21 families suffering from hearing impairment belonging to diverse ethnicities was ascertained from different regions of Pakistan and Azad Kashmir. After sample collection, twenty-one families with hearing loss were studied for linkage analysis, out of which four families were linked with DFNB1 locus. All these four families i.e. family 03, family 04, family 06 and family 09 were then subjected to Sanger sequencing to identify the causative variants in GJB2 gene. Sequencing data revealed the one most frequent variants c.231G>A (p.W77X) in these families. This variant was previously identified as recurrent in Pakistani population. However, this mutation c.231G>A (p.W77X) having founder effect is reported for the first
time in the present study in Pakistani and Kashmiri families. The evolutionary conservation of amino acid Trp77 (W77) as described in Table-1 was also compared in other GJB2 orthologs and is conserved along with other amino acids present in the vicinity. A map of single nucleotide polymorphism (SNP) flanking the causative mutations was also constructed and the established haplotype (CGGAGG), suggested the common founder for the c.231G>A transition in GJB2 gene (Table-2). The detrimental effects of this recurrent mutation c.231G>A (p.W77X) were recorded using different bioinformatics tools like mutation taster (Pathogenic), likelihood ratio test (LRT) is deleterious and genomic evolutionary rate profiling (GERP) score was 5.33 and was classified as disease causing.

**Family 03**
This family 03 was registered from Azad Kashmir region of Pakistan having 3 individuals with hearing loss in main loop and 2 individuals in second loop. All affected (IV:3, IV:4, IV:5, IV:6 and IV:8) and normal individual (IV:1, IV:2 and IV:7) along with parents (III:1 and III:2, III:3 and III:4) were enrolled in this research (Figure-1A).

**Family 04**
This consanguineous Punjabi family was ascertained from a city in Upper part of the Punjab province, Mandi Bahauddin and had three affected individuals from a consanguineous reunion in single sib ship. Three affected individuals (IV:1, IV:4, IV:5), two unaffected individuals (IV:2, IV:5) and their parents (III: 1, III:2) were enrolled. The disease phenotype was segregating in all three affected individuals of the pedigree (Figure-1B).

**Family 06**
Another family was hailed from upper region of Punjab province. The family was collected from Chakwal and was multigenerational and consanguineous. In this family three affected individuals (IV:2, IV:3 and V:1), four normal individuals (IV:1, IV:4, IV:5 andV:2 ) and their parents (III:1 and III:2) were tested for linkage analysis and mutational analysis of DFNB1 locus and GJB2 gene respectively. Clinical history of this family revealed non syndromic hearing impairment in affected individuals while all other individuals of the family were phenotypically normal (Figure-1C).

**Family 09**
This family belonging to Punjabi ethnicity was registered from South Punjab province city named Rahim Yar Khan. Three affected individuals (IV:1, IV:3 and IV:5) two normal siblings (IV:2 and IV:4) and their parents (III:1 and III:2) were registered. All seven individuals enrolled from this family were tested genetically. All affected individuals were clinically evaluated and revealed no other clinical manifestation other than hearing impairment (Figure-1D).

**Haplotype Analysis:**
The Haplotype of all the affected individuals in four families showed linkage to DFNB1 locus. All the affected individuals were homozygous to markers GJB2, D13S175 and D13S1275 and normal individuals were heterozygous as it carries different combination of alleles from both parents. A maximum 2 point Lod Score (Zmax) was calculated with markers D13S1275 and D13S175 at recombination fraction (θ) = 0 and was found above 3.0 for all four families.

**Mutational Analysis:**
Sanger sequencing of GJB2 gene revealed a nonsense mutation due to transition from G to A (c.251G>A) (Figure-1E & F). Due to this alteration the amino acid tryptophan TGG at position W77 is changed into stop codon TAG resulting in truncated protein (p.Trp77*). This single base pair change G>A at position 231 (p.W77X) shows penetrance with hearing loss phenotype in family 03, 04, 06 and 09 (Table 3).

**Discussion**
The most prevalent sensorineural disease in humans with genetic contributors is hearing impairment. Among all forms of HI, autosomal recessive non-syndromic form is most frequent which accounts for 70% of genetic causes [11]. The most common cause among them is due to DFNB1/ GJB2 mutations. The GJB2 gene at DFNB1 locus encodes a gap junction protein Connexin26 and is responsible for 20% of all childhood deafness.

Hearing loss is presented with extreme allelic and locus heterogeneity. The results of the screening studies carried out worldwide suggest that GJB2 is more frequently mutated gene than any other counterpart. The two most commonly prevalent causative variants of GJB2 gene in Pakistani population are c.71G>A (p.W24X) and c.231G>A (p.W77X) which is in line with the results of present study and four families presented with same c.231G>A (p.W77X) mutation [12,13]. The most frequent GJB2 mutation in European, Middle Eastern and African ancestries is c.35delC (p.G12V) [6]. However, c.235delC (p.L79Cfs) is the major contributing GJB2 mutation in East Asian lineages [14]. The findings of this study are consistent with the previous studies carried out on consanguineous Pakistani families indicating that GJB2 gene is frequently mutated gene for HI in various ethnicities. However, this is the first study to report founder effect of GJB2 gene in a Kashmiri family.
Figure 1: Pedigree of family 03 (A), family 04 (B), family 06 (C), family 09 (D), and chromatograms (E & F) of the mutation c. 231G>A, (p.W77X).

Table 1: Evolutionary conservation of amino acid Tryptophan (W77)

<table>
<thead>
<tr>
<th>Species</th>
<th>I</th>
<th>R</th>
<th>L</th>
<th>W</th>
<th>A</th>
<th>L</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Rhesus</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Mouse</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Dog</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Elephant</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Chicken</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>X_tropicalis</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>C</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>I</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
<tr>
<td>Lampray</td>
<td>V</td>
<td>R</td>
<td>L</td>
<td>W</td>
<td>A</td>
<td>L</td>
<td>Q</td>
</tr>
</tbody>
</table>

Table 2: SNP Haplotypes of hearing impaired Individuals of four families

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Individual</th>
<th>rs529500747</th>
<th>rs1801002</th>
<th>rs104894596</th>
<th>rs2274084</th>
<th>rs80338944</th>
<th>rs2274083</th>
<th>rs111033196</th>
<th>rs111033186</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 03</td>
<td>IV:1</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Family 04</td>
<td>IV:5</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Family 06</td>
<td>IV:2</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Family 09</td>
<td>IV:5</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

Table 3: Mutational analysis of four families segregating hearing impairment

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Ethnicity</th>
<th>Area</th>
<th>Gene</th>
<th>Variant</th>
<th>Protein Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 03</td>
<td>Kashmiri</td>
<td>Azad Kashmir</td>
<td>GJB2</td>
<td>c.231G&gt;A</td>
<td>(p.Trp77*)</td>
</tr>
<tr>
<td>Family 04</td>
<td>Saraiki</td>
<td>Mandi Bhauddin</td>
<td>GJB2</td>
<td>c.231G&gt;A</td>
<td>(p.Trp77*)</td>
</tr>
<tr>
<td>Family 06</td>
<td>Punjabi</td>
<td>Chakwal</td>
<td>GJB2</td>
<td>c.231G&gt;A</td>
<td>(p.Trp77*)</td>
</tr>
<tr>
<td>Family 09</td>
<td>Punjabi</td>
<td>Rahim Yar Khan</td>
<td>GJB2</td>
<td>c.231G&gt;A</td>
<td>(p.Trp77*)</td>
</tr>
</tbody>
</table>
This study is a modest contribution to overall spectrum of GJB2 mutations worldwide. This also emphasize the need of a comprehensive screening program in all clans and ethnicities in Pakistan which will provide us population specific knowledge for clinical diagnosis and genetic counseling of these families as described in previous studies. The identification of this mutation c.231G>A (p.W77X) along with other prevalent mutations in GJB2 gene will be helpful in an early diagnosis of the disorder in infants, which in turn will allow the implementation of different interventional approaches, for example, provision of hearing aids and rehabilitation (learning the sign language). This will further help in avoiding a time period with little or no auditory stimulation in the early childhood. The identification of these mutations can be easily carried out by sequencing in clinical setup for screening purposes and be especially useful in countries with high rate of consanguinity, like Pakistan.

**Abbreviations**

GERP: Genomic evolutionary rate profiling
GJB2: Gap junction beta 2
HI: Hearing impairment
LOD: Logarithm of the odds
LRT: Likelihood ratio test
SNP: Single nucleotide polymorphism
STR: Short tandem repeats

**Declarations**

**Ethics approval and consent to participate:**
Ethical approval was obtained from the institutional review board, National Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan. After explanation of the objectives of this project, written informed consent was obtained from all participating individuals.

**Consent for publication:**
Written informed consent to publish findings of this study was acquired from each participant.

**Availability of data and material:**
All the data generated or analyzed during this study have been included in this manuscript.

**Funding:**
This study was supported by the indigenous fellowship of Higher Education Commission (HEC), Islamabad Pakistan. The funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Conflict of Interest Statement**
The authors declare that there is no conflict of interest regarding the publication of this paper.

**Acknowledgement**
We are highly thankful to Higher Education Commission (HEC) of Pakistan for funding this work and Centre of Excellence in Molecular Biology for providing research facilities. The authors are also thankful to all the participating families for their cooperation in this study.

**Authors’ Contribution**
HT, KZ, SK: study concept and design; implementation of experiments; data acquisition; analysis and interpretation of data; drafting of the manuscript. AA: study concept and design; analysis and interpretation of data; revising the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

**References**

