Genetic Determinants of Non-syndromic Hearing Impairment

Genetische determinanten van niet-syndromale slechthorendheid

Regie Lyn Pastor Santos

Printing of this thesis was realized with financial support from the Laboratory of Statistical Genetics (Leal Laboratory), Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas USA.

Genetic Determinants of Non-syndromic Hearing Impairment
Santos, Regie Lyn Pastor
Thesis, Erasmus Medical Center, Erasmus University Rotterdam
With summary in English and Dutch

Cover design and layout: Dennis A. Santos

Printed by RainTree Trading & Publishing Inc., Quezon City, Philippines

Genetic Determinants of Non-syndromic Hearing Impairment

Genetische determinanten van niet-syndromale slechthorendheid

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
donderdag 14 september 2006 om 11 uur

door

Regie Lyn Pastor Santos

geboren te Caloocan City, Philippines

PROMOTIECOMMISSIE

Promotor:	Prof.dr. C.M. van Duijn
Overige leden:	Prof.dr. A. Uitterlinden
	Prof.dr. J.G.G. Borst
	Prof.dr. G. van Camp
Co-promotoren:	Prof.dr. S.M. Leal
	Dr. Y.S. Aulchenko

Para kay Mama --

The enemy was cut off,
lost the chance to stare and grope at it.
I felt the knife sear,
though it is not my body
but my genes.
Did we win the DNA lottery?
Only time will tell.

Contents

Chapter 1
Introduction

The deprivation of sound has a tremendous impact not only on individuals but also on society as a whole. The delay or failure to acquire age-appropriate language and speech perception skills in childhood or loss of hearing later in life has pervasive effects, particularly on a person's psychosocial and economic well-being.

Previously it was estimated that congenital hearing impairment (HI) affects 1-2 per 1000 newborns [1-2]. However, more recent estimates from Universal Newborn Hearing Screening programs in the United States show that the prevalence of congenital HI among neonates is 2-4 per 1000 [3]. Nevertheless it is clear that this is still a gross underestimate given the high rates of loss to follow-up (23-52%); therefore HI with postnatal onset is not fully accounted for. It should be noted that HI increases as the population ages. While 18% of the population has mild to profound HI between the ages 45-64, this figure jumps to 46% for individuals older than 74 years [4].

In developed countries where the frequency of HI due to preventable causes such as infection and drug ototoxicity has been largely minimized, the proportion of congenital hearing impairment (HI) due to genetic factors has increased. Past estimates have indicated that roughly half of congenital HI cases are expected to be genetic [1, 5]. Of the genetic HI cases,

30% are syndromic. For non-syndromic (NS) HI, 77% of the cases are due to autosomal recessive (AR) inheritance, 22% display autosomal dominant (AD) inheritance and about 1% of cases show X-linked inheritance [5].

The most notable aspect in the study of NSHI is this trait's extensive genetic heterogeneity. To date about 120 NSHI loci have been mapped, of which 51 are autosomal dominant or ADNSHI (with the loci designated as *DFNA* followed by the locus number), 67 autosomal recessive or ARNSHI (*DFNB#*), five X-linked (*DFN#*) and one Y-linked (*DFNY#*) [6]. Two loci for modifier genes have also been identified: the first being *DFNM1*, which dominantly suppresses *DFNB26* [7], while *DFNM2* influences expression of maternally inherited deafness associated with the 1555G>A substitution in mitochondrial 12S ribosomal RNA [8]. For some NSHI loci, the HI was due to different mutations in genes which cause syndromic HI. Examples include (but are not limited to) Usher syndrome (due to mutations in *MYO7A*, *CDH23*, *USH1C* or *PCDH15*), Pendred syndrome (mutations in *SLC26A4*) and Wolfram syndrome (mutations in *WFS1*) [9-11]. NSHI loci and genes are spread throughout the human genome (Figure 1.1).

Currently many laboratories and clinical centers worldwide are involved in gene mapping for hearing impairment. A common goal is to develop a genetic screening program that can lead to earlier diagnosis and management of genetic HI. The longer the diagnosis of HI is delayed, the greater the impact on cognitive development. It has been shown that children who were diagnosed at less than 6 months of age and treated for HI within the first year of life have language development that is comparable to their normal-hearing peers [12].

At present, many academic hospitals in developed countries offer screening for mutations in the *GJB2* gene for children with ARNSHI. *GJB2* is the most prevalent ARNSHI gene known to date, with as much as 50% of prelingual ARNSHI cases in populations of European Caucasian and Ashkenazi Jewish descent having biallelic *GJB2* mutations [13]. The following mutations were shown to be highly prevalent in specific populations: 35delG among Europeans, Middle Easterners and Tunisians; 167delT among Ashkenazim; R143W in Ghanaians; and 235delC in East Asians [14-18]. Based on the literature on *GJB2*, it is clear that different populations carry different functional variants and polymorphisms.

In some cases, the initial study which was done to identify an HI gene may have used a special population or sample (for example, a large consanguineous family). Therefore the first sequence variants which were identified may be specific to the sample or family being studied and are

not necessarily prevalent in the general population. In the case of *GJB2*, the p.M34T sequence variant was first to be identified in a pedigree with NSHI and palmoplantar keratoderma [19], but it was later shown that p.M34T either causes milder hearing impairment (based on genotype-phenotype correlation) or is a benign polymorphism (based on prevalence studies) [20-22]. Given that this variant occurs at an evolutionarily conserved residue and that repeated studies on cell expression systems result in the failure of p.M34T-transfected cells to form functional channels [23-24], we are thus inclined to believe that this variant truly causes hearing impairment, although it is not as prevalent and severe in phenotype as other *GJB2* variants. Thus for prevalence estimates of HI gene mutations to be reliable, functional analyses are important to first assess the value of sequence variants in disease causation prior to inclusion of the said variant in prevalence estimation.

Aside from *GJB2*, although not nearly as well studied, considerable work has been done on the NSHI genes *GJB6*, *WFS1*, *CDH23*, *SLC26A4*, *TECTA* and *TMPRSS3*. Table 1.1 presents the published NSHI genes, with the following headings: the gene name; number according to the Online Mendelian Inheritance in Man database; the protein product; the name of the syndrome and/or NSHI locus; the cytogenetic position of the gene; the primary defect which the mutation causes; and, the type of protein molecule.

This list of identified NSHI genes, however, is by no means complete, because it is predicted that there may be as many as 300 HI genes [25]. For the mapped NSHI loci, only 37 genes have been identified. This means that aside from the ~80 loci for which genes are yet to be isolated, more than half of the predicted number of HI genes still need to be localized and then identified.

Figure 1.1 Diagram of the location (*red lines*) of published loci and genes for non-syndromic hearing impairment. The gene is indicated if it has been isolated in hearing-impaired humans. Loci numbers are cited if the causative gene remains unknown.

Table 1.1 Known non-syndromic hearing impairment genes

Gene	OMIM [a]	Encoded Product	Loci	Position	Primary Defect	Type of Molecule
ACTG1	102560	Actin, gamma-1	*DFNA20/26* [b]	17q25.3	Stereocilia, cuticular plate, zonula adherens	Cytoskeletal protein
CDH23	605516	Cadherin 23	*DFNB12,* USH1D	10q22.1	Stereocilia	Hair bundle protein
CLDN14	605608	Claudin-14	*DFNB29*	21q22.13	Tight junctions of reticular lamina	Tight junction transmembrane protein
COCH	601369	Cochlin	*DFNA9*	14q12	Spiral ligament/ limbus	Extracellular matrix component
COL11A2	120290	Collagen, type XI	*DFNA13,* STL3, OSMED, WZS	6p21.32	Tectorial membrane	Extracellular matrix component
DFNA5	600994	DFNA5	*DFNA5*	7p15.3	Unknown	Unknown
DIAPH1	602121	Diaphanous-1	*DFNA1*	5q31.3	Hair cells / stereocilia	Regulator of actin cytoskeleton
ESPN	606351	Espin	*DFNB36*	1p36.31	Stereocilia	Actin-bundling protein
EYA4	601316	Eyes absent 4	*DFNA10*	6q23.2	Unknown	Transcriptional coactivator
GJB2 [c,d]	121011	Connexin 26	*DFNA3, DFNB1*	13q12.11	Spiral ligament/limbus, stria vascularis, supporting cells	Gap junction protein
GJB3	603324	Connexin 31	*DFNA2*	1p34.3	Spiral limbus	Gap junction protein
GJB6 [c,d]	604418	Connexin 30	*DFNA3, DFNB1*	13q12.11	Spiral ligament/limbus, stria vascularis, supporting cells	Gap junction protein
KCNQ4	603537	K^+ channel voltage-gated, subfamily Q, member 4	*DFNA2*	1p34.2	Outer hair cells	K^+ channel subunit
MYH9	160775	Myosin heavy polypeptide 9	*DFNA17,* FTNS	22q12.3	Epithelial layer of cochlear duct, spiral limbus/ligament, Reissner membrane	Nonmuscle myosin heavy chain
MYH14	608568	Myosin heavy polypeptide 14	*DFNA4*	19q13.33	Stria vascularis, supporting cells	Nonmuscle myosin heavy chain
MYO1A	601478	Myosin IA	*DFNA48*	12q13.3	Unknown	Motor protein
MYO3A	606808	Myosin IIIA	*DFNB30*	10p12.1	Hair cells	Motor protein
MYO6 [c]	600970	Myosin VI	*DFNA22, DFNB37*	6q14.1	Cuticular plate	Motor protein
MYO7A [c]	276903	Myosin VIIA	*DFNA11, DFNB2,* USH1B	11q13.5	Stereocilia, hair cell bodies	Motor protein
MYO15A	602666	Myosin XVA	*DFNB3*	17p11.2	Cuticular plate, stereocilia	Motor protein

OTOA	607038	Otoancorin	*DFNB22*	16p12.2	Tectorial membrane	Membrane-anchored protein
OTOF	603681	Otoferlin	*DFNB9*	2p23.3	Inner hair cells	Vesicle trafficking protein
PCDH15	605514	Protocadherin 15	*DFNB23,* USH1F	10q21.1	Stereocilia	Hair bundle protein
PDS /SLC26A4	274600	Pendrin	*DFNB4,* PDS, EVA	7q31.1	Stria vascularis	Iodide/chloride transporter
POU4F3	602460	POU domain, class 4, transcription factor 3	*DFNA15*	5q32	Hair cells	Transcription factor
POU3F4	300039	POU domain, class 3, transcription factor 4	*DFN3*	Xq21.1	Spiral ligament	Transcription factor
STRC	606440	Stereocilin	*DFNB16*	15q15	Stereocilia	Hair bundle protein
TECTA [c]	602574	Alpha-tectorin	*DFNA8/12* [b], *DFNB21*	11q23.3	Tectorial membrane	Extracellular matrix component
TFCP2L3	608576	Transcription factor CP2-like 3	*DFNA28*	8q22.3	Epithelial tissues of scala media	Transcription factor
TMC1 [c]	606706	Transmembrane channel-like 1	*DFNA36, DFNB7/11* [b]	9q21.13	Hair cells	Transmembrane protein
TMHS / LHFPL5	609427	Tetraspan membrane protein of hair cell stereocilia / LHFP-like protein 5	*DFNB67*	6p21.31	Stereocilia, hair cells	Transmembrane protein
TMIE	607237	Transmembrane inner ear	*DFNB6*	3p21.31	Stereocilia, hair cells	Transmembrane protein
TMPRSS3	605511	Transmembrane protease, serine 3	*DFNB8/10* [b]	21q22.3	Spiral ligament, supporting cells, stria vascularis	Transmembrane serine protease
TRIOBP	609761	TRIO and F-actin binding protein	*DNFB28*	22q13.1	Hair cells	Regulator of actin cytoskeleton organization
USH1C	605242	Harmonin	*DFNB18,* USH1C	11p15.1	Stereocilia	Hair bundle protein
WFS1	606201	Wolframin	*DFNA6/14* [b], WFS1	4p16.3	Canalicular reticulum	Transmembrane protein
WHRN	607928	Cip98	*DFNB31*	9q32	Stereocilia	Hair bundle protein

[a] Number from Online Mendelian Inheritance of Man database (http://www.ncbi.nlm.nih.gov/omim/).

[b] Two loci were mapped to the same region, but when the gene was identified, it was found that it is the same for the two loci. This occurs when two or more independent laboratories mapped the same locus so that more than one *DFN#* are erroneously assigned.

[c] Different mutations in the same gene can cause ARNSHI and ADNSHI.

[d] The *GJB2* and *GJB6* genes are tightly linked and cannot be distinguished based on linkage analysis.

An outstanding resource for genetic studies is a collection of large extended pedigrees which express the HI trait across generations. Populations with high rates of endogamy are particularly useful since families from these populations are very informative for homozygosity mapping. In fact, many of the ARNSHI loci were mapped in single large, consanguineous families. Due to country-specific cultural norms and close communities, the Pakistani population has one of the highest rates of consanguinity in the world (61.2% in 1990-91, with average inbreeding coefficient (F) = 0.03 [26]). Of the 67 ARNSHI loci, at least 29 were mapped in Pakistani families [27-42; see Chapters 2 and 5 for other loci]. Because of the availability of genotypic data from more than 200 Pakistani families of which >90% are consanguineous, this thesis contains articles on the identification of novel ARNSHI loci within the Pakistani population.

When a locus or genetic variant is identified, genotype-phenotype correlation may be performed to determine the severity of the phenotype for a particular genotype or haplotype and whether there is potential genetic or environmental modification. Genotype-phenotype studies of genetic variants add to the huge knowledge base that is required in order to predict clinical outcomes and facilitate risk assessment and genetic counseling. In addition, for NSHI loci for which the gene has not been identified, phenotypic description of individuals who carry the locus-specific haplotype may facilitate accrual of subjects with the same phenotype and gene identification through further fine mapping.

Eventually hundreds of HI genes may be localized or identified, but this knowledge will not be clinically useful if we cannot perform an efficient strategy for genetic screening. Although technology in the future may provide a "gene chip" for NSHI that will make genetic screening faster and easier, the question remains: Which genes and mutations should be screened for? To understand the public health significance of the various HI gene mutations in different populations, the spectrum of functional variants for each gene must first be examined. No matter what advanced technology is used to carry out genetic screening, without this information the sensitivity and specificity of the screen will be low, which will greatly reduce the utility of genetic testing.

This author was fortunate to work with a huge dataset of consanguineous Pakistani families with non-syndromic hearing impairment, which increased in number from 192 to 260 within a span of two years. Upon joining the Leal lab, 168 families have undergone genome scan but only about half have been completely analyzed, while seven novel loci were

already submitted for publication. Thus part of the thesis work entailed linkage analysis of genome scan and/or fine-mapping markers, and organizing the families into computer files that would facilitate identification of their status within the project pipeline. It was later found that ~40% of the families were linked or had some evidence of linkage to markers that flank known HI genes. Laboratory work for this thesis therefore included sequencing for both candidate genes for NSHI loci and known NSHI genes. Because of the fast pace at which new loci and genes were being discovered, this author also strived to understand how these new genes could possibly impact clinical practice. Towards this end, the determination of functionality of sequence variants in NSHI genes, estimation of prevalence rates of functional variants, and genotype-phenotype correlation through available audiometric data seemed most relevant.

The objectives of this thesis are:

1. *To identify novel loci for non-syndromic hearing impairment;*
2. *To test candidate genes within the genetic intervals of NSHI loci;*
3. *To detect novel variants in known NSHI genes and determine their potential functionality;*
4. *To estimate gene-specific prevalence rates among HI individuals; and*
5. *To correlate genotype with audiometric phenotype.*

Chapter 2 contains articles on five novel ARNSHI loci: *DFNB47*, *DFNB55*, *DFNB62*, *DFNB65* and *DFNB68*. These loci were mapped in consanguineous Pakistani families via two-point and multipoint linkage analyses, with further definition of intervals through homozygosity mapping. Though no NSHI gene for these loci has yet been found, the laboratory continues its current work of sequencing candidate genes.

In Chapter 3, the most commonly studied NSHI gene, *GJB2*, is presented, with a description of allelic frequencies and spectrum of variants within Pakistan. Two other ARNSHI genes, *TMC1* and *TMIE*, were also studied for novel variants in the Pakistani population. These two genes were selected based on the number of families that were mapped to the gene locus, the ease of sequencing due to gene size and G-C content, and the scarcity of published literature on the gene. Due to the large number of families that were ascertained from Pakistan, it was possible to estimate gene-specific prevalence rates for individuals with ARNSHI. A novel functional variant was also identified for the *TMIE* gene in a Jordanian family. However, due to a small number of families from Jordan, it was not possible

to estimate the prevalence of *TMIE* variants among HI individuals in Jordan. Functional variants were identified through the use of bioinformatics tools, which included multiple sequence alignment and transmembrane domain prediction, and allele frequency comparison between cases and controls.

Lastly Chapter 4 contains two articles which describe audiometric phenotype based on genotype. Better audiometric facilities in Northern Europe have resulted in huge repositories of phenotypic data from which decades-old records can be easily retrieved. Thus the first article in this chapter presents the *GJB2* phenotype in a Dutch patient population, while the second article characterizes in detail the audiometric profile of a multigenerational Swiss-German family that was previously mapped to the ADNSHI locus *DFNA24*.

REFERENCES

1. Morton NE. Genetic epidemiology of hearing impairment. Ann NY Acad Sci 1991; 630: 16-31.
2. Reardon W. Genetic deafness. J Med Genet 1992; 29: 521-526.
3. White KR. Early hearing detection and intervention programs: opportunities for genetic services. Am J Med Genet 2004; 130A: 29-36.
4. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2002. Vital and Health Statistics 2004; 10: 38. URL: http://www.cdc.gov/nchs/data/series/sr_10/sr10_222acc.pdf
5. Cohen M, Gorlin R. Chapter 3: Epidemiology, Etiology, and Genetic Patterns. In: Gorlin R, Toriello H, Cohen M. Hereditary Hearing Loss and its Syndromes. New York: Oxford University Press, 1995. pp 19-21.
6. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2006. URL: http://webhost.ua.ac.be/hhh/
7. Riazuddin S, Castelein CM, Ahmed ZM, Lalwani AK, Mastroianni MA, Naz S, Smith TN, Liburd NA, Friedman TB, Griffith AJ, Riazuddin S, Wilcox ER. Dominant modifier DFNM1 suppresses recessive deafness DFNB26. Nat Genet 2000; 26: 431-434.
8. Bykhovskaya Y, Estivill X, Taylor K, Hang T, Hamon M, Casano RA, Yang H, Rotter JI, Shohat M, Fischel-Ghodsian N. Candidate locus for a nuclear modifier gene for maternally inherited deafness. Am J Hum Genet 2000; 66: 1905-1910.
9. Ahmed ZM, Riazuddin S, Riazuddin S, Wilcox ER. The molecular genetics of Usher syndrome. Clin Genet 2003; 63: 431-444.
10. Usami S, Abe S, Weston MD, Shinkawa H, Van Camp G, Kimberling WJ. Non-syndromic hearing loss associated with enlarged vestibular aqueduct is caused by PDS mutations. Hum Genet 1999; 104: 188-192.
11. Bespalova IN, Van Camp G, Bom SJ, Brown DJ, Cryns K, DeWan AT, Erson AE, Flothmann K, Kunst HP, Kurnool P, Sivakumaran TA, Cremers CW, Leal SM, Burmeister

M, Lesperance MM. Mutations in the Wolfram syndrome 1 gene (WFS1) are a common cause of low frequency sensorineural hearing loss. Hum Mol Genet 2001; 10: 2501-2508.
12. Yoshinaga-Itano C, Sedey AL, Coulter DK, Mehl AL. Language of early- and later-identified children with hearing loss. Pediatrics 1998; 102: 1161-1171.
13. Zelante L, Gasparini P, Estivill X, Melchionda S, D'Agruma L, Govea N, Mila M, Monica MD, Lutfi J, Shohat M, Mansfield E, Delgrosso K, Rappaport E, Surrey S, Fortina P. Connexin26 mutations associated with the most common form of non-syndromic neurosensory autosomal recessive deafness (DFNB1) in Mediterraneans. Hum Mol Genet 1997; 6: 1605-1609.
14. Gasparini P, Rabionet R, Barbujani G, Melchionda S, Petersen M, Brondum-Nielsen K, Metspalu A, Oitmaa E, Pisano M, Fortina P, Zelante L, Estivill X. High carrier frequency of the 35delG deafness mutation in European populations. Genetic Analysis Consortium of GJB2 35delG. Eur J Hum Genet 2000; 8: 19-23.
15. Sobe T, Vreugde S, Shahin H, Berlin M, Davis N, Kanaan M, Yaron Y, Orr-Urtreger A, Frydman M, Shohat M, Avraham KB. The prevalence and expression of inherited connexin 26 mutations associated with nonsyndromic hearing loss in the Israeli population. Hum Genet 2000; 106: 50-57.
16. Masmoudi S, Elgaied-Boulila A, Kassab I, Ben Arab S, Blanchard S, Bouzouita JE, Drira M, Kassab A, Hachicha S, Petit C, Ayadi H. Determination of the frequency of connexin26 mutations in inherited sensorineural deafness and carrier rates in the Tunisian population using DGGE. J Med Genet 2000; 37: E39.
17. Hamelmann C, Amedofu GK, Albrecht K, Muntau B, Gelhaus A, Brobby GW, Horstmann RD. Pattern of connexin 26 (GJB2) mutations causing sensorineural hearing impairment in Ghana. Hum Mutat 2001; 18: 84-85.
18. Yan D, Park HJ, Ouyang XM, Pandya A, Doi K, Erdenetungalag R, Du LL, Matsushiro N, Nance WE, Griffith AJ, Liu XZ. Evidence of a founder effect for the 235delC mutation of GJB2 (connexin 26) in east Asians. Hum Genet 2003; 114: 44-50.
19. Kelsell DP, Dunlop J, Stevens HP, Lench NJ, Liang JN, Parry G, Mueller RF, Leigh IM. Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. Nature 1997; 387: 80-83.
20. Cucci RA, Prasad S, Kelley PM, Green GE, Storm K, Willocx S, Cohn ES, Van Camp G, Smith RJ. The M34T allele variant of connexin 26. Genet Test 2000; 4: 335-344.
21. Skerrett IM, Di WL, Kasperek EM, Kelsell DP, Nicholson BJ. Aberrant gating, but a normal expression pattern, underlies the recessive phenotype of the deafness mutant Connexin26 M34T. FASEB J 2004; 18: 860-862.
22. Feldmann D, Denoyelle F, Loundon N, Weil D, Garabedian EN, Couderc R, Joannard A, Schmerber S, Delobel B, Leman J, Journel H, Catros H, Ferrec C, Drouin-Garraud V, Obstoy MF, Moati L, Petit C, Marlin S. Clinical evidence of the nonpathogenic nature of the M34T variant in the connexin 26 gene. Eur J Hum Genet 2004; 12: 279-284.
23. Skerrett IM, Di WL, Kasperek EM, Kelsell DP, Nicholson BJ. Aberrant gating, but a normal expression pattern, underlies the recessive phenotype of the deafness mutant Connexin26M34T. FASEB J 2004; 18: 860-862.
24. Palmada M, Schmalisch K, Bohmer C, Schug N, Pfister M, Lang F, Blin N. Loss of function mutations of the GJB2 gene detected in patients with DFNB1-associated hearing impairment. Neurobiol Dis 2006; 22: 112-118.
25. Nance WE. The genetics of deafness. Ment Retard Dev Disabil Res Rev 2003; 9: 109-119.

26. National Institute of Population Studies. Pakistan demographic and health survey 1990/91. Demographic and Health Surveys. Islamabad, Pakistan: NIPS, 1992.
27. Naz S, Giguere CM, Kohrman DC, Mitchem KL, Riazuddin S, Morell RJ, Ramesh A,Srisailpathy S, Deshmukh D, Riazuddin S, Griffith AJ, Friedman TB, Smith RJ, Wilcox ER. Mutations in a novel gene, TMIE, are associated with hearing loss linked to the DFNB6 locus. Am J Hum Genet 2002; 71: 632-636.
28. Kurima K, Peters LM, Yang Y, Riazuddin S, Ahmed ZM, Naz S, Arnaud D, Drury S, Mo J, Makishima T, Ghosh M, Menon PS, Deshmukh D, Oddoux C, Ostrer H, Khan S, Riazuddin S, Deininger PL, Hampton LL, Sullivan SL, Battey JF Jr, Keats BJ, Wilcox ER, Friedman TB, Griffith AJ. Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. Nat Genet 2002; 30: 277-284.
29. Veske A, Oehlmann R, Younus F, Mohyuddin A, Muller-Myhsok B, Mehdi SQ, Gal A. Autosomal recessive non-syndromic deafness locus (DFNB8) maps on chromosome 21q22 in a large consanguineous kindred from Pakistan. Hum Mol Genet 1996; 5: 165-168.
30. Scott HS, Kudoh J, Wattenhofer M, Shibuya K, Berry A, Chrast R, Guipponi M, Wang J, Kawasaki K, Asakawa S, Minoshima S, Younus F, Mehdi SQ, Radhakrishna U, Papasavvas MP, Gehrig C, Rossier C, Korostishevsky M, Gal A, Shimizu N, Bonne-Tamir B, Antonarakis SE. Insertion of beta-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. Nat Genet 2001; 27: 59-63.
31. Bork JM, Peters LM, Riazuddin S, Bernstein SL, Ahmed ZM, Ness SL, Polomeno R, Ramesh A, Schloss M, Srisailpathy CR, Wayne S, Bellman S, Desmukh D, Ahmed Z, Khan SN, Kaloustian VM, Li XC, Lalwani A, Riazuddin S, Bitner-Glindzicz M, Nance WE, Liu XZ, Wistow G, Smith RJ, Griffith AJ, Wilcox ER, Friedman TB, Morell RJ. Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. Am J Hum Genet 2001; 68: 26-37.
32. Masmoudi S, Charfedine I, Rebeh IB, Rebai A, Tlili A, Ghorbel AM, Belguith H, Petit C, Drira M, Ayadi H. Refined mapping of the autosomal recessive non-syndromic deafness locus DFNB13 using eight novel microsatellite markers. Clin Genet 2004; 66: 358-364.
33. Campbell DA, McHale DP, Brown KA, Moynihan LM, Houseman M, Karbani G, Parry G, Janjua AH, Newton V, al-Gazali L, Markham AF, Lench NJ, Mueller RF. A new locus for non-syndromal, autosomal recessive, sensorineural hearing loss (DFNB16) maps to human chromosome 15q21-q22. J Med Genet 1997; 34: 1015-1017.
34. Ahmed ZM, Smith TN, Riazuddin S, Makishima T, Ghosh M, Bokhari S, Menon PS, Deshmukh D, Griffith AJ, Riazuddin S, Friedman TB, Wilcox ER. Nonsyndromic recessive deafness DFNB18 and Usher syndrome type IC are allelic mutations of USHIC. Hum Genet 2002; 110: 527-531.
35. Moynihan L, Houseman M, Newton V, Mueller R, Lench N. DFNB20: a novel locus for autosomal recessive, non-syndromal sensorineural hearing loss maps to chromosome 11q25-qter. Eur J Hum Genet 1999; 7: 243-246.
36. Wilcox ER, Burton QL, Naz S, Riazuddin S, Smith TN, Ploplis B, Belyantseva I, Ben-Yosef T, Liburd NA, Morell RJ, Kachar B, Wu DK, Griffith AJ, Riazuddin S, Friedman TB. Mutations in the gene encoding tight junction claudin-14 cause autosomal recessive deafness DFNB29. Cell 2001; 104: 165-172

37. Naz S, Griffith AJ, Riazuddin S, Hampton LL, Battey JF Jr, Khan SN, Riazuddin S, Wilcox ER, Friedman TB. Mutations of ESPN cause autosomal recessive deafness and vestibular dysfunction. J Med Genet 2004; 41: 591-595.
38. Ahmed ZM, Morell RJ, Riazuddin S, Gropman A, Shaukat S, Ahmad MM, Mohiddin SA, Fananapazir L, Caruso RC, Husnain T, Khan SN, Riazuddin S, Griffith AJ, Friedman TB, Wilcox ER. Mutations of MYO6 are associated with recessive deafness, DFNB37. Am J Hum Genet 2003; 72: 1315-1322.
39. Ahmad J, Khan SN, Khan SY, Ramzan K, Riazuddin S, Ahmed ZM, Wilcox ER, Friedman TB, Riazuddin S. DFNB48, a new nonsyndromic recessive deafness locus, maps to chromosome 15q23-q25.1. Hum Genet 2005; 116: 407-412.
40. Ramzan K, Shaikh RS, Ahmad J, Khan SN, Riazuddin S, Ahmed ZM, Friedman TB, Wilcox ER, Riazuddin S. A new locus for nonsyndromic deafness DFNB49 maps to chromosome 5q12.3-q14.1. Hum Genet 2005; 116: 17-22.
41. Shaikh RS, Ramzan K, Nazli S, Sattar S, Khan SN, Riazuddin S, Ahmed ZM, Friedman TB, Riazuddin S. A new locus for nonsyndromic deafness DFNB51 maps to chromosome 11p13-p12. Am J Med Genet A 2005; 138: 392-395.
42. Shabbir MI, Ahmed ZM, Khan SY, Riazuddin S, Waryah AM, Khan SN, Camps RD, Ghosh M, Kabra M, Belyantseva IA, Friedman TB, Riazuddin S. Mutations of human TMHS cause recessively inherited nonsyndromic hearing loss. J Med Genet 2006; Epub ahead of print.

Chapter 2
Novel Hearing Impairment Loci

Chapter 2.1

A novel autosomal recessive non-syndromic hearing impairment locus (*DFNB47*) maps to chromosome 2p25.1-p24.3

Hassan MJ, Santos RLP, Rafiq MA, Chahrour MH, Pham TL, Wajid M, Hijab N, Wambangco M, Lee K, Ansar M, Yan K, Ahmad W, Leal SM. Reprinted from Human Genetics 118(5): 605-610, with kind permission of Springer Science and Business Media.

SUMMARY

Hereditary hearing impairment (HI) displays extensive genetic heterogeneity. Autosomal recessive (AR) forms of prelingual HI account for ~75% of cases with a genetic etiology. A novel AR non-syndromic HI locus (*DFNB47*) was mapped to chromosome 2p25.1-p24.3, in two distantly related Pakistani kindreds. Genome scan and fine mapping were carried out using microsatellite markers. Multipoint linkage analysis resulted in a maximum LOD score of 4.7 at markers D2S1400 and D2S262. The three-unit support interval was bounded by D2S330 and D2S131. The region of homozygosity was found within the three-unit support interval and flanked by markers D2S2952 and D2S131, which corresponds to 13.2 cM according to the Rutgers combined linkage-physical map. This region contains 5.3

Mb according to the sequence-based physical map. Three candidate genes, *KCNF1*, *ID2* and *ATP6V1C2* were sequenced, and were found to be negative for functional sequence variants.

INTRODUCTION

Genetic hearing impairment (HI) can be classified as either syndromic or non-syndromic. Autosomal recessive (AR) inheritance predominates in hereditary non-syndromic HI (NSHI), and accounts for ~75% of the cases, whereas autosomal dominant forms are observed in 15% of cases. Mitochondrial and X-linked NSHI, on the other hand, are less frequent [1]. HI in AR forms is usually due to a sensorineural defect [2], and is generally prelingual, not progressive and all frequencies are affected with severe to profound HI. NSHI is the most heterogeneous trait known. To date 46 loci for ARNSHI have been mapped and 21 genes have been isolated [3]. A large number of genes with different functions are involved in the etiology of HI because of the complexity of the inner ear, and the various mechanisms that can lead to the HI phenotype [4]. This article describes the mapping of a novel ARNSHI locus, *DFNB47*, to chromosome 2p25.1-p24.3 in two distantly related consanguineous Pakistani kindreds.

MATERIALS AND METHODS

Family history

Approval was obtained from the Institutional Review Boards of the Quaid-I-Azam University and Baylor College of Medicine and Affiliated Hospitals prior to the onset of the study. Informed consent was obtained from all family members who participated in the study. A number of pedigrees were ascertained including pedigrees 4052a and 4052b, which are distantly related consanguineous families from Sadiq Abad in Pakistan. Information obtained during interviews with multiple family members was used to construct the pedigrees and to clarify consanguineous relationships. Although various pedigree members reported that the two family branches were related, the exact relationship could not be specified. In the two pedigrees, 4052a and 4052b (Figure 2.1), the HI displayed an AR mode of inheritance. Affected individuals manifest prelingual profound HI and use sign language for communication. Physical examination of hearing-impaired individuals did not reveal any syndromic features, such as maxillofacial or limb deformities, visual loss and mental deficiency. In addition, no gross vestibular involvement was noted in the clinical history and physical examination. Due to the remote location of the family, further vestibular

testing, electroretinography, and temporal bone imaging were not feasible. An audiogram from hearing-impaired individual 14 of pedigree 4052a shows bilateral, profound HI that affects all frequencies (Figure 2.2).

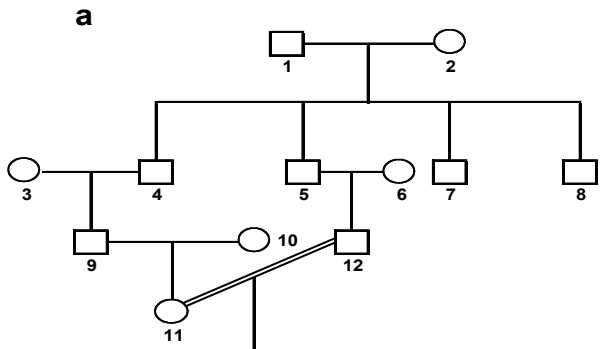
Extraction of genomic DNA and genotyping

Venous blood samples were collected from a total of ten individuals, including five who are hearing-impaired. Genomic DNA was extracted from whole blood following a standard protocol [5], quantified by spectrophotometric readings at optical density 260, and diluted to 40 ηg/μl for polymerase chain reaction (PCR) amplification. A genome scan was carried out on all the DNA samples at the Center for Inherited Disease Research (CIDR). A total of 396 fluorescently labeled short tandem repeat (STR) markers were genotyped. These markers were spaced ~10 cM apart and are located on the 22 autosomes and the X and Y chromosomes.

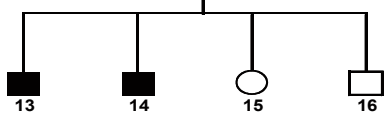
For fine mapping, PCR for microsatellite markers were performed according to standard procedure in a total volume of 25 μl with 40 ηg of genomic DNA, 0.3 μl of each primer, 200 μM dNTP and 1x PCR buffer (Fermentas Life Sciences, Burlington, ON, Canada). PCR was carried out for 35 cycles: 95°C for 1 min, 57°C for 1 min and 72°C for 1 min in a thermal cycler (PerkinElmer Life and Analytical Sciences, Boston, MA, USA). PCR products were resolved on 8% non-denaturing polyacrylamide gel and genotypes were assigned by visual inspection.

Linkage analysis

The order of genome scan markers and fine mapping markers were determined based on the National Center for Biotechnology Information (NCBI) Build 34 sequence-based physical map [6]. The Rutgers combined linkage-physical map of the human genome [7] was utilized for genetic map distances in multipoint linkage analysis for the fine mapping and genome scan markers. The fine-mapping marker D2S1861 was not found on the NCBI sequence-based physical map and was placed on the sequence-based physical map using e-PCR [8], then the genetic map position was deduced by interpolation on the Rutgers combined linkage-physical map. PEDCHECK [9] was used to identify Mendelian inconsistencies while the MERLIN [10] program was utilized to detect potential genotyping errors that did not produce a Mendelian inconsistency. Haplotypes were constructed using SIMWALK2 [11-12]. Two-point linkage analysis was carried out on all autosomal markers from the genome scan using the MLINK program of the FASTLINK computer package [13]. Multipoint linkage analysis was



D2S1861	1 1	2 1
D2S330	2 2	2 1
D2S2952	8 4	8 4
D2S2164	2 1	2 1
D2S423	2 1	2 1
D2S1400	1 2	3 2
D2S262	3 1	4 1
D2S131	2 1	3 2
D2S2267	1 1	1 1
D2S149	2 2	1 2



D2S1861	1 1	1 1
D2S330	2 1	2 1
D2S2952	4 4	4 4
D2S2164	1 1	1 1
D2S423	1 1	1 1
D2S1400	2 2	2 2
D2S262	1 1	1 1
D2S131	1 2	2 2
D2S2267	1 1	1 1
D2S149	2 2	2 2

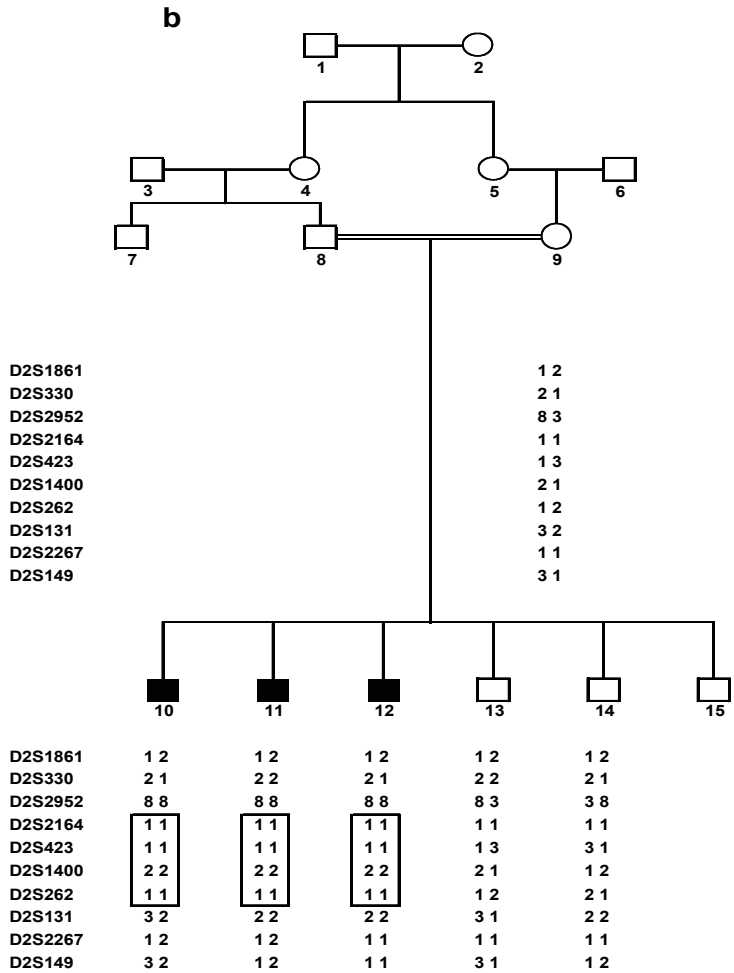


Figure 2.1 Drawing of pedigrees 4052a and 4052b which segregate *DFNB47*. Black symbols represent individuals with hearing impairment (HI) due to *DFNB47*. Clear symbols represent unaffected individuals. Haplotypes for the most closely linked STRs are shown below each symbol. The haplotype for *DFNB47* is displayed in a box.

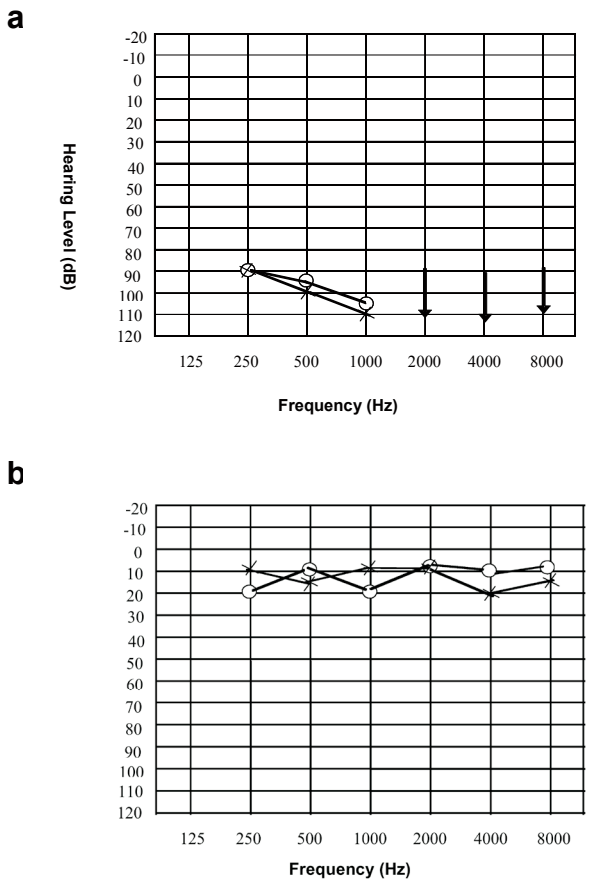


Figure 2.2 Audiograms from (a) hearing-impaired individual 14 of pedigree 4052a, and (b) unaffected individual 16 of pedigree 4052a. Circles and crosses represent air conduction for right and left ear, respectively. Arrows represent residual hearing. The audiogram from the hearing-impaired family member shows that *DFNB47* is associated with bilateral, profound HI that affects all frequencies.

performed using ALLEGRO [14]. An AR mode of inheritance with complete penetrance and a disease allele frequency of 0.001 were used for the analysis. For the genome scan markers, allele frequencies were estimated from the founders and reconstructed genotypes of founders from the two pedigrees (4052a and 4052b) and 43 additional pedigrees that underwent a genome scan at CIDR at the same time. Equal allele frequencies were used for the fine mapping markers, because it was not possible to estimate allele frequencies from the founders, since these markers were only genotyped in the two families. A sensitivity analysis was carried out in order to evaluate whether a false positive result had occurred due to using incorrect allele frequencies [15]. Multipoint linkage analysis was performed by varying the allele frequency for the allele segregating with the disease allele from 0.2 to 0.8 for the fine mapping markers.

Sequencing of candidate genes

Primers were designed for the exons and 1000 base pairs upstream of the first exon (promoter region) of the *KCNF1* (MIM 603787; NM_002236), *ID2* (MIM 600386; NM_002166) and *ATP6V1C2* (NM_144583) genes using Primer3 software [16]. DNA from hearing-impaired individuals 14 of pedigree 4052a and 12 of pedigree 4052b, plus DNA from unaffected individual 13 of family 4052b, were diluted to 5 μg/ηl, amplified by PCR under standard conditions, then purified with ExoSAP-IT (USB Corp., Cleveland, OH, USA). Sequencing was performed with the BigDye Terminator v3.1 Cycle Sequencing Kit together with an Applied Biosystems 3700 DNA Analyzer (Applera Corporation, Foster City, CA, USA). Sequence variants were identified via Sequencher™ Version 4.1.4 software (Gene Codes Corporation, Ann Arbor, MI, USA).

RESULTS

Two-point linkage analysis of the genome scan markers gave a maximum LOD score of 1.3 (θ=0) at marker D2S1360 for pedigree 4052a, and a LOD score of 1.5 (θ=0) at marker D2S2952 for pedigree 4052b. The maximum multipoint LOD score for the genome scan markers was 3.0 when the scores for the two branches of the family were summed, and it was obtained at marker D2S2952. In order to fine map the region on chromosome 2, twelve additional markers were selected from the Marshfield genetic map [17]. Two markers (D2S1861, D2S330) were proximal to genome scan marker D2S2952, another two (D2S2164, D2S423) were between D2S2952 and genome scan marker D2S1400, while eight markers (D2S262, D2S131,

D2S2267, D2S149, D2S2346, D2S272, D2S320 and D2S2375) lie distal to D2S1400. Analysis of the marker genotypes within this region with PEDCHECK and MERLIN did not elucidate any genotyping errors. Genotypes for the fine mapping markers were analyzed using two-point and multipoint linkage analysis. The maximum two-point LOD score was at marker D2S2952 with a value of 3.0 at θ=0 (Table 2.1). Multipoint linkage analysis gave a maximum LOD score of 4.7 at markers D2S1400 and D2S262. When the marker allele frequencies were varied for the fine mapping markers from 0.2 to 0.8, the maximum multipoint LOD score varied from 4.7 to 4.5 respectively, and remained at markers D2S1400 and D2S262. The three-unit support interval is flanked by markers D2S330 and D2S131. This region is 17.9 cM according to the Rutgers combined linkage-physical map of the human genome, and corresponds to 6.6 Mb on the sequence-based physical map (Table 2.1).

Haplotypes were constructed to determine the critical linkage interval. A historic recombination event between markers D2S262 and D2S131 defined the centromeric boundary of this interval, and it was observed in hearing-impaired individual 13 of pedigree 4052a and individuals 11 and 12 of pedigree 4052b. All hearing-impaired individuals were homozygous at genome scan marker D2S2952, but for different alleles for each pedigree (Figure 2.1). Because these two pedigrees are historically related and are thus assumed to inherit the disease haplotype from common founders, the telomeric boundary is then assigned between markers D2S2952 and D2S2164. The region of homozygosity was therefore contained within the three-unit support interval and flanked by markers D2S2952 and D2S131. This narrowed down the *DFNB47* interval to a physical map distance of 5.3 Mb, and to 13.2 cM according to the Rutgers combined linkage-physical map of the human genome.

Three candidate genes within the *DFNB47* interval, namely *KCNF1*, *ID2* and *ATP6V1C2*, were screened in three family members, and were found to be negative for functional sequence variants.

DISCUSSION

The linkage data presented here map *DFNB47* to a 13.2 cM-interval on chromosomal region 2p25.1-p24.3, according to the Rutgers combined linkage-physical map of the human genome. Two loci for ARNSHI have previously been localized to chromosome 2, *DFNB9* (2p22-p23) and *DFNB27* (2q23-q31), while two autosomal dominant NSHI loci, *DFNA16* (2q24) and *DFNA43* (2p12), have been mapped. Only the gene for *DFNB9*, *OTOF*

(MIM 603681), has been identified [3]. *OTOF* (NM_194248; 26,654,606-26,756,101) is positioned ~20 Mb distal to the *DFNB47* interval according to Build 34 of the human genome sequence [6]. By interpolation onto the Rutgers map, this places *OTOF* approximately 17 cM centromeric to *DFNB47*.

To date 23 known genes lie in the 5.3 Mb-region that contains *DFNB47*. One of the genes in this region, *KCNF1*, is a strong candidate for *DFNB47*. This gene codes for potassium voltage-gated channel subfamily F member 1. Potassium ion channels are a diverse family of plasma membrane proteins that play an essential role in various cellular processes, including maintenance of membrane potential and cell signalling [18]. *KCNQ4* (MIM 603537), a voltage-gated K^+ channel gene expressed in the cochlea, has been previously mapped to the *DFNA2* locus (1p34.2) for a form of autosomal dominant NSHI [19]. Voltage-gated K^+ channel genes have been shown to be responsible for various hereditary diseases. For instance, mutations in the *KVLQT1* gene (MIM 607542) (a voltage-gated K^+ channel gene) result in Jervell and Lange-Nielson syndrome (JLNS) and Long QT syndrome, which are inherited AR diseases, with congenital HI being one of their characteristics [20]. JLNS can also result from mutations in another voltage-gated K^+ channel gene, *KCNE1* (MIM 176261). Although *KCNF1* is the strongest candidate gene within the *DFNB47* interval, this gene was screened in family 4052 and was found to be negative for sequence variants.

Another good candidate gene is inhibitor of DNA binding 2 (*ID2*), which is a member of the ID family of genes that promotes cell proliferation. In embryonic mouse, *ID2* expression was detected in the vestibular and acoustic ganglia, and also in the epithelium of the otic vesicle and surrounding mesenchyme [21]. However, no functional sequence variants were found in *ID2* for family 4052. Other genes that are expressed in the inner ear [22-23] include: (1) cleavage and polyadenylation specific factor 3 (*CPSF3*; MIM 606029), an important regulator of viral and cellular gene expression [24]; (2) tyrosine 3/tryptophan 5 –monooxygenase (*YWHAQ*; MIM 609009), which is also expressed in the spinal cord of patients with amyotrophic lateral sclerosis [25]; and (3) ornithine decarboxylase 1 (*ODC1*; MIM 165640), the rate-limiting enzyme in polyamine synthesis.

ATP6V1C2 (ATPase, H^+ transporting, lysosomal 42kDa, V1) is an isoform of an H^+-ATPase subunit [26], which has been implicated in AR distal renal tubular acidosis (dRTA). This gene was screened in eight dRTA families, two of whom had sensorineural HI, but no functional sequence

Table 2.1 Two-point LOD score results between *DFNB47* and chromosome 2 fine mapping markers

Marker [a]	Genetic map position [b]	Physical map position [c]	LOD score at *theta*=							
			0.0	0.01	0.02	0.04	0.05	0.1	0.2	0.3
D2S1861	10.17	4,955,741	- 1.14	-0.49	-0.27	-0.06	0	0.14	0.14	0.08
D2S330	**13.94**	**6,802,851**	**- ∞**	**-3.40**	**-2.65**	**-1.87**	**-1.62**	**-0.91**	**-0.34**	**-0.13**
D2S2952	*18.57*	*8,099,709*	*2.99*	*2.91*	*2.84*	*2.68*	*2.60*	*2.20*	*1.45*	*0.79*
D2S2164	19.11	8,248,730	1.12	1.09	1.05	0.98	0.94	0.77	0.45	0.21
D2S423	23.72	9,962,317	2.56	2.49	2.42	2.28	2.21	1.87	1.22	0.65
D2S1400	*28.51*	*11,628,666*	*2.52*	*2.45*	*2.38*	*2.24*	*2.18*	*1.84*	*1.21*	*0.66*
D2S262	28.51	11,920,777	2.82	2.74	2.67	2.52	2.45	2.08	1.38	0.77
D2S131	**31.80**	**13,389,278**	**- ∞**	**-2.71**	**-1.85**	**-1.06**	**-0.82**	**-0.18**	**0.19**	**0.21**
D2S2267	33.08	13,757,406	- ∞	-2.18	-1.66	-1.14	-0.98	-0.53	-0.20	-0.08
D2S149	33.25	14,414,735	- ∞	-3.33	-2.47	-1.65	-1.39	-0.69	-0.18	-0.02

LOD scores are sums of two-point LOD scores for pedigrees 4052a and 4052b.

[a] Genome scan markers are shown in italics. Markers in bold type flank the three-unit support interval.

[b] Sex-averaged Kosambi map distance (cM) from the Rutgers combined linkage-physical map [7].

[c] Sequence-based physical map distance in bases according to Build 34 of the human reference sequence [6].

variants were reported to have been found [26]. Likewise, in family 4052, *ATP6V1C2* had negative results after sequencing.

No functional sequence variants were identified in the three candidate genes *KCNF1*, *ID2* and *ATP6V1C2*. Because only the exonic and promoter regions were sequenced, the possibility of a functional variant in the intronic regions cannot be ruled out. On the other hand, 20 other known genes and 7 expressed sequence tags that encode hypothetical proteins are contained within the *DFNB47* interval. Further fine-mapping and sequencing work are required in order to identify the *DFNB47* gene which causes ARNSHI.

Acknowledgements We wish to thank the family members for their invaluable participation and cooperation. This work was supported by the Higher Education Commission, Pakistan, The American Hearing Research Foundation, and the National Institutes of Health National Institute of Deafness and other Communication Disorders grant R01-DC03594. Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, Contract Number N01-HG-65403.

REFERENCES

1. Morton NE. Genetic epidemiology of hearing impairment. Ann N Y Acad Sci 1991; 630: 16–31.
2. Petit C. Genes responsible for human hereditary deafness: symphony of a thousand. Nat Genet 1996; 14: 385–391.
3. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
4. Steel KP, Bussoli TJ. Deafness genes: expressions of surprise. Trends Genet 1999; 15: 207-211.
5. Grimberg J, Nawoschik S, Bellusico L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from Blood. Nucleic Acids Res 1989; 17: 8390.
6. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860-921.
7. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
8. Schuler GD. Sequence mapping by electronic PCR. Genome Res 1997; 7: 541-550.
9. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63: 259-266.
10. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30: 97-101.
11. Weeks DE, Sobel E, O'Connell JR, Lange K. Computer programs for multilocus haplotyping of general pedigrees. Am J Hum Genet 1995; 56: 1506-1507.

12. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996; 58: 1323-1337.
13. Cottingham R, Indury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993; 53: 252-263.
14. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12-13.
15. Freimer NB, Sandkuijl LA, Blower SM. Incorrect specification of marker allele frequencies: effects on linkage analysis. Am J Hum Genet 1993; 52: 1102-1110.
16. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. Bioinformatics methods and protocols: methods in molecular biology. New Jersey: Humana Press, 2000. pp 365-386.
17. Broman K, Murray JC, Scheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex specific variation in recombination. Am J Hum Genet 1998; 63: 861-869.
18. Su K, Kyaw H, Fan P, Zeng Z, Shell BK, Carter KC, Li Y. Isolation, characterization, and mapping of two human potassium channels. Biochem Biophys Res Commun 1997; 241: 675-681.
19. Kubisch C, Schroeder BC, Friedrich T, Lutjohann B, El-Amraoui A, Marlin S, Petit C, Jentsch TJ. KCNQ4, a novel potassium channel expressed in sensory outer hair cells, is mutated in dominant deafness. Cell 1999; 96: 437-446.
20. Ilhan A, Tuncer C, Komsuoglu SS, Kali S. Jervell and Lange-Nielsen syndrome: neurologic and cardiologic evaluation. Pediatr Neurol 1999; 21: 809-813.
21. Jen Y, Manova K, Benezra R. Each member of the Id gene family exhibits a unique expression pattern in mouse gastrulation and neurogenesis. Dev Dyn 1997; 208: 92-106.
22. The Hearing Research Group at Brigham & Women's Hospital. Human Cochlear cDNA library and EST database. 2002. URL: http://hearing.bwh.harvard.edu/estinfo.HTM
23. Holme RH, Bussoli TJ, Steel KP. Table of gene expression in the developing ear. 2002. URL: http://www.ihr.mrc.ac.uk/Hereditary/ genetable/search.shtml
24. Calzado MA, Sancho R, Munoz E. Human immunodeficiency virus type 1 Tat increases the expression of cleavage and polyadenylation specificity factor 73-kilodalton subunit modulating cellular and viral expression. J Virol 2004; 78: 6846-6854.
25. Malaspina A, Kaushik N, de Belleroche J. A 14-3-3 mRNA is up-regulated in amyotrophic lateral sclerosis spinal cord. J Neurochem 2000; 75: 2511-2520.
26. Smith AN, Borthwick KJ, Karet FE. Molecular cloning and characterization of novel tissue-specific isoforms of the human vacuolar H(+)-ATPase C, G and d subunits, and their evaluation in autosomal recessive distal renal tubular acidosis. Gene 2002; 297: 169-177.

Chapter 2.2

Localization of a novel autosomal recessive non-syndromic hearing impairment locus *DFNB55* to chromosome 4q12-q13.2

Irshad S, Santos RLP, Muhammad D, Lee K, McArthur N, Haque S, Ahmad W, Leal SM. Reprinted from Clinical Genetics 68(3): 262-267.

SUMMARY

Hereditary hearing impairment (HI) is the most genetically heterogeneous trait known in humans. So far, 54 autosomal recessive nonsyndromic hearing impairment (ARNSHI) loci have been mapped, and 21 ARNSHI genes have been identified. Here is reported the mapping of a novel ARNSHI locus, *DFNB55*, to chromosome 4q12-q13.2 in a consanguineous Pakistani family. A maximum multipoint LOD score of 3.5 was obtained at marker D4S2638. The region of homozygosity and the three-unit support interval are flanked by markers D4S2978 and D4S2367. The region spans 8.2 cM on the Rutgers combined linkage-physical map and contains 11.5 Mb. *DFNB55* represents the third ARNSHI locus mapped to chromosome 4.

INTRODUCTION

Hearing disorders produce significant health problems in the world population. The vast majority of genetic hearing impairment (HI) is designated as non-syndromic, which is further categorized by mode of inheritance: approximately 77% of cases are autosomal recessive; 22% are autosomal dominant; 1% are X-linked; and <1% are due to mitochondrial inheritance [1]. Autosomal recessive non-syndromic HI (ARNSHI) is usually clinically homogeneous, is non-progressive in nature and exhibits a high degree of genetic heterogeneity. To date, 54 loci have been mapped for ARNSHI and 21 genes have been identified [2]. The majority of these genes were mapped and identified by studying large consanguineous families. Identification of genetic defects causing HI contributes to the understanding of the molecular basis that is important to hearing function. This article describes a new ARNSHI locus, *DFNB55*, which maps to 4q12-q13.2 and was localized in a consanguineous Pakistani family.

MATERIALS AND METHODS

Family history

Before the start of the study, approval was obtained from the Quaid-i-Azam University and Baylor College of Medicine Institutional Review

Boards. Signed informed consent was obtained from all family members who participated in the study. The pedigree structure is based upon interviews with multiple family members. Personal interviews with key figures in the kindred clarified the consanguineous relationships. For the analysis correct specification of consanguineous relationships is extremely important, as misspecification of familial consanguineous relations will increase type I and II error. Family 4153 is from the Sind province of Pakistan. This consanguineous pedigree provides convincing evidence of autosomal recessive mode of inheritance (Figure 2.3). Clinical findings in this family are consistent with the diagnosis of ARNSHI. Medical history and physical examination of the affected individuals were performed by trained otolaryngologists affiliated with government hospitals. All affected individuals have a history of prelingual profound HI involving all frequencies and use sign language for communication. The hearing-impaired family members underwent a clinical examination for mental retardation, defects in ear morphology, dysmorphic facial features, eye disorders including night blindness and tunnel vision, and other clinical features that could indicate that HI was syndromic. There was no evidence in this kindred that HI belonged to a syndrome or that there was gross vestibular involvement.

Extraction of genomic DNA and genotyping

Venous blood samples were obtained from 7 family members including 4 individuals who are hearing-impaired. Genomic DNA was extracted from whole blood following a standard protocol [3]. Since the *GJB2* gene is the most frequent cause of ARNSHI, all non-syndromic HI families are screened for mutations in this gene before undergoing a genome scan. Those families which are positive for functional *GJB2* variants are excluded from the genome scan. In family 4153 the *GJB2* gene was sequenced in two hearing-impaired family members, V-5. and V-6. [4], and no functional variants or benign polymorphisms were observed. A genome scan was carried out on seven DNA samples at the National Heart, Lung and Blood Institute (NHLBI) Mammalian Genotyping Service (Center for Medical Genetics, Marshfield, WI, USA). A total of 410 short tandem repeat polymorphism (STR) markers with an average heterozygosity of 0.75 were genotyped. These markers are spaced approximately 10 cM apart and are located on the 22 autosomes and the X and Y chromosomes. After the completion of the genome scan, three additional unaffected (V-6., V-7. and V-8.) family members were ascertained and their DNA samples were used for fine mapping of the *DFNB55* locus.

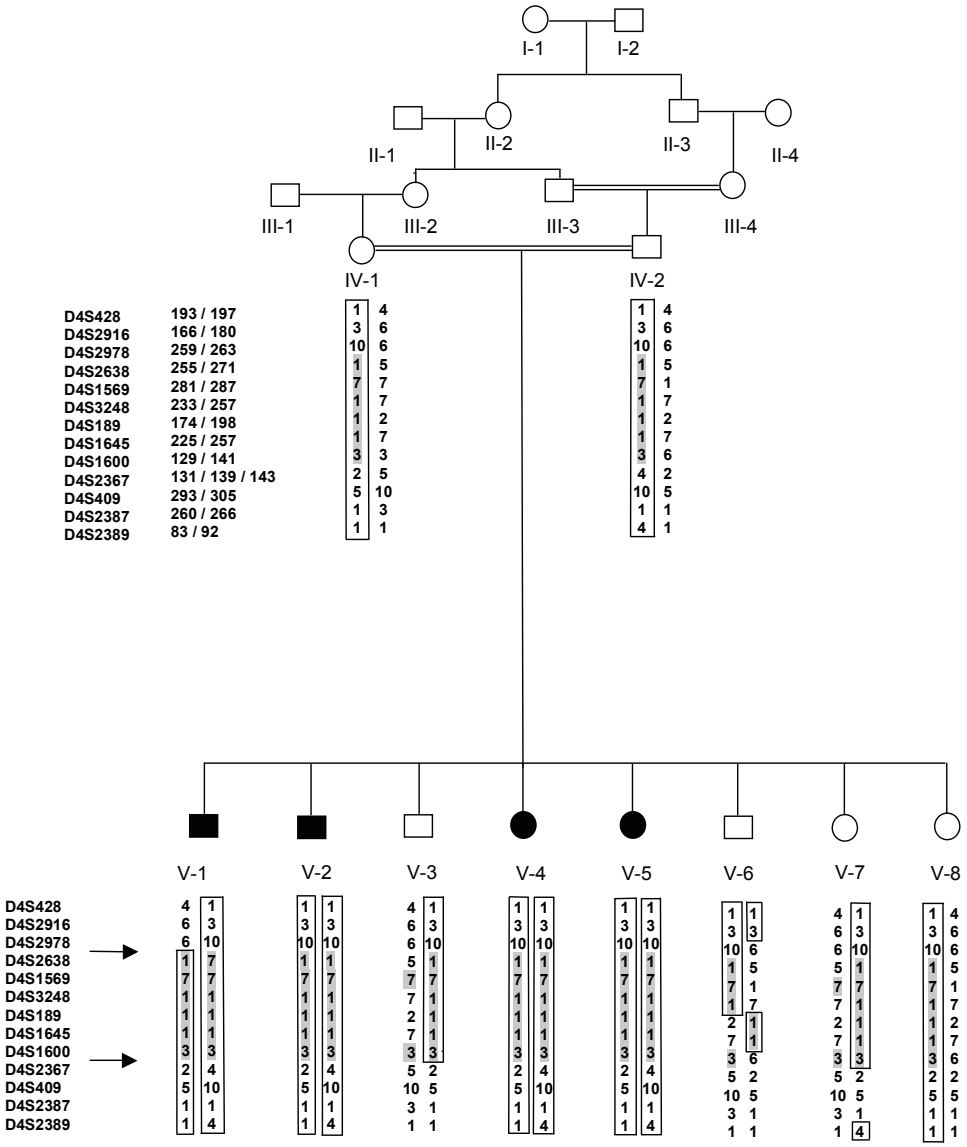


Figure 2.3 Pedigree drawing of family 4153. Black symbols represent individuals with hearing impairment. Clear symbols represent unaffected individuals. The sex of some of the family members was changed to protect the anonymity of the family. Allele sizes in kilobases are indicated beside the upper column of microsatellite markers. Alleles are numbered according to the CEPH Genotype database V10.0 [5]. Haplotypes are shown below each individual for whom genotypes are available. For individuals in generation V maternal haplotypes are displayed on the left-hand side and paternal haplotypes on the right-hand side. Arrows and shaded areas indicate the region of homozygosity.

For fine mapping, polymerase chain reactions (PCR) for microsatellite markers were performed according to standard procedure in a total volume of 25 μl with 40 ηg of genomic DNA, 0.3 μl of each primer, 200 μM of dNTP and 1x PCR buffer (Fermentas Life Sciences, Burlington, ON, Canada). PCR was carried out for 35 cycles: 95°C for 1 min, 57°C for 1 min and 72°C for 1 min in a thermal cycler (PerkinElmer Life and Analytical Sciences, Boston, MA, USA). PCR products were resolved on 8% non-denaturing polyacrylamide gel and genotypes were assigned by visual inspection. Alleles were numbered according to the CEPH Genotype database V10.0 [5].

Linkage analysis

The National Center for Biotechnology Information (NCBI) Build 34 sequence-based physical map was used to determine the order of the genome scan markers and fine mapping markers [6]. Several of the genome scan markers were not found on the NCBI sequence-based physical map and were placed on the sequence-based physical map using e-PCR [7]. Genetic map distances according to the Rutgers combined linkage-physical map of the human genome were used to carry out the multipoint linkage analysis for the fine map and genome scan markers [8]. For those genome scan markers for which no genetic map position was available, interpolation was used to place these markers on the Rutgers combined linkage-physical map. PEDCHECK [9] was used to identify Mendelian inconsistencies while the MERLIN [10] program was used to detect potential genotyping errors that did not produce a Mendelian inconsistency. Haplotypes were constructed using SIMWALK2 [11-12]. Two-point linkage analysis was carried out using the MLINK program of the FASTLINK computer package for the genome scan and fine mapping marker loci [13]. Multipoint linkage analysis was performed using ALLEGRO [14]. An autosomal recessive mode of inheritance with complete penetrance and a disease allele frequency of 0.001 were used. For the genome scan, marker allele frequencies were estimated from the founders and reconstructed genotypes of founders from this family and 34 additional families from Pakistan that underwent a genome scan at the same time at the NHLBI Mammalian Genotyping Service. For the fine mapping markers, it was not possible to estimate allele frequencies from the founders because these markers were genotyped only in this family. False positive results can be obtained when analyzing the data using too low of an allele frequency for the allele segregating with the disease locus [15]. Therefore, a sensitivity analysis was carried out for the multipoint linkage analysis by varying the allele frequency of the marker alleles that

are segregating with the disease locus from 0.2 to 0.8 for the fine mapping markers. This was performed in order to determine whether the multipoint LOD scores are robust to allele frequency misspecification.

RESULTS

From the genome scan data a maximum two-point LOD score of 1.8 (θ=0) was obtained at marker D4S3248 and a maximum multipoint LOD score of 2.6 was obtained at marker D4S3248. In order to establish linkage and fine-map the *DFNB55* locus on chromosome 4, fifteen additional polymorphic microsatellite markers were selected from the Marshfield genetic map [16]. Eight markers are proximal to D4S3248 (D4S2632, D4S1627, D4S2996, D4S428, D4S2916, D4S2978, D4S2638 and D4S1569) and seven of the markers distal to D4S3248 (D4S189, D4S1645, D4S1600, D4S409, D4S2387, D4S2389 and D4S1517). These markers and the genome scan markers in the region were genotyped in the seven family members who were originally included in the genome scan and in three additional family members that were ascertained after completion of the genome scan. Analysis of the marker genotypes within this region with PEDCHECK and MERLIN did not elucidate any genotyping errors. For the genotype data on the additional marker loci and family members, a maximum two-point LOD score of 2.7 (θ=0) was obtained with marker D4S1645 (Table 2.2).

A maximum multipoint LOD score of 3.5 was obtained at marker D4S2638. The three-unit multipoint support interval is an 8.2 cM region according to the Rutgers combined linkage-physical map of the human genome and spans from marker D4S2978 to marker D4S2367 [8]. Haplotypes were then constructed to determine the critical recombination events (Figure 2.3). The region of homozygosity in HI individuals was also flanked by markers D4S2978 and D4S2367. This region corresponds to a physical map distance of 11.5 Mb [6].

When the marker allele frequencies for the alleles segregating with HI were varied for the fine mapping markers from 0.2 to 0.4, the maximum multipoint LOD score remained to be 3.5 at marker D4S2638. When the allele frequencies were varied between 0.6 and 0.8, the maximum LOD score occurred at marker D4S2638 but decreased to 3.4.

DISCUSSION

In this study a novel ARNSHI locus, *DFNB55*, was mapped to an 8.2 cM interval on chromosome 4q12-q13.2. Seven nonsyndromic hearing impair-

Table 2.2 Two point LOD score results between the *DFNB55* locus and chromosome 4 markers

Marker [a]	Genetic map position [b]	Physical map position [c]	LOD score at *theta*=							
			0.0	0.01	0.02	0.04	0.05	0.1	0.2	0.3
D4S428	71.42	55,549,948	-∞	-1.12	-0.58	-0.09	0.05	0.37	0.42	0.26
D4S2916	72.35	55,728,673	-∞	-1.12	-0.58	-0.09	0.05	0.37	0.42	0.26
D4S2978	*73.97*	*56,831,962*	-∞	*0.57*	*0.82*	*1.00*	*1.04*	*1.04*	*0.77*	*0.42*
D4S2638	75.85	58,503,365	2.63	2.57	2.51	2.38	2.32	2.00	1.37	0.76
D4S1569	76.81	59,564,470	1.13	1.10	1.08	1.03	1.00	0.88	0.62	0.37
D4S3248	**76.81**	**60,024,835**	**2.63**	**2.57**	**2.51**	**2.38**	**2.32**	**2.00**	**1.37**	**0.76**
D4S189	77.32	60,668,220	2.63	2.57	2.51	2.38	2.32	2.00	1.37	0.76
D4S1645	77.37	61,986,341	2.75	2.68	2.62	2.48	2.42	2.08	1.43	0.79
D4S1600	78.57	62,900,873	1.13	1.10	1.08	1.03	1.00	0.88	0.62	0.37
D4S2367	***82.19***	***68,286,182***	***0.19***	***1.04***	***1.27***	***1.46***	***1.50***	***1.54***	***1.27***	***0.82***
D4S409	82.19	68,327,249	-1.65	-0.79	-0.54	-0.30	-0.24	-0.07	-0.01	-0.02
D4S2387	82.60	69,105,934	1.37	1.35	1.32	1.27	1.24	1.10	0.82	0.52
D4S2389	85.86	73,110,824	-0.93	-0.18	0.06	0.28	0.34	0.49	0.46	0.29

[a] Markers displayed in italics flank the haplotype. Genome scan markers are shown in bold type.
[b] Cumulative sex-averaged Kosambi cM genetic map distances from the Rutgers combined linkage-physical map of the human genome [8].
[c] Sequence-based physical map distance in bases according to Build 34 of the human reference sequence [6].

ment (NSHI) loci, *DFNA6* (4p16.3) [17], *DFNA14* (4p16) [17], *DFNA24* (4q35) [18], *DFNA27* (4q12) [19], *DFNA39* (4q21.3) [20], *DFNB25* (4p15.3-q12) [21] and *DFNB26* (4q31) [22], have been previously mapped to chromosome 4. For these seven loci, two genes, *WFS1* for *DNA6/DFNA14* [17] and *DSPP* for *DFNA39* [20], have been identified.

The genetic region for *DFNB55* overlaps with the autosomal dominant NSHI locus *DFNA27* [19]. The *DFNA27* locus maps between markers D4S428 (71.42 cM) and D4S392 (83.58 cM) and thus the genetic interval for *DFNB55* is within the genetic interval for *DFNA27*. Although two genes in close proximity might cause *DFNB55* and *DFNA27*, it is also possible that NSHI in *DFNB55* and *DFNA27* families is caused by different mutations in the same gene. It has been observed that different mutations in the same gene can cause both autosomal dominant and recessive NSHI (e.g. *GJB2*, *MYO7A*, *TECTA* and *TMC1*) [2]. It should also be noted that the *DFNB55* locus does not overlap with the *DFNB25* locus (Richard Smith, personal communication).

The *DFNB55* interval contains 14 known genes, and a large number of hypothetical genes and expressed sequenced tags (ESTs). Several of the known genes in this region are expressed in the inner ear [23-24], namely: ephrin receptor *EPHA5* (MIM 600004), *PPAT* (MIM 172450), *POLR2B* (MIM 180661) and *IGFBP7* (MIM 602867). In particular, *EPHA5*, a member of a family of ligands that are implicated in the development and spatial patterning of cortical pathways to target organs such as the retina and the olfactory bulb [25-26], has been immunolocalized to the apical and basal portions of supporting cells which flank utricular hair cells in the rat [27]. Likewise *EPHA5* was detected in lateral wall fibrocytes and interdental cells in the spiral limbus of postnatal and adult gerbil cochlea [28]. All exons of *EPHA5* were sequenced in one hearing and two hearing-impaired members of family 4153, and were found to be negative for functional sequence variants.

An additional strong candidate gene found within the *DFNB55* region is RE1-silencing transcription factor (*REST* [MIM 600571]), which is a transcriptional repressor of neural genes in non-neural cells [29]. The *REST* gene was found to be expressed in supporting cells but not in hair cells of chick auditory epithelium [30]. Additionally *REST* mRNA is upregulated in both supporting cells and hair cells following gentamicin damage to the inner ear. The coding regions of the *REST* gene were sequenced in one unaffected and two hearing-impaired family members and no functional variant was identified.

Acknowledgements We thank the family members for their invaluable participation and cooperation. This work was supported by the NHLBI Mammalian Genotyping Service, the Higher Education Commission, Islamabad, Pakistan, and the National Institutes of Health National Institute of Deafness and other Communication Disorders grant R01-DC03594.

REFERENCES

1. Morton NE. Genetic epidemiology of hearing impairment. Ann NY Acad Sci 1991; 630: 16-31.
2. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
3. Grimberg J, Nawoschik S, Bellusico L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
4. Santos RLP, Wajid M, Pham TL, Hussan J, Ali G, Ahmad W, Leal SM. Low prevalence of Connexin 26 (GJB2) variants in Pakistani families with autosomal recessive non-syndromic hearing impairment. Clin Genet 2005; 67: 61-68.
5. Murray JC, Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Manion F, Quillen J, Sheffield VC, Sunden S, Duyk GM, Weissenbach J, Gyapay G, Dib C, Morrissette J, Lathrop GM, Vignal A, White R, Matsunami N, Gerken S, Melis R, Albertsen H, Plaetke R, Odelberg S, Ward D, Dausset J, Cohen D, Cann H. A comprehensive human linkage map with centimorgan density. Cooperative Human Linkage Center (CHLC). Science 1994; 265: 2049-2054. CEPH Genotype database V10.0 – November 2004. URL: http://www.cephb.fr/cephdb/
6. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860-921. July 2003 reference sequence as viewed in: Genome Bioinformatics Group of UC Santa Cruz. UCSC Genome Browser. 2005. URL: http://genome.ucsc.edu/cgi-bin/hgGateway.
7. Schuler GD. Sequence mapping by electronic PCR. Genome Res 1997; 7:541-550.
8. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
9. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63: 259-266.
10. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30: 97-101.
11. Weeks DE, Sobel E, O'Connell JR, Lange K. Computer programs for multilocus haplotyping of general pedigrees. Am J Hum Genet 1995; 56: 1506-1507.
12. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996; 58: 1323-1337.
13. Cottingham R, Indury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993; 53: 252-263.
14. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12-13.
15. Freimer NB, Sandkuijl LA, Blower S. Incorrect specification of marker allele frequencies: effects on linkage analysis. Am J Hum Genet 1993; 52: 1102-1110.
16. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 1998; 63: 861-869.
17. Bespalova IN, Van Camp G, Bom SJ, Brown DJ, Cryns K, DeWan AT, Erson AE, Flothmann

K, Kunst HP, Kurnool P, Sivakumaran TA, Cremers CW, Leal SM, Burmeister M, Lesperance MM. Mutations in the Wolfram syndrome 1 gene (WFS1) are a common cause of low frequency sensorineural hearing loss. Hum Mol Genet 2001; 10: 2501-2508.

18. Hafner FM, Salam AA, Linder TE, Balmer D, Baumer A, Schinzel AA, Spillmann T, Leal SM. A novel locus (DFNA24) for prelingual nonprogressive autosomal dominant nonsyndromic hearing loss maps to 4q35-qter in a large Swiss German kindred. Am J Hum Genet 2000; 66: 1437–1442.
19. Fridell RA, Boger EA, San Agustin T, Brownstein MJ, Friedman TB, Morell RJ. *DFNA27*, a new locus for autosomal dominant hearing impairment on chromosome 4. Program Nr: 1388. URL: http://genetics.faseb.org/genetics/ashg99/f1388.htm. [Abstract]
20. Xiao S, Yu C, Chou X, Yuan W, Wang Y, Bu L, Fu G, Qian M, Yang J, Shi Y, Hu L, Han B, Wang Z, Huang W, Liu J, Chen Z, Zhao G, Kong X. Dentinogenesis imperfecta 1 with or without progressive hearing loss is associated with distinct mutations in DSPP. Nat Genet 2001; 27: 129-130.
21. Odeh H, Hagiwara N, Skynner M, Mitchem KL, Beyer LA, Allen ND, Brilliant MH, Lebart MC, Dolan DF, Raphael Y, Kohrman DC. Characterization of two transgene insertional mutations at pirouette, a mouse deafness locus. Audiol Neurootol 2004; 9: 303-314.
22. Riazuddin S, Castelein CM, Ahmed ZM, Lalwani AK, Mastroianni MA, Naz S, Smith TN, Liburd NA, Friedman TB, Griffith AJ, Riazuddin S, Wilcox ER. Dominant modifier DFNM1 suppresses recessive deafness DFNB26. Nat Genet 2000; 26: 431-434.
23. The Hearing Research Group at Brigham & Women's Hospital. Human Cochlear cDNA library and EST database. 2002. URL: http://hearing.bwh.harvard.edu/estinfo.HTM
24. Holme RH, Bussoli TJ, Steel KP. Table of gene expression in the developing ear. 2002. URL: http://www.ihr.mrc.ac.uk/Hereditary/ genetable/index.shtml
25. Feldheim DA, Nakamoto M, Osterfield M, Gale NW, DeChiara TM, Rohatgi R, Yancopoulos GD, Flanagan JG. Loss-of-function analysis of EphA receptors in retinotectal mapping. J Neurosci 2004; 24: 2542-2550.
26. St John JA, Pasquale EB, Key B. EphA receptors and ephrin-A ligands exhibit highly regulated spatial and temporal expression patterns in the developing olfactory system. Brain Res Dev Brain Res 2002; 138: 1-14.
27. Matsunaga T, Greene MI, Davis JG. Distinct expression patterns of Eph receptors and ephrins relate to the structural organization of the adult rat peripheral vestibular system. Eur J Neurosci 2000; 12: 1599-1616.
28. Bianchi LM, Liu H. Comparison of ephrin-A ligand and EphA receptor distribution in the developing inner ear. Anat Rec 1999; 254: 127-134.
29. Schoenherr CJ, Anderson DJ. The neuron-*REST*rictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science 1995; 267; 1360-1363.
30. Roberson DW, Alosi JA, Mercola M, Comanche DA. *REST* mRNA expression in normal and regenerating avian auditory epithelium. Hear Res 2002; 172: 62-72.

Chapter 2.3

The mapping of *DFNB62*, a new locus for autosomal recessive non-syndromic hearing impairment, to chromosome 12p13.2-p11.23

Ali G, Santos RLP, John P, Wambangco MAL, Lee K, Ahmad W, Leal SM. Reprinted from Clinical Genetics 69(5): 429-433.

SUMMARY

Autosomal recessive non-syndromic hearing impairment (ARNSHI) is the most common form of prelingual inherited hearing impairment. Here is described the mapping of a novel ARNSHI locus in a consanguineous Pakistani family with profound congenital hearing impairment. Two-point and multipoint linkage analyses were performed for the genome scan and fine mapping markers. Haplotypes were constructed to determine the region of homozygosity. At $\theta = 0$ the maximum two-point LOD score of 4.0 was obtained at marker AAC040. A maximum multipoint LOD score of 5.3 was derived at marker D12S320, with the three-unit support interval demarcated by D12S89 and D12S1042. The region of homozygosity is flanked by markers D12S358 and D12S1042, which corresponds to 22.4 cM according to the Rutgers combined linkage-physical map of the human genome and spans 15.0 Mb on the sequence-based physical map. A novel ARNSHI locus *DFNB62* was mapped to chromosome 12p13.2-p11.23. *DFNB62* represents the second ARNSHI locus to map to chromosome 12.

INTRODUCTION

Hearing impairment (HI) is the most common hereditary sensorineural disease in humans. Among cases of hereditary non-syndromic hearing impairment (NSHI), the autosomal recessive form accounts for approximately 77% [1] and usually exhibits a more severe hearing phenotype which is prelingual in onset [2]. It has been estimated that at least 1% of human protein-coding genes are involved in the hearing process [3]. Thus far >60 ARNSHI loci have been mapped and 21 ARNSHI genes have been identified [4]. In most cases ARNSHI loci have been mapped either in families with consanguineous marriages or in endogamous populations [3]. In this report, a consanguineous Pakistani family segregating a novel ARNSHI locus, *DFNB62*, is described. This locus was mapped to a 15 Mb region on chromosome 12p13.2-p11.23.

MATERIALS AND METHODS

Family history

The study was approved by the Institutional Review Boards of Quaid-I-Azam University and Baylor College of Medicine. Signed informed consent was obtained from all family members who participated in the study. The family described here resides in the Punjab province of Pakistan. Owing to strict social customs, the family members rarely marry outside their community. Pedigree 4134 was constructed based on information obtained from family members. All affected individuals have a history of prelingual profound HI involving all frequencies and use sign language for communication. The transmission of HI within the family is consistent with autosomal recessive inheritance (Figure 2.4). Regardless of age, affected family members display the same level of profound HI, implying that the hearing loss is not progressive. Medical history and physical examination of the affected individuals were performed by trained otolaryngologists affiliated with government hospitals. The hearing-impaired individuals underwent careful examination for balance problems, mental retardation, defects in ear morphology, dysmorphic facial features, eye disorders including night blindness and tunnel vision, and were not found to have any syndromic features or vestibular disorders. Due to logistic constraints, test battery for vestibular disease and temporal bone imaging are not available. The audiogram of affected individual VI-6 is shown in Figure 2.5, thus demonstrating that *DFNB62* causes profound hearing impairment that affects all frequencies.

DNA isolation and genotyping

Genomic DNA was extracted from whole blood by a standard protocol [5], quantified by spectrophotometry at optical density 260, and diluted to 40 ηg/μl for polymerase chain reaction (PCR) amplification. A genome scan was carried out on all DNA samples at the National Heart, Lung and Blood Institute (NHLBI) Mammalian Genotyping Service (Center for Medical Genetics, Marshfield, WI, USA). A total of 410 fluorescently labeled short tandem repeat (STR) markers were genotyped. These markers are spaced ~10 cM apart and are located on 22 autosomes and the X and Y chromosomes. For fine mapping, PCR amplification of polymorphic STR markers was performed by using 40 ηg of genomic DNA in 25 μl reaction mixture and resolved on 8% non-denaturing polyacrylamide gel. Genotypes were assigned by visual inspection.

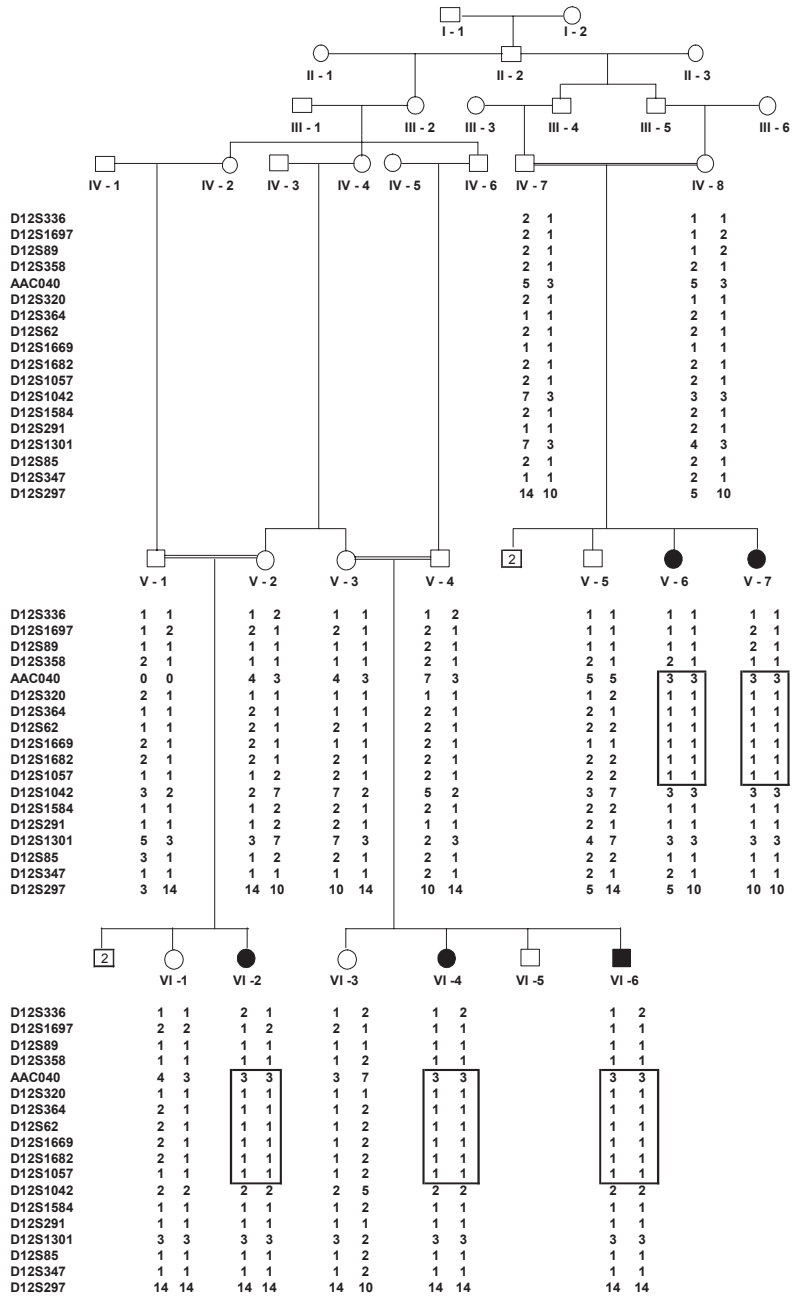


Figure 2.4. Pedigree of family 4134 with autosomal recessive non-syndromic hearing impairment due to *DFNB62*. Filled symbols represent hearing-impaired individuals. Clear symbols represent unaffected family members. For genotyped individuals haplotypes are shown beneath each symbol, with the disease-associated haplotype in a box. The genotype for individual V-1 at marker AAC040 is unknown, and has been designated "0 0".

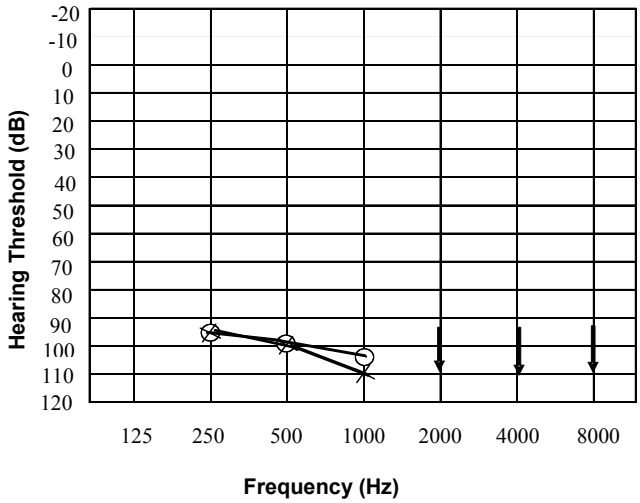


Figure 2.5 Audiogram of affected individual VI-6, demonstrating profound hearing impairment that involves all frequencies in both ears. Circles and crosses represent air conduction for right and left ear, respectively.

Linkage analysis

The pedigree and genotype data were checked with PEDCHECK [6] and MERLIN [7] software. Two-point linkage analysis was carried out using the MLINK program of the FASTLINK computer package [8]. Multipoint linkage analysis was performed using ALLEGRO [9]. The order of the genome scan and fine mapping markers were determined following the National Center for Biotechnology Information (NCBI) Build 34 sequence-based physical map [10] then genetic map distances were derived from the Rutgers combined linkage-physical map of the human genome [11]. Haplotypes were constructed using SIMWALK2 [12-13]. An autosomal recessive mode of inheritance with complete penetrance and a disease allele frequency of 0.001 were used for the analysis. For the genome scan markers, allele frequencies were estimated from the founders and reconstructed genotypes of founders from pedigree 4134 and 35 other pedigrees that underwent a genome scan at the same time. Equal allele frequencies were used for the fine mapping markers, because it was not possible to estimate allele frequencies from the founders, since these markers were only genotyped in this family. To evaluate whether false positive results were obtained due to incorrect allele frequencies [14], a sensitivity analysis was carried out. This was done by carrying out the multipoint linkage analysis

using varying allele frequencies between 0.2 and 0.7 for the fine mapping markers for the alleles that was segregating with the disease allele.

RESULTS

Two-point linkage analysis of the genome scan markers produced a maximum LOD score of 4.0 ($\theta = 0$) at marker AAC040 while the maximum multipoint LOD score was 4.8 at the same marker. In order to fine map the *DFNB62* locus, thirty-one additional markers were selected from the Marshfield [15] and deCode genetic maps [16]. Nineteen of these markers were informative for linkage (Table 2.3). After genotyping these markers, the data was reanalyzed using two-point and multipoint linkage analysis. The maximum two-point LOD score of 4.0 ($\theta = 0$) remained at marker AAC040 (Table 2.3). A maximum multipoint LOD score of 5.3 was achieved at D12S320. The three-unit support interval extended from marker D12S89 to marker D12S1042, spanning a 24.3 cM region according to the Rutgers combined linkage-physical map of the human genome [11]. This interval includes 15.7 Mb on the sequence-based physical map [10].

Using SIMWALK2, haplotypes were constructed to determine the critical recombination events (Figure 2.4). The disease haplotype (region of homozygosity) is flanked by markers D12S358 and D12S1042 and is smaller than the three-unit support interval. It is 22.4 cM long and contains 15.0 Mb. The critical recombination defining the co-segregating interval occurred in the affected individuals. The telomeric boundary of this interval was defined by a recombination between markers D12S358 and AAC040 observed in individual V-6. The affected individuals in both branches of the family were homozygous at genome scan marker D12S1042 but for different alleles in each family branch. Therefore the centromeric boundary of the region of homozygosity was assigned between markers D12S1057 and D12S1042.

DISCUSSION

Large consanguineous families such as those found in the Pakistani population have been instrumental in mapping autosomal recessive HI loci. In this study evidence for linkage was found in a consanguineous Pakistani family to the novel HI locus, *DFNB62*. The *DFNB62* interval was mapped to a 22.4 cM region on chromosome 12p13.2-p11.23. Four loci for NSHI, *DFNB50* (12q23) [4], *DFNA25* (12q22-q24.11) [17], *DFNA41* (12q24.32-qter) [18] and *DFNA48* (12q13.13-q15) [19-20], have been previously mapped on chromosome 12. For these four loci, only one gene, *MYO1A*,

Table 2.3 Two point LOD score results between the *DFNB62* locus and chromosome 12 markers

Marker [a]	Genetic map position [b]	Physical map position [c]	LOD score at *theta*=							
			0.0	0.01	0.02	0.04	0.05	0.1	0.2	0.3
D12S336	24.51	9,385,296	-4.60	-2.16	-1.62	-1.09	-0.92	-0.47	-0.14	-0.05
D12S1697	26.72	11,685,470	-3.47	-1.19	-0.67	-0.20	-0.07	0.25	0.30	0.18
D12S89	27.00	11,793,257	-1.43	-0.37	-0.12	0.10	0.16	0.29	0.27	0.17
D12S358	**28.89**	**12,530,590**	**1.75**	**1.70**	**1.65**	**1.55**	**1.50**	**1.26**	**0.82**	**0.44**
AAC040	*32.19*	*13,065,278*	*4.02*	*3.94*	*3.85*	*3.67*	*3.58*	*3.13*	*2.21*	*1.30*
D12S320	32.19	13,513,310	1.15	1.12	1.08	1.01	0.98	0.82	0.53	0.29
D12S364	32.19	13,724,570	1.65	1.60	1.56	1.47	1.42	1.21	0.80	0.45
D12S62	34.32	15,219,370	2.31	2.24	2.18	2.06	2.00	1.71	1.14	0.63
D12S1669	37.83	19,429,661	1.62	1.58	1.53	1.44	1.40	1.18	0.77	0.42
D12S1682	40.34	20,570,897	2.31	2.24	2.18	2.05	1.98	1.67	1.06	0.53
D12S1057	46.46	24,568,387	2.31	2.24	2.18	2.06	2.00	1.71	1.14	0.63
D12S1042	***51.27***	***27,538,599***	-∞	***0.29***	***0.69***	***1.07***	***1.17***	***1.36***	***1.18***	***0.78***
D12S1584	55.41	31,610,488	2.31	2.24	2.18	2.06	2.00	1.71	1.14	0.63
D12S291	59.08	41,688,390	1.55	1.51	1.48	1.40	1.36	1.17	0.80	0.48
D12S1301	*59.08*	*42,348,809*	-∞	*2.69*	*2.91*	*3.03*	*3.03*	*2.87*	*2.21*	*1.41*
D12S85	61.52	45,622,954	2.98	2.91	2.84	2.69	2.61	2.24	1.52	0.86
D12S347	65.70	50,298,255	1.58	1.53	1.49	1.41	1.37	1.17	0.80	0.46
D12S297	*66.82*	*50,899,108*	-∞	*-1.86*	*-1.15*	*-0.46*	*-0.25*	*0.28*	*0.49*	*0.38*

[a] Markers displayed in bold flank the haplotype. Genome scan markers are shown in italics.
[b] Cumulative sex-averaged Kosambi cM genetic map distances from the Rutgers combined linkage-physical map of the human genome [12].
[c] Sequence-based physical map distance in bases according to Build 34 of the human reference sequence [11].

has been identified for *DFNA48* [20]. *MYO1A* (MIM 601478) is 28.2 Mb distal to the *DFNB62* locus, and was excluded as the cause of ARNSHI in this family by both linkage analysis and direct sequencing.

Currently the *DFNB62* linkage interval contains 75 known genes, 15 of which encode hypothetical proteins. Among the known genes is a gene for Keutel syndrome, which is caused by mutations in matrix Gla protein (*MGP*; MIM 154870). Patients with Keutel syndrome are found mostly in consanguineous families and have the following as cardinal features: midface hypoplasia, flat nasal bridge, cartilage calcification and brachytelephalangism or "drumstick" fingers [21]. Additional clinical signs include heart defects (e.g. pulmonary artery stenosis), developmental delay and respiratory abnormalities. About 70% of cases report hearing impairment with or without otitis media. Family 4134 was screened for the *MGP* gene via direct sequencing and no functional variants were discovered.

Besides *MGP*, there are at least 9 other genes in the *DFNB62* interval that are expressed in the inner ear [22]. Notable among these is *EMP1* (epithelial membrane protein 1; MIM 602333), a member of the peripheral myelin protein 22 (*PMP22*) family. Mutations in *PMP22* (MIM 601097) have been shown to cause autosomal dominant Charcot-Marie-Tooth (*CMT*) disease with deafness [23], presumably due to dysfunctional Schwann cells thus resulting in hypomyelination of the auditory nerve. In one *CMT* family, otoacoustic emission testing suggested cochlear abnormalities in addition to the auditory nerve dysfunction seen with auditory brainstem response [24]. *EMP1* has not been associated with inner ear disorders, but it is highly expressed in the developing murine nervous system including the peripheral (cranial) nerves [25]. The *EMP1* gene was negative for functional variants in one unaffected and two affected individuals from family 4134.

Identification of the *DFNB62* gene will enhance our knowledge of the molecular mechanisms behind the hearing impairment phenotype.

Acknowledgements We wish to thank the family members for their invaluable participation and cooperation. The work was funded by the Higher Education Commission, Islamabad, Pakistan, the American Hearing Research Foundation and the NIH National Institute of Deafness and other Communication Disorders grant R01-DC03594. Genotyping services were provided by the NHLBI Mammalian Genotyping Service.

REFERENCES

1. Morton NE. Genetic epidemiology of hearing impairment. Ann N Y Acad Sci 1991; 630: 16–31.
2. Cohen MM, Gorlin RJ. Epidemiology, etiology and genetic patterns. p. 43–61. In: Gorlin RJ, Toriello HV, Cohen MM. Hereditary hearing loss and its syndromes. Oxford: Oxford University Press, 1995.
3. Friedman TB, Griffith AJ. Human nonsyndromic sensorineural deafness. Annu Rev Genomics Hum Genet 2003; 4: 341–402.
4. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
5. Grimberg J, Nawoschik S, Bellusico L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
6. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63: 259-266.
7. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30: 97-101.
8. Cottingham R, Indury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993; 53: 252–263.
9. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12–13.
10. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860–921.
11. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
12. Weeks DE, Sobel E, O'Connell JR, Lange K. Computer programs for multilocus haplotyping of general pedigrees. Am J Hum Genet 1995; 56: 1506-1507.
13. Sobel E, Lange K. Descent graphs in pedigree analysis: Applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996; 58: 1323–1337.
14. Freimer NB, Sandkuijl LA, Blower SM. Incorrect specification of marker allele frequencies: effects on linkage analysis. Am J Hum Genet 1993; 52: 1102-1110.
15. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 1998; 63: 861–869.
16. Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. A high-resolution recombination map of the human genome. Nat Genet 2002; 31: 241–247.
17. Greene CC, McMillan PM, Barker SE, Kurnool P, Lomax MI, Burmeister M, Lesperance MM. DFNA25, a novel locus for dominant nonsyndromic hereditary hearing impairment, maps to 12q21-24. Am J Hum Genet 2001; 68: 254-260.
18. Blanton SH, Liang CY, Cai MW, Pandya A, Du LL, Landa B, Mummalanni S, Li KS, Chen ZY, Qin XN, Liu YF, Balkany T, Nance WE, Liu XZ. A novel locus for autosomal dominant non-syndromic deafness (DFNA41) maps to chromosome 12q24-qter. J Med Genet 2002; 39: 567-570.

19. D'Adamo P, Pinna M, Capobianco S, Cesarani A, D'Eustacchio A, Fogu P, Carella M, Seri M, Gasparini PA. A novel autosomal dominant non-syndromic deafness locus (DFNA48) maps to 12q13-q14 in a large Italian family. Hum Genet 2003; 112: 319-320.
20. Donaudy F, Ferrara A, Esposito L, Hertzano R, Ben-David O, Bell RE, Melchionda S, Zelante L, Avraham KB, Gasparini P. Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. Am J Hum Genet 2003; 72: 1571-1577.
21. Hur DJ, Raymond GV, Kahler SG, Riegert-Johnson DL, Cohen BA, Boyadjiev SA. A novel MGP mutation in a consanguineous family: review of the clinical and molecular characteristics of Keutel syndrome. Am J Med Genet A 2005; 135: 36-40.
22. The Hearing Research Group at Brigham & Women's Hospital. Human Cochlear cDNA library and EST database. 2002. URL: http://hearing.bwh.harvard.edu/estinfo.HTM
23. Sambuughin N, de Bantel A, McWilliams S, Sivakumar K. Deafness and CMT disease associated with a novel four amino acid deletion in the PMP22 gene. Neurology 2003; 60: 506-508.
24. Kovach MJ, Campbell KC, Herman K, Waggoner B, Gelber D, Hughes LF, Kimonis VE. Anticipation in a unique family with Charcot-Marie-Tooth syndrome and deafness: delineation of the clinical features and review of the literature. Am J Med Genet 2002; 108: 295-303.
25. Wulf P, Suter U. Embryonic expression of epithelial membrane protein 1 in early neurons. Brain Res Dev Brain Res 1999; 116: 169-180.

Chapter 2.4

Localization of a novel autosomal recessive non-syndromic hearing impairment locus *DFNB65* to chromosome 20q13.2-q13.32

Tariq A, Santos RLP, Khan MN, Lee K, Hassan MJ, Ahmad W, Leal SM. Reprinted from Journal of Molecular Medicine 84(3): 226-231, with kind permission of Springer Science and Business Media.

SUMMARY

Autosomal recessive non-syndromic hearing impairment (ARNSHI) is the most frequent form of prelingual hereditary hearing loss in humans. Between 75 and 80% of all non-syndromic deafness is inherited in an autosomal recessive pattern. Using linkage analysis, we have mapped a novel gene responsible for this form of non-syndromic hearing impairment, *DFNB65*, in a consanguineous family from the Azad Jammu and Kashmir regions, which border Pakistan and India. A maximum multipoint LOD score of 3.3 was obtained at marker D20S840. The three-unit support interval is contained between markers D20S902 and D20S430, while the region of homozygosity is flanked by markers D20S480 and D20S430. The novel locus maps to a 10.5 cM region on chromosome 20q13.2-q13.32 and corresponds to a physical map distance of 4.3 Mb. *DFNB65* represents the first ARNSHI locus to map to chromosome 20.

INTRODUCTION

It is estimated that 1-2 per 1000 newborns are affected with profound hearing impairment (HI) [1]. For congenital HI, roughly half of the cases are expected to be genetic. Of the genetic cases of HI, 30% are known to be syndromic, with the remaining cases displaying non-syndromic hearing impairment (NSHI). For NSHI, 70-75% of the cases are due to autosomal recessive (AR) inheritance, 15-20% display autosomal dominant inheritance, 2-3% of cases show X-linked inheritance and less than 1% of cases are due to mitochondrial inheritance [2]. Thus far, for ARNSHI more than 60 loci have been mapped and 21 genes identified [3]. This extreme heterogeneity reflects involvement of different molecular mechanisms within the auditory system that malfunction to cause HI. In the current study, we investigated a consanguineous family segregating ARNSHI and mapped the gene responsible for the defect to chromosome 20q13.2-q13.32.

MATERIALS AND METHODS

Family history

Before the start of the study, approval was obtained from the Quaid-I-Azam University and the Baylor College of Medicine Institutional Review Boards. Informed consent was obtained from all family members who agreed to participate in the study. Family 4020 (Figure 2.6), which is from the city of Rawlakot in Azad Jammu and Kashmir region that borders Pakistan and India, provided convincing evidence of autosomal recessive mode of inheritance. All affected individuals have a history of prelingual profound HI and use sign language for communication. In addition, all available medical records of the affected individuals were scrutinized for evidence of childhood illness. Each HI individual underwent a physical and ophthalmologic examination. Defects in ear morphology, dysmorphic facial features, eye disorders including night blindness and tunnel vision, limb deformities, mental retardation, and other clinical features that could indicate that HI was syndromic were not identified. The clinical history and physical examination of affected individuals indicated no vestibular involvement.

Figure 2.7 exhibits the audiometric profile of hearing-impaired individuals IV-2 and IV-3, aged 17 and 15 respectively, who have profound HI at all frequencies in the two ears. The audiograms display a flat configuration.

Extraction of genomic DNA and genotyping

Peripheral blood samples were obtained from six family members, four of whom were hearing-impaired. Genomic DNA was extracted from whole blood following a standard protocol [4], quantified by spectrophotometric reading at optical density 260, and diluted to 40 ηg/μl for polymerase chain reaction (PCR) amplification. A genome scan was carried out on six DNA samples at the Center for Inherited Disease Research (CIDR). A total of 405 fluorescently labeled short tandem repeat (STR) markers were genotyped. These markers are spaced at approximately 10 cM apart and are located on 22 autosomes and X and Y chromosomes. After the completion of the genome scan, two additional unaffected (IV-1 and IV-5) family members were ascertained and their DNA samples were used for fine mapping of the *DFNB65* locus.

For fine mapping, microsatellite markers were amplified by PCR according to standard procedure in a total volume of 25 μl with 40 ηg genomic DNA, 0.3 μl of primer (Invitrogen Corp., Carlsbad, CA, USA), 200 μM dNTP and

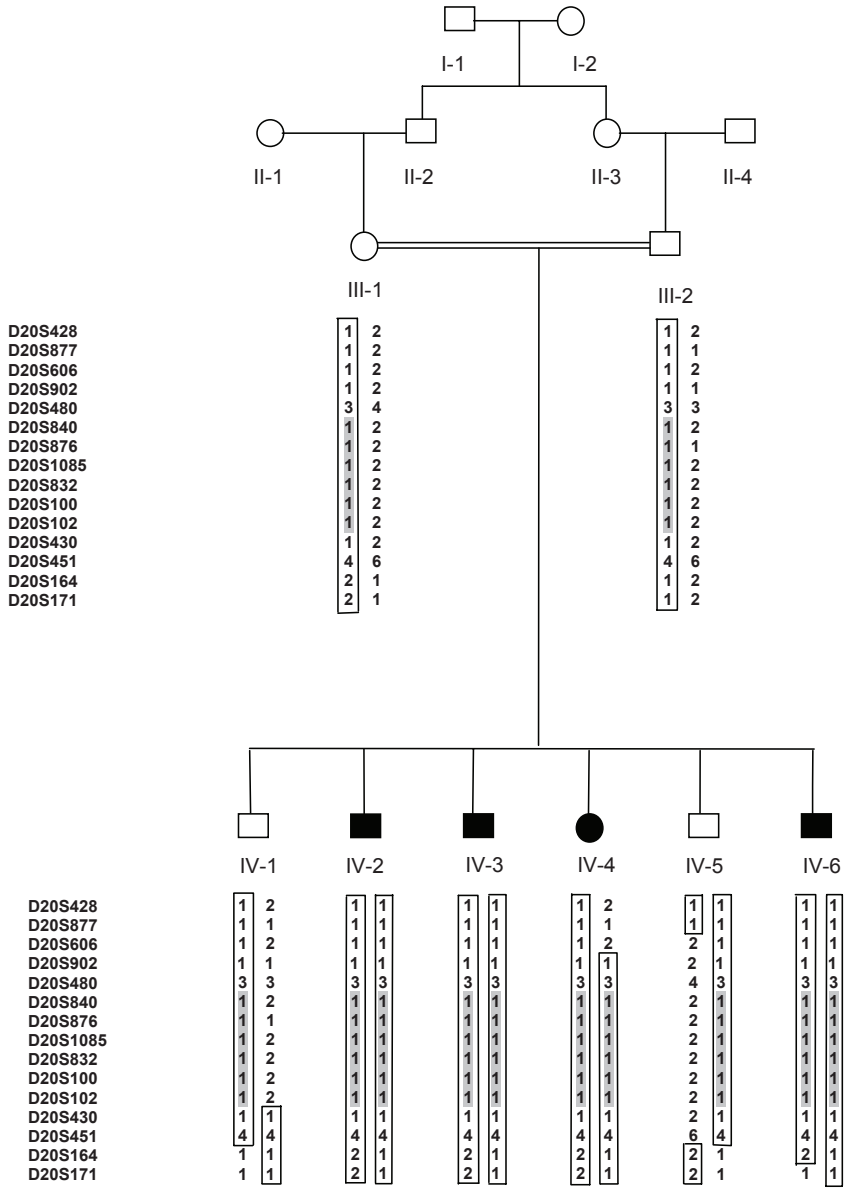


Figure 2.6 Pedigree drawing of family 4020. Black symbols represent individuals with hearing impairment. Clear symbols represent unaffected individuals. The gender of some of the family members was changed to protect the anonymity of the family. Haplotypes are shown below each individual for whom genotypes are available. For individuals in generation IV, maternal haplotypes are displayed on the left-hand side and paternal haplotypes on the right-hand side. Shaded areas indicate the region of homozygosity.

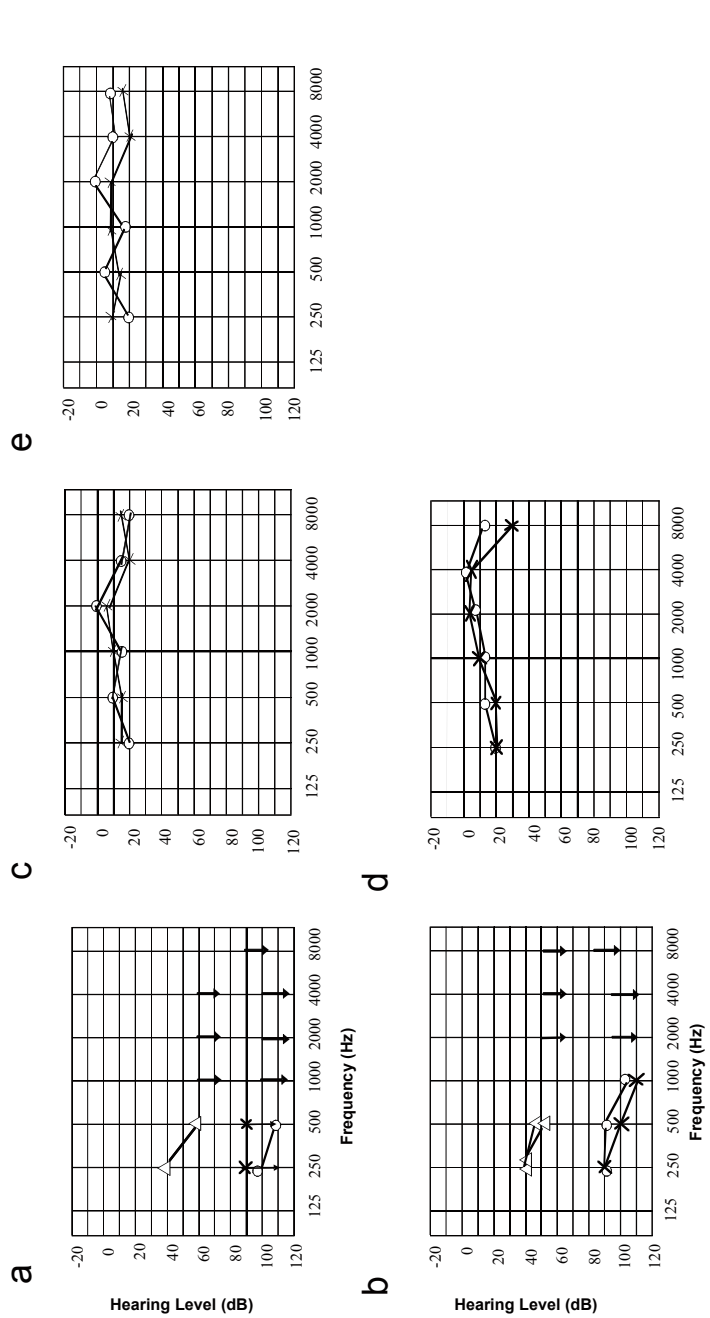


Figure 2.7 Audiograms of hearing-impaired individuals (a) IV-2 and (b) IV-3 of family 4020, demonstrating profound hearing impairment involving all frequencies in both ears. The normal audiograms of family members (c) III-2, (d) IV-1 and (e) IV-5 are also shown. Individual IV-5 is aged 21 years, while IV-2 is 17 and IV-3 15 years of age. Circles and crosses represent air conduction for right and left ear, respectively. Triangles represent masked air conduction thresholds.

1 unit of Taq DNA Polymerase (Fermentas Life Sciences, Burlington, ON, Canada) in a thermal cycler (Whatman Biometra, Goettingen, Germany). PCR products were resolved on 8% non-denaturing polyacrylamide gel.

Linkage analysis

The National Center for Biotechnology Information (NCBI) Build 34 sequence-based physical map was used to determine the order of the genome scan markers and fine mapping markers [5]. Genetic map distances according to the Rutgers combined linkage-physical map of the human genome [6] were used to carry out the multipoint linkage analysis for the fine map and genome scan markers. For those genome scan markers where no genetic map position was available, interpolation was performed to place these markers on the Rutgers combined linkage-physical map. PEDCHECK [7] was used to identify Mendelian inconsistencies while the MERLIN [8] program was utilized to detect potential genotyping errors that did not produce a Mendelian inconsistency. Haplotypes were constructed using SIMWALK2 [9-10]. Two-point linkage analysis was carried out using the MLINK program of the FASTLINK computer package [11] and multipoint linkage analysis was performed using ALLEGRO [12]. An autosomal recessive mode of inheritance with complete penetrance and a disease allele frequency of 0.001 were assumed. Marker allele frequencies were estimated from the founders and the reconstructed genotypes of founders from this family and 32 additional Pakistani families that underwent a genome scan at the same time at CIDR. Equal allele frequencies were used for fine mapping markers, because it was not possible to estimate allele frequencies from the founders, because these markers were only genotyped in this family. Because false positive results can be obtained when analyzing the data using too low of an allele frequency for the allele segregating with the disease locus [13], a sensitivity analysis was carried out for the multipoint linkage analysis by varying the allele frequency of the allele that is segregating with the disease locus from 0.2 to 0.7 for the fine mapping markers.

RESULTS

Two point linkage analysis of the genome scan markers generated maximum LOD scores at $\theta=0$ of 1.2 and 1.9 with markers D20S480 and D20S451, respectively. A maximum multipoint LOD score of 1.4 was obtained at D20S451. Figure 2.8 is a genome-wide representation of the multipoint LOD scores for the 22 autosomes. Only on chromosome 20 was a LOD score greater than 1.0 achieved.

To fine-map this region on chromosome 20, 33 additional markers located in the vicinity of D20S480 and D20S451 were selected from the Marshfield genetic map [14]. Sixteen of the markers were uninformative, although all markers that were selected had a reported heterozygosity of > 0.7. Of the 17 fine mapping markers that are informative for linkage, eight markers (including D20S428, D20S877, D20S606, and D20S902) lie between genome scan markers D20S481 and D20S480, while seven markers (D20S840, D20S876, D20S1085, D20S832, D20S100, D20S102, and D20S430) are between D20S480 and genome scan marker D20S451. In addition, D20S164 and D20S171, two of the fine mapping markers, are distal to the genome scan marker D20S451. Analysis of the genome scan and fine mapping marker genotypes within this region with PEDCHECK and MERLIN did not elucidate any genotyping errors. Table 2.4 summarizes the two-point LOD scores obtained after fine mapping. The maximum two-point LOD score of 2.4 was obtained with five markers (D20S840, D20S1085, D20S832, D20S100 and D20S102) at θ=0. Multipoint linkage analysis resulted in a maximum LOD score of 3.3 at marker D20S840 (Figure 2.9). The three-unit support interval was between markers D20S902 and D20S430, spanning 11.7 cM according to the Rutgers combined linkage-physical map of the human genome [6] and corresponding to a physical map distance of 4.4 Mb (Table 2.4).

Haplotypes were constructed to determine the critical recombination events in the family (Figure 2.6). All hearing-impaired individuals were homozygous for markers D20S902 and D20S480, but unaffected individual IV-1 is also homozygous for the same markers. Thus, the centromeric boundary of the interval is assigned between markers D20S480 and D20S840. The telomeric boundary of the interval corresponds to a recombination event between markers D20S102 and D20S430, which occurred in unaffected individual IV-1. The region of homozygosity that is flanked by markers D20S480 and D20S430 contains 4.3 Mb and is 10.5 cM long according to the Rutgers combined linkage-physical map of the human genome (Table 2.4). The region of homozygosity is therefore narrower than the three-unit support interval and most likely includes the gene for *DFNB65*.

When the marker allele frequencies for the alleles segregating with HI were varied for the fine mapping markers between 0.2 and 0.4, the maximum multipoint LOD score remained to be 3.3 at marker D20S840. When the allele frequencies were increased, the maximum LOD score still occurred at marker D20S840 but decreased to 3.2 and 3.1 at allele frequencies of 0.6 and 0.7, respectively.

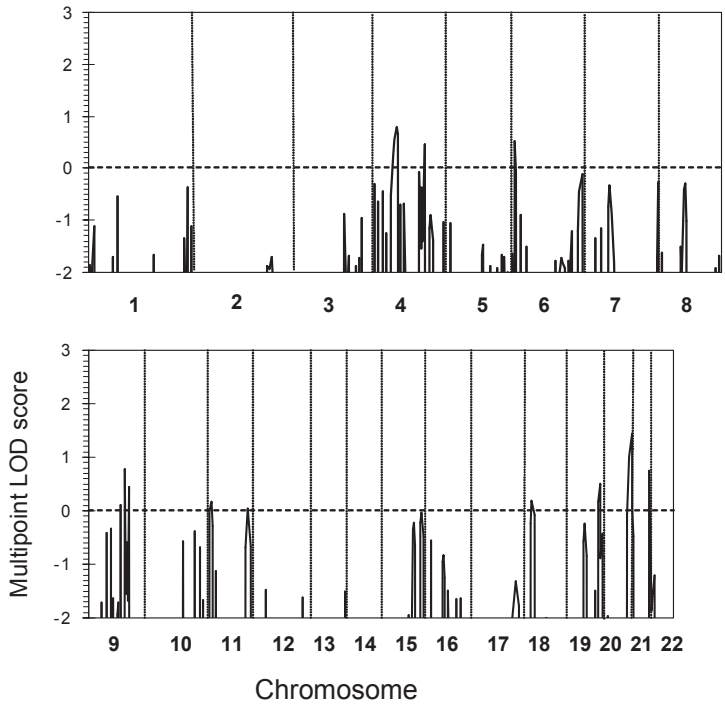


Figure 2.8 Genome-wide plot of multipoint LOD scores from the genome scan genotypes of family 4020. The LOD score is plotted against the genetic map distance in centimorgans (cM) per chromosome. Only on chromosome 20 at marker D20S451 was a multipoint LOD score of more than 1.0 (LOD = 1.4) derived.

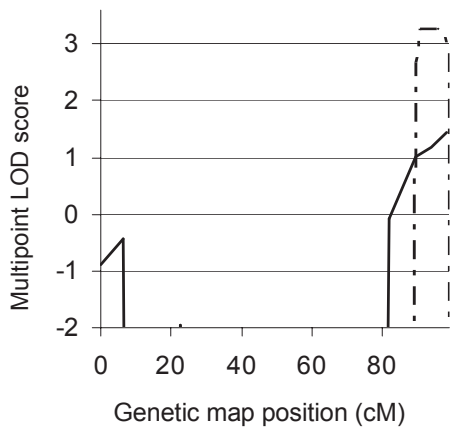


Figure 2.9 The plot of the multipoint LOD scores including the fine mapping markers on chromosome 20 (broken line) is superimposed on the LOD score plot from the genome scan markers (solid line).

Table 2.4 Two point LOD score results between the *DFNB65* locus and chromosome 20 markers after fine mapping

Marker [a]	Genetic map position [b]	Physical map position [c]	LOD score at *theta*=							
			0.0	0.01	0.02	0.04	0.05	0.1	0.2	0.3
D20S428	80.34	51,588,497	-∞	-1.39	-0.83	-0.32	-0.17	0.20	0.34	0.24
D20S877	80.49	51,799,558	0.85	0.83	0.81	0.78	0.76	0.67	0.48	0.29
D20S606	81.41	52,247,558	-∞	0.31	0.56	0.77	0.81	0.88	0.70	0.41
D20S902	81.41	52,436,690	1.15	1.13	1.10	1.05	1.03	0.91	0.66	0.41
D20S480 [d]	***82.62***	***52,542,629***	***1.31***	***1.28***	***1.26***	***1.20***	***1.17***	***1.03***	***0.75***	***0.46***
D20S840	82.80	52,832,756	2.35	2.30	2.25	2.15	2.09	1.83	1.29	0.75
D20S876	83.13	53,033,103	1.15	1.13	1.10	1.05	1.03	0.91	0.66	0.41
D20S1085	84.67	53,366,902	2.35	2.30	2.25	2.15	2.09	1.83	1.29	0.75
D20S832	86.80	54,580,765	2.35	2.30	2.25	2.15	2.09	1.83	1.29	0.75
D20S100	88.19	54,999,475	2.35	2.30	2.25	2.15	2.09	1.83	1.29	0.75
D20S102	89.21	55,676,797	2.35	2.30	2.25	2.15	2.09	1.83	1.29	0.75
D20S430	**93.11**	**56,835,225**	**-∞**	**0.61**	**0.86**	**1.06**	**1.10**	**1.15**	**0.93**	**0.57**
D20S451 [d]	*94.79*	*57,349,950*	-∞	*0.81*	*1.06*	*1.25*	*1.29*	*1.31*	*1.04*	*0.63*
D20S164	95.96	57,738,776	-1.17	-0.49	-0.28	-0.08	-0.02	0.07	0.02	-0.03
D20S171	98.00	58,493,299	-∞	-2.19	-1.66	-1.15	-1.00	-0.56	-0.23	-0.09

[a] Markers in bold type flank the haplotype. Genome scan markers are shown in italics.
[b] Cumulative sex-averaged Kosambi cM genetic map distances from the Rutgers combined linkage-physical map of the human genome [6].
[c] Sequence-based physical map distances in bases according to Build 34 of the human reference sequence [5].
[d] The two unaffected individuals who were not included in the genome scan were also genotyped for these genome scan markers. Therefore the results reported in this table for the genome scan markers are different from those obtained from the initial genome scan.

Table 2.5 *DFNB65* candidate genes [a]

Gene	Protein	MIM	Physical map position	Cytogenetic band	Description
ZNF217	zinc finger protein 217	602967	52,869,034	20q13.2	Putative oncogene
BCAS1	breast carcinoma amplified sequence 1	602968	53,245,939	20q13.2	Putative oncogene
CYP24A1 [b]	Cytochrome P450, family 24 precursor	126065	53,455,410	20q13.2	Has role in calcium metabolism and blood cell differentiation
PFDN4	prefoldin 4	604898	53,509,924	20q13.2	Molecular chaperone; putative oncogene
DOK5	DOK5 protein isoform b	608334	53,777,679	10q13.2	Docking protein
CBLN4	cerebellin 4 precursor	--	55,257,918	20q13.31	Neuromodulatory glycoprotein in adrenal medulla
MC3R [c]	melanocortin 3 receptor	155540	55,509,211	20q13.31	Associated with obesity
STK6	serine/threonine protein kinase 6	602687	55,629,867	20q13.31	Has role in spindle formation and chromosomal segregation
CSTF1	cleavage stimulation factor subunit 1	600369	55,653,041	20q13.31	Ensures correct polyadenylation of mRNA
TFAP2C [d]	transcription factor AP-2 gamma	601602	55,889,780	20q13.31	Regulator of gene expression for ectodermal tissue development
BMP7	bone morphogenetic protein 7 precursor	112267	56,430,977	20q13.31	Member of transforming growth factor-beta superfamily
SPO11	meiotic recombination protein SPO11 isoform a	605114	56,590,253	20q13.31-q13.32	Required for meiotic double strand break formation
RAE1	RAE1 (RNA export 1, S.pombe) homolog	603343	56,611,567	20q13.32	May function in nucleocytoplasmic transport and attaching cytoplasmic mRNA-binding protein to cytoskeleton
RNPC1	RNA-binding region containing protein 1 isoform	--	56,651,885	20q13.32	Probable RNA-binding protein
HMG1L1	high-mobility group (nonhistone chromosomal)	--	56,748,868	20q13.32	May be involved in tumorigenesis
CTCFL	CCCTC-binding factor-like protein	607022	56,757,645	20q13.32	Putative oncogene
PCK1	cytosolic phosphoenol-pyruvate carboxykinase 1	261680	56,821,558	20q13.32	Main target for regulation of gluconeogenesis

[a] This list does not include five genes that encode hypothetical proteins: *C20orf108 (LOC116151), C20orf32 (HEF-like protein), C20orf43 (LOC51507), C20orf106 (LOC200232)* and *LOC388799*.
[b] The *Cyp24a1* rat manifests with albuminuria and hyperlipidemia [15].
[c] *Mc3r*$^{-/-}$ mice have increased fat mass [16].
[d] The *Tfap2c* knockout is embryonic lethal in mice [17].

DISCUSSION

The linkage data presented here suggest that a novel gene for ARNSHI is located on chromosome 20q13.2-q13.32. This is the first HI locus identified on human chromosome 20. Many of the known NSHI loci have been identified in single large families in which linkage can be established independently. However, due to the limited number of meiosis within each family, the genetic interval for the HI locus is often large, which makes gene identification difficult.

The *DFNB65* interval spans 4.3 Mb and contains 17 known genes (Table 2.5). Of these, bone morphogenetic protein 7 (*BMP7*, MIM 112267) has extensive expression in the otic placode of embryonic chick, which eventually becomes restricted to cochlear supporting cells over time [18]. This makes the *BMP7* gene the best candidate of the known genes in the *DFNB65* locus. However sequencing of DNA for the *BMP7* gene from one unaffected and two hearing-impaired individuals from family 4020 resulted in no functional sequence variants within the promoter and exonic regions.

Other genes within the interval that are included in expression databases for the inner ear [19-20] are breast carcinoma amplified sequence 1 (*BCAS1*, MIM 602967) and prefoldin 4 (*PFDN4*, MIM 604898). These two genes have been implicated in several soft tissue cancers (e.g. breast, prostate, colon), which may indicate the oncogenic potential of these genes, although their specific functions remain unknown.

Acknowledgements We wish to thank the family members for their invaluable participation and cooperation. This research was funded by the National Institutes of Health - National Institute of Deafness and other Communication Disorders grant R01-DC03594-06A1 and the Higher Education Commission, Islamabad, Pakistan. Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, Contract Number N01-HG-65403.

REFERENCES

1. Parving A. Epidemiology of hearing loss and aetiological diagnosis of hearing impairment in childhood. International Journal of Pediatric Otorhinolaryngology 1983; 7: 29-38.
2. Morton NE. Genetic epidemiology of hearing impairment. Ann NY Acad Sci 1991; 630: 16-31.
3. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
4. Grimberg J, Nawoschik S, Bellusico L, McKee R, Trucks A, Eisenberg A. A simple and

efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
5. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860-921. July 2003 reference sequence as viewed in: Genome Bioinformatics Group of UC Santa Cruz. UCSC Genome Browser. URL: http://genome.ucsc.edu/cgi-bin/hgGateway
6. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
7. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63: 259-266.
8. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30: 97-101.
9. Weeks DE, Sobel E, O'Connell JR, Lange K. Computer programs for multilocus haplotyping of general pedigrees. Am J Hum Genet 1995; 56: 1506-1507.
10. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996; 58: 1323-1337.
11. Cottingham Jr RW, Indury RM, Schaffer AA. Faster sequential genetic linkage computation. Am J Hum Genet 1993; 53: 252-263.
12. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12-13.
13. Freimer NB, Sandkuijl LA, Blower S. Incorrect specification of marker allele frequencies: effects on linkage analysis. Am J Hum Genet 1993; 52: 1102-1110.
14. Broman K, Murray JC, Scheffield VC, White RL, Weber JL. Comprehensive human genetics maps: Individual and sex specific variation in recombination. Am J Hum Genet 1998; 63: 861-869.
15. Kasuga H, Hosogane N, Matsuoka K, Mori I, Sakura Y, Shimakawa K, Shinki T, Suda T, Taketomi S. Characterization of transgenic rats constitutively expressing vitamin D-24-hydroxylase gene. Biochem Biophys Res Commun 2002; 297: 1332-1338.
16. Chen AS, Marsh DJ, Trumbauer ME, Frazier EG, Guan XM, Yu H, Rosenblum CI, Vongs A, Feng Y, Cao L, Metzger JM, Strack AM, Camacho RE, Mellin TN, Nunes CN, Min W, Fisher J, Gopal-Truter S, MacIntyre DE, Chen HY, Van der Ploeg LH. Inactivation of the mouse melanocortin-3 receptor results in increased fat mass and reduced lean body mass. Nat Genet 2000; 26: 97-102.
17. Werling U, Schorle H. Transcription factor gene AP-2 gamma essential for early murine development. Mol Cell Biol 2002; 22: 3149-3156.
18. Oh SH, Johnson R, Wu DK. Differential expression of bone morphogenetic proteins in the developing vestibular and auditory sensory organs. J Neurosci 1996; 16: 6463-6475.
19. The Hearing Research Group at Brigham & Women's Hospital. Human Cochlear cDNA library and EST database. 2002. URL: http://hearing.bwh.harvard.edu/estinfo.HTM
20. Holme RH, Bussoli TJ, Steel KP. Table of gene expression in the developing ear. 2002. URL: http://www.ihr.mrc.ac.uk/Hereditary/genetable/index.shtml

Chapter 2.5

DFNB68, a novel autosomal recessive non-syndromic hearing impairment locus at chromosomal region 19p13.2

Santos RLP, Hassan MJ*, Sikandar S, Lee K, Ali G, Martin PE, Wambangco MAL, Ahmad W, Leal SM. Reprinted from Human Genetics 120(1): 85-92, with kind permission of Springer Science and Business Media.*

** These authors contributed equally to this work.*

SUMMARY

From a large collection of families with autosomal recessive non-syndromic hearing impairment (NSHI) from Pakistan, linkage has been established for two unrelated consanguineous families to 19p13.2. This new locus was assigned the name *DFNB68*. A 10 cM genome scan and additional fine mapping were carried out using microsatellite marker loci. Linkage was established for both families to *DFNB68* with maximum multipoint LOD scores of 4.8 and 4.6. The overlap of the homozygous regions between the two families was bounded by D19S586 and D19S584, which limits the locus interval to 1.9 cM and contains 1.4 Mb. The genes *CTL2*, *KEAP1* and *CDKN2D* were screened but were negative for functional sequence variants.

INTRODUCTION

The extreme genetic heterogeneity for hearing impairment (HI) testifies to the inherent complexity of the mammalian inner ear. As more HI genes are identified, the elucidation of the function of the proteins that these genes encode contributes greatly to the understanding of cochlear mechanisms and their role in disease causation. Currently, 37 genes have been identified for NSHI; however, for more than 60% of the mapped NSHI loci a gene has yet to be identified [1]. Many of these loci with an unidentified gene were mapped in single families, some of which are consanguineous. The size of the mapped interval plus the number of candidate genes within the interval, if large, may make the task of finding the causative gene daunting. The identification of linkage of multiple NSHI families to a single locus validates the linkage finding and can make gene identification more attainable due to a smaller physical interval for the NSHI locus.

The *DFNB68* locus is within the chromosome 19 region for *DFNB15*. The *DFNB15* locus interval was previously mapped in a consanguineous Indian family to two chromosomal locations, 19p13.3-p13.1 and 3q21.3-

q25.2. Linkage to both of these regions was based on a non-significant LOD score of 2.8 with a genetic region of ≥32 cM intervals on both chromosomes 3 and 19 [2]. Presented here are two Pakistani families with autosomal recessive NSHI that exclusively maps to chromosome 19p13.2, with highly significant multipoint LOD scores of 4.8 and 4.6. The overlapping region of homozygosity for these two families is 1.9 cM. Since it is not clear whether the *DFNB15* locus is on chromosome 3 or 19, the NSHI locus segregating in these two Pakistani families has been designated as novel NSHI locus *DFNB68*.

MATERIALS AND METHODS

Family history

Prior to onset the study was approved by the Institutional Review Boards of the Quaid-I-Azam University and Baylor College of Medicine and Affiliated Hospitals. Informed consent was secured from all individuals who agreed to participate in the research. Family 4100 resides in the southern part of Punjab province, while family 4154 hails from Sind province. For both kindreds, family members rarely marry outside the community, and consequently consanguineous unions are common. The pedigree drawings (Figure 2.10) provided convincing evidence of autosomal recessive mode of inheritance in both families, and consanguineous loops accounted for all the affected persons being homozygous for the mutant allele. Clinical findings in these families are consistent with the diagnosis of autosomal recessive NSHI. All affected individuals have a history of prelingual profound HI and communicate by sign language. Affected individuals underwent examination for defects in ear morphology, dysmorphic facial features, eye disorders including night blindness and tunnel vision, limb deformities, mental retardation and other clinical features that could indicate that HI was syndromic, and were all negative for findings. No gross vestibular involvement was noted from the clinical history and physical examination.

Extraction of genomic DNA and genotyping

Peripheral blood samples were taken from ten members in family 4100 including five hearing-impaired individuals, and eleven members from family 4154, six of whom had HI. Genomic DNA was extracted from whole blood following a standard protocol [3], quantified by spectrophotometric reading at optical density 260, and diluted to 40ηg/μl for amplification by polymerase chain reaction (PCR). Genome scans were carried out at the Center for Inherited Disease Research (CIDR) and at the National

Heart, Lung and Blood Institute (NHLBI) Mammalian Genotyping Service (Marshfield, WI, USA). An average of 399 fluorescently labeled short tandem repeat (STR) markers were genotyped across the 22 autosomes and X and Y chromosomes at approximately 10 cM apart. For fine mapping, microsatellite markers were PCR-amplified according to standard procedure in a total volume of 25µl with 40ηg genomic DNA, 240ηM of primer (Invitrogen Corp., Carlsbad, CA, USA), 200µM dNTP, and 1 unit of Taq DNA Polymerase (Fermentas Life Sciences, Burlington, ON, Canada) in GeneAmp® PCR System 9700 (Applied Biosystems, Applera Corp., Foster City, CA, USA). PCR products were resolved on 8% non-denaturing polyacrylamide gel and genotypes were assigned by visual inspection.

Linkage analysis

PEDCHECK [4] was used to identify Mendelian inconsistencies while the MERLIN [5] program was utilized to detect potential genotyping errors that did not produce a Mendelian inconsistency. For both genome scan and fine-mapping markers, two-point linkage analysis was carried out with the MLINK program of the FASTLINK computer package [6] and multipoint linkage analysis was performed using ALLEGRO [7]. An autosomal recessive mode of inheritance with complete penetrance and a disease allele frequency of 0.001 were assumed. Genome scan marker allele frequencies were estimated from the founders and reconstructed genotypes of founders from these pedigrees and additional pedigrees that underwent a genome scan at the same time. For the fine-mapping markers, it was not possible to estimate allele frequencies from the founders, because these markers were only genotyped in these two families, so equal allele frequencies were applied initially. Since false positive results can be obtained when analyzing the data using too low of an allele frequency for an allele segregating with the disease locus [8], a sensitivity analysis was carried out for the multipoint linkage analysis by varying the allele frequency of the allele that is segregating with the disease locus from 0.2 to 0.6 for the fine-mapping markers. In order to determine the order of fine-mapping and genome scan markers, the physical position of each marker was determined from the National Center for Biotechnology Information (NCBI) Build 34 sequence-based physical map [9]. Genetic map distances were then derived from the Rutgers combined linkage-physical map of the human genome [10], either directly or by interpolation. After linkage analyses, haplotypes were constructed via SIMWALK2 [11-12].

Candidate gene sequencing

Using Primer3 software [13], primers were designed for the exons of the following genes: *MYO1F* (MIM 601480; NM_012335); *KEAP1* (MIM 606016; NM_012289); *SLC44A2*/*CTL2* (MIM 606106; NM_020428); and *CDKN2D* (MIM 600927; NM_001800). From each family, DNA from one unaffected and two hearing-impaired individuals were diluted to 5μg/ηl, amplified by PCR under standard conditions, and purified with ExoSAP-IT® (USB Corp., Cleveland, Ohio, USA). Sequencing was performed with the BigDye® Terminator v3.1 Cycle Sequencing Kit together with an Applied Biosystems 3700 DNA Analyzer (Applera Corp., Foster City, CA, USA). Sequence variants were identified via Sequencher™ Version 4.1.4 software (Gene Codes Corp., Ann Arbor, MI, USA).

RESULTS

Two-point linkage analysis of the genome scan markers generated a maximum LOD score at θ=0 of 2.2 for family 4100 and 3.3 for family 4154; both LOD scores were derived at marker D19S586. For the multipoint analysis, the maximum LOD score was obtained also at genome scan marker D19S586 for both families, with a score of 3.0 for family 4100 and 3.2 for family 4154.

In order to fine-map the region, 23 additional markers from the Marshfield genetic map [14] plus five markers from the Rutgers combined linkage-physical map [10] were selected and genotyped in both families. The results of the two-point linkage analyses are presented in Table 2.6, which shows the LOD scores that were derived at each marker at θ=0. The highest two-point LOD score for family 4100 was 2.3 at markers D19S583~D19S581 and D19S558, and for family 4154 it was 3.3 at D19S586. Note that some of the markers were uninformative for one of the two families (Table 2.6); these markers were removed while performing multipoint linkage analyses and haplotype reconstruction for the family for which the markers were uninformative.

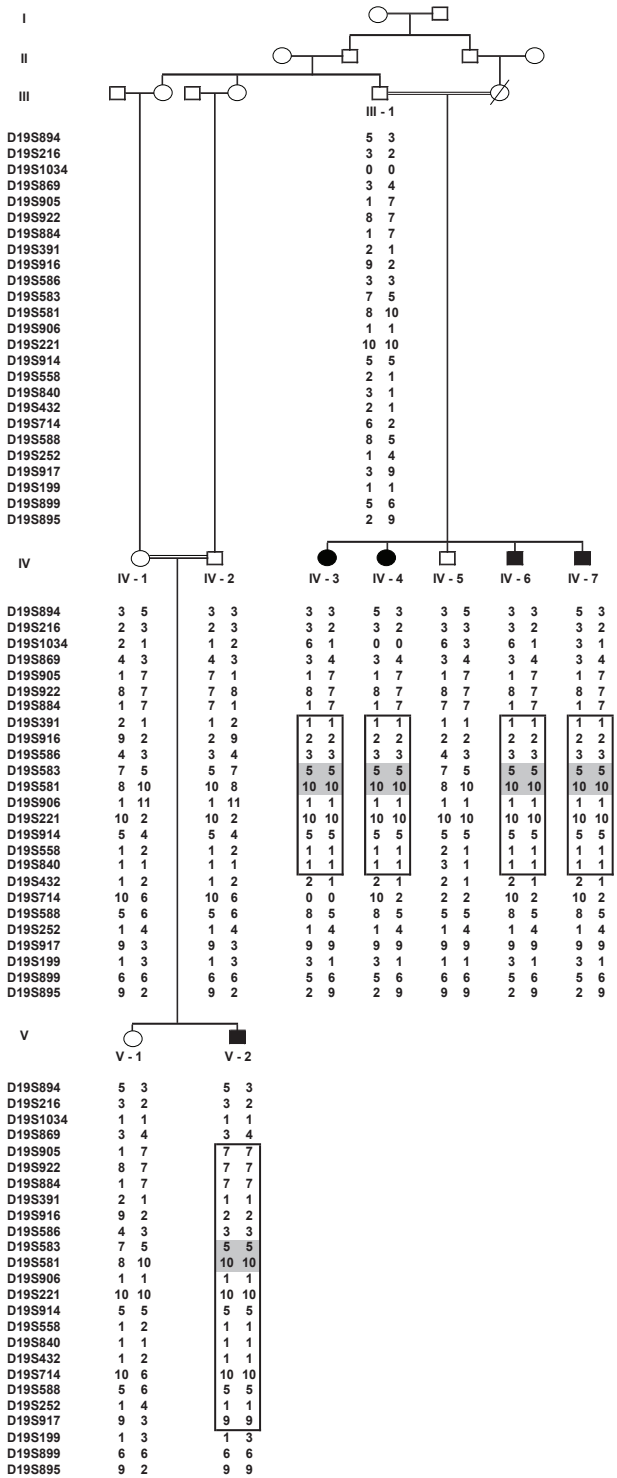
For family 4100, two peaks in the multipoint LOD scores can be seen: one with a maximum LOD score of 4.8 at D19S581, and a second peak with a value of 4.7 at marker D19S558. For the first peak, the 3-unit support interval was delimited by markers D19S884 and D19S906. The second peak was bordered by D19S914 and D19S432. When sensitivity analysis was carried out by varying the frequencies of the fine mapping marker alleles which segregate with the disease phenotype between 0.2 and 0.6, the maximum LOD score fluctuated from 5.1 to 3.2 but remained at the same markers D19S581 and D19S558.

On the other hand, for family 4154, the maximum LOD score after multipoint linkage analysis was 4.6 at markers D19S432~D19S714~D19S252, with the 3-unit support interval between markers D19S1034 and D19S199. When sensitivity analyses were performed at marker allele frequencies 0.2-0.6, the maximum multipoint LOD score was maintained at 4.5-4.6, not only at markers D19S432~D19S714~D19S252 but also at markers D19S581~D19S584.

The three-unit support interval for family 4154 contains 10.8 Mb according to the human reference sequence and has a length of 21.8 cM based on the Rutgers combined linkage-physical map. In comparison, the three-unit support interval for family 4100 is smaller, with 8.2 cM/3.7 Mb for the first peak and 5.8 cM/2.7 Mb for the second peak, both of which are completely contained within the interval for family 4154.

To further delineate the boundaries of the *DFNB68* locus, haplotypes were reconstructed for each family (Figure 2.10). In four hearing-impaired siblings from family 4100, the region of homozygosity is bounded by markers D19S884 and D19S432 (Figure 2.10a), and therefore spans the two three-unit support interval (includes 7.5 Mb and is 14.3 cM long). The region of homozygosity in family 4154 is the same as the three-unit support interval, and is delimited at the proximal end by D19S1034 in all affected family members, and at the distal end by marker D19S199 due to a historic recombination event in individual V-6 (Figure 2.10b). Upon closer inspection, however, the unaffected individuals from both families are also homozygous for several of the markers within the region of homozygosity. If we take into consideration solely those markers that are homozygous only in the hearing-impaired but not among unaffected relatives, the true region of homozygosity in family 4100 is bounded by D19S586 and D19S906, while in family 4154 the interval is bordered by D19S905 and D19S584. The overlap between these two regions from the two families is flanked by markers D19S586 and D19S584, which can now be assigned as the limits of the *DFNB68* locus. The genetic interval for the *DFNB68* locus is 1.9 cM and contains 1.4 Mb.

Within the 1.4 Mb region of homozygosity, there are 40 known genes, of which four encode hypothetical proteins (Build 34 of the human reference sequence as seen from the University of California Santa Cruz (UCSC) Genome Browser). Five genes were found within published inner ear databases [15-16], namely: *DNMT1* (MIM 126375); *ICAM1* (MIM 147840); *KEAP1*; *CTL2*; and *SMARCA4* (MIM 603254) (Figure 2.11). The genes *KEAP1* and *CTL2* were screened by sequencing in the two families, but no functional


I
II
III
III - 1
IV
IV - 1
IV - 2
IV - 3
IV - 4
IV - 5
IV - 6
IV - 7
V
V - 1
V - 2
D19S894
D19S216
D19S1034
D19S869
D19S905
D19S922
D19S884
D19S391
D19S916
D19S586
D19S583
D19S581
D19S906
D19S221
D19S914
D19S558
D19S840
D19S432
D19S714
D19S588
D19S252
D19S917
D19S199
D19S899
D19S895

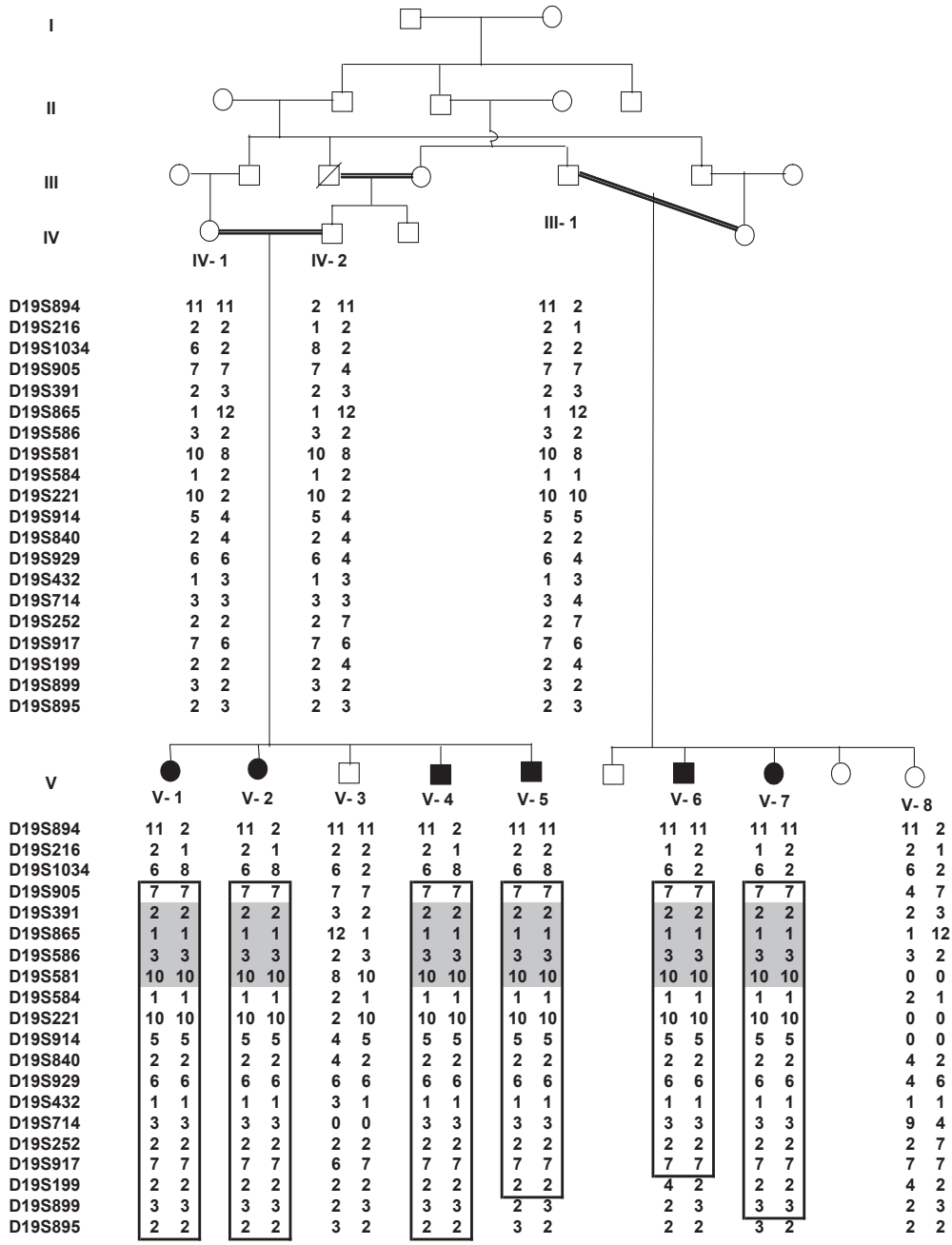


Figure 2.10 Drawing of pedigrees 4100 (a) and 4154 (b). Filled symbols represent individuals with hearing impairment, while clear symbols are for unaffected family members. Haplotypes are shown beneath each genotyped individual with maternal haplotypes on the left-hand side and paternal haplotypes on the right. The boxed area indicates the region of homozygosity for each individual. Shaded genotypes are for markers that completely segregate with the hearing impairment status. Genotypes are marked as "0 0" are unknown for the individual.

Table 2.6 Two-point LOD score results between *DFNB68* and chromosome 19 fine-mapping markers

Marker [a]	Genetic map position [b]	Physical map position [c]	LOD score at *theta*= 0 [d]	
			4100	4154
D19S591	*9.26*	*3,026,844*	*-0.48*	*-1.22*
D19S894	14.64	4,343,406	-∞	-∞
D19S216	16.55	4,900,357	-2.14	-∞
D19S1034	*18.80*	*6,064,253*	-∞	*-1.05*
D19S869	22.31	7,025,187	-2.44	--
D19S905	23.56	7,575,847	-0.85	1.73
D19S922	24.43	8,043,696	-0.85	**--**
D19S884	24.43	8,055,955	-0.87	**--**
D19S391	26.21	8,524,537	1.73	3.15
D19S916	27.99	8,978,841	1.73	--
D19S865	28.79	9,039,342	**--**	3.15
D19S586	***29.67***	***9,665,793***	***2.29***	***3.28***
D19S583	29.67	9,919,343	2.30	--
D19S581	31.59	10,556,822	2.30	3.02
D19S584	**31.59**	**11,063,672**	**--**	**2.88**
D19S906	32.66	11,787,026	-∞	--
D19S221	32.91	12,573,742	-∞	2.62
D19S914	32.91	12,802,614	-∞	2.62
D19S558	33.79	13,174,061	2.30	--
D19S840	34.62	13,701,918	2.08	2.88
D19S929	37.99	15,021,440	--	1.88
D19S432	38.71	15,542,990	-0.85	2.63
D19S714	*38.71*	*15,589,133*	*0.09*	*1.32*
D19S588	38.71	15,589,156	-0.5	--
D19S252	38.71	15,598,394	-0.85	1.88
D19S917	39.72	16,222,017	1.73	2.63
D19S199	40.56	16,829,262	-2.12	-∞
D19S899	41.58	17,095,083	-1.07	-∞
D19S895	45.68	18,614,110	-0.87	-∞
D19S931	48.74	33,315,122	--	-∞
D19S919	49.31	34,602,836	0.54	-∞
D19S433	*50.73*	*35,108,867*	-∞	-∞
D19S396	51.01	35,447,110	--	-∞

[a] Genome scan markers are shown in italics. Markers in bold flank the haplotype. Five markers – D19S583, D19S581, D19S906, D19S221 and D19S914 – were chosen from the Rutgers combined linkage-physical map [10]. The rest of the fine-mapping markers were selected from the Marshfield genetic map [14].

[b] Sex-averaged Kosambi map distance (cM) from the Rutgers map [10].

[c] Sequence-based physical map distance in bases according to Build 34 of the human reference sequence [9].

[d] LOD scores were kept blank at markers that were non-informative for the family.

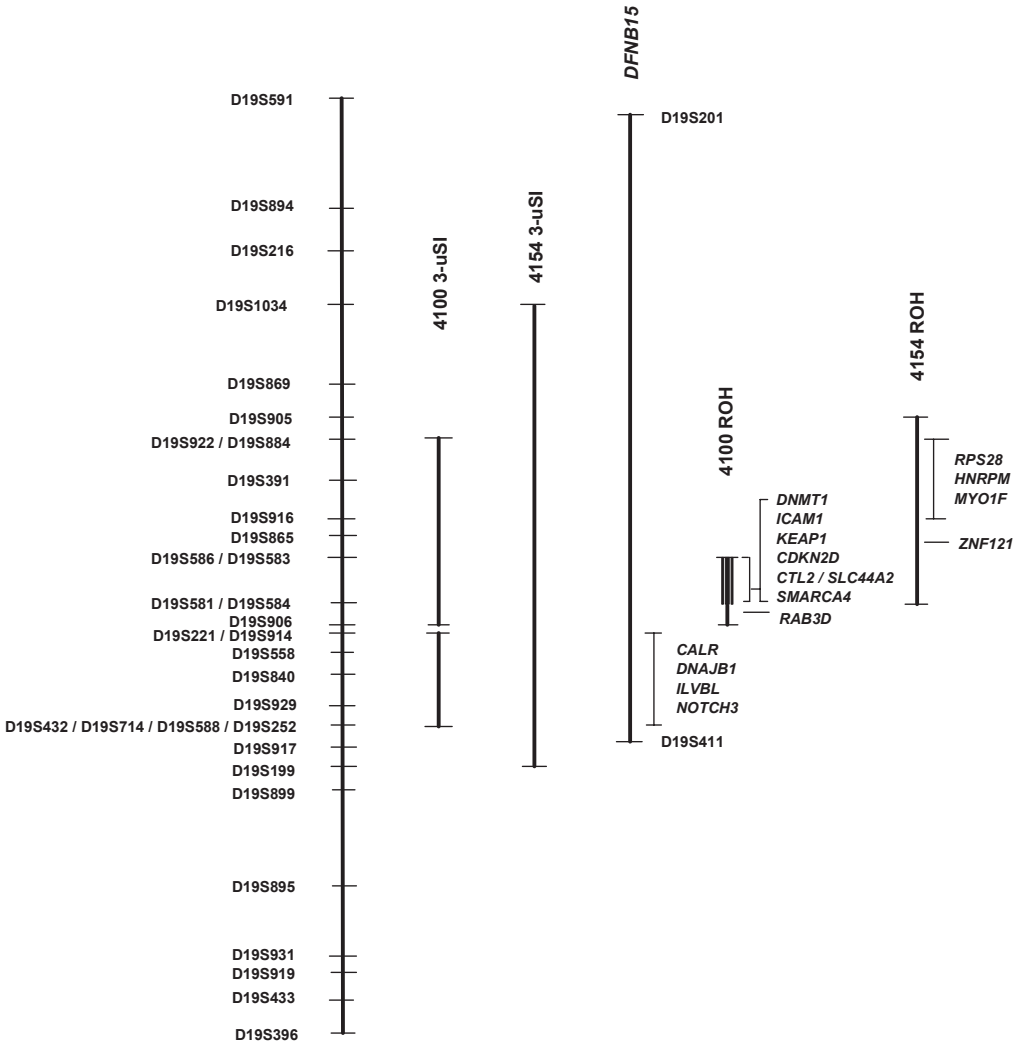


Figure 2.11 Schematic representation of the three-unit support interval (3-u SI) and region of homozygosity (ROH) for the *DFNB68* locus for families 4100 and 4154. Markers are scaled according to genetic map position in the Rutgers map [10]. The *DFNB15* interval at chromosome 19 is also shown for comparison. Candidate genes that are found in inner ear expression databases are listed. The region of overlap between the ROH of the two families that completely segregate with hearing impairment (thick line on 4100 ROH) is flanked by D19S586 and D19S584, and includes the genes *DNMT1, ICAM1, KEAP1, CDKN2D, CTL2* and *SMARCA4*.

sequence variants were identified. The *CDKN2D* gene, which is homologous to hearing loss gene *Ink4d* in mice [17], was also sequenced; however no functional variants were found in the coding exons of *CDKN2D*.

DISCUSSION

The *DFNB15* locus was previously mapped to two regions, 3q21.3-q25.2 and 19p13.3-p13.1 [2]. However it has not been established which one of these regions actually holds the causative HI gene, or if there is digenic inheritance for the family involved. One caveat was that the maximum LOD score that could ever be derived for the *DFNB15* family was less than 3.0, and analyses of both regions on chromosomes 3 and 19 both resulted in a LOD score of 2.8. The new locus described here is contained within the *DFNB15* interval on chromosome 19 (Figure 2.11). Additionally for the two families reported here there is no evidence of linkage to chromosome 3q or to any other chromosome, discounting any possibility of digenic inheritance. Nevertheless this does not bolster or refute the hypothesis that an HI gene in the 19p13 region caused the HI in the *DFNB15* family. Due to the lack of evidence that the HI locus in this article is the same as *DFNB15*, this novel locus on 19p13.2 was assigned as *DFNB68*. Only one candidate gene, *MYO1F*, a member of the unconventional myosin family (which includes HI genes *MYO1A* (MIM 601478), *MYO6* (MIM 600970), *MYO7A* (MIM 276903), and *MYO15* (MIM 602666)), was reported to have been sequenced in the *DFNB15* family, but no mutation was found [18]. Likewise the *MYO1F* gene had no functional sequence variants in families 4100 and 4154. This is not surprising given that *MYO1F* lies outside the 1.9 cM region of *DFNB68* (Figure 2.11).

Although high LOD scores were obtained for the two families at the more proximal marker D19S391 and at distal markers D19S558 and D19S840 (Table 2.6), it was discovered that at these markers the families were dissimilar in terms of the allele that was homozygous among affected individuals in each family. On the other hand, hearing-impaired individuals in both families are homozygous for the same alleles at contiguous markers from D19S586 to D19S914, though some of these markers were non-informative for each family. Based on the segregation of alleles among hearing-impaired and hearing members of each pedigree, the assigned region of homozygosity for each pedigree decreased in length (Figure 2.10) and from these smaller intervals the region of overlap was designated as the *DFNB68* locus. Currently we are concentrating our search for candidate genes within the 1.4 Mb region of overlap seen in the two families, which is

between markers D19S586 and D19S584 (Figure 2.11). It is not possible to conclude if both families have HI due to the same functional variant, due to different functional variants within the same gene, or even due to two different genes. Nevertheless the observation of the same haplotypes in both families within the 1.9 cM region of homozygosity suggests that they may segregate the same variant.

Of the 40 genes within the overlapping regions of homozygosity, two candidate genes in the inner ear databases [15-16], *KEAP1* and *CTL2*, were selected for sequencing in the *DFNB68* families. *KEAP1* (Kelch-like Ech-associated protein 1) is a functional repressor of transcription factor *NRF2* (MIM 600492), which is a regulator of detoxifying and antioxidant genes. It was discovered that the Kelch repeat domain of the *KEAP1* protein associates with the SH3 domain of *MYO7A* in specialized adhesion junctions in the testis [19]. Moreover both *KEAP1* and *MYO7A* are expressed in cell bodies of cochlear inner hair cells, and it was proposed that *KEAP1* facilitates the association of *MYO7A* with the actin cytoskeleton of hair cell stereocilia, thus promoting cell adhesion that is important for stereocilia bundle organization. Screened family members of 4100 and 4154 were negative for functional sequence variants in the coding regions of the *KEAP1* gene.

The *CTL2* (choline-transporter like protein 2) gene encodes a highly conserved transmembrane protein which is the target of antibody-mediated hair cell loss [20]. The protein is expressed in cochlear and vestibular supporting cells, and is purported to function in these cells as a choline transporter. However DNA sequencing in families that were linked to *DFNB68* revealed no functional variants in the *CTL2* gene.

The *CDKN2D* (cyclic-dependent kinase inhibitor 2D) gene is the human homologue of mouse gene *Ink4d*, which is required to maintain the inner hair cells in the post-mitotic state [17]. In *Ink4d* knockout mice, reentry of the inner hair cells into the cell cycle resulted in apoptotic cell death and subsequently progressive hearing loss. Although *CDKN2D* was not identified in the inner ear databases, other kinase inhibitors (*CDKN1A, CDKN1B*) were also found to be expressed in the inner ear [15]. Members of families 4100 and 4154 did not carry functional variants within the coding exons of *CDKN2D*.

Three other genes within the *DFNB68* interval, *DNMT1*, *ICAM1* and *SMARCA4*, were found in the inner ear databases. For these genes there is currently no additional information that is linked to hearing physiology. There is, however, a report of *ICAM1* mRNA expression by cultured spiral

ligament fibrocytes after cytokine (TNF-α) stimulation [21], which may suggest a role for *ICAM1* in immune-mediated inner ear disease. Based on a positional cloning approach, the genes within the *DFNB68* interval will be investigated further for a possible role in HI causation.

Acknowledgements We thank the two families who participated in this research. We are also grateful to Xuan Zhang for assistance with the figures. This work was supported by the Higher Education Commission, Pakistan, and National Institutes of Health National Institute of Deafness and other Communication Disorders grant R01-DC03594. Genotyping services were provided by the Center for Inherited Disease Research (CIDR) and the National Heart, Lung and Blood Institute Mammalian Genotyping Service (Marshfield, WI, USA). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, Contract Number N01-HG-65403.

REFERENCES

1. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
2. Chen A, Wayne S, Bell A, Ramesh A, Srisailapathy CRS, Scott DA, Sheffield VC, Van Hauwe P, Zbar RIS, Ashley J, Lovett M, Van Camp G, Smith RJH. New gene for autosomal recessive non-syndromic hearing loss maps to either chromosome 3q or 19p. Am J Med Genet 1997; 71: 467-471.
3. Grimberg J, Nawoschik S, Bellusico L, McKee R, Trucks A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
4. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63: 259-266.
5. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin - rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002; 30: 97-101.
6. Cottingham Jr RW, Indury RM, Schaffer AA. Faster sequential genetic linkage computation. Am J Hum Genet 1993; 53: 252-263.
7. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12-13.
8. Freimer NB, Sandkuijl LA, Blower S. Incorrect specification of marker allele frequencies: effects on linkage analysis. Am J Hum Genet 1993; 52: 1102-1110.
9. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860-921. July 2003 reference sequence as viewed in: Genome Bioinformatics Group of UC Santa Cruz. UCSC Genome Browser. URL: http://genome.ucsc.edu/cgi-bin/hgGateway
10. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
11. Weeks DE, Sobel E, O'Connell JR, Lange K. Computer programs for multilocus haplotyping of general pedigrees. Am J Hum Genet 1995; 56: 1506-1507.
12. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996; 58: 1323-1337.

13. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. Bioinformatics Methods and Protocols: Methods in Molecular Biology. NJ: Humana Press, 2000. pp 365-386.
14. Broman K, Murray JC, Scheffield VC, White RL, Weber JL. Comprehensive human genetics maps: Individual and sex specific variation in recombination. Am J Hum Genet 1998; 63: 861-869.
15. The Hearing Research Group at Brigham & Women's Hospital. Human Cochlear cDNA library and EST database. 2002. URL: http://hearing.bwh.harvard.edu/estinfo.HTM
16. Holme RH, Bussoli TJ, Steel KP. Table of gene expression in the developing ear. 2002. URL: http://www.ihr.mrc.ac.uk/Hereditary/ genetable/index.shtml
17. Chen P, Zindy F, Abdala C, Liu F, Li X, Roussel MF, Segil N. Progressive hearing loss in mice lacking the cyclin-dependent kinase inhibitor *Ink4d*. Nat Cell Biol 2003; 5: 422-426.
18. Chen AH, Stephan DA, Hasson T, Fukushima K, Nelissen CM, Chen AF, Jun AI, Ramesh A, Van Camp G, Smith RJ. *MYO1F* as a candidate gene for nonsyndromic deafness, *DFNB15*. Arch Otolaryngol Head Neck Surg. 2001; 127: 921-925.
19. Velichkova M, Guttman J, Warren C, Eng L, Kline K, Vogl AW, Hasson T. A human homologue of Drosophila kelch associates with myosin-VIIa in specialized adhesion junctions. Cell Motil Cytoskeleton 2002; 51: 147-164.
20. Nair TS, Kozma KE, Hoefling NL, Kommareddi PK, Ueda Y, Gong TW, Lomax MI, Lansford CD, Telian SA, Satar B, Arts HA, El-Kashlan HK, Berryhill WE, Raphael Y, Carey TE. Identification and characterization of choline transporter-like protein 2, an inner ear glycoprotein of 68 and 72 kDa that is the target of antibody-induced hearing loss. J Neurosci 2004; 24: 1772-1779.
21. Ichimiya I, Yoshida K, Suzuki M, Mogi G. Expression of adhesion molecules by cultured spiral ligament fibrocytes stimulated with proinflammatory cytokines. Ann Otol Rhinol Laryngol 2003; 112: 722-728.

Chapter 3
Functional Variants in Known Hearing Impairment Genes

Chapter 3.1

Low prevalence of Connexin 26 (*GJB2*) variants in Pakistani families with autosomal recessive non-syndromic hearing impairment

Santos RLP, Wajid M, Pham TL, Hussan J, Ali G, Ahmad W, Leal SM. Reprinted from Clinical Genetics 67(1): 61-68. Copyright © Blackwell Munskgaard 2005.

SUMMARY

The Pakistani population has become an important resource for research on autosomal recessive non-syndromic hearing impairment (ARNSHI) due to the availability of large extended and highly consanguineous pedigrees. Here is presented the first report on the prevalence of gap junction beta-2 (*GJB2*) variants in Pakistan. One hundred and ninety-six unrelated Pakistani families with ARNSHI were recruited for a study on the genetics of NSHI. DNA sequencing of the *GJB2* coding region was done on two affected individuals per family. Evolutionary conservation and predicted effect on the protein product were studied in order to hypothesize whether or not a variant was potentially deleterious. Homozygous putatively functional *GJB2* variants were identified in 6.1% of

families. None of the putatively functional *GJB2* variants were observed in the compound heterozygous state. The six putatively causative variants noted were 231G>A(W77X), 71G>A(W24X), 167delT, 95G>A(R32H), 358_360delGAG(delE120), and 269T>C(L90P), with 231G>A(W77X) and 71G>A(W24X) being the most common. In addition, five benign polymorphisms, 380G>A(R127H), 457G>A(V153I), 493C>T(R165W), 79G>A(V27I) and 341A>G(E114G), were identified within this population. In a few individuals, benign polymorphisms were observed to occur on the same haplotype, namely [457G>A(V153I);493C>T(R165W)] and [79G>A(V27I);341A>G(E114G)]. The spectrum of *GJB2* sequence variants in Pakistan may reflect shared origins of hearing impairment alleles within the Indian subcontinent. The high degree of consanguinity within Pakistan may have maintained the *GJB2* prevalence at a much lower rate than within India and other populations.

INTRODUCTION

In 1997 Kelsell et al. [1] identified the first non-syndromic hearing impairment (NSHI) gene, the gap junction beta-2 gene (*GJB2*), which encodes for connexin 26 (*CX26*). Since then 21 genes have been identified and 37 loci mapped for autosomal recessive (AR) NSHI [2]. At present there are >100 known sequence variants for *GJB2*, of which 56 are associated with ARNSHI [3]. Some *GJB2* mutations have high prevalence rates in specific populations, namely: 35delG among people of European descent, 167delT in the Ashkenazim, 235delC among East Asians, and 427C>T(R143W) in the Ghanaian population [4-7].

The elevated frequency of *GJB2* alleles in large populations has been attributed to relaxed selection and assortative mating in the past two centuries, i.e. a rapid increase in survival and reproductive fitness and higher rates of intermarriage among the deaf [8]. Alternatively, being a carrier might confer a selective advantage for better survival, such as thicker epidermis and greater salt concentration in sweat [9]. On the other hand, consanguinity was associated with a decreased risk of hearing impairment (HI) due to *GJB2* [10].

The Pakistani population has become an important resource for research on ARNSHI due to the availability of large extended pedigrees. In addition, about 60% of marriages are consanguineous, of which 80% are first cousin unions [11]. Here we describe the allele frequency and spectrum of *GJB2* sequence variants in Pakistan using data on families with ARNSHI.

MATERIALS AND METHODS

The study was approved by the Quaid-I-Azam University Institutional Review Board and by the Institutional Review Board for Human Subject Research for Baylor College of Medicine and Affiliated Hospitals. Informed consent was obtained from all family members who participated in the study.

One hundred and ninety-six unrelated Pakistani families with at least two ARNSHI individuals were ascertained from various regions of Pakistan. Medical and family history and information on pedigree structure were obtained from multiple family members. Pure tone audiometry at 2500-8000 Hz was performed for selected subjects. All hearing-impaired (HI) family members underwent physical examination. No clinical features, including mental retardation, that would indicate that the HI was part of a syndrome, were observed. In addition, no gross vestibular involvement was noted. The HI phenotype was prelingual, severe to profound, and was not known to be caused by inflammatory middle ear disease or specific environmental factors.

DNA was isolated from venous blood samples following a standard protocol [12] and stored at -20°C. Genomic DNA was diluted to 5ηg/μl and optimized under standard PCR conditions with primers 5'-TGTGCATTCGTCTTTTCCAG-3' and 5'-GGGAAATGCTAGCGACTGAG-3'. The PCR products were then separated by electrophoresis to check proper amplification with genomic DNA. When the desired *GJB2* coding region was amplified, DNA from at least two affected individuals from each family were diluted to 5 ηg/μl, amplified, and purified with ExoSAP-IT® (USB Corp., Cleveland, Ohio, USA, 1999-2004). The second exon of *GJB2* was sequenced using either the forward or the reverse primer and the BigDye® Terminator v3.1 Cycle Sequencing Kit together with an Applied Biosystems 3700 DNA Analyzer (Applera Corp., Foster City, CA, USA, 2004). Single nucleotide polymorphisms (SNPs) were identified via SEQUENCHER™ Version 4.1.4 software (Gene Codes Corp., Ann Arbor, MI, USA, 1991-2002).

To determine the evolutionary conservation of identified substitutions, the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB) was used to look for homologs of the *CX26* protein. ExPASy uses the NCBI BLASTP 1.5.4-Paracel program [13] to search the ExPASy/UniProt database. To perform the BLASTP search, the default settings were used, except that the threshold for expected random matches (E) was set to a more conservative value of

10^{-4} in order to minimize false positive results. Of the 256 query matches (score range 50-470), one match per species was chosen for further alignment. The match that was selected for each of 36 non-human species was least similar to human *CX26* protein sequence in order to decrease the bias for mammalian *CX26* protein sequences. All 37 proteins, including the human *CX26* sequence, were then submitted for multiple sequence alignment via ClustalW [14] at the European Bioinformatics Institute (EBI) using default settings.

The web server PolyPhen from the European Molecular Biology Laboratory (EMBL) was used to assess the functionality of identified substitutions. It annotates non-synonymous SNPs through the identification of sequence changes in comparison with the HGVbase v.12 database [15]. It uses both mathematical algorithms and an empirically-derived set of prediction rules to predict the phenotypic effect of the substitution.

For all point estimates 95% confidence intervals (CI) were calculated using the binomial distribution. To determine the effect of ethno-linguistic origin on the distribution of *GJB2* alleles, chi-square tests were applied.

RESULTS

Of 196 families studied, 94% presented with consanguineous matings, with an average of three consanguineous matings per pedigree. In total 430 affected subjects were screened for *GJB2* variations. Forty-eight of the families had 11 different variants within the *GJB2* gene. Six of the identified variants are putatively functional variants while five other identified polymorphisms are presumably benign (Table 3.1). No 35delG, 235delC or 427C>T(R143W) alleles were detected. One pedigree presented with HI individuals who were homozygous for 167delT. The allele frequency in this study population for both putatively functional and benign *GJB2* variants is 16.6% (95%CI: 13.0, 20.6) while the allele frequency for functional variants alone is 6.6% (95%CI: 4.4, 9.6). There were two families for which only heterozygous individuals were identified with putatively causative variants 71G>A (W24X) and 269T>C (L90P). It was considered that these families did not have HI due to *GJB2*. Therefore only 6.1% (95%CI: 3.2, 10.4) of the families had *GJB2*-related HI. In two of these families, one affected family member was homozygous for a functional *GJB2* variant while another HI family member was homozygous wild type. This suggests that for these families there is either intrafamilial heterogeneity and/or phenocopies present.

Table 3.1 Gap junction beta-2 (*GJB2*) sequence variations identified in Pakistani families

Sequence variation	Number of families with homozygous/ heterozygous genotype [a]	Evolutionary conservation [b]	Functional effect [c]	Allele frequencies among controls in other populations [d]
71G>A (W24X)	4 homozygous, 1 heterozygous	Conserved	Possibly damaging	0.0007 (French [16]), 0.01 (Indian [17])
231G>A (W77X)	5 homozygous	Highly conserved	Possibly damaging	Not available
95G>A (R32H)	1 homozygous	Highly conserved	Possibly damaging	0.0001 (French [16])
167delT	1 homozygous	-	Frameshift → truncated	0.001 (Palestinian [18]), 0.04 (Ashkenazim [5])
358_360delGAG (delE120)	1 homozygous	-	Frameshift → truncated	0.0009 (French [16])
269T>C (L90P)	1 heterozygous	Semi-conserved	Possibly damaging	Not available
380G>A (R127H)	1 homozygous, 18 heterozygous	Not conserved	Benign	0.007 (French [16]), 0.18 (Indian [17])
457G>A (V153I)	1 homozygous, 12 heterozygous	Not conserved	Benign	0.004 (French [16]), 0.06 (Indian [17])
457G>A (V153I); 493C>T(R165W)	1 homozygous, 2 heterozygous	Not conserved	Benign	R165W = 0 (Indian [17])
79G>A(V27I); 341A>G(E114G)	1 heterozygous	V27I: Conserved; E114G: Not conserved	Benign	V27I = 0 (Indian [17]), 0.8 (Korean [19]); E114G= 0 (Indian [17]), 0.4 (Korean [19]); Haplotype = 0.05 (Japanese [20])

[a] Only benign polymorphisms 380G>A (R127H) and 457G>A (V153I) were found in the compound heterozygous state.

[b] Human *GJB2* sequence compared with various connexin and connexin-like protein sequences of 36 species via the NCBI BLASTP [13] program and the ExPASy/UniProt databases. Species include four primates, 14 other mammals, one avian, two reptiles, three amphibians, nine bony fishes, one cartilaginous fish and two urochordates.

[c] Possible effect of amino acid substitutions on the protein structure and function were predicted via the web server PolyPhen [15].

[d] References cited in brackets.

Table 3.2 Pakistani families with compound heterozygosity (CH) or haplotypes (Hap) of gap junction beta-2 (*GJB2*) polymorphisms

Family	Genotype of HI Individual	CH or Hap	Genotype of Parents with Normal Hearing	Genotype of Other HI Relatives
4055	[R127H] het + [V153I] het	CH	Mother: [R127H] het + [V153]-Wt Father: NA	Half-sib by father: Wt
4063	[R127H] het + [V153I] het	CH	Mother: [R127]-Wt + [V153I] het Father: [R127H] het + [V153]-Wt	Sib: Wt
4078	[V153I;R165W] hom	Hap	Not tested	A Sib and a 3rd-cousin: [V153I] hom + [R165W] het
4126	[V153I;R165W] het	Hap	Mother: [V153I;R165W] het Father: Wt	Sib: Wt
4193	[V27I;E114G] het	Hap	Mother: Wt Father: [V27I;E114G] het	Sib: Wt
4218	[V153I;R165W] het	Hap	Mother: [R127H] + [V153;R165]-Wt Father: NA	1st cousin once removed: Wt

HI, hearing-impaired; *het,* heterozygous; *hom,* homozygous; *Wt,* wild-type; *NA*, DNA not available. Nucleotide changes: 79G>A = V27I; 341A>G = E114G; 380G>A = R127H; 457G>A = V153I; 493C>T = R165W.

Table 3.2 enumerates the families who had an affected individual who was heterozygous for two different putatively benign variants. For these individuals their normal-hearing parents were sequenced for *GJB2* in order to determine the haplotype for the HI individual. It was found that the individuals who had polymorphisms 380G>A(R127H) and 457G>A(V153I) were compound heterozygous while the [79G>A(V27I);341A>G(E114G)] and [457G>A(V153I);493C>T(R165W)] polymorphisms occurred on the same haplotype. The [79G>A(V27I);341A>G(E114G)] haplotype was previously reported in the Japanese population [20] while the [457G>A(V153I);493C>T(R165W)] haplotype was observed in a large Sri Lankan pedigree [21] (Table 3.2).

When multiple sequence alignment was done using non-*CX26* proteins from 36 species, residues for 95G>A(R32H) and 231G>A(W77X) were shown to be highly conserved. Residues for 71G>A(W24X) and 79G>A(V27I) were also conserved. The 269T>C(L90P) variant was at a semi-conserved residue, while the other variants, 41A>G(E114G), 380G>A(R127H), 457G>A(V153I) and 493C>T(R165W) were at residues that were not conserved.

The results of the homology search were mostly supported by the functional annotation from the PolyPhen web server, wherein the *GJB2* variants 71G>A(W24X), 95G>A(R32H), 231G>A(W77X), and 269T>C(L90P) that occurred at conserved residues were deemed possibly damaging, while those at non-conserved residues, variants 341A>G(E114G), 380G>A(R127H), 457G>A(V153I) and 493C>T(R165W) were considered benign. The only exception was 79G>A(V27I), which occurred at a conserved residue based on the homology search but was predicted to be functionally benign. This can be explained by its location at a transmembrane region which, for hydrophobic residues, is variable in conservation according to a predicted hydrophobic and transmembrane matrix (PHAT) [22]. In contrast 231G>A(W77X), 95G>A(R32H) and 71G>A(W24X), which also occur at the transmembrane region, were considered conserved and damaging due to polarity of residues. The 269T>C(L90P) substitution at the transmembrane region, though with a hydrophobic residue, results in a negative PHAT matrix score and was thus considered possibly damaging.

Of the families that were studied, 70% come from the Punjab province, the largest province in Pakistan. The native languages of these families are Punjabi and Sairiki. The remaining 30% of the families in the study were equally representative of the other provinces and their corresponding language groups. No significant association was found between the presence of a particular *GJB2* variant and the region of origin or linguistic background. Interestingly, four of five 71G>A(W24X) families spoke Sairiki.

DISCUSSION

Allele frequency and spectrum of GJB2 variants

The prevalence of putatively functional *GJB2* variants is in general 3-4 times higher in Europeans and Middle Easterners than in Pakistanis (Table 3.3). There was one exception in the Middle East, Oman, where the prevalence of HI-associated *GJB2* variants was reported to be 0.0 [95% confidence interval (CI): 0, 0.04]. This observation could be due to the mixed population and that only 35delG and 167delT were screened [28]. The prevalence rates of functional *GJB2* variants within Africa varied greatly depending on country, with Kenya and Sudan having lower prevalence rates than Pakistan, while in Tunisia and Ghana the prevalence rates were twice higher than those observed in Pakistan [7, 29-30]. When 95% CI were estimated for the prevalence rates of the putatively functional variants, there was overlap of the 95% CI for the Pakistani prevalence rate and the CIs of the prevalence rate for several Mid-east, African, and East Asian Countries and Brazil [19-21, 25-27, 29-34]. It should be noted that most of the studies from these countries had small sample sizes.

Table 3.3 Overview of gap junction beta-2 (*GJB2*) variants and prevalence among the hearing-impaired in different populations

Place of origin	Probands (*n*)	Frequencies of homozygous and compound heterozygous probands with putatively functional *GJB2* variants (95%CI)[a]	Most common putatively functional *GJB2* variant(s)	References
EUROPE [b-1]	2146	0.28 (0.26, 0.30)	35delG	23
MIDDLE EAST				
Israel	64	0.39 (0.27, 0.52)	35delG, 167delT	24
Palestine [c]	48	0.23 (0.12, 0.37)	35delG, 167delT	18
Lebanon [b-2]	48	0.33 (0.20, 0.48)	35delG	25
Jordan	68	0.16 (0.08, 0.27)	35delG	26
Iran [b-3, c]	83	0.11 (0.05, 0.20)	35delG	27
Oman [b-4]	95	0 (0, 0.04)	-	28
AFRICA				
Ghana [c]	365	0.16 (0.12, 0.20)	427C>T (R143W)	7
Tunisia [b-5]	70	0.17 (0.09, 0.28)	35delG	29
Sudan [c]	183	0.03 (0.009, 0.06)	35delG	30
Kenya [c]	406	0 (0, 0.009)	-	30
SOUTH AMERICA				
Brazil [c]	62	0.11 (0.05, 0.22)	35delG	31
EAST ASIA				
Japan	35+39	0.18 (0.10, 0.28)	235delC	20, 32
China [c]	118	0.16 (0.10, 0.24)	235delC	33
Korea	147	0.05 (0.02, 0.10)	235delC	19
Thailand	12	0.17 (0.02, 0.48)	235delC	34
INDIAN SUBCONTINENT				
British Asian [b-6]	13	0.23 (0.05, 0.54)	231G>A (W77X)	21
India [b-7, c]	45+215	0.17 (0.13, 0.22)	71G>A (W24X)	17,35
Pakistan	**196**	**0.06 (0.03, 0.10)**	**231G>A (W77X), 71G>A (W24X)**	**This article**

[a] Frequencies are based on number of probands and sporadic *HI* individuals tested. If subjects with prelingual and postlingual HI were specified, only prelingually *HI* individuals were included.

[b] Mutation detection strategies that were not by direct sequencing of the second exon of *GJB2* as follows:

1. Greece – initial screening for 35delG by allele-specific polymerase chain reaction (ASPCR) then sequencing for 35delG heterozygotes
2. initial screening via polymerase chain reaction (PCR)-mediated site-directed mutagenesis and *BsiYI* digestion then sequencing for non-35delG subjects
3. initial screening for 35delG by ASPCR then single strand conformational polymorphism (SSCP) for 35delG heterozygotes
4. screening only for 35delG and 167delT by restriction fragment length polymorphism (RFLP) method
5. initial denaturing gradient gel electrophoresis (DGGE) screening then sequencing for samples with shifts
6. denaturing high performance liquid chromatography (DHPLC) screening then sequencing if with altered elution peaks
7. initial ASPCR analysis then confirmation by sequencing [35]

[c] Only these samples (including [17] from India) were tested for exon 1 of *GJB2*.

More importantly, however, the spectra of sequence variations in these studies were essentially different from those observed in Pakistan (Table 3.3). The commonly described mutations 35delG, 427C>T(R143W) and 235delC were not found in the Pakistani sample. Conversely, the 231G>A(W77X) variant has only been reported in people originating from the Indian subcontinent [17, 21, 35]. The 71G>A(W24X) variant has been observed in Europeans, but its prevalence is about three times higher in Pakistanis and at least 20 times higher in Indians [16-17, 23, 35]. In addition 380G>A(R127H) and 457G>A(V153I) had much higher frequencies among Indians and Pakistanis than in Europeans. The 493C>T(R165W) polymorphism has only been observed in families from the Indian subcontinent [17, 21].

Thus, the spectrum of *GJB2* variants in Pakistan may reflect shared origins of HI alleles within the Indian subcontinent. For example, the Slovak Romanies who migrated from India demonstrated a similar pattern of *GJB2* allele distribution, with the 71G>A(W24X) and 380G>A(R127H) variants as most common [36]. Although there was similarity in the *GJB2* variants that were observed in Pakistanis and Indians, absolute frequencies were not similar. The allele frequency of *GJB2* sequence variants is two times higher in India than in Pakistan [17]. For putatively functional variants, the prevalence rate in India is three times higher than in Pakistan. Notably the 71G>A(W24X) variant is seven times more common in India than in Pakistan whereas 231G>A(W77X) comprised a bigger proportion of *GJB2* alleles among Pakistanis [17, 35]. It could be hypothesized that the elevated consanguinity rates in Pakistan versus the rest of the subcontinent have depressed the *GJB2* prevalence in this country.

Identified sequence variants

Based on the criteria of evolutionary conservation, predicted effect of amino acid substitutions on protein, and published frequencies among controls in different populations, there is strong evidence that 71G>A(W24X), 231G>A(W77X), 95G>A(R32H) and 358_360delGAG(delE120) are damaging mutations, whereas 380G>A(R127H), 457G>A(V153I), 493C>T(R165W) and [79G>A(V27I);341A>G(E114G)] are benign polymorphisms (Table 3.1). This was corroborated by articles on profound hearing loss among 71G>A(W24X) and 231G>A(W77X) homozygotes, and in a patient with the [95G>A(R32H)+35delG] genotype [1, 16, 37]. For 167delT homozygotes, HI is usually prelingual, bilateral and non-progressive, but severity may range from mild to profound [5, 24]. The mutations

358_360delGAG(delE120) and 269T>C(L90P) failed to form functional gap junction channels in cellular studies [38-40] though the phenotype was less severe than in 35delG homozygotes [41-42].

On the other hand 380G>A(R127H) and 341A>G(E114G) were not different from wild type in functional studies on transfected HeLa cells [38-39, 43]. Furthermore, 380G>A(R127H) and 457G>A(V153I) were mostly observed in the heterozygous state among the hearing-impaired, and in addition occurred with a relatively high frequency in the hearing control population [16-17]. Polymorphisms 79G>A(V27I) and 341A>G(E114G) have been observed independently and as a haplotype. These variants have high frequencies among East Asians [19-20]. Also the 79G>A(V27I) polymorphism was observed in a hearing individual in the homozygous state [44]. The 493C>T(R165W) variant has not been noted among hearing controls [17], nevertheless its predicted effect on the protein product point to its benignity.

All putatively functional *GJB2* variants occurred on the wild type haplotype, but for putatively benign variants there were a few cases in which individuals were observed to have two *GJB2* polymorphisms. Two individuals were identified to be compound heterozygous for the polymorphisms [380G>A(R127H)]+[457G>A(V153I)]. Higher allele frequencies for these SNPs among individuals from the Indian subcontinent as compared to other populations tend to support the existence of a founder effect [17, 35]. However, mutational hotspots cannot be ruled out. Similarly the observation in the Pakistani sample of haplotypes [457G>A(V153I);493C>T(R165W)] and [79G>A(V27I);341A>G(E114G)] which were reported in populations both within and outside the Indian subcontinent may suggest inheritance from common ancestors [20-21].

Effect of consanguinity on allele frequency

It is hypothesized that the increased *GJB2* frequency in some countries is due to assortative mating among the deaf [8]. Only five of our families (2.6%) had matings between HI subjects, and within four of these families the couples were related. These unions may not have occurred due to the couples' hearing-impaired status but because of their kinship. The lower rate of HI due to *GJB2* variants in this sample of families with ARNSHI may be due to the increased occurrence of other recessive deafness loci in multiple inbred lineages that have been maintained for generations. Currently 14 loci and 9 genes for ARNSHI have been reported in Pakistani pedigrees [2]. Notably, consanguinity is about twice as frequent in Pakistan

than in India, which lends further support to the idea that the lower *GJB2* prevalence in Pakistan compared to India may be due to a higher consanguinity rate. Additionally random genetic drift may explain the difference in *GJB2* prevalence between India and Pakistan. Ascertainment of more HI families may allow deeper scrutiny of the ethno-linguistic background, pedigree structures and genetic constitution of the Pakistani population and further elucidation of the relationship between inbreeding and genetic hearing impairment.

Acknowledgements We are grateful to the families who allowed us to conduct this research. We also thank Dr. Kim Worley for her useful comments and suggestions. This study was supported by the Higher Education Commission, Pakistan and NIH-National Institute of Deafness and Other Communication Disorders Grant DC03594.

DATABASE INFORMATION

On AR genes and loci: Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2004. URL: http://webhost.ua.ac.be/hhh/
On *GJB2* mutations and polymorphisms: Calvo J, Rabionet R, Gasparini P, Estivill X. The Connexin-deafness Homepage. 2004. URL: http://www.crg.es/deafness
On *GJB2* sequence: National Center for Biotechnology Information. URL: http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi? db=nucleotide&val=42558282 Nucleotide NM_004004
For homology search: ExPASy Proteomics Server. 2004. URL: http://us.expasy.org/tools/blast/
For multiple sequence alignment: EMBL-EBI European Bioinformatics Institute, ClustalW. 2004. URL: http://www.ebi.ac.uk/clustalw/
For prediction of effect of non-synonymous SNPs on protein: Polyphen. 2004. URL: http://www.bork.embl-heidelberg.de/PolyPhen/
For prevalence rates of consanguinity: Bittles A. consang.net. 2004. URL: http://www.consang.net/

REFERENCES

1. Kelsell DP, Dunlop J, Stevens HP, Lench NJ, Liang JN, Parry G, Mueller RF, Leigh IM. Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. Nature 1997; 387: 80-83.
2. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2004. URL: http://dnalab-www.uia.ac.be/dnalab/hhh/
3. Calvo J, Rabionet R, Gasparini P, Estivill X. Connexins and deafness Homepage. 2004. URL: http://www.crg.es/deafness
4. Gasparini P, Rabionet R, Barbujani G, Melchionda S, Petersen M, Brondum-Nielsen K, Metspalu A, Oitmaa E, Pisano M, Fortina P, Zelante L, Estivill X. High carrier frequency

of the 35delG deafness mutation in European populations. Genetic Analysis Consortium of *GJB2* 35delG. Eur J Hum Genet 2000; 8: 19-23.
5. Lerer I, Sagi M, Malamud E, Levi H, Raas-Rothschild A, Abeliovich D. Contribution of connexin 26 mutations to nonsyndromic deafness in Ashkenazi patients and the variable phenotypic effect of the mutation 167delT. Am J Med Genet 2000; 95: 53-56.
6. Yan D, Park HJ, Ouyang XM, Pandya A, Doi K, Erdenetungalag R, Du LL, Matsushiro N, Nance WE, Griffith AJ, Liu XZ. Evidence of a founder effect for the 235delC mutation of *GJB2* (connexin 26) in east Asians. Hum Genet 2003; 114: 44-50.
7. Hamelmann C, Amedofu GK, Albrecht K Muntau B, Gelhaus A, Brobby GW, Horstmann RD. Pattern of connexin 26 (*GJB2*) mutations causing sensorineural hearing impairment in Ghana. Hum Mutat 2001; 18: 84-85.
8. Nance WE, Kearsey MJ. Relevance of Connexin Deafness (DFNB1) to Human Evolution. Am J Hum Genet 2004; 74: 1081-1087.
9. Meyer CG, Amedofu GK, Brandner JM, Pohland D, Timmann C, Horstmann RD. Selection for deafness? Nat Med 2002; 8:1332-1333.
10. Tekin M, Duman T, Bogoclu G, Incesulu A, Comak E, Ilhan I, Akar N. Spectrum of *GJB2* mutations in Turkey comprises both Caucasian and Oriental variants: roles of parental consanguinity and assortative mating. Hum Mutat 2003; 21: 552-553.
11. Hussain R, Bittles AH. The prevalence and demographic characteristics of consanguineous marriages in Pakistan. J Biosoc Sci 1998; 30: 261-275. [Abstract]
12. Grimberg J, Nawoschik S, Bellusico L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
13. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389-3402.
14. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-4680.
15. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002; 30: 3894-3900.
16. Roux AF, Pallares-Ruiz N, Vielle A, Faugere V, Templin C, Leprevost D, Artieres F, Lina G, Molinari N, Blanchet P, Mondain M, Claustres M. Molecular epidemiology of DFNB1 deafness in France. BMC Med Genet 2004; 5: 5.
17. RamShankar M, Girirajan S, Dagan O, Ravi Shankar HM, Jalvi R, Rangasayee R, Avraham KB, Anand A. Contribution of connexin26 (*GJB2*) mutations and founder effect to non-syndromic hearing loss in India. J Med Genet 2003; 40: e68.
18. Shahin H, Walsh T, Sobe T, Lynch E, King MC, Avraham KB, Kanaan M. Genetics of congenital deafness in the Palestinian population: multiple connexin 26 alleles with shared origins in the Middle East. Hum Genet 2002; 110: 284-289.
19. Park HJ, Hahn SH, Chun YM, Park K, Kim HN. Connexin26 mutations associated with nonsyndromic hearing loss. Laryngoscope 2000; 110: 1535-1538.
20. Kudo T, Ikeda K, Kure S, Matsubara Y, Oshima T, Watanabe K, Kawase T, Narisawa K, Takasaka T. Novel mutations in the connexin 26 gene (*GJB2*) responsible for childhood deafness in the Japanese population. Am J Med Genet 2000; 90: 141-145.
21. Rickard S, Kelsell DP, Sirimana T, Rajput K, MacArdle B, Bitner-Glindzicz M. Recurrent

mutations in the deafness gene *GJB2* (connexin 26) in British Asian families. J Med Genet 2001; 38: 530-533.
22. Ng PC, Henikoff JG, Henikoff S. PHAT: a transmembrane-specific substitution matrix. Bioinformatics 2000; 16: 760-766.
23. Pampanos A, Economides J, Iliadou V, Neou P, Leotsakos P, Voyiatzis N, Eleftheriades N, Tsakanikos M, Antoniadi T, Hatzaki A, Konstantopoulou I, Yannoukakos D, Gronskov K, Brondum-Nielsen K, Grigoriadou M, Gyftodimou J, Iliades T, Skevas A, Petersen MB. Prevalence of *GJB2* mutations in prelingual deafness in the Greek population. Int J Pediatr Otorhinolaryngol 2002; 65: 101-108.
24. Sobe T, Vreugde S, Shahin H, Berlin M, Davis N, Kanaan M, Yaron Y, Orr-Urtreger A, Frydman M, Shohat M, Avraham KB. The prevalence and expression of inherited connexin 26 mutations associated with nonsyndromic hearing loss in the Israeli population. Hum Genet 2000; 106: 50-57.
25. Mustapha M, Salem N, Delague V, Chouery E, Ghassibeh M, Rai M, Loiselet J, Petit C, Megarbane A. Autosomal recessive non-syndromic hearing loss in the Lebanese population: prevalence of the 30delG mutation and report of two novel mutations in the connexin 26 (*GJB2*) gene. J Med Genet 2001; 38: E36.
26. Medlej-Hashim M, Mustapha M, Chouery E, Weil D, Parronaud J, Salem N, Delague V, Loiselet J, Lathrop M, Petit C, Megarbane A. Non-syndromic recessive deafness in Jordan: mapping of a new locus to chromosome 9q34.3 and prevalence of DFNB1 mutations. Eur J Hum Genet 2002; 10: 391-394.
27. Najmabadi H, Cucci RA, Sahebjam S, Kouchakian N, Farhadi M, Kahrizi K, Arzhangi S, Daneshmandan N, Javan K, Smith RJ. *GJB2* mutations in Iranians with autosomal recessive non-syndromic sensorineural hearing loss. Hum Mutat 2002; 19: 572.
28. Simsek M, Al-Wardy N, Al-Khayat A, Shanmugakonar M, Al-Bulushi T, Al-Khabory M, Al-Mujeni S, Al-Harthi S. Absence of deafness-associated connexin-26 (*GJB2*) gene mutations in the Omani population. Hum Mutat 2001; 18: 545-546.
29. Masmoudi S, Elgaied-Boulila A, Kassab I, Ben Arab S, Blanchard S, Bouzouita JE, Drira M, Kassab A, Hachicha S, Petit C, Ayadi H. Determination of the frequency of connexin26 mutations in inherited sensorineural deafness and carrier rates in the Tunisian population using DGGE. J Med Genet 2000; 37: E39.
30. Gasmelseed NM, Schmidt M, Magzoub MM, Macharia M, Elmustafa OM, Ototo B, Winkler E, Ruge G, Horstmann RD, Meyer CG. Low frequency of deafness-associated *GJB2* variants in Kenya and Sudan and novel *GJB2* variants. Hum Mutat 2004; 23: 206-207.
31. Oliveira CA, Maciel-Guerra AT, Sartorato EL. Deafness resulting from mutations in the *GJB2* (connexin 26) gene in Brazilian patients. Clin Genet 2002; 61: 354-358.
32. Abe S, Usami S, Shinkawa H, Kelley PM, Kimberling WJ. Prevalent connexin 26 gene (*GJB2*) mutations in Japanese. J Med Genet 2000; 37: 41-43.
33. Liu XZ, Xia XJ, Ke XM, Ouyang XM, Du LL, Liu YH, Angeli S, Telischi FF, Nance WE, Balkany T, Xu LR. The prevalence of connexin 26 (*GJB2*) mutations in the Chinese population. Hum Genet 2002; 111: 394-397.
34. Kudo T, Ikeda K, Oshima T, Kure S, Tammasaeng M, Prasansuk S, Matsubara Y. *GJB2* (connexin 26) mutations and childhood deafness in Thailand. Otol Neurotol 2001; 22: 858-861.
35. Maheshwari M, Vijaya R, Ghosh M, Shastri S, Kabra M, Menon PS. Screening of families

with autosomal recessive non-syndromic hearing impairment (ARNSHI) for mutations in *GJB2* gene: Indian scenario. Am J Med Genet 2003; 120A: 180-184.
36. Minarik G, Ferak V, Ferakova E, Ficek A, Polakova H, Kadasi L. High frequency of *GJB2* mutation W24X among Slovak Romany (Gypsy) patients with non-syndromic hearing loss (NSHL). Gen Physiol Biophys 2003; 22: 549-556. [Abstract]
37. Marlin S, Garabedian EN, Roger G, Moatti L, Matha N, Lewin P, Petit C, Denoyelle F. Connexin 26 gene mutations in congenitally deaf children: pitfalls for genetic counseling. Arch Otolaryngol Head Neck Surg 2001; 127: 927-933.
38. D'Andrea P, Veronesi V, Bicego M, Melchionda S, Zelante L, Di Iorio E, Bruzzone R, Gasparini P. Hearing loss: frequency and functional studies of the most common connexin26 alleles. Biochem Biophys Res Commun 2002; 296: 685-691.
39. Thonnissen E, Rabionet R, Arbones ML, Estivill X, Willecke K, Ott T. Human connexin26 (*GJB2*) deafness mutations affect the function of gap junction channels at different levels of protein expression. Hum Genet 2002; 111: 190-197.
40. Bruzzone R, Veronesi V, Gomes D, Bicego M, Duval N, Marlin S, Petit C, D'Andrea P, White TW. Loss-of-function and residual channel activity of connexin26 mutations associated with non-syndromic deafness. FEBS Lett 2003; 533: 79-88.
41. Gabriel H, Kupsch P, Sudendey J, Winterhager E, Jahnke K, Lautermann J. Mutations in the connexin26/*GJB2* gene are the most common event in non-syndromic hearing loss among the German population. Hum Mutat 2001; 17: 521-522.
42. Cryns K, Orzan E, Murgia A, Huygen PL, Moreno F, del Castillo I, Chamberlin GP, Azaiez H, Prasad S, Cucci RA, Leonardi E, Snoeckx RL, Govaerts PJ, Van de Heyning PH, Van de Heyning CM, Smith RJ, Van Camp G. A genotype-phenotype correlation for *GJB2* (connexin 26) deafness. J Med Genet 2004; 41: 147-154.
43. Choung YH, Moon SK, Park HJ. Functional study of *GJB2* in hereditary hearing loss. Laryngoscope 2002; 112: 1667-1671.
44. Wilcox SA, Saunders K, Osborn AH, Arnold A, Wunderlich J, Kelly T, Collins V, Wilcox LJ, McKinlay Gardner RJ, Kamarinos M, Cone-Wesson B, Williamson R, Dahl HH. High frequency hearing loss correlated with mutations in the *GJB2* gene. Hum Genet 2000; 106: 399-405.

Chapter 3.2

Novel sequence variants in the *TMC1* gene in Pakistani families with autosomal recessive hearing impairment

Santos RLP, Wajid M, Khan MN, McArthur N, Pham TL, Bhatti A, Lee K, Irshad S, Mir A, Yan K, Chahrour MH, Ansar M, Ahmad W, Leal SM. Reprinted from Human Mutation 26(4): 396. Copyright © Wiley-Liss Inc 2005.

SUMMARY

Though many hearing impairment genes have been identified, only a few of these genes have been screened in population studies. For this study, 168 Pakistani families with autosomal recessive hearing impairment not due to mutations in the *GJB2* (*CX26*) gene underwent a genome scan. Two-point and multipoint parametric linkage analyses were carried out. Twelve families had two-point or multipoint LOD scores of 1.4 or greater within the transmembrane cochlear expressed gene 1 (*TMC1*) region and were subjected to further screening with direct DNA sequencing. Five novel putatively functional non-synonymous sequence variants, c.830A>G (p.Y277C), c.1114G>A (p.V372M), c.1334G>A (p.R445H), c.2004T>G (p.S668R) and c.2035G>A (p.E679K), were found to segregate within seven families, but were not observed in 234 Pakistani control chromosomes. The variants c.830A>G (p.Y277C), c.1114G>A (p.V372M) and c.1334G>A (p.R445H) occurred at highly conserved regions and were predicted to lie within hydrophobic transmembrane domains, while non-synonymous variants c.2004T>G (p.S668R) and c.2035G>A (p.E679K) occurred in extracellular regions that were not highly conserved. There is evidence that the c.2004T>G (p.S668R) variant may have occurred at a phosphorylation site. One family has the known splice site mutation c.536 -8T>A. The prevalence of non-syndromic hearing impairment due to *TMC1* in this Pakistani population is 4.4% (95%CI: 1.9, 8.6%). The *TMC1* protein might have an important function in K^+ channels of inner hair cells, which would be consistent with the hypothetical structure of protein domains in which sequence variants were identified.

INTRODUCTION

Due to the highly complex structure and function of the human inner ear, it is not surprising that sensorineural hearing impairment (HI) is genetically heterogeneous. Currently, 37 non-syndromic hearing impairment (NSHI) genes have been identified, of which 21 are associated with autosomal recessive (AR) NSHI [1]. However, of the ARNSHI genes, only a few [e.g. *GJB2*

(MIM# 121011), *CDH23* (MIM# 605516) and *WFS1* (MIM# 606201)] have been well-characterized in terms of prevalence and sequence variants found in different populations. Additional studies of genes that are involved in NSHI within various populations are necessary to better understand their prevalence, spectrum of sequence variants and public health significance. In this study, *TMC1* (MIM# 606706) was studied within the Pakistani population.

A total of 168 Pakistani families with ARNSHI that were negative for *GJB2* mutations underwent a 10 cM genome scan. Twelve families had two-point or multipoint LOD scores of $\geq$1.4 within the 9q21.13 region where *TMC1* is physically mapped. These families were selected for sequencing of the *TMC1* gene. Thereafter the prevalence of *TMC1* variants in this population was estimated. Also, by locating the residues at predicted transmembrane domains and studying evolutionary conservation in multiple sequence alignment, the possible effects on the *TMC1* protein product of both novel and previously reported sequence variants were examined.

MATERIALS AND METHODS

Ascertainment of study subjects

The study was approved by the Quaid-I-Azam University Institutional Review Board and by the Institutional Review Board for Human Subject Research for Baylor College of Medicine and Affiliated Hospitals. Informed consent was obtained from all family members who participated in the study.

For this study, 192 unrelated Pakistani families with at least two ARNSHI individuals were ascertained from various regions of Pakistan. Medical and family history and information on pedigree structure was obtained from multiple family members. Pure tone audiometry at 250-8000 Hz was performed for selected subjects. All hearing-impaired family members underwent physical examination. Within these families, no clinical features that would indicate that the HI was part of a syndrome were observed. In addition, no gross vestibular involvement was noted. The HI phenotype was prelingual, severe to profound, and was not known to be caused by inflammatory middle ear disease or specific environmental factors. Additionally, 117 unrelated hearing individuals without a family history of hearing impairment were ascertained from Pakistan.

Genome scan

DNA was isolated from venous blood samples following a standard protocol [2], quantified by spectrophotometry at optical density 260, and

stored at -20°C. Of the 192 families, twelve families were positive for *GJB2* (GenBank accession # NM_004004.3) mutations. Twelve other families were not included for genome scan due to an insufficient number of DNA samples in order to carry out a linkage analysis. DNA samples from 126 families were diluted to 40 ηg/μl and sent to the Center for Inherited Disease Research (CIDR; URL: http://www.cidr.jhmi.edu/) for genome scan, while diluted DNA samples from 42 families were sent to the National Heart, Lung and Blood Institute (NHLBI) Mammalian Genotyping Service (Center for Medical Genetics, Marshfield, WI, USA; URL: http://research.marshfieldclinic.org/genetics/) for genotyping. From 2002 to 2004, samples were sent in six batches, with an average of 395 short tandem repeat (STR) markers spaced at ~10cM apart for each genome scan that was done on all 22 autosomes and the X and Y chromosomes.

Linkage analysis

Linkage analyses were performed under a fully penetrant autosomal recessive model with a disease allele frequency of 0.001. The MLINK program of the FASTLINK computer package was utilized for two-point linkage analysis [3], while multipoint analysis was performed using ALLEGRO [4] with map distances from the Marshfield genetic map [5]. Some of the families were too large to analyze in their entirety using ALLEGRO and were therefore broken into two or more branches for the analysis, then the LOD scores from each branch of the family were summed. Marker allele frequencies were estimated from the data by means of both observed and reconstructed genotypes of founders from each pedigree and the other pedigrees in the same genome scan.

Sequencing TMC1

Polymerase chain reaction (PCR) primers were designed for 24 exons plus 1000 base pairs of the promoter region of *TMC1* (GenBank accession # NM_138691.2) via Primer3 software [6]. DNA from one unaffected and two affected individuals from each family were diluted to 5ηg/μl, amplified, and purified with ExoSAP-IT® (USB Corp., Cleveland, OH, USA; URL: http://www.usbweb.com/). Sequencing with the appropriate primers was performed with the BigDye® Terminator v3.1 Cycle Sequencing Kit together with an Applied Biosystems 3700 DNA Analyzer (Applera Corp., Foster City, CA, USA; URL: http://www.appliedbiosystems.com/). Sequence variants were identified via Sequencher™ Version 4.1.4 software (Gene Codes Corp., Ann Arbor, MI, USA; URL: http://www.genecodes.com/se-

quencher/). When a sequence variant was found, DNA samples from the rest of the family were sequenced for the exon in which the variant was identified. Similarly, 117 control individuals from Pakistan were screened for the same exon.

Protein sequence analysis

To determine the evolutionary conservation of identified substitutions, the ExPASy (Expert Protein Analysis System; URL: http://us.expasy.org/) proteomics server of the Swiss Institute of Bioinformatics (SIB) was used to look for homologues of the *TMC1* protein. ExPASy uses the NCBI BLASTP 1.5.4-Paracel program [7] to search the ExPASy/UniProt database. To perform the BLASTP search, the default settings were used, except that the threshold for expected random matches (E) was set to a more conservative value of 10^{-4} in order to minimize false positive results. Of the 81 query matches (score range 58-1196), one match per species was chosen for further alignment. Nine proteins, including the human *TMC1* sequence, were then submitted for multiple sequence alignment via ClustalW [8] at the European Bioinformatics Institute (EBI; URL: http://www.ebi.ac.uk/clustalw/) using default settings.

Membrane-spanning domains of the *TMC1* protein were predicted through the TMHMM Server v.2.0 [9]. An independent study demonstrated that TMHMM v.2.0 had 92% true positive predictions, which was higher than all other transmembrane protein prediction programs [10]. To check for motifs or patterns within the protein sequence, the PROSITE database of SIB and the European Molecular Biology Laboratory (EMBL) - EBI was scanned [11]. The web server PolyPhen (from EBI) was used to assess the functional effect of identified substitutions [12].

RESULTS

Of 168 families that underwent two-point and/or multipoint linkage analysis, twelve families had two-point or multipoint LOD scores of 1.4 or greater between markers flanking the *TMC1* gene. Table 3.4 shows the LOD scores per family, along with ethno-linguistic information and sequence variants that were found to segregate within families. These sequence variants were found in the homozygous state in the hearing-impaired individuals but not among unaffected family members.

Eight families were positive for putatively functional sequence variants. Of the identified variants, c.830A>G (p.Y277C), which occurred at a highly conserved residue, was deemed to be damaging because of a change from a

Table 3.4 Families screened for *TMC1* mutations [a]

Family	Place of origin [b]	Language	2-point LOD	Multipoint LOD	Sequence variant
4008	DG Khan, Punjab	Sairiki	3.1	4.0	c.2035G>A (p.E679K)
4027	Mian Wali, NWFP-Punjab border	Punjabi	2.5; 2.6 at chr.11	1.1; 2.0 at chr.11	None
4033	Sarghoda, Punjab	Punjabi	2.2; 1.2 at chr.2	2.9; 2.1 at chr.2	c.1334G>A (p.R445H)
4049	Bahwalpur, Punjab	Sairiki	4.7	4.8	c.830A>G (p.Y277C)
4070	Kashmir, AJK	Kashmiri	2.4	3.0	c.2004T>G (p.S668R)
4090	Chistian, Punjab	Punjabi	1.4; 2.1 at chr.10	1.8; -0.6 at chr.10	c.536 -8T>A [c]
4119	Sadiqabad, Punjab	Sairiki	2.4	2.6	c.1114G>A (p.V372M)
4138	Kotli Kshmir, AJK	Kashmiri	2.6	3.3	c.2004T>G (p.S668R)
4156	Larkhana, Sind	Sindi	2.2; 2.3 at chr.8	3.1; 0 at chr.8	None
4160	Dhandi, Sind	Sindi	3.0	3.6	c.1114G>A (p.V372M)
4165	Taj Ghar, Punjab	Sairiki	1.5; 2.0 at chr.17	2.7; 2.8 at chr.17	None
4173	Patoki, Punjab	Punjabi	1.6; 1.7 at chr.1	1.6; 1.9 at chr.1	None

[a] *TMC1* (GenBank accession # NM_138691.2).
[b] *NWFP* Northwestern Frontier Province, *AJK* Azad Jammun and Kashmir
[c] The traditional nomenclature for c.536 -8T>A was IVS10 -8T>A.

Table 3.5 *TMC1* sequence variants found to segregate within Pakistani families

Exon	Nucleotide change	Amino acid change	Domain [a]	Evolutionary conservation [b]	Predicted functional effect [c]	Allele frequencies among control chromosomes
11	c.536 -8T>A	Splice acceptor	N/A	N/A	Exon 11 skipped	0 / 234
13	c.830A>G	p.Y277C	TM2	Identical	Probably damaging	0 / 234
15	c.1114G>A	p.V372M	TM3	Conserved	Benign	0 / 234
16	c.1334G>A	p.R445H	TM4	Identical	Benign	0 / 234
21	c.2004T>G	p.S668R	EC	Non-conserved	Benign; PKC phosphorylation site [d]	0 / 234
21	c.2035G>A	p.E679K	EC	Non-conserved	Benign	0 / 234

[a] TMHMM v.2.0 predicted six membrane-spanning domains for the *TMC1* protein [9]. *TM,* transmembrane; *EC,* extracellular; *N/A,* not applicable.

[b] Human *TMC1* sequence compared with *TMC1*-like protein sequences of 8 species which were identified via the NCBI BLASTP [7] program and the ExPASy/UniProt databases. Species include one mammal, one amphibian, two bony fishes, three insects and one nematode. Comparison done via ClustalW [8]. The serine residue at position 668 and glutamic acid at 679 are conserved in *M. musculus Tmc1* and *F. rubripes Tmc2-related* proteins. Additionally, S668 is also conserved in *A. gambiae TmcA-like* protein and *D. melanogaster CG3280-PA*.

[c] Possible effect of amino acid substitutions on the protein structure and function were predicted via the web server PolyPhen [12].

[d] The p.S668R sequence change was found to occur at a possible protein kinase C (PKC) phosphorylation site according to the PROSITE database [11].

hydroxyl to a sulfhydryl side chain in the amino acid that was located within the second transmembrane (TM) domain (Table 3.5). The c.1114G>A (p.V372M) and c.1334G>A (p.R445H) variants were also at highly conserved TM residues, but the amino acid was changed to another of the same subclass. Residues for variants c.2004T>G (p.S668R) and c.2035G>A (p.E679K) were less conserved and occurred within the extracellular region. These last four sequence changes were considered benign by the Polyphen website, though scanning through PROSITE showed that the serine residue at position 668 may belong to a protein kinase C phosphorylation site. The nucleotide change c.2004T>C removes this putative phosphorylation site. Also c.1114G>A (p.V372M) and c.2004T>G (p.S668R) were each found to segregate with the NSHI status in two families. One family has the known splice site mutation c.536 -8T>A [13]. All the reported sequence variants were not found in control subjects, who were linguistically matched for the Punjabi- and Sairiki-speaking families. Due to the limited samples from the remote Sindi and Kashmiri regions, most of the controls for the c.2004T>G variant are Punjabi, while for the c.1114G>A variant most of the controls speak Sairiki.

Several polymorphisms were also observed in the twelve families (Table 3.6). None of these variants segregated with hearing impairment status within the families. Two of the variants were non-synonymous substitutions, but were also considered benign polymorphisms because they did not segregate with HI status. These substitutions also occurred at non-conserved non-transmembrane residues.

Aside from c.536 -8T>A, other *TMC1* sequence variants that were identified in previous research were not observed in the families in this study, but were also analyzed in terms of evolutionary conservation and occurrence at transmembrane domains and known motifs (Table 3.7).

DISCUSSION

Putatively functional *TMC1* variants were observed in eight of the 168 Pakistani families that underwent a genome scan. Twelve additional families in the study did not undergo a genome scan but are known to have functional variants in the *GJB2* gene [Section 3.2]. Thus, when *GJB2*-affected families were included, the prevalence of *TMC1* mutations in this Pakistani population was 4.4% (95%CI: 1.9, 8.6). This is a lower prevalence rate than the previously reported prevalence of $5.4 \pm 3.0\%$ among consanguineous Indian and Pakistani families [13]. However, the difference in these prevalence rates is not statistically significant.

Table 3.6 *TMC1* polymorphisms that do not segregate within families

Exon	Nucleotide change
1	g.74G>A (-467 from the ATG)
3	g.[236 -67G>A (+) 322A>G (-219 from the ATG)]
4	g.346 -11delT
6	c.45C>T (synonymous)
8	c.237 -44G>A
8	c.241G>A (p.E81K) [a]
10	c.[535 +101A>G (+) 535 +106_113delCAAACAAA]
14	c.1029 +85T>C
16	c.1404 +32A>G
17	c.1457T>C (p.M486T) [b]
22	c.2208 +22A>G

[a] p.E81K does not segregate with the HI status in three families. Additionally p.E81K is at the intracellular glutamic acid (E)-rich region and is non-conserved.
[b] p.M486T was found in the heterozygous state in two hearing individuals from two different families but not among the hearing-impaired. The residue is predicted to be intracellular and is non-conserved.

Among 12 families in which the *TMC1* gene was sequenced, a potentially functional variant was not identified in four families (Table 3.4). Three of these families (4027, 4165 and 4173) had LOD scores of suggestive linkage (1.9-2.8) in other chromosomal regions, which may contain the gene that causes the HI in these families. In family 4156, the two-point LOD score was relatively high for chromosome 2, but the multipoint LOD score at chromosome 2 was zero, while a multipoint LOD score of 3.1 was obtained in the *TMC1* region. For this family, it is possible that the functional variant(s) are within the intronic regions of *TMC1*, or in another gene that lies close to the *TMC1* gene.

The predicted structure of *TMC1* bears similarity to that of the α-subunit of voltage-dependent K^+ channels, which has six α-helical TM segments and intracellular N and C termini [15]. The first four TM domains of the K^+ channel α-subunit act as voltage sensors for activation gating [16], whereas the intervening segment between TM5 and TM6 appears to confer channel selectivity [15]. Three highly conserved *TMC1* sequence variants in this study – c.830A>G (p.Y277C), c.1114G>A (p.V372M) and c.1334G>A (p.R445H) – lie within the central portion of hydrophobic TM segments. Meanwhile, the *TMC1* variants c.2004T>G (p.S668R) and c.2035G>A (p.E679K) were found between TM5 and TM6 among a cluster

Table 3.7 Previously published sequence variants in the *TMC1* gene [a]

Exon	Nucleotide change	Protein change	Domain	Evolutionary conservation	Predicted functional effect
Intron 3-5	IVS3_IVS5 del27kb [b]	Transcription initiation site removed	N/A	N/A	Exons 4 and 5 deleted
7	c. 100C>T	p.R34X	IC	Non-conserved	Truncated protein; Occurs at glutamic-acid rich N-terminus region; cAMP- and cGMP-dependent protein kinase phosphorylation site
8	c.295delA	Frameshift	N/A	N/A	Truncated protein
13	c.884 +1G>A [c]	Splice donor	N/A	N/A	Exon 13 skipped
15	c. 1165C>T	p.R389X	EC	Non-conserved	Truncated protein
17	c. 1534C>T	p.R512X	IC	Non-conserved	Truncated protein
19	c. 1714G>A [d]	p.D572N	IC	Non-conserved	Benign; Casein kinase II phosphorylation site
20	c. 1960A>G	p.M654V	TM5	Non-conserved	Benign

[a] *TMC1* sequence changes were reported by Kurima et al. [13] except for p.R389X at exon 15 [14]. Membrane-spanning domains, evolutionary conservation and functional effect were determined as described in Table 3.5.
[b] Because the article did not specify the exact position at which this deletion occurred, it was not possible to rename this sequence variant according to the current nomenclature recommendations.
[c] The traditional nomenclature for c.884 +1G>A was IVS13 +1G>A.
[d] c.1714G>A (p.D572N) was observed in a North American family with dominant hearing impairment.

of less conserved residues. Though these amino acid residues were non-conserved when human *TMC1* was aligned with eight proteins from other species, alignment with other human and murine TMC proteins shows that these residues are conserved for members of the TMC subfamily A, *TMC1*, *TMC2* and *TMC3* [17-18]. Additionally, the serine residue at position 668 is a putative protein kinase C phosphorylation site. Therefore, even though the Polyphen server labeled four of five novel substitutions as benign, we believe that all five variants are functional. It should be noted that the Polyphen server successfully predicts only 82% of mutations in the SwissProt protein database [12].

The *deafness (dn)* mouse is the homologous model for recessive *TMC1* mutations, while the *Beethoven (Bth)* mouse is identified with dominant *TMC1* mutations that cause postlingual hearing impairment [13, 19]. In the mouse *Tmc1* expression has been found to localize to cochlear hair cells [13]. In microscopic sections of *dn* mouse cochlea, *Tmc1* mutations caused the greatest damage to inner hair cells (IHC), with progressive degeneration that begins at postnatal day 13, which is the time of onset of hearing function in mice [19-20]. In functional studies, the cochlear microphonic, which is a measure of the electric potential that is generated by the organ of Corti after acoustic stimulation, was never detected in *dn* mice from postnatal day 12 onwards [21-22]. These studies further support the findings of bilateral profound prelingual hearing impairment due to recessive homozygous *TMC1* mutations, and implicate the IHC as the main site of hearing dysfunction.

It was suggested that *TMC1* may be an ion channel or transporter which mediates K^+ homeostasis in the inner ear [17]. Because K^+ transport across hair cell membranes requires very rapid response, the use of second messenger or cascade mechanisms has been deemed highly unlikely, and the gating-spring model for mechanoelectrical transduction is largely favored [23]. This model proposed that mechanical forces brought about by bending of stereocilia and tension on the tip links directly activate ion channels. If it is true that *TMC1* is an ion channel that is mainly localized in the IHC, then it might be involved in the most basic auditory process of hair-cell transduction.

In conclusion, five novel sequence variants of the *TMC1* gene for hearing impairment were identified in this study. These variants were found in *TMC1* regions that might be critical to function of K^+ channels in hair cells. In the highly consanguineous Pakistani population, the prevalence of putatively functional *TMC1* variants was found to be low. However, the true

impact of genetic defects in *TMC1* cannot be known until more studies are done in other populations.

Acknowledgements We wish to thank the family members for their invaluable participation and cooperation. We are also grateful to RA Harris for his useful comments and suggestions. This work was made possible through the National Institutes of Health - National Institute of Deafness and other Communication Disorders grant R01-DC03594. Genotyping services were provided by the Center for Inherited Disease Research (CIDR) and the National Heart, Lung and Blood Institute Mammalian Genotyping Service. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, Contract Number N01-HG-65403.

REFERENCES

1. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
2. Grimberg J, Nawoschik S, Bellusico L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
3. Cottingham RW, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993; 53: 252-263.
4. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12-13.
5. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 1998; 63: 861-869.
6. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. Bioinformatics Methods and Protocols: Methods in Molecular Biology. NJ: Humana Press, 2000. pp 365-386.
7. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389-3402.
8. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-4680.
9. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol 2001; 305: 567-580.
10. Moller S, Croning MD, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 2001; 17: 646-653.
11. Sigrist CJ, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform 2002; 3: 265-274.
12. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002; 30: 3894-3900.

13. Kurima K, Peters LM, Yang Y, Riazuddin S, Ahmed ZM, Naz S, Arnaud D, Drury S, Mo J, Makishima T, Ghosh M, Menon PS, Deshmukh D, Oddoux C, Ostrer H, Khan S, Riazuddin S, Deininger PL, Hampton LL, Sullivan SL, Battey JF Jr, Keats BJ, Wilcox ER, Friedman TB, Griffith AJ. Dominant and recessive deafness caused by mutations of a novel gene, *TMC1*, required for cochlear hair-cell function. Nat Genet 2002; 30: 277-284.
14. Meyer CG, Gasmelseed NM, Mergani A, Magzoub MMA, Muntau B, Thorsten T, Horstmann RD. Novel *TMC1* structural and splice variants associated with congenital nonsyndromic deafness in a Sudanese pedigree. Hum Mutat 2005; 25: 100.
15. Hanlon MR, Wallace BA. Structure and function of voltage-dependent ion channel regulatory beta subunits. Biochemistry 2002; 41: 2886-2894.
16. Li-Smerin Y, Hackos DH, Swartz KJ. α-helical structural elements within the voltage-sensing domains of a K(+) channel. J Gen Physiol 2000; 115: 33-50.
17. Keresztes G, Mutai H, Heller S. TMC and EVER genes belong to a larger novel family, the TMC gene family encoding transmembrane proteins. BMC Genomics 2003; 4: 24.
18. Kurima K, Yang Y, Sorber K, Griffith AJ. Characterization of the transmembrane channel-like (TMC) gene family: functional clues from hearing loss and epidermodysplasia verruciformis. Genomics 2003; 82: 300-308.
19. Vreugde S, Erven A, Kros CJ, Marcotti W, Fuchs H, Kurima K, Wilcox ER, Friedman TB, Griffith AJ, Balling R, Hrabe De Angelis M, Avraham KB, Steel KP. Beethoven, a mouse model for dominant, progressive hearing loss DFNA36. Nat Genet 2002; 30: 257-258.
20. Pujol R, Shnerson A, Lenoir M, Deol MS. Early degeneration of sensory and ganglion cells in the inner ear of mice with uncomplicated genetic deafness (dn): preliminary observations. Hear Res 1983; 12: 57-63.
21. Steel KP, Bock GR. The nature of inherited deafness in deafness mice. Nature 1980; 288: 159-161.
22. Bock GR, Steel KP. Inner ear pathology in the deafness mutant mouse. Acta Otolaryngol 1983; 96: 39-47.
23. Corey DP, Hudspeth AJ. Kinetics of the receptor current in bullfrog saccular hair cells. J Neurosci 1983; 3: 962-976.

Chapter 3.3

Novel sequence variants in the *TMIE* gene in families with autosomal recessive non-syndromic hearing impairment.

Santos RLP, El-Shanti H, Sikandar S, Lee K, Bhatti A, Yan K, Chahrour MH, McArthur N, Pham TL, Mahasneh AA, Ahmad W, Leal SM. Reprinted from Journal of Molecular Medicine 84(3): 226-231, with kind permission of Springer Science and Business Media.

SUMMARY

To date 37 genes have been identified for non-syndromic hearing impairment (NSHI). Identifying the functional sequence variants within these genes and knowing their population-specific frequencies is of public health value, in particular for genetic screening for NSHI. To determine putatively functional sequence variants in the transmembrane inner ear (*TMIE*) gene in Pakistani and Jordanian families with autosomal recessive (AR) NSHI, four Jordanian and 168 Pakistani families with ARNSHI that is not due to *GJB2* (*CX26*) were submitted to a genome scan. Two-point and multipoint parametric linkage analyses were performed, and families with logarithmic odds (LOD) scores of 1.0 or greater within the *TMIE* region underwent further DNA sequencing. The evolutionary conservation and location in predicted protein domains of amino acid residues where sequence variants occurred were studied to elucidate the possible effects of these sequence variants on function. Of seven families that were screened for *TMIE*, putatively functional sequence variants were found to segregate with hearing impairment in four families but were not seen in not less than 110 ethnically matched control chromosomes. The previously reported c.241C>T (p.R81C) variant was observed in two Pakistani families. Two novel variants, c.92A>G (p.E31G) and the splice site mutation c.212 -2A>C, were identified in one Pakistani and one Jordanian family, respectively. The c.92A>G (p.E31G) variant occurred at a residue that is conserved in the mouse and is predicted to be extracellular. Conservation and potential functionality of previously published mutations were also examined. The prevalence of functional *TMIE* variants in Pakistani families is 1.7% (95% CI: 0.3, 4.8). Further studies on the spectrum, prevalence rates and functional effect of sequence variants in the *TMIE* gene in other populations should demonstrate the true importance of this gene as a cause of hearing impairment.

INTRODUCTION

In the past decade there has been great progress in the mapping and identification of hearing impairment (HI) genes in different populations worldwide. To date more than 100 HI loci have been mapped and 37 HI genes have been identified [1]. It is predicted that there may be as many as 300 genes involved in the HI phenotype [2]. The large-scale effort to identify HI genes was partially driven by the hope that genetic screening will contribute greatly to earlier diagnosis of genetic HI, and allow for early therapeutic and rehabilitative management. However, the clinical usefulness of knowing hundreds of HI genes will remain limited until individual genes are studied for their spectrum of sequence variants, prevalence rates of mutations in various populations, and functional significance.

The transmembrane inner ear (*TMIE*; MIM# 607237) gene is one of the more recently identified HI genes [3], the function of which is yet unknown. In this study, two novel sequence variants of the *TMIE* gene were found in Pakistani and Jordanian families with autosomal recessive non-syndromic (ARNS) HI. Additionally the prevalence of putatively functional *TMIE* variants was estimated in Pakistani families with ARNSHI. The occurrence of sequence variants in predicted protein domains and at evolutionarily conserved residues were studied for both previously published and novel *TMIE* variants in order to determine whether these variants are involved in the etiology of HI.

MATERIALS AND METHODS

Ascertainment of study subjects

The study was approved by the Institutional Review Boards of Quaid-I-Azam University, Jordan University of Science and Technology, and Baylor College of Medicine and Affiliated Hospitals. Informed consent was obtained from all family members who participated in the study.

For this study, 192 unrelated Pakistani families with at least two ARNSHI individuals were ascertained from various regions of Pakistan. In addition, five families from Jordan were recruited for ARNSHI research. Medical and family history and information on pedigree structure was obtained from multiple family members. Pure tone audiometry at 250-8,000 Hz was performed for selected subjects. All hearing-impaired family members underwent physical examination. The clinical history indicates that the hearing impairment in all four families is prelingual and probably congenital. There is no evidence that HI is due to inflammatory middle ear disease or specific environmental factors. Within these families, no clinical features that would indicate that the HI is part of a syndrome, including mental retardation, were observed. The audiograms clearly demonstrate that the hearing impairment

due to *TMIE* is bilateral, profound and affects all frequencies (Figure 3.1). No additional diagnostic procedures for vestibular and temporal bone abnormalities were performed, although from the clinical history and physical examination, there were no reports of delay in ambulation during development or of positive findings during vestibular testing and otoscopy.

Genome scan

DNA was isolated from venous blood samples following a standard protocol [4], quantified by spectrophotometry at optical density 260, and stored at -20°C. Of the 197 families, one Jordanian and twelve Pakistani families were positive for *GJB2* (MIM# 121011; GenBank accession# NM_004004.3) mutations. Twelve other Pakistani families were not included for genome scan due to an insufficient number of DNA samples in order to carry out a linkage analysis. DNA samples from 128 families were diluted to 40 ηg/μl and sent to the Center for Inherited Disease Research (CIDR) for genome scan, while diluted DNA samples from 44 families were sent to the National Heart, Lung and Blood Institute (NHLBI) Mammalian Genotyping Service (Center for Medical Genetics, Marshfield, WI, USA) for genotyping. From 2000 to 2004, samples were sent in six batches, with an average of 395 short tandem repeat (STR) markers spaced at ~10 cM apart for each genome scan that was done on all 22 autosomes and the X and Y chromosomes.

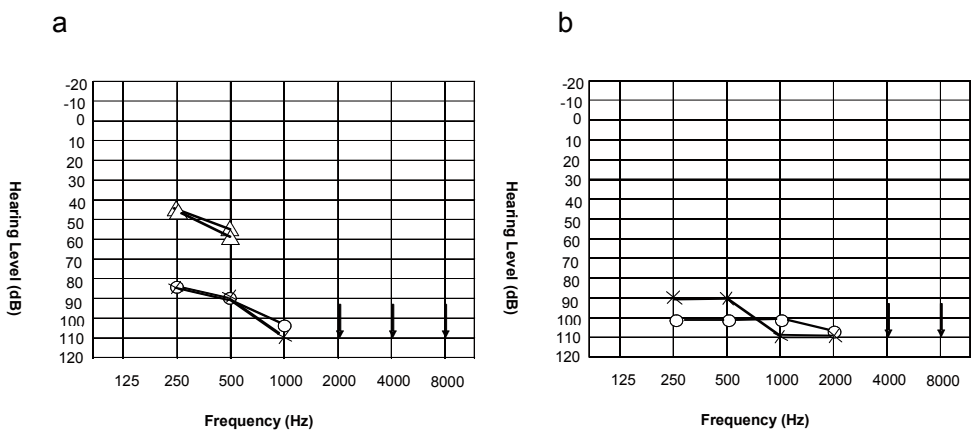


Figure 3.1 Audiograms of hearing-impaired individuals from families (a) 4043 and (b) 4139. Circles and crosses represent air conduction for right and left ear, respectively. Triangles represent masked air conduction thresholds. Both audiograms exhibit bilateral, profound hearing impairment that affects all frequencies. Family 4043 has the p.E31G variant while family 4139 carries the p.R81C substitution.

Linkage analysis

Linkage analyses were performed under a fully penetrant autosomal recessive model with a disease allele frequency of 0.001. The MLINK program of the FASTLINK computer package was utilized for two-point linkage analysis [5], while multipoint analysis was performed using ALLEGRO [6] with map distances from the Marshfield genetic map [7], DeCode genetic map [8] or estimated using Map-O-Mat [9]. Some of the families were too large to analyze in their entirety using ALLEGRO and were therefore broken into two or more branches for the analysis, then the LOD scores from each branch of the family were summed. Marker allele frequencies were estimated from the marker data by means of both observed and reconstructed genotypes of founders from each pedigree and other pedigrees from the same population that were genotyped in the same genome scan.

Sequencing TMIE

The *TMIE* (GenBank accession# NM_147196.1) gene was sequenced in one hearing and two hearing-impaired members of families showing linkage to the 3p21.31 region after two-point and multipoint linkage analyses. Primer3 software [10] was used to design primers for the four exons and 1000 bp of promoter region of the *TMIE* gene. After DNA amplification and purification with ExoSAP-IT® (USB Corp., Cleveland, OH, USA), sequencing was performed with the BigDye® Terminator v3.1 Cycle Sequencing Kit together with an Applied Biosystems 3700 DNA Analyzer (Applera Corp., Foster City, CA, USA). Sequence variants were identified via Sequencher™ Version 4.1.4 software (Gene Codes Corp., Ann Arbor, MI, USA). When a potentially functional sequence variant was found, the exon in which the variant was found was sequenced in all other family members for whom DNA was available. When the identified sequence variant was shown to segregate with HI status within a family, a minimum of 55 unrelated ethnically-matched control individuals were also screened for the same exon.

Protein sequence analysis

For the protein analyses, transmembrane domains were predicted with TMHMM v.2.0 [11]. Similarity search was performed through the ExPASy/UniProt database and the NCBI BLASTP 1.5.4-Paracel program [12] using default settings, then the identified eukaryotic proteins were aligned with ClustalW [13]. The PROSITE database [14] was scanned for known patterns and motifs.

RESULTS

Of 172 Pakistani and Jordanian families that underwent two-point and/or multipoint linkage analyses, seven families had two-point and multipoint LOD scores of 1.0 or greater between markers flanking the *TMIE* gene. Table 3.8 shows the LOD scores per family, along with ethno-linguistic information and sequence variants that were found in homozygosity in hearing-impaired individuals but not in unaffected family members.

Four families were positive for putatively functional sequence variants. Two Pakistani families had a previously published variant, c.241C>T (p.R81C) [3], which occurs at a highly conserved cytoplasmic amino acid residue (Table 3.9). Two other families possess novel sequence variants. One Pakistani family had an A to G transition at position 92, thus resulting in the replacement of the charged polar amino acid side chain by a hydrogen molecule (p.E31G). The amino acid residue is predicted to occur at the extracellular region and is conserved in the mouse *Tmie* protein but not in *Tmie*-like proteins of *Drosophila melanogaster* and *Caenorhabditis elegans*. The splice site mutation c.212 -2A>C that was found in a Jordanian family is predicted to cause skipping of the third exon and subsequent removal of part of the second transmembrane segment and half of the long C-terminal tail of the protein. Neither p.R81C nor p.E31G was identified with a known protein sequence motif. All three variants were not observed in at least 110 control chromosomes.

Several polymorphisms were observed in the seven families, but none of these variants segregated with HI status within the families (Table 3.10). Four *TMIE* sequence variants that were identified in a previous study [3] were not found in the families that are reported here but were also examined in terms of conservation and occurrence of amino acid residues at functional sites (Table 3.11). Three missense changes at arginine residues, including the known variant p.R81C, were predicted to occur within the intracellular domain at the C-terminal tail. Two substitutions, p.R81C and p.R84W, are at highly conserved residues, while p.R92W occurs at a putative tyrosine kinase phosphorylation site.

DISCUSSION

Putatively functional *TMIE* variants were observed in three of the 168 Pakistani families that underwent a genome scan. Twelve additional Pakistani families in the study did not undergo a genome scan but are known to have functional variants in the *GJB2* gene [Section 3.1]. The prevalence of functional *TMIE* variants in these Pakistani families is therefore 1.7% (95%

Table 3.8 Families screened for *TMIE* mutations

Family	Place of origin	Language	Two-point LOD	Multipoint LOD	Sequence variant
4009	DG Khan, Punjab	Sairiki	1.3; 2.9 at chr.7	2.2; 6.6 at chr.7	None
4019	Rawalpindi, Punjab	Punjabi	3.5	5.5	c.241C>T (p.R81C)
4043	Rahim Yar Khan, Punjab	Punjabi	1.5	2.8	c.92A>G (p.E31G)
4051	Bahwalpur, Punjab	Sairiki	1.8; 1.0 at chr.10	1.8; 1.9 at chr.10	None
4139	Kotli Kshmir, AJK	Kashmiri	1.9	2.0	c.241C>T (p.R81C)
4197	Quetta, Baloochistan	Pushto	1.1; 1.6 at chr.4	1.8; 1.8 at chr.4	None
3002	Jordan	Arabic	4.1	6.1	c.212 -2A>C

TMIE (GenBank accession # NM_147196.1). *AJK,* Azad Jammu and Kashmir.

Table 3.9 *TMIE* sequence variants found to segregate with hearing impairment status within families

Exon	Nucleotide change	Protein change	Domain[a]	Evolutionary conservation[b]	Allele frequencies among control chromosomes[c]
1	c.92A>G	p.E31G	EC	Conserved in mouse only	0 / 110
3	c.212 -2A>C	Exon skipping predicted	N/A	N/A	0 / 176
3	c.241C>T	p.R81C	IC	Identical	0 / 122

IC, intracellular; *EC*, extracellular; *N/A*, not applicable.

[a] TMHMM v.2.0 predicted two membrane-spanning domains for the *TMIE* protein [11].

[b] Human *TMIE* sequence compared with mouse *Tmc1*, *CG15130-PA* protein of *Drosophila melanogaster* and hypothetical protein *Y39A1C.1* of *Caenorhabditis elegans*.

[c] Control individuals were matched by country of origin. The c.212 -2A>C mutation was also negative in 122 Pakistani control chromosomes, and the c.241C>T (p.R81C) variant was not found in 176 control chromosomes from Jordan.

Table 3.10 *TMIE* polymorphisms that do not segregate with hearing impairment status within families

Exon	Nucleotide change
Promoter	g.1 -1017T>C
Promoter	g.[1 -804T>C (+) 1 -74A>T]
Promoter	g.1 -300C>A
2	c.94 -46C>A
2	c.94 -25C>T
3	c.212 -63G>A
3	c.321G>A (synonymous)
4	c.388_390dupAAG (p.K9dup)
4	g.1598G>A (at +975 of the TGA)

Table 3.11 Previously published sequence variants in the *TMIE* gene [a]

Exon	Nucleotide change	Protein change	Domain	Evolutionary conservation
Intron1-exon 2	c.94 -2_98del AGCCCAGinsC [b]	Exon skipping predicted	N/A	N/A
2	c.125_126dupCGCC [b]	Frameshift results in truncation	N/A	N/A
3	c.250C>T	p.R84W	IC	Identical
3	c.274C>T	p.R92W [c]	IC	Non-conserved

N/A, not applicable.

[a] *TMIE* sequence variants are from Naz et al. [3]. Membrane-spanning domains, evolutionary conservation and functional effect were determined as described in Table 3.9.

[b] These sequence variants were originally reported as IVS1-2_98del AGCCCAGinsC and 125-126insCGCC. The nomenclature has been modified in accordance with recommendations from the Human Genome Variation Society (URL: http://www.genomic.unimelb.edu.au/mdi/).

[c] The arginine residue at position 92 belongs to a putative tyrosine kinase phosphorylation site according to the PROSITE database [14].

CI: 0.3, 4.8). This is the first report of the population-specific prevalence of *TMIE* variants in individuals with ARNSHI. In this population of Pakistani families with ARNSHI, the prevalence of putatively functional variants in the *GJB2* gene is 6.1% (95% CI: 3.2, 10.4) [Section 3.1], while the prevalence of putatively functional *TMC1* (MIM# 606706) variants is 4.4% (95%CI: 1.9, 8.6) [Section 3.2]. The importance of the *TMIE* gene as a cause of HI cannot be known until other populations are studied. Due to insufficient number of families, it is not possible to accurately estimate the prevalence rate of functional *TMIE* variants in the Jordanian population.

Among seven families in which the *TMIE* gene was sequenced, a potentially functional variant was not identified in three families (Table 3.8). A conservative criterion of a LOD score of 1.0 or greater was used in choosing the families for sequencing. It is therefore not surprising that, for families 4009, 4051, and 4197 which had multipoint LOD scores of 6.6, 1.9, and 1.8, respectively to other chromosomal regions, the *TMIE* gene could not be implicated as the cause of HI. For family 4009, there is highly significant evidence of linkage to *DFNB44* [15]; the gene that causes HI in this family almost certainly lies in this region. It was probably overly conservative to rule out the involvement of *TMIE* in the HI phenotype in this family by sequencing of the gene. For both families 4051 and 4197, a multipoint LOD score of 1.8 was obtained within the interval of the *TMIE* gene. However, a multipoint LOD score of 1.9 was achieved on chromosome 10 for family 4051, while a LOD score of 1.8 was attained on chromosome 4 for family 4197. In these two families it was not possible to a priori rule out *TMIE* being involved in HI since there is no statistically significant evidence to map the gene segregating in this family to any chromosomal region.

Although the function of *TMIE* in the inner ear remains unknown, there are indications that the protein may have a role in both stereocilia maturation and auditory nerve function. The deaf *Tmie*-mutant *spinner* mouse was shown to have progressive hair cell degeneration with irregular apical surfaces of hair cells and extra rows of maturing stereocilia in inner hair cells [16]. Also the absence of auditory brainstem response in *spinner* mice correlates with spiral ganglion degeneration [16-17]. In another mouse model called *circling*, the mapped locus also includes the *Tmie* gene [18-19] and the *circling* mouse also exhibited hair cell axonal and spiral ganglion degeneration [20]. There is a similar pattern of histologic features in *Cdh23*$^{+}$*/Pcdh15*$^{+}$ and *UshIC*$^{+/+}$ mice [21-22]; additionally, the *CDH23*, *PCDH15*, and *USH1C* proteins have been identified as stereocilia bundling proteins. However it is difficult to surmise whether the hair cell and/or

spiral ganglion loss in the *Tmie* mouse mutants were due to primary stereocilia defects or to secondary responses to another type of inner ear malfunction. The occurrence of putatively functional sequence variants in the extracellular domain and the cytoplasmic carboxyl tail of the *TMIE* protein may indicate the importance of these segments to cochlear function, but the lack of protein structures that are similar to *TMIE* inhibits prediction of how these sequence variants may cause hearing impairment. More information on related protein structures should help in building a good model for the pathophysiologic role of *TMIE* in the inner ear.

Acknowledgements We are thankful to all family members who participated in this study. This work was made possible through grants from the National Institutes of Health - National Institute of Deafness and other Communication Disorders (NIH-NIDCD R01-DC03594), the Deanship of Research, Jordan University of Science and Technology, Irbid, Jordan, and Higher Education Commission, Islamabad, Pakistan. Genotyping services were provided by CIDR and the NHLBI Mammalian Genotyping Service. CIDR is fully funded through a federal contract from the NIH to The Johns Hopkins University, Contract Number N01-HG-65403.

REFERENCES

1. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
2. Nance WE. The genetics of deafness. Ment Retard Dev Disabil Res Rev 2003; 9: 109-119.
3. Naz S, Giguere CM, Kohrman DC, Mitchem KL, Riazuddin S, Morell RJ, Ramesh A,Srisailpathy S, Deshmukh D, Riazuddin S, Griffith AJ, Friedman TB, Smith RJ, Wilcox ER. Mutations in a novel gene, *TMIE*, are associated with hearing loss linked to the DFNB6 locus. Am J Hum Genet 2002; 71: 632-636.
4. Grimberg J, Nawoschik S, Bellusico L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. Nucleic Acid Res 1989; 17: 83-90.
5. Cottingham RW, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993; 53: 252-263.
6. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2000; 25: 12-13.
7. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps individual and sex-specific variation in recombination. Am J Hum Genet 1998; 63: 861-869.
8. Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. A high-resolution recombination map of the human genome. Nat Genet 2002; 31: 241-247.
9. Kong X, Matise TC. MAP-O-MAT: internet-based linkage mapping. Bioinformatics 2005; 21: 557-559.

10. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. Bioinformatics Methods and Protocols: Methods in Molecular Biology. NJ: Humana Press, 2000. pp 365-386.
11. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol 2001; 305: 567-580.
12. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25: 3389-3402.
13. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-4680.
14. Sigrist CJ, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform 2002; 3: 265-274.
15. Ansar M, Chahrour MH, Amin Ud Din M, Arshad M, Haque S, Pham TL, Yan K, Ahmad W, Leal SM. DFNB44, a novel autosomal recessive non-syndromic hearing impairment locus, maps to chromosome 7p14.1-q11.22. Hum Hered 2004; 57: 195-199.
16. Mitchem KL, Hibbard E, Beyer LA, Bosom K, Dootz GA, Dolan DF, Johnson KR, Raphael Y, Kohrman DC. Mutation of the novel gene *TMIE* results in sensory cell defects in the inner ear of spinner, a mouse model of human hearing loss DFNB6. Hum Mol Genet 2002; 11: 1887-1898.
17. Deol MS, Robins MW. The spinner mouse. J Heredity 1962; 53: 133-136.
18. Cho KI, Lee JW, Kim KS, Lee EJ, Suh JG, Lee HJ, Kim HT, Hong SH, Chung WH, Chang KT, Hyun BH, Oh YS, Ryoo ZY. Fine mapping of the circling (cir) gene on the distal portion of mouse chromosome 9. Comp Med 2003; 53: 642-648.
19. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. Nature 2002; 20: 520-562. May 2004 reference sequence as viewed in: Genome Bioinformatics Group of UC Santa Cruz. UCSC Genome Browser. URL: http://genome.ucsc.edu/cgi-bin/hgGateway
20. Lee JW, Lee EJ, Hong SH, Chung WH, Lee HT, Lee TW, Lee JR, Kim HT, Suh JG, Kim TY, Ryoo ZY. Circling mouse: possible animal model for deafness. Comp Med 2001; 51: 550-554.
21. Zheng QY, Yan D, Ouyang XM, Du LL, Yu H, Chang B, Johnson KR, Liu XZ. Digenic inheritance of deafness caused by mutations in genes encoding cadherin 23 and protocadherin 15 in mice and humans. Hum Mol Genet 2005; 14: 103-111.
22. Johnson KR, Gagnon LH, Webb LS, Peters LL, Hawes NL, Chang B, Zheng QY. Mouse models of USH1C and DFNB18: phenotypic and molecular analyses of two new spontaneous mutations of the Ush1c gene. Hum Mol Genet 2003; 12: 3075-3086.

Chapter 4
Correlation of Genotype with Audiometric Phenotype

Chapter 4.1

Hearing impairment in Dutch patients with Connexin 26 (*GJB2*) and Connexin 30 (*GJB6*) mutations*

Santos RLP, Aulchenko YS, Huygen PLM, van der Donk KP, de Wijs IJ, Kemperman MH, Admiraal RJC, Kremer H, Hoefsloot LH, Cremers CWRJ. Reprinted from International Journal of Pediatric Otorhinolaryngology 69(2): 165-174.

SUMMARY

Despite the identification of mutations in the connexin 26 (*GJB2*) gene as the most common cause of recessive non-syndromic hearing loss, the pattern of hearing impairment with these mutations remains inconsistent. Recently a deletion encompassing the *GJB6* gene was identified and hypothesized to also contribute to hearing loss. We hereby describe the hearing impairment in Dutch patients with biallelic connexin 26 (*GJB2*) and *GJB2* + connexin 30 (*GJB6*) mutations. The audiograms of patients who were screened for *GJB2* and *GJB6* mutations were analysed retrospectively.

* Note: Nucleotide sequence data reported are available in the GenBank database under the following accession numbers: M86849 for exon 2 of connexin 26; U43932 for exon 1, *GJB2*; and AJ005585 for connexin 30.

Standard statistical testing was done for symmetry and shape, while repeated measurement analysis was used to assess the relation between mutation and severity. Progression was also studied via linear regression analysis. Of 222 hearing-impaired individuals, 35 exhibited sequence variations; of these 19 had audiograms for study. Hearing loss in patients with biallelic "radical" (i.e. deletions, nonsense and splice site) mutations was significantly worse than in the wild type and heterozygotes (SAS proc GENMOD, p=0.013). The presence of at least one missense mutation in compound heterozygotes tends to lead to better hearing thresholds compared to biallelic radical mutations (p=0.08). One patient with the [35delG]+[del(*GJB6*-D13S1830)] genotype was severely impaired. Non-progressive hearing impairment was demonstrated in five 35delG homozygotes in individual longitudinal analyses. However a patient with the [299A>C]+[416G>A] genotype showed significant threshold progression in the lower frequencies. Findings on asymmetry and shape were inconclusive. Our data support the hypothesis that severity is a function of genotype and its effect on the amino acid sequence. A bigger cohort is required to establish non-progressivity more definitively.

INTRODUCTION

The connexin 26 (*GJB2*) gene is the most studied determinant of recessive non-syndromic hearing loss. The continued interest in this gene lies in its high carrier frequency (0.5-5.4%) across different ethnic groups [1-2]. There are more than 100 sequence variations known [3] yet one mutation – 35delG – comprises about 70-85% of allelic mutations among Northern Europeans [4-7]. In addition, the gene is small (5.5 kb long) and only one of two exons has the coding sequence. All these favor the development of an effective genetic screening method for recessive deafness that is expected to have a high impact on clinical practice and genetic counselling.

Yet a phenotypic picture for *GJB2* remains incomplete, which makes genetic counselling difficult. The presence of two *GJB2* mutations (that is, homozygous or compound heterozygous; from hereon referred to as biallelic) has been associated with hearing impairment that affects all frequencies and that is prelingual, symmetric, sensorineural, nonprogressive, and of variable severity even within sibships [4-5, 8-9]. The first four characteristics seem consistent across populations. Progression has been described in some research, including some description of sudden progressive hearing loss, but was tested for statistical significance in only a few studies [6, 10]. Severity has been extensively discussed and there is evidence that it

varies as a function of genotype. Genetic screening was previously limited to 35delG until more studies showed varying patterns of severity in compound heterozygotes and non-35delG homozygotes [6, 11-12]. The trend now is to sequence the coding exon and the splice sites of the gene, so that different combinations of mutations may aid in predicting the severity of hearing impairment. In a relatively large study, the hearing loss in 35delG homozygotes still ranged widely from moderate to profound [10], which may be due to yet undiscovered gene-gene or gene-environment interactions.

A deletion in the connexin 30 (*GJB6*) gene can also cause hearing impairment, either as a homozygous deletion or in combination with a *GJB2* mutation [13-16]. Several genetic centers have retrospectively analysed the DNA of proven *GJB2* heterozygotes for del(*GJB6*-D13S1830), with the hope that the variability of hearing loss in *GJB2* heterozygotes may be explained by interaction between these genes. The prevailing hypotheses are either: (1) that the connexin genes have a digenic mode of inheritance, so that a dosage effect accounts for the inner ear dysfunction; and/or (2) that the deletion may include a regulatory segment for *GJB2*, thus inhibiting normal gene expression [13-16]. However there is very little known about the phenotype.

We retrieved the audiometric records of hearing-impaired patients who were referred from different regions of the Netherlands to the University Medical Centre St. Radboud Nijmegen (UMCN) for *GJB2* and *GJB6* screening. Our aim was to describe the hearing impairment, particularly the severity and the progression, in Dutch children that have biallelic *GJB2* or *GJB6* mutations. Furthermore severity was analysed as a function of both the genotype and the type of mutation according to the change in amino acid sequence.

MATERIALS AND METHODS

From November 1998 to January 2003, 264 individuals with apparently recessive hearing impairment were referred from various genetic and otolaryngologic clinics in the Netherlands to the UMCN Human Genetics Department for *GJB2* and *GJB6* testing. Part of this sample (11 patients) was described previously [17-18]. In the analysis we included patients who met the following criteria: (1) no evidence of other causes of genetic hearing loss, syndromic or non-syndromic; (2) Dutch Caucasian origin; and (3) available audiometric records from below 18 years of age and onwards. For multi-affected families, only one hearing-impaired individual (pro-

band) was sampled. Interestingly there were patients with other genetic diseases who also carried connexin 26 mutations: a patient with hereditary neuropathy and polymorphisms [79G>A]+[341A>G]; a child with a 416G>A allele and *DFNA3*; and, a 35delG carrier with Usher2A syndrome. Another excluded patient was a member of a previously reported family with both *GJB2* and *GJB6* mutations segregating [19]. Eight adult patients with mutations, including two who had [35delG]+[del(*GJB6*-D13S1830)] and [167delT]+[del(*GJB6*-D13S1830)], were not included in the sample for analysis.

Genomic DNA was isolated from peripheral blood using standard salt extraction methods. For sequencing of the entire *GJB2* coding region, exon 2 was amplified with the primers 5'-CCTATGACAAACTAAGTTGGTTC-3' and 5'-GCCTCATCCCTCTCATGCTGTC-3'. Sequence analysis was performed with amplification primers and additional internal primers (forward primers 5'- GGGGAGATGAGCAGGCCGAC-3' and 5'-CGGCTGGTGAAGTGCAACGC-3'; reverse primers 5'-CGTAGCACACGTTCTTGCAGCCTG–3' and 5'- CGATGCGGACCTTCTGGGTTTTG–3') using the ABI PRISM Big Dye Terminator cycle sequencing V2.0 ready reaction kit and ABI PRISM 3700, or 3730 DNA analyser (Applied Biosystems Inc.). For the analysis of the splice donor site of exon 1, primers 5'-TCCGTAACTTTCCCAGTCTCCGAGGGAAGAGG-3' and 5'- CCCAAGGACGTGTGTTGGTCCAGCCCC-3' were used for amplification and sequence analysis.

The presence of the deletion encompassing part of the *GJB6* gene was analysed essentially as described by del Castillo et al. [14]. In brief, a multiplex PCR that detects both the wild type and the mutant allele was employed. The nomenclature system that was proposed by den Dunnen and Antonarakis (2001) was used to denote all sequence variations [20].

When available, documentation of medical history, physical examination and additional laboratory and audiovestibular tests were reviewed. For UMCN patients (n=106), questionnaires were administered to further screen for other causes of hearing loss, such as recurrent middle ear infections, prolonged use of antibiotics/medication, noise exposure, problems surrounding delivery at birth, head trauma and meningitis. Parents' consent was obtained in all pediatric cases; the rest of the patients also gave written informed consent.

Only pure tone audiometric results that were acquired through standard methods (soundproof booth with a calibrated machine, using ISO standards) were used. For cross-sectional analysis, the latest audiogram

was employed. Air-bone gap was checked to rule out middle ear problems at the time of testing. Asymmetry was defined as >10 dB difference in at least two frequencies between the two ears [21]. To measure hearing loss severity, the air conduction pure tone average (PTA) of both ears was computed using three frequencies (500, 1,000 and 2,000 Hz). Out-of-scale measurements were treated as missing values. Hearing loss was then categorized according to the guidelines of the European Work Group on Genetics of Hearing Impairment: 21-40 dB as mild, 41-70 dB as moderate, 71-95 dB as severe, and more than 95 dB as profound. The criteria set by Liu and Xu were followed for audiogram shape [22]. Using SPSS 11.0.1 (SPSS Inc., 1989-2001) software, Fisher's exact test, tests between means and nonparametric tests were utilized to check relationships between variables as appropriate. *P*-values of less than or equal to 0.05 were considered significant.

Because an audiogram measures thresholds for several frequencies (usually four to six) per ear, each test per frequency may be treated as a repeated measurement. Moreover, a patient may have several audiograms available. To avoid multiple repeated measures, only the last audiometric test was analysed using the SAS/STAT Version 8 procedure GENMOD (SAS Institute Inc., Cary, NC USA, 1999) in order to establish the relation between hearing threshold, mutation, gender and age at audiometry. GENMOD applies Generalized Estimation Equations that are less sensitive to violations of the assumption of normality. An autoregressive covariance structure was assumed. The [35delG]+[del(*GJB6*-D13S1830)] genotype was excluded from this analysis.

For longitudinal analysis of progression, linear regression analysis was done for patients with at least three consecutive measurements using GraphPad Prism version 3 software (GraphPad Software Inc., San Diego, CA USA, 2002). This method of analysis was outlined previously [23].

RESULTS

Out of 264 referred individuals, 222 were of Dutch descent. Of these, 35 patients with no other identifiable cause of genetic hearing impairment screened positive for connexin 26 mutations and were pediatric and/or had audiometric measurements in childhood. Ten were heterozygous for the following polymorphisms, with frequencies indicated in parentheses if occurring more than once: 35delG (2); 101T>C (3); 42C>G; 249C>G; 355 G>A; 457G>A; and 478G>A. There were 13 35delG homozygotes and 8 compound heterozygotes ([35delG]+[IVS1+1G>A] (2); [35delG]+[109G>A];

[35delG]+[229T>C]; [35delG]+[313_326del]; [35delG]+[449delT]; [101T>C]+[427C>T]; [299A>C]+[416G>A]). Four had connexin 30 plus connexin 26 mutations ([35delG]+[del(*GJB6*-D13S1830)] (3); [313_326del]+[del(*GJB6*-D13S1830)]). Presence of mutations was equally distributed between genders.

Table 4.1 Severity, shape and progression in patients with mutations

Mutation [a]	Protein change [b]	Severity	Audiometric shape
A. Biallelic mutations			
[35delG]+[35delG]	R+R	Severe [c]	Gently sloping/U-shaped
[35delG]+[35delG] [d]	R+R	Profound [c]	Residual
[35delG]+[35delG]	R+R	Profound [c]	Flat / gently sloping
[35delG]+[35delG]	R+R	Profound	Residual
[35delG]+[35delG]	R+R	Profound	Residual
[35delG]+[35delG]	R+R	Profound	Residual
[35delG]+[35delG]	R+R	Profound	U-shaped / flat
[35delG]+[35delG]	R+R	Profound	Gently sloping/residual
[35delG]+[35delG]	R+R	Severe	Gently sloping
[35delG]+ [IVS1+1G>A]	R+R	Profound [c]	Residual
[35delG]+ [IVS1+1G>A]	R+R	Severe	Gently sloping / flat
[35delG]+[313_326del]	R+R	Profound	Residual
[35delG]+[109G>A]	R+M	Moderate	Sharply sloping
[101T>C]+[427C>T]	M+M	Moderate	Flat
[299A>C]+[416G>A]	M+M	Profound [e]	Residual
[35delG]+[del(*GJB6*-D13S1830)]	R+R	Severe	Flat/ascending
B. Heterozygous mutations			
[35delG]+ Wt	R+Wt	Severe	Sharply sloping / gently sloping
[457G>A]+ Wt	M+Wt	Moderate	Ascending
[478G>A]+ Wt	M+Wt	Moderate	Ascending

[a] *Wt,* wild type.
[b] *R,* mutation with radical change in protein sequence; *M,* missense mutation.
[c] Nonprogressive (Figure 3.1).
[d] Progressive (Figure 3.2).
[e] Patient with monozygotic twin.
Expected amino acid changes with substitutions: 109G>A = V37I; 101T>C = M34T; 427C>T = R143W; 299A>C = H100P (novel); 416G>A = S139N; 457G>A = V153I; 478G>A = G160S. The last two were identified as polymorphisms, while 101T>C remains controversial.

Table 4.2 Severity according to genotype

Genotype [a]	*N*	Mean PTA (dB) [b]	Severity
[35delG]+[35delG]	9	101.5 (s.d.=11.6)	Severe to profound
[35delG]+[IVS1+1G>A]	2	97.9	Severe to profound
[35delG]+[313_326del]	1	106.7	Profound
[35delG]+ [del(*GJB6*-D13S1830)]	1	81.7	Severe
[35delG]+[109G>A]	1	45.0	Moderate
[101T>C]+[427C>T]	1	40.8	Moderate
[299A>C]+[416G>A]	1	107.5	Profound
[35delG]+Wt	1	82.5	Severe
[457G>A]+Wt	1	70.0	Moderate
[478G>A]+Wt	1	60.8	Moderate
Wt	22	91.8 (s.d.=17.1)	Moderate to profound

[a] *Wt*, wild type.
[b] Values computed from air conduction pure tone averages (0.5-2 kHz) of both ears from latest audiogram per individual.

Of the 35 patients with mutations, we were able to retrieve standard audiometric records for 19 (Table 4.1), 12 of whom were from the UMCN. Twenty-two more Dutch patients from the UMCN were negative for *GJB2* or *GJB6* mutations and other deafness-causing mutations screened but had pediatric audiograms available. Table 4.2 shows the mean thresholds for each genotype.

For the patients with *GJB2* mutations, the average age at first audiometry was 4.4 $\pm$ 0.59 years (s.d. = 2.58). For the last audiometry the average age was 12.3 $\pm$ 1.95 (s.d. = 8.51). The proportion of cases that had other siblings affected comprised 57.9% (11/19) of those with mutations and audiometry. Most of the mutations (29 of 35 alleles) were expected to change the protein sequence radically; these included deletions, nonsense and splice site mutations. Only six alleles had missense changes.

Five of nine 35delG homozygotes demonstrated asymmetric hearing impairment. The patient with the 109G>A mutation and two heterozygotes also showed asymmetric hearing loss.

The audiometric shape was highly variable among the biallelic mutations and even between ears, but was usually residual, flat or gently sloping (Table 4.1). Two 35delG homozygotes had U-shaped curves in one ear. Specifically the [35delG]+[109G>A] genotype resulted in sharply

sloping audiograms for both ears. Ascending shapes were found only in heterozygotes.

One [35delG]+[del(*GJB6*-D13S1830)] patient had asymmetric hearing loss: moderate in one ear and severe in the other, with a mean PTA of 81.7dB. Moreover the audiogram shape was flat on the right and ascending on the left. The audiologic records of two [35delG]+[del(*GJB6*-D13S1830)] individuals were not available. The child with the [313_326del]+[del(*GJB6*-D13S1830)] mutations had profound hearing loss as shown by auditory brainstem response (ABR).

Twenty-two UMCN patients who were negative for *GJB2* and *GJB6* and whose audiograms were also available were used as a comparison group to test if a specific picture of the *GJB2* phenotype emerges. Age, gender and familiality of hearing loss were not significantly different between the mutation and comparison groups. Differences in shape, air-bone gap and symmetry between the two groups were also not significant.

Table 4.3 Generalized Estimation Equations (GEE) parameter estimates for mutations grouped according to function [a]

	Estimate ± Se	*Z*-value	*p*-value
Model A:			
Radical + radical	Reference		
Radical + missense	-37.9 ± 5.06	-7.5	<0.0001
Missense + missense	-23.6 ± 22.55	-1.0	0.3
Radical + Wt	-8.9 ± 5.12	-1.7	0.08
Missense + Wt	-34.9 ± 3.86	-9.1	<0.0001
Wild type	-8.7 ± 4.63	-1.9	0.06
Type 3 Wald: $\beta_{mutation}$, d.f. = 5; Chi-square statistic = 535.6; *p*-value <0.0001; β_{age} and β_{gender} not significant.			
Model B:			
Radical + radical	Reference		
Missense + missense or radical	-29.2 ± 16.69	-1.8	0.08
Wt and mutation + Wt	-12.1 ± 4.86	-2.5	0.01
Type 3 Wald: $\beta_{mutation}$, d.f. = 2; Chi-square statistic = 6.9; *p*-value = 0.03; β_{age} and β_{gender} not significant.			

[a] Regression equation: db ~ intercept + $\beta_{mutation}$ • mutation + β_{age} • age + β_{gender} • gender.

Repeated measures analysis on patients with and without *GJB2* mutations showed a significant effect of the mutations, but not age and gender, on the hearing thresholds (SAS proc GENMOD, Wald type 3 statistic = 478.3, $p<0.0001$). Based on the parameter estimates, the severity of hearing impairment in subjects with the [35delG]+[35delG], [35delG]+[IVS1+1G>A] and [35delG]+Wt genotypes were not significantly different from the wild type, but those with [35delG]+[109G>A], [101T>C]+[427C>T] and polymorphisms had significantly lower thresholds (data not shown). On the other hand, threshold estimates for [35delG]+[313-326del] and [299A>C]+[416G>A] were significantly higher than for wild type. When the mutations were re-categorized into 35delG and non-35delG genotypes, these differences became non-significant. However when the grouping was changed according to amino acid sequence variation, i.e. "radical" [24] and missense, different combinations of mutations have a significant effect on the thresholds (SAS proc GENMOD Wald type 3 statistic = 535.6, $p<0.0001$, Model A in Table 4.3). After further reclassification into three groups (i.e. (1) radical + radical mutations, (2) at least one missense mutation, biallelic, (3) wild type and heterozygotes), the association between mutation and hearing threshold remained significant (Wald type 3 statistic = 6.9, p = 0.03, Model B in Table 4.3). A combination of radical mutations would be significantly worse than wild type and heterozygotes (p = 0.01). The presence of at least one missense mutation also showed a marked improvement in threshold; however the analysis showed a p-value of 0.08.

We pooled the data on severity of hearing impairment with three other studies from Northern Europe that used the same categorization [5-7]. Only the genotypes similar to those in this study were included. With a sample of 115, a significant association between severity and genotype was seen (Fisher's exact test, p = 0.009). Grouping into 35delG and non-35delG genotypes also resulted in significant association with severity (Fisher's, p = 0.006, Table 4.4), while grouping according to amino acid sequence change was highly significant when tested versus severity (Fisher's, $p<0.003$). It should be noted, however, that this significance is mainly derived from the comparison between 35delG homozygotes and 35delG heterozygotes.

Progression was examined in five 35delG homozygotes who had 7-22.7 years of follow-up (mean follow-up = 15.2 years), and was found to be not significant. Two of the patients are identical twins. When data from all five patients were put together, the age range tested was from 2.7 to 30 years old. Figure 4.1 shows the plot of one of the twins.

Table 4.4 Pooled data on severity according to genotype and amino acid change [a]

	n	Severity, *n* (%)
A. Genotype [b]		
[35delG]+[35delG]	70	45 (64.3%) profound [moderate to profound]
[35delG] + non-35delG	9	6 (66.7%) profound [moderate to profound]
Non-35delG + non-35delG	2	Moderate-profound
[35delG] + Wt	30	15 (50%) profound [mild to profound]
Non-35delG + Wt	4	Mild-severe
B. Amino acid change		
Radical + radical	78	51 (65.4%) profound [moderate to profound]
Radical + missense	1	Moderate
Missense + missense	2	Moderate-profound
Radical + Wt	30	15 (50%) profound [mild to profound]
Missense + Wt	4	Mild to severe

[a] Additional data from references [5-7].
[b] *Wt*, wild type.

For the five 35delG homozygotes, hearing impairment was found to be non-progressive. However the patient with the [299A>C]+[416G>A] genotype had significant progression at the low frequencies (2.2 dB/y at 0.25 kHz, 0.9 dB/y at 0.5 and 2.2 dB/y at 1 kHz; Figure 4.2).

DISCUSSION

In summary, we did not detect progression of hearing impairment with 35delG homozygotes. We also support the results of previous research that severity is a function of genotype and its effect on the amino acid sequence. As opposed to the severe or profound hearing loss that was found with deletions and splice site mutations, the genotypes with missense mutations tend to be moderate in hearing impairment.

From our sample of hearing-impaired individuals 15.8% (35/222) were positive for mutations, while 37 (66.1%) of 56 mutated *GJB2* alleles were 35delG. These percentages are consistent with literature [4-6, 8-10, 25]. Our average age at first audiometry may seem late for the diagnosis of *GJB2* deafness, which is usually prelingual. However earlier measurements by free field and/or observation audiometry were excluded from analysis.

We found patients with syndromic or non-syndromic hearing loss combined with *GJB2* mutations. Extra-auditory signs previously reported in

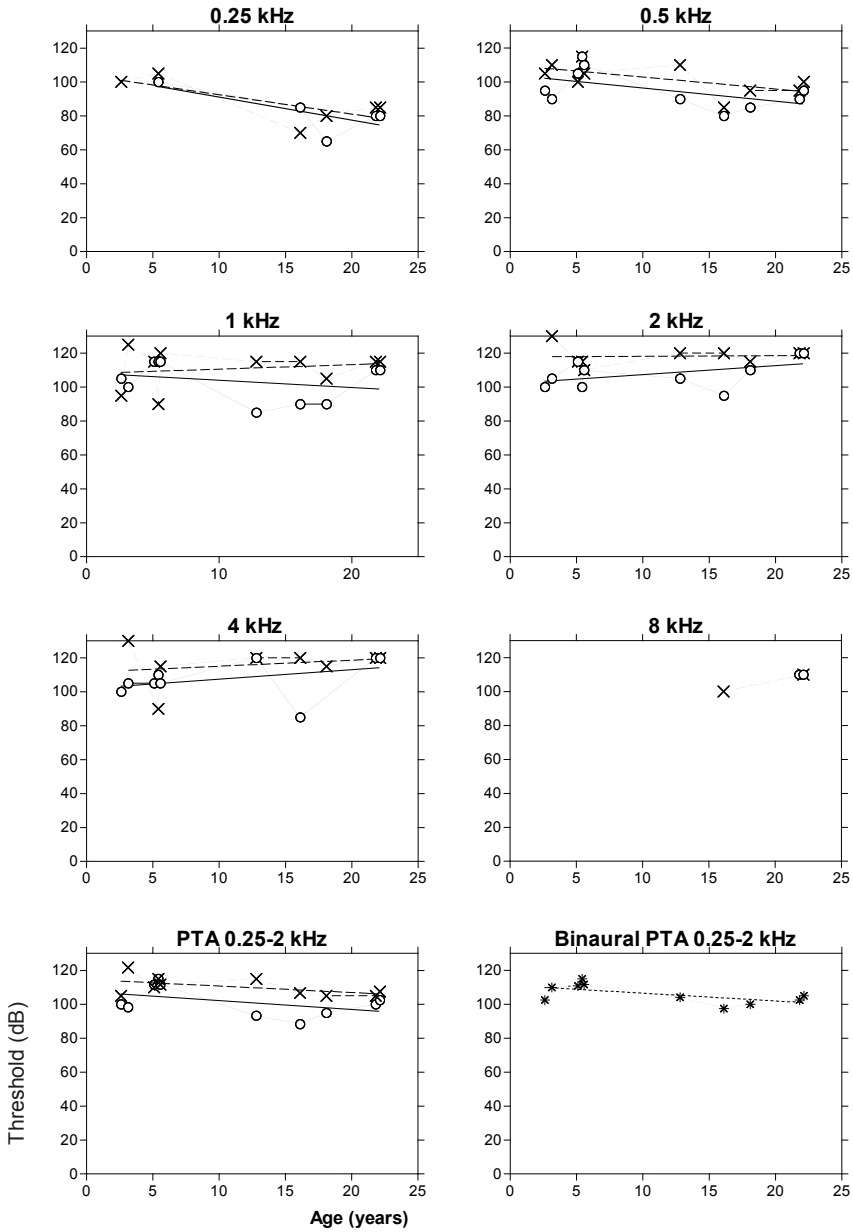


Figure 4.1 Serial threshold measurements in a homozygous 35delG patient not showing any significant progression. Air conduction threshold (circles, right; crosses, left) is plotted against age with connecting hairlines and regression lines (solid lines, right; dashed lines, left). Bottom panels show pure tone average (PTA) for each side and binaural PTA (asterisks and dotted lines).

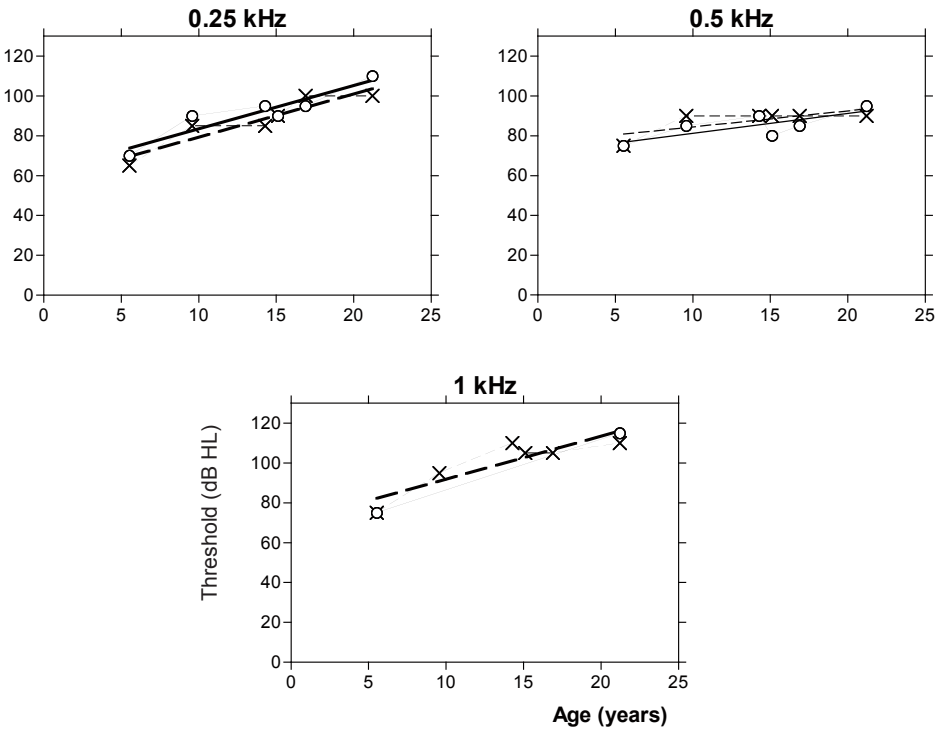


Figure 4.2 Measurable air conduction thresholds plotted against age in a patient with the [299A>C]+[416G>A] genotype. Same symbols as in Figure 4.1. Bold regression lines indicate significant progression. Progression was 2.2 dB/year in both ears at 0.25 kHz, 1.0 dB/year in the right ear and 0.8 dB/year in the left ear at 0.5 kHz, and 2.2 dB/year in the left ear at 1 kHz. Threshold data for the other frequencies and PTA data are not shown because they covered too many out-of-scale measurements to allow for reliable analysis.

GJB2 patients were speculated to be entirely due to *GJB2* gene dysfunction [26]. Occurrence of other genetic diseases should always be considered in these cases that might warrant further screening. This also paves the way for the study of yet undiscovered gene-gene interactions.

Surprisingly six patients with biallelic mutations had asymmetric hearing loss based on Mazzoli et al.'s criteria of >10 dB difference in ≥ 2 frequencies [21]. Interaural differences in *GJB2* patients were claimed to be rare [4], and no difference in symmetry was reported between those who screened positive and negative for *GJB2* mutations [5]. When we redefined asymmetry as >15 dB difference in at least two frequencies, in only two patients does the hearing loss remain asymmetric. Further use of Mazzoli et al.'s criteria for asymmetry should demonstrate its applicability to datasets.

Our results on audiometric curve shape agree with past research that there is no pathognomonic pattern for *GJB2*, and that flat and sloping

curves are common [4, 9, 27]. However, residual hearing occurred more often in our sample. Also we found U-shaped curves, although only in one ear; ascending shapes were seen only in heterozygotes. Our data on progression do not support the possibility of temporal changes in shape as previously proposed [27]. Better ascertainment of patients with milder hearing impairment may provide more clues concerning shape in the *GJB2* genotype.

Certain genotypes, specifically the [35delG]+[109G>A] and [101T>C]+[427C>T], had milder phenotypes, as compared to [35delG]+[35delG], [35delG]+[IVS1+1G>A] and [35delG]+[313-326del]. This was further confirmed by repeated measurements analysis. Recent research put forward the idea that the wide variability in severity for *GJB2* mutations is dependent on the genotype [6, 10]. Looking into the effect of these genotypes on the amino acid sequence strengthens this hypothesis even more. The phenotype for the 35delG homozygote remained variable, as in past research, which ranged from moderate to profound. Compared to Cryns et al. however, our [35delG]+[109G>A] and [35delG]+[IVS1+1G>A] patients exhibited more severe hearing impairment [10]. Cryns et al. also did not include the 101T>C alleles in their analysis. Though much controversy has surrounded this mutation, several studies suggested that the 101T>C is a recessive mutation that may cause milder forms of hearing loss [7, 11-12, 28-29].

The genotype [299A>C]+[416G>A], though expected to produce a missense change, produced hearing impairment which is worse than in 35delG homozygotes and with progression at the lower frequencies (Figure 4.2). We did not encounter previous reports of this genotype. The affected patient's younger sibling with the same genotype has severe hearing impairment based on click brainstem audiometry. More individuals may be needed to study the phenotype for this specific genotype.

In a multicenter study, heterozygous *GJB6* mutations were found in 0-29 patients among different populations and even rarer, del(*GJB6*-D13S1830) homozygotes in <0.5% [30]. We found six del(*GJB6*-D13S1830) heterozygotes in 222 (2.7%) hearing-impaired individuals, which further testifies that this mutation is not as rare as previously thought. To identify the *GJB6* deletion, the haplotypes of prelingually and profoundly deaf patients were analysed in previous reports [13, 15]. In a study of several multi-affected families from West Flanders, the [35delG]+[del(*GJB6*-D13S1830)] genotype was also shown to cause profound deafness [19]. Unfortunately we had only one patient with available audiometry, whose hearing loss was

severe rather than profound. More research, especially in populations, is required to confirm these findings.

In past research *GJB2* mutations were shown to be nonprogressive [4-5, 9, 10]. Although 88% of their sample was non-progressive, Janecke et al. reported progressivity of hearing loss associated with sudden sensorineural hearing loss in 3 patients [6]. We demonstrated non-progressive hearing impairment for 35delG homozygotes at age 2 1/2 and above.

Our data on progression agrees with the hypothesis that hearing loss in most *GJB2* mutations is prelingual [4]. In studies in mice, hearing impairment began at about the onset of inner ear function [31], which corresponds to the 20th week of gestation in humans. The hair cell damage and death that ensued was attributed to low endolymphatic K^+ potentials. The question as to whether hearing impairment progresses prenatally or postnatally may be important, because intervention prior to complete degeneration of the inner ear might result in better habilitation [31]. Since our number of cases is small, we hope that our data can be part of a bigger sample to establish nonprogression in connexin 26 mutations.

To conclude, we presented additional evidence that *GJB2* mutations cause hearing impairment that in general does not progress, and that varies in severity depending on the genotype and its effect on the amino acid sequence. *GJB6* mutations were relatively frequent, and may cause a severe phenotype. Further use of new criteria for symmetry and more data on audiogram shape may be required to better establish the genotype-phenotype correlation for the connexin genes.

Acknowledgements The authors would like to thank the participating patients and families and the doctors who provided clinical information. We also thank Prof. Theo Stijnen for statistical advice, and Prof. CM van Duijn for general support. This study was sponsored in part by the Heinsius Houboult Foundation (to C.C.) and the Netherlands organization for international cooperation in higher education (Nuffic) (to R.S.).

REFERENCES

1. Rabionet R, Gasparini P, Estivill X. Molecular genetics of hearing impairment due to mutations in gap junction genes encoding beta connexins. Hum Mutat 2000; 16: 190-202.
2. Van Laer L, Coucke P, Mueller RF, Caethoven G, Flothmann K, Prasad SD, Chamberlin GP, Houseman M, Taylor GR, Van de Heyning CM, Fransen E, Rowland J, Cuccci RA, Smith RJ, Van Camp G. A common founder for the 35delG *GJB2* gene mutation in connexin 26 hearing impairment. J Med Genet 2001; 38: 515-518.
3. Calvo J, Rabionet R, Gasparini P, Estivill X. Connexins and Deafness Homepage. 2004. URL: http://www.crg.es/deafness.

4. Denoyelle F, Marlin S, Weil D, Moatti L, Chauvin P, Garabedian EN, Petit C. Clinical features of the prevalent form of childhood deafness, DFNB1, due to a connexin-26 gene defect: implications for genetic counselling. Lancet 1999; 353: 1298-1303.
5. Mueller RF, Nehammer A, Middleton A, Houseman M, Taylor GR, Bitner-Glindzicz M, Van Camp G, Parker M, Young ID, Davis A, Newton VE, Lench NJ. Congenital non-syndromal sensorineural hearing impairment due to connexin 26 gene mutations – molecular and audiological findings. Int J Pedia Otorhinolaryngol 1999; 50: 3-13.
6. Janecke AR, Hirst-Stadlmann A, Gunther B, Utermann B, Muller T, Loffler J, Utermann G, Nekahm-Heis D. Progressive hearing loss, and recurrent sudden sensorineural hearing loss associated with *GJB2* mutations – phenotypic spectrum and frequencies of *GJB2* mutations in Austria. Hum Genet 2002; 111: 145-153.
7. Zoll B, Petersen L, Lange K, Gabriel P, Kiese-Himmel C, Rausch P, Berger J, Pasche B, Meins M, Gross M, Berger R, Kruse E, Kunz J, Sperling K, Laccone F. Evaluation of CX26/*GJB2* in German hearing impaired persons: mutation spectrum and detection of disequilibrium between M34T (c.101T>C) and -493del10. Hum Mutat 2003; 21: 98.
8. Cohn ES, Kelley PM, Fowler TW, Gorga MP, Lefkowitz DM, Kuehn HJ, Schaefer GB, Gobar LS, Hahn FJ, Harris DJ, Kimberling WJ. Clinical studies of families with hearing loss attributable to mutations in the connexin 26 gene (*GJB2*/DFNB1). Pediatrics 1999; 103: 546-550.
9. Murgia A, Orzan E, Polli R, Martella M, Vinanzi C, Leonardi E, Arslan E, Zacchello F. CX26 deafness: mutation analysis and clinical variability. J Med Genet 1999; 36: 829-832.
10. Cryns K, Orzan E, Murgia A, Huygen PLM, Moreno F, del Castillo I, Chamberlin GP, Azaiez H, Prasad S, Cucci RA, Leonardi E, Snoeckx RL, Govaerts PJ, Van de Heyning PH, Van de Heyning CM, Smith RJH, Van Camp G. A genotype-phenotype correlation for *GJB2* (connexin 26) deafness. J Med Genet 2004; 41: 147-154.
11. Cucci RA, Prasad S, Kelley PM, Green GE, Storm K, Willocx S, Cohn ES, Van Camp G, Smith RJ. The M34T allele variant of connexin 26. Genet Test 2000; 4: 335-344.
12. Houseman MJ, Ellis LA, Pagnamenta A, Di W, Rickard S, Osborn AH, Dahl HH, Taylor GR, Bitner-Glindzicz M, Reardon W, Mueller RF, Kelsell DP. Genetic analysis of the connexin-26 M34T variant: identification of genotype M34T/M34T segregating with mild-moderate non-syndromic sensorineural hearing loss. J Med Genet 2001; 38: 20-25.
13. Lerer I, Sagi M, Ben-Neriah Z, Wang T, Levi H, Abeliovich D. A deletion mutation in *GJB6* cooperating with a *GJB2* mutation in trans in non-syndromic deafness: a novel founder mutation in Ashkenazi Jews. Hum Mutat 2001; 18: 160.
14. del Castillo I, Villamar M, Moreno-Pelayo MA, del Castillo FJ, Alvarez A, Telleria D, Menendez I, Moreno F. A deletion involving the connexion 30 gene in nonsyndromic hearing impairment. N Engl J Med 2002; 346: 243-249.
15. Pallares-Ruiz N, Blanchet P, Mondain M, Claustres M, Roux AF. A large deletion including most of *GJB6* in recessive non syndromic deafness: a digenic effect? Eur J Hum Genet 2002; 10: 72-76.
16. Teubner B, Michel V, Pesch J, Lautermann J, Cohen-Salmon M, Sohl G, Jahnke K, Winterhager E, Herberhold C, Hardelin JP, Petit C, Willecke K. Connexin30 (*GJB6*)-deficiency causes severe hearing impairment and lack of endocochlear potential. Hum Mol Gen 2003; 12: 13-21.
17. Kemperman MH, Hoefsloot LH, Cremers CWRJ. Hearing loss and connexin 26. J Royal Soc Med 2002; 95: 171-177.

18. Hoefsloot LH, Kemperman MH, Cremers CWRJ. Van gen naar ziekte; slechthorendheid en connexine 26. Nederlands Tijdschrift Geneeskunde 2002; 146: 259-261.
19. Stinckens C, Kremer H, van Wijk E, Hoefsloot LH, Huygen PLM, Standaert L, Fryns JP, Cremers CWRJ. Longitudinal phenotypic analysis in patients with connexin 26 (*GJB2*) (DFNB1) and connexin 30 (*GJB6*) mutations. Ann Otol Rhinol Laryngol 2004; 113: 587-593.
20. den Dunnen JT, Antonarakis E. Nomenclature for the description of human sequence variations. Hum Genet 2001; 109: 121-124.
21. Mazzoli M, Van Camp G, Newton V, Giarbini N, Declau F, Parving A. Recommendations for the description of genetic and audiological data for families with nonsyndromic hereditary hearing impairment. In: Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2003. URL: http://dnalab-www.uia.ac.be/dnalab/hhh/
22. Liu X, Xu L. Nonsyndromic hearing loss: an analysis of audiograms. Ann Otol Rhinol Laryngol 1994; 103: 428-433.
23. De Leenheer EMR, Huygen PLM, Coucke PJ, Admiraal RJC, Van Camp G, Cremers CWRJ. Longitudinal and cross-sectional phenotype analysis in a new, large Dutch DFNA2/KCNQ4 family. Ann Otol Rhinol Laryngol 2002; 111: 267-274.
24. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Genet Supp 2003; 33: 228-237.
25. Pampanos A, Economides J, Iliadou V, Neou P, Leotsakos P, Voyiatzis N, Eleftheriades N, Tsakanikos M, Antoniadi T, Hatzaki A, Konstantopoulou I, Yannoukakos D, Gronskov K, Brondum-Nielsen K, Grigoriadou M, Gyftodimou J, Iliades T, Skevas A, Petersen MB. Prevalence of *GJB2* mutations in prelingual deafness in the Greek population. Int J Pedia Otolaryngol 2002; 65: 101-108.
26. Gualandi F, Ravani A, Berto A, Sensi A, Trabanelli C, Falciano F, Trevisi P, Mazzoli M, Tibiletti MG, Cristofari E, Burdo S, Ferlini A, Martini A, Calzolari E. Exploring the clinical and epidemiological complexity of *GJB2*- linked deafness. Am J Med Genet 2002; 112: 38-45.
27. Wilcox SA, Saunders K, Osborn AH, Arnold A, Wunderlich J, Kelly T, Collins V, Wilcox LJ, McKinlay Gardner RJ, Kamarinos M, Cone-Wesson B, Williamson R, Dahl HH. High frequency hearing loss correlated with mutations in the *GJB2* gene. Hum Genet 2000; 106: 399-405.
28. Griffith AJ, Chowdhry AA, Kurima K, Hood LJ, Keats B, Berlin CI, Morell RJ, Friedman TB. Autosomal recessive nonsyndromic neurosensory deafness at DFNB1 not associated with the compound-heterozygous *GJB2* (connexin 26) genotype M34T/167delT. Am J Hum Genet 2000; 67: 745-749.
29. Kenna MA, Wu BL, Cotanche DA, Korf BR, Rehm HL. Connexin 26 studies in patients with sensorineural hearing loss. Arch Otolaryngol Head Neck Surg 2001; 127: 1037-1042.
30. del Castillo I, Moreno-Pelayo MA, del Castillo FJ, Brownstein Z, Marlin S, Adina Q, Cockburn DJ, Pandya A, Siemering KR, Chamberlin GP, Ballana E, Wuyts W, Maciel-Guerra AT, Alvarez A, Villamar M, Shohat M, Abeliovich D, Dahl HHM, Estivill X, Gasparini P, Hutchin T, Nance WE, Sartorato EL, Smith RJH, Van Camp G, Avraham KB, Petit C, Moreno F. Prevalence and evolutionary origins of the del(*GJB6*-D13S1830) mutation in the DFNB1 locus in hearing-impaired subjects: a multicenter study. Am J Hum Genet 2003; 73: 1452-1458.
31. Cohen-Salmon M, Ott T, Michel V, Hardelin JP, Perfettini I, Eybalin M, Wu T, Marcus DC, Wangemann P, Willecke K, Petit C. Targeted ablation of connexin26 in the inner ear epithelial gap junction network causes hearing impairment and cell death. Curr Biol 2002; 12: 1106-1111.

Chapter 4.2

Phenotypic characterization of *DFNA24*: prelingual progressive sensorineural hearing impairment

Santos RLP, Häfner FM, Huygen PLM, Linder TE, Schinzel AA, Spillmann T, Leal SM. Reprinted from Audiology & Neurotology 11(5): 269-275.

SUMMARY

This article describes the hearing impairment (HI) phenotype which segregates in a large multi-generation Swiss German family with autosomal dominant non-syndromic HI. The locus segregating within this pedigree is located on chromosome 4q35-qter and is designated as *DFNA24*. For this pedigree audiometric data on 24 hearing-impaired family members are available. It was demonstrated that within this kindred the HI is sensorineural, bilateral, prelingual in onset, and progressive throughout life. Age-related typical audiograms depict steeply down-sloping curves, with moderate high-frequency HI at birth, then steady progression to moderate HI in the low frequencies, severe HI at mid-frequencies and profound HI at high frequencies by age 70. Annual threshold deterioration was ~0.5 dB/year at 1-2 kHz after correction for presbycusis.

INTRODUCTION

Because of the extensive genetic heterogeneity in hearing impairment (HI), the importance of detailed phenotypic characterization for each gene or locus must be emphasized. Currently, for the approximately 120 loci which have been mapped for non-syndromic HI, only 37 genes have been identified [1]. Finding the causative gene for each locus, documenting the functional mutations, and the estimation of population-specific prevalences for each gene and/or mutation will ultimately help in genetic screening, clinical management and physiologic understanding of the inner ear. However, the process of gene finding is often hindered by huge genetic intervals from which to screen candidate genes and the lack of validation regarding the significance of sequence variants in the same gene from other families or populations.

The *DFNA24* locus was first identified in a Swiss-German family with a history of autosomal dominant non-syndromic HI which dates back to the mid-19th century [2]. To date the causative gene for this locus has not been identified, and no other family or population has been linked to the same locus, inhibiting further fine-mapping efforts and validation of potential

sequence variants. The aim of this study is to describe the audiometric phenotype associated with the *DFNA24* locus, and hopefully aid in the identification of other families or individuals with the same phenotype that may also have HI due to *DFNA24*. Additionally, awareness of this phenotype could aid in the clinical management and genetic counseling of patients that have HI due to this locus.

MATERIALS AND METHODS

Prior to data collection, the study was approved by the institutional review board of the University Hospital of Zurich. All procedures were carried out according to the Declaration of Helsinki. Informed consent was obtained from all adult study participants and parents of participants who were below 18 years of age.

The pedigree structure was reconstructed through interviews with family members (Figure 4.3). DNA and audiometric data were available for the last three of six generations. In total, 47 family members were genotyped (25 hearing-impaired and 15 hearing, plus seven spouses), and of these individuals 38 had audiologic testing (25 hearing-impaired, 12 hearing, and one spouse who was homozygous for the 35delG variant of the *GJB2* gene). Through careful medical history and physical examination, there was no evidence for the existence of syndromic abnormalities or vestibular disorders within the family, or that the HI was due to environmental risk factors.

For most family members, HI was first noted in early childhood while in primary school and fitted with hearing aids. At least eight individuals in generations V and VI were diagnosed and fitted with hearing aids at ages 3-7 years (born 1957 to 1991), while four family members were known to be hearing-impaired in childhood but were fitted with hearing aids in adulthood (born 1922 to 1953, with ages at first hearing aid at 33-62). Individual V-18 was noted to be hearing-impaired while in kindergarten but started using amplification in high school. Eleven individuals received education in regular schools. Only three family members attended schools for the deaf. There is no evidence of anticipation within the pedigree because there was not a greater severity of HI or earlier age of onset among individuals in later generations. The earlier use of hearing aids in later generations is most likely due to advances in screening for hearing impairment and hearing aid technology.

Spouse V-25 is homozygous for the 35delG variant in the *GJB2* gene, thus her affection status was made "unknown" for the linkage analysis. Her

children were also tested and they were heterozygous for 35delG, therefore the hearing impairment in 2 out of 4 children who inherited the *DFNA24* haplotype was assumed to be due to *DFNA24* and not *GJB2*. The six other spouses who were genotyped but did not have audiograms were assumed "unaffected" in the linkage analysis.

The genetic analysis for the family under study is reported elsewhere [2]. A maximum multipoint LOD score of 11.6 was obtained at marker D4S1652. The linkage interval was between markers GATA129D03 and D4S1523/UT1366 at chromosomal region 4q35-qter, and the *DFNA24* haplotype was found within the same linkage interval. This corresponds to a genetic length of 7.3 cM according to the Rutgers combined linkage-physical map [3] and contains 8.1 Mb based on Build 34 of the human reference sequence [4]. All individuals who carry the *DFNA24* haplotype are marked in the pedigree drawing with a (+) sign (Figure 4.3).

Results of previous audiometric testing were collected when possible. For 36 family members, audiometric measurements were performed in a soundproof compartment under ISO standards, at frequencies 0.125-8 kHz for air conduction and 0.25-4 kHz for bone conduction. Selected individuals underwent tympanometry, stapedius reflex and acoustic reflex testing. Three affected relatives were found to have concomitant middle ear problems. Stapedius reflex and acoustic reflex testing indicated loss of inner ear function in a few affected individuals who were tested. For the audiogram analyses, immeasurable air conduction levels were fixed at an arbitrary value of 130 dB. After a visual check on symmetry of the last-visit audiograms of all the affected individuals (Figure 4.4), the binaural mean air conduction threshold was calculated and used for further analyses. These analyses comprised linear regression analysis of threshold on age for each sound frequency. It was tested whether the slope of the regression line, referred to as annual threshold deterioration (ATD), expressed in dB per year, differed significantly from zero. To test if the intercept differed significantly from zero, it was checked whether the estimated 95% confidence interval (95%CI) for the Y-intercept did not include zero. The test for significant slope was repeated after "correcting" the air conduction threshold data by subtracting the median (P50) threshold for normal presbycusis according to the patient's age and sex as indicated by the ISO 7029 norm [5]. The relative prevalence of significant slopes found among the total number of evaluated frequencies was used to test whether this prevalence was significantly high (i.e. $p < 0.05$) according to the appropriate binomial distribution. As outlined in a previous report [6], the derived regression

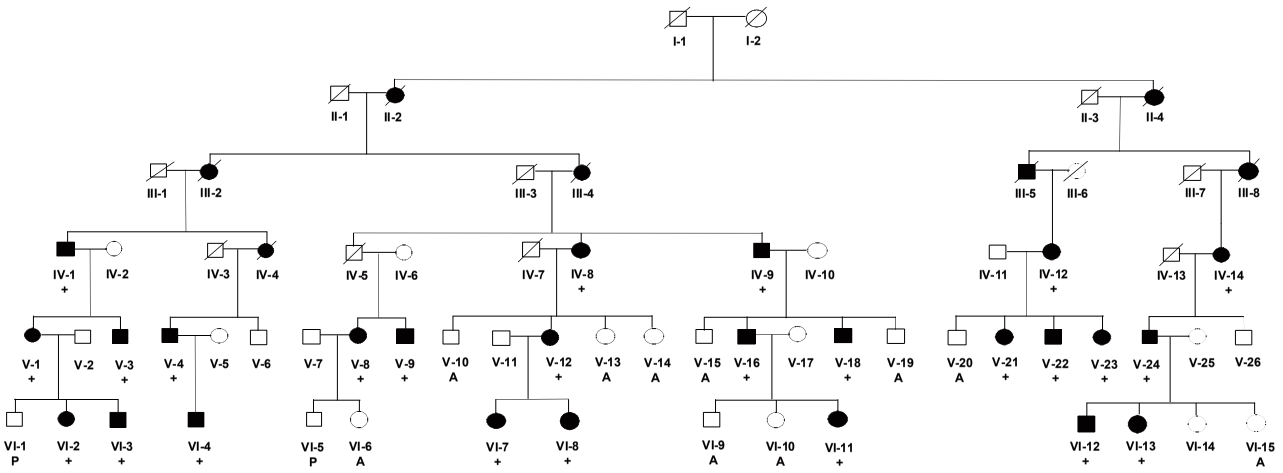


Figure 4.3 Pedigree drawing of *DFNA24* family. The order of birth and gender of some individuals were modified to protect the family's anonymity. Filled symbols indicate hearing impairment, while clear symbols denote normal hearing or high-frequency impairment due to presbycusis. Family members who carry the *DFNA24* haplotype are indicated by a (+) sign. Audiograms were available for all hearing-impaired individuals and for hearing family members that are marked as (A). Two individuals marked with (P) have normal hearing based on pediatric audiology.

equations were used to obtain air conduction threshold estimates by age in decades in order to construct age-related typical audiograms (ARTA).

RESULTS

The audiogram data showed that the HI due to *DFNA24* is sensorineural. Air-bone gaps (>15 dB averaged at 0.5-2 kHz) were noted only in three individuals whose last-visit audiograms were taken in childhood; in all three family members, a middle ear problem suggestive of effusion was documented. The audiograms from these three individuals were excluded from further analyses.

A random selection of individual last-visit audiograms is presented in Figure 4.4. Both ears in all cases showed fair symmetry in air conduction threshold (Figure 4.4), so that the binaural mean threshold could be used as a key parameter. It is clear that HI, most pronounced at the higher frequencies, was already present at the youngest ages, and showed progressive deterioration at all frequencies with advancing age. Across all ages the shape of the audiogram remained down-sloping, with greater affectation of the higher frequencies compared to the lower and middle frequencies.

Figure 4.5 shows the cross-sectional regression analysis of air conduction threshold on age. Threshold saturation played a role at frequencies of >3 kHz. The plots in Figure 3 show a gradual increase in threshold with advancing age. The ATD increased from ~0.5-0.6 dB/year at 0.125-0.5 kHz to ~0.9 dB/year at 4-8 kHz. The Y-intercept differed significantly from zero at all frequencies above 0.125 kHz; estimated values increased from 4 dB at 0.125 kHz to 65 dB at 8 kHz (Figure 4.5). In other words, all frequencies were affected in individuals with HI due to *DFNA24*, presumably at birth, but hearing at the higher frequencies was more severely impaired from the start and deteriorated at a more rapid rate. Unfortunately, there was an insufficient number of longitudinal data available to allow for reliable checks on individual progression.

To further determine whether there was progression of HI beyond presbycusis, cross-sectional linear regression analysis of the air conduction threshold minus the P50 for presbycusis for ages <70 years was performed for the frequencies 0.25, 0.5, 1, 2, 4 and 8 kHz. The "corrected" ATD varied from ~0.2 dB/year at 0.25 kHz to around 0.4-0.5 dB/year at 1-2 kHz. At 4 and 8 kHz, the corrected ATD was only 0.4 and 0.2 dB/year, respectively (data not shown). However, it should be realized that at the latter frequencies the HI in many individuals was beyond the measurement limits of the audiometer so that no corrected threshold values could

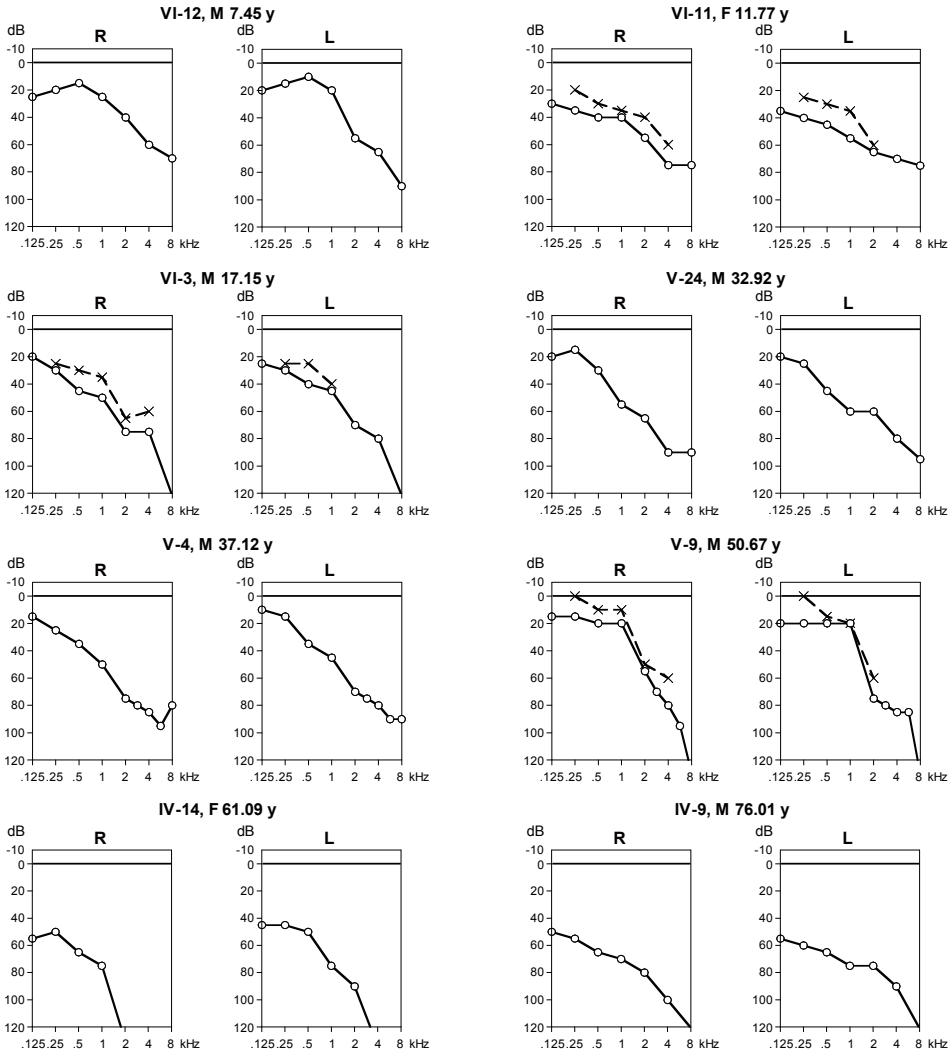


Figure 4.4 Audiograms of a selection of hearing-impaired members of *DFNA24* family ordered from top left to bottom right by age at last visit. These individuals were randomly selected by ordering all affected individuals by age and then choosing the audiogram from every third person in the list. Above the audiogram panels, pedigree ID, gender (M male, F female), age (in years) and laterality (R right, L left) are indicated. Open circles denote air conduction thresholds (in dB), crosses are bone conduction thresholds.

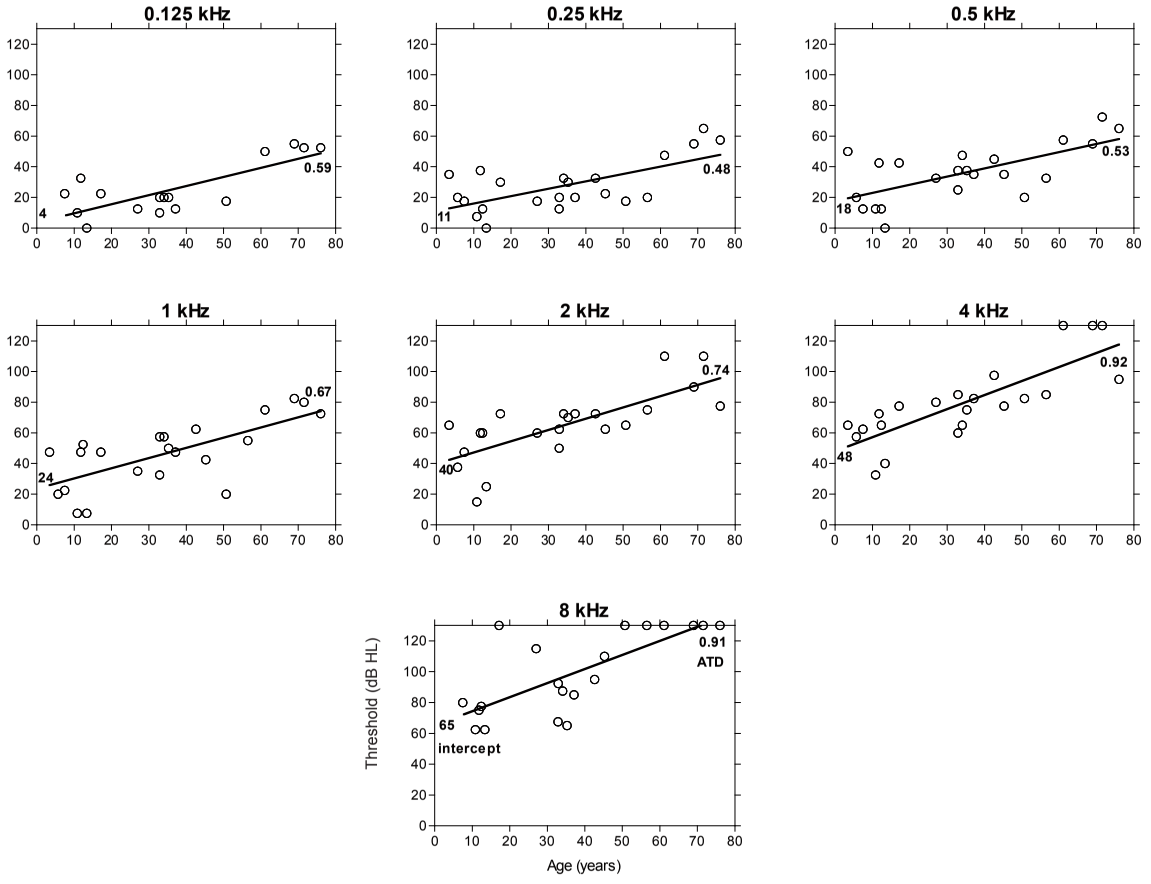


Figure 4.5 Cross-sectional analysis of binaural mean air conduction threshold (dB HL) on age (in years) at the frequencies 0.125-8 kHz for which n>10 measurements were available. Each frequency panel includes the regression line with a bold figure on the right end indicating the annual threshold deterioration or ATD (dB/year) and a bold figure on the left end indicating the Y-intercept (dB). All slopes and the Y-intercepts of > 4 dB are significantly different from zero.

be calculated from the crude threshold measurements that were out of scale. Therefore there was an increasing bias associated with the loss of data at these frequencies (Figure 4.5). Progression beyond presbycusis was significant in two out of six frequencies (1 and 2 kHz), which implies a significantly high prevalence (binomial tail probability, $p = 0.009$). Therefore, progression beyond presbycusis was distinctly present at the frequencies for which reliable, unbiased evaluation could be achieved; however, deterioration in excess of normal presbycusis was limited to ~0.5 dB/year at these frequencies.

Figure 4.6 shows the ARTA that were derived from the regression results, and demonstrates that HI based on average hearing thresholds gradually progressed from mild (~10-30dB at 0.25-1 kHz) to moderate (~40-60 dB at 2-8 kHz) impairment in childhood, to moderate (~40-70 dB at 0.25-1 kHz) to profound HI (90 dB or more at 2-8 kHz) by age 70. The degree of congenital HI that could be extrapolated from the cross-sectional regression analysis was significant and substantial, predominantly at the high frequencies (at age 0, > 60 dB at 8 kHz).

Pure tone audiometry records were available for ten out of fifteen family members who were not spouses and were therefore at risk. Linkage analysis later revealed that these individuals did not carry the *DFNA24* haplotype. For two other pedigree members who did not inherit the *DFNA24* haplotype, normal hearing status is based on pediatric audiology (marked as "P" in Figure 4.3). For the remaining three non-carriers, audiometry has not been performed; based upon self-report, these individuals were assumed "unaffected" in linkage analysis. Of those who had audiometric records, four non-carriers (aged 4-10) had normal hearing at all frequencies. In 6 individuals aged >28, some amount of HI (average threshold 20-48 dB or 0-28 dB hearing loss) was detected at the high frequencies (4-12 kHz). However, testing based on P95 thresholds of ISO 7029 standards for presbycusis according to the individual's age and sex showed that these individuals did not have significant HI beyond presbycusis (data not shown). Therefore family members who did not carry the *DFNA24* haplotype either had normal hearing (age <10) or isolated high-frequency hearing impairment that may be attributable to presbycusis (ages 28-52). This is in contrast to carriers of the *DFNA24* haplotype in which HI is detected at the mid-to-high frequencies by age 10 and at all frequencies by age 20. As stated above, the HI in haplotype carriers is significantly greater than what is expected from presbycusis, particularly in the low frequencies.

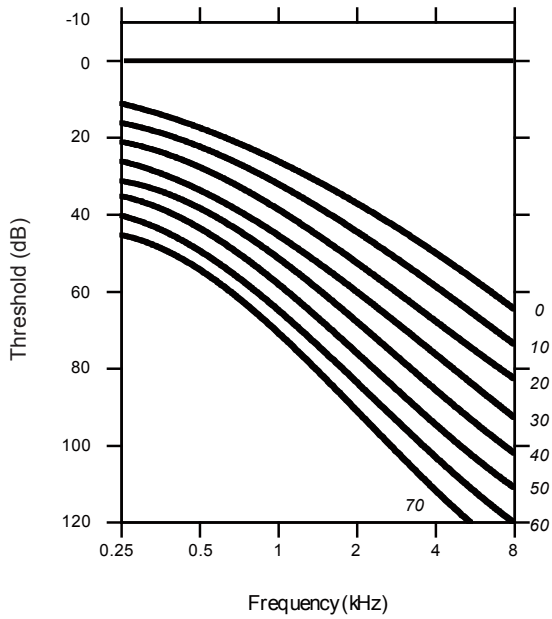


Figure 4.6 Age-related typical audiograms (ARTA) for the *DFNA24* trait. Age (in years) is in italics on the right side of the audiogram curve.

DISCUSSION

In summary, the *DFNA24* locus is associated with HI that is sensorineural, bilateral and steeply down-sloping in audiometric configuration. When presbycusis was taken into account, there was a limited but significant degree of excess in hearing deterioration, amounting to ~0.5 dB/year at 1-2 kHz, i.e. the frequencies that could be evaluated without tangible saturation bias. Thus it was established that, similar to most autosomal dominant non-syndromic HI phenotypes, *DFNA24* causes progressive deterioration of hearing with increasing age. However the HI is prelingual, with moderate high-frequency loss expected at birth. The down-sloping shape of the audiogram is maintained throughout life, with steady progression to moderate impairment at the middle frequencies and severe impairment at high frequencies towards early adulthood, and then to moderate HI in the low frequencies, severe HI in mid-frequencies, and residual hearing at high frequencies in elderly age.

The *DFNA24* phenotype clearly belongs to the category of high-frequency (down-sloping) progressive HI, with the phenotypes of *DFNA2*, *DFNA5* and *DFNA7* being most similar. Eleven of 15 documented *DFNA2* families harbor a mutation in the *KCNQ4* gene (MIM 603537); the HI in most of these families was also presumably congenital and, after correc-

tion for presbycusis, had an ATD averaged for all frequencies at 0.65 dB/year [7]. *DFNA5* and *DFNA7* also have down-sloping audiometric shapes, but both have postlingual onset and much more rapid progression particularly in the first decades of life [8-9]. Therefore of the three loci, the *DFNA24* phenotype is closest in its description to that of *DFNA2/KCNQ4*, save for a seemingly slower progression rate of HI due to *DFNA24*. It is interesting to speculate whether the pathology involved in this group of high-frequency progressive HI phenotypes involves a common pathway, but there is no evidence available at present to forward such a hypothesis. Recently it was suggested that dysfunction of the voltage-gated potassium channel *KCNQ4* in the basal inner hair cells and spiral ganglion neurons affect electrical signaling in the inner ear [10]. As regards to *DFNA5*, *Dfna5*$^{-/-}$ mice did not demonstrate significant HI or vestibular dysfunction [11]. The causative genes within the *DFNA7* locus [12] and the gene-poor *DFNA24* region are yet unknown.

As more HI loci are mapped, the distinction among the different loci or genes in terms of functional effect is easily blurred. It has been shown previously that reference to already published HI phenotypes, based on a specific audiometric pattern, can actually lead to faster genetic screening [7]. Detailed phenotypic description of more families or patients that are identified with a specific non-syndromic HI locus will contribute to better clinical management, whether it be in the prevention of speech delay or disability through aggressive monitoring and rehabilitation after early diagnosis, in genetic counseling of affected families, or in greater awareness of environmental or occupational hazards that may compound an existing genetic susceptibility to development of HI. Because only the family reported here has been associated with *DFNA24*, there is a possibility that the phenotype has been influenced by genetic or environmental factors that are specific to this family and may therefore not exactly reflect the phenotype of other families or individuals that have hearing impairment due to the same gene or variant. It is hoped that the publication of the description of the *DFNA24* phenotype will facilitate the reporting of other families with HI due to *DFNA24* that will allow not only further fine-tuning of the phenotype but, more importantly, lead to the identification of a novel HI gene.

Acknowledgements We thank the family members who participated in this study. This work was supported by the National Institutes of Health – National Institute of Deafness and Communication Disorders grant R01-DC03594 (to S.M.L.).

REFERENCES

1. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2005. URL: http://webhost.ua.ac.be/hhh/
2. Häfner FM, Salam AA, Linder T, Balmer D, Baumer A, Schinzel A, Spillmann T, Leal SM. A novel locus (*DFNA24*) for prelingual nonprogressive autosomal dominant nonsyndromic hearing loss maps to 4q35-qter in a large Swiss German kindred. Am J Hum Genet 2000; 66: 1437-1442.
3. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
4. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860-921.
5. ISO 7029. Acoustics: Threshold of hearing by air conduction as a function of age and sex for otologically normal persons. Geneva: International Organization for Standardization, 1984.
6. Huygen PLM, Pennings RJE, Cremers CWRJ. Characterizing and distinguishing progressive phenotypes in nonsyndromic autosomal dominant hearing impairment. Audiol Med 2003; 1: 37-46.
7. Topsakal V, Pennings RJE, te Brinke H, Hamel B, Huygen PLM, Kremer H, Cremers CWRJ. Phenotype determination guides swift genotyping of a DFNA2/KCNQ4 family with a hot spot mutation (W276S). Otol Neurotol 2005; 26: 52-58.
8. De Leenheer EMR, van Zuijlen DA, Van Laer L, Van Camp G, Huygen PLM, Huizing EH, Cremers CWRJ. Clinical features of DFNA5. Adv Otorhinolaryngol 2002; 61: 53-39.
9. Huygen PLM, Bom SJH, Van Camp G, Cremers CWRJ. Clinical presentation of the DFNA loci where causative genes have not yet been cloned. DFNA4, DFNA6/14, DFNA7, DFNA16, DFNA20 and DFNA21. Adv Otorhinolaryngol 2002; 61: 98-106.
10. Beisel KW, Rocha-Sanchez SM, Morris KA, Nie L, Feng F, Kachar B, Yamoah EN, Fritzsch B. Differential expression of KCNQ4 in inner hair cells and sensory neurons is the basis of progressive high-frequency hearing loss. J Neurosci 2005; 25: 9285-9293.
11. Van Laer L, Pfister M, Thys S, Vrijens K, Mueller M, Umans L, Serneels L, Van Nassauw L, Kooy F, Smith RJ, Timmermans JP, Van Leuven F, Van Camp G. Mice lacking Dfna5 show a diverging number of cochlear fourth row outer hair cells. Neurobiol Dis 2005; 19: 386-399.
12. Fagerheim T, Nilssen Ø, Raeymaekers P, Brox V, Moum T, Elverland HH, Teig E, Omland HH, Fostad GK, Tranebjærg L. Identification of a new locus for autosomal dominant non-syndromic hearing impairment (DFNA7) in a large Norwegian family. Hum Mol Genet 1996; 5: 1187-1191.

Chapter 5
Discussion

Several areas of active research in NSHI were tackled in this thesis. In Chapter 2, the localization of new NSHI genes and testing of candidate genes within the mapped NSHI loci are described. Five novel loci were mapped in consanguineous families with ARNSHI through linkage analyses using homozygosity mapping. This analysis is based on certain assumptions; these assumptions and the choice of methodology are discussed in the first part of this chapter. Follow-up studies on some of the loci with single-nucleotide polymorphism (SNP) fine-mapping are also presented. The second part of this chapter discusses possible approaches for candidate gene testing, including the criteria for prioritization of candidate genes for sequencing and the sequencing method.

In Chapter 3 we reported novel variants in known NSHI genes and estimated NSHI gene-specific prevalence rates within Pakistan. The methods for determining the functional significance of identified variants are crucial to these studies. Therefore in the third section of the discussion, the various means by which variants are deemed functional or benign are outlined. After deciding which variants are functional, the NSHI gene-specific prevalence rate is estimated (Chapter 3). Section 4 of the discussion relays the different prevalence estimates that were derived in this thesis. The factors that may influence the prevalence rates are discussed.

Genotype-phenotype correlation through the study of audiometric data was described in Chapter 4. However genetic and phenotypic heterogeneity may cause false associations. The importance of careful examination for both syndromic and otologic features is underlined in the fifth section of this chapter. Likewise the verification of specific genotypes as disease-causing is stressed.

At the end of the discussion, future avenues for NSHI research are explored.

IDENTIFICATION OF NOVEL HEARING IMPAIRMENT LOCI

NSHI may be caused by a mutation in any of a great number of genes. In most cases individual NSHI mutations are rare and fully penetrant; a large part of these mutations are recessive. Therefore the mapping of NSHI loci necessitates the ascertainment of consanguineous families that are large enough to independently establish linkage. In Pakistan, there is a long-standing tradition of consanguinity among families, with several generations of first-cousin marriages customarily, which is encouraged to preserve wealth and status within the kindred. At the same time, a history of consanguinity over multiple generations increases the possibility of segregating a rare mutation within the family. The collaboration between the Leal and Ahmad laboratories has been nurtured over the years, resulting in the accrual of families that segregate mostly autosomal recessive NSHI.

Currently, the Leal laboratory has in hand a large collection of Pakistani families (N = 260), of which 14 are *GJB2*-positive, and 203 were submitted for genome scan with approximately 400 microsatellite markers spaced every 10 cM across the autosomes and sex chromosomes. Forty-four families have undergone fine-mapping in order to zoom in on a smaller genetic interval. Most of the fine-mapping was carried out using STR markers with one marker placed every cM. Previously this collection of families has led to the identification of seven novel loci: *DFNB35* [1], *DFNB38* [2], *DFNB39* [3], *DFNB42* [4], *DFNB44* [5], *DFNB45*, and *DFNB46* [6]. In this thesis, five new loci – *DFNB47* [Section 2.1], *DFNB55* [Section 2.2], *DFNB62* [Section 2.3], *DFNB65* [Section 2.4], and *DFNB68* [Section 2.5] – are described.

All the genome scan and fine mapping data for these loci were analyzed using parametric linkage analysis. This type of analysis remains the method of choice in the genetic study of extended pedigrees which segregate NSHI since for the majority of families studied the genetic mode of inheritance is Mendelian. When the underlying genetic model is known, parametric linkage analysis is more powerful than nonparametric linkage analysis. Parametric linkage analysis uses genotypic and phenotypic information from

all informative family members and provides estimates of recombination fractions. Additionally linkage can be tested for in the presence of linkage admixture (genetic heterogeneity) [7].

To run parametric linkage analyses, the disease model, disease allele frequency, disease penetrance and marker allele frequencies must be specified. For all pedigrees in this thesis in which novel loci were mapped, the segregation of HI was compatible with autosomal recessive inheritance. The disease allele frequency for the analysis was set to 0.001, since in most cases the disease allele frequency is expected to be very low. A fully penetrant autosomal recessive inheritance model with no phenocopies was used in the analysis. Careful examination by otolaryngologists from Pakistan decreases the possibility of misclassification of HI due to environmental or non-genetic factors as genetic.

In families wherein genotype data are not available for many founders and some family members, there is a danger that misspecified marker allele frequencies may lead to spurious results, specifically when equal allele frequencies are assumed [8-9]. When available, the allele frequencies that were estimated from families that were genotyped in the same genome scan (*N* families = 34 to 47) were used. The problem arises when fine-mapping STRs are genotyped only for one or two families. In these cases, we performed sensitivity analysis [9] in which marker allele frequencies for the marker that is assumed to segregate with the disease allele were varied between 0.2 and 0.8. This analysis elucidates the presence and extent of false inflation of LOD scores due to incorrect allele frequencies.

For the mapping of NSHI loci, we first carried out two-point linkage analysis using the MLINK program from the FASTLINK computer package [10], followed by multipoint linkage analysis. Two-point linkage analysis has the advantage of being robust to allele frequency misspecification and to the effects of linkage disequilibrium (LD) [8]. On the other hand, multipoint linkage analysis has some attractive features. Firstly, it has potentially higher power: by using linkage information from several markers, it is sometimes possible to establish linkage when linkage could not be established using a single marker. Secondly, it has better location accuracy such that it can better refine the genetic region for a NSHI locus. Because of more accurate localization, it can therefore facilitate the exclusion of other NSHI loci when two-point analysis cannot.

In linkage analysis programs, there is a trade-off between the size of the pedigree and the number of marker loci which can be analyzed for multipoint analysis. The FASTLINK, LINKAGE [8, 10] and VITESSE

[11] programs use the Elston-Stewart algorithm, which produces exact LOD scores. Although this algorithm can handle large pedigrees, it can only handle a limited number of consanguinity loops and marker loci for multipoint analysis. The number of marker loci which can be handled for multipoint analysis is dependent on the number of alleles at each locus to be analyzed. Generally programs cannot handle a maxhap greater than 500, with maxhap being the product of the number of alleles at each locus. For STR markers that have many alleles, this limit is reached with only a few loci included in the analysis. Thus for programs that employ the Elston-Stewart algorithm, large pedigrees are accommodated but only small numbers of ($\leq$8) marker loci can be analyzed.

For programs which use the Lander-Green algorithm (e.g. GENEHUNTER [12], MERLIN [13], ALLEGRO [14]), thousands of marker loci can be theoretically analyzed when carrying out multipoint analysis but the analysis is constrained by pedigree size. The pedigree size is measured by the maxbit, which is computed as $2n - f$, where n is the number of non-founders, and f is the number of founders. Thus the limiting factor is the number of non-founders within the pedigree. GENEHUNTER is generally limited to pedigrees with $\leq$24 bits. MERLIN can deal with larger pedigrees, but is usually limited to a maxbit of about 28-30 when multiple pedigree members with missing genotypic information are present. The original version of ALLEGRO is faster than GENEHUNTER and can handle pedigrees with up to approximately 28 bits. The reason for the different number of maxbits that these programs can handle is due to different techniques used to optimize the Lander-Green algorithm.

SIMWALK2, which uses the Markov chain-Monte Carlo method [15], can handle many loci and large pedigrees of a few hundred members and dozens of loops. However the program slows down very much in a case when a very dense marker map is studied. Also SIMWALK2 produces only approximate results.

Given the different programs that are available and the size and complexity of the families under study, we performed multipoint analysis using the ALLEGRO program.

Although in this thesis it was possible to analyze all of the pedigrees in their entirety when carrying out two-point analysis using MLINK of the FASTLINK/LINKAGE computer program, some of the large pedigrees needed to be split into two or more parts when multipoint linkage analysis was performed with ALLEGRO. When the pedigree is split in two or more parts, the resulting LOD scores from each pedigree section must be

summed. The disadvantage of splitting the pedigrees when carrying out linkage analysis is that there is a potential reduction of linkage information, i.e. LOD score is underestimated. When a pedigree must be split due to size, the results obtained from ALLEGRO can be compared to those obtained from FASTLINK and SIMWALK2. With the release of a newer version ALLEGRO2 [16], which promises to accommodate families with as many as 39 bits, larger families can be analyzed in their entirety.

One concern with the analyses of highly consanguineous families is the possibility of hidden/unknown consanguinity. Because these families usually exist within close communities, the pedigree structure may be much more inbred than what the family members can remember in the absence of detailed records that go back through centuries. If families are much more inbred than is reported, underestimation of relationship coefficients can increase both type I and II error [17]. Recently a program was designed to estimate inbreeding coefficients from genomic marker data [18]. However the current requirements for reliability of results from this program are genome-wide marker genotypes at a density of ≤ 5 cM and accurate marker allele frequency estimates from a control population, both of which are not available for the families described in this thesis. Thus at present there are no feasible measures for hidden consanguinity in the families under study.

Like most methods for multipoint analysis, ALLEGRO assumes inter-marker linkage equilibrium, which may result in erroneously inflated LOD scores specifically with missing parental data [19]. When there is LD between markers, the haplotype frequencies cease to be the product of the allele frequencies, thus increasing type I error rates if linkage equilibrium is incorrectly assumed. Recently the MERLIN program [13] has a newly added feature that allows for inter-marker LD. However correct specification of haplotype frequencies is required for the analysis. This cannot be provided with our current data because only a few families are genotyped for fine mapping markers within each locus. In any case our multipoint LOD scores for our NSHI loci were significant (range between 3.3 and 5.3), so the existence of a NSHI gene within the locus interval cannot be denied.

For haplotype reconstruction, GENEHUNTER, MERLIN and SIMWALK2 are the most commonly used programs. SIMWALK2 was utilized in the analyses of the loci that are reported in this thesis. Although none of the methods existing can completely resolve haplotype ambiguity [20], in contrast to the other programs, SIMWALK2 gives an indication if during haplotype assignment multiple heterozygosity occurs, so that the user

is aware if the phase cannot be determined and haplotypes cannot be assigned with certainty. There is a danger of haplotype misassignment when small families and many SNP loci with high heterozygosity are involved [20]. In comparison to SNPs, STRs, which were used in our studies involving extended pedigree structures, have much lower misassignment rates.

Table 5.1 Loci that were fine-mapped using SNPs

Locus	Chr.	Old interval with STRs only			New interval with SNPs and STRs		
		Max. LOD	Length (cM)	Length (Mb)	Max. LOD	Length (cM)	Length (Mb)
DFNB35	14q	5.3	11.19	7.85	6.0	10.18	6.91
DFNB43/48	15q	4.6	6.33	6.91	4.4	1.66	1.59
DFNB46	18p	3.8	16.90	5.13	4.8	16.90	5.59
DFNB47	2p	4.7	13.23	5.29	4.7	0.54	0.12

SNP, single nucleotide polymorphism; *Chr.,* chromosomal region; *STR,* short tandem repeat.

Recent studies compared the informativeness of dense maps of SNP markers to maps of STR markers in the context of whole-genome linkage analysis [21]. SNPs are biallelic, have lower heterozygosity and carry less information than STRs. However they are: less prone to genotyping error; have much greater density across the genome and may therefore result in improved localization; and SNP genotyping can be more automated than genotyping of STR markers, thus making them more cost-effective to genotype than STR markers. Assuming an accurate genetic map and linkage equilibrium, it has been shown that, when parental genotypes are unavailable, the increased density provided by SNP genotyping can increase linkage information by as much as 20% [21]. Table 5.1 presents follow-up studies on four loci that were previously mapped by the Leal lab using STRs only (for genome scan and fine-mapping) and then reanalyzed using SNPs combined with STRs (4-22 markers per cM). After completing a genome scan and fine mapping with STR marker loci, the SNP marker loci were genotyped in order to try to reduce the size of the genetic interval for the NSHI loci and thus the number of candidate genes to be sequenced. However, because of the density of markers and the assumption of linkage equilibrium by ALLEGRO, the effects of inter-marker LD cannot be completely ruled out.

To obtain accurate results in multipoint linkage analysis, it is essential that the correct order of markers and precise genetic inter-marker distances are specified. After the completion of the human genome project [22], the physical map position and order of markers are publicly available from National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) and UCSC Genome Bioinformatics (http://genome.cse.ucsc.edu/). Recently, the Rutgers combined linkage-physical map of the human genome was released [23]. Since this map was constructed using both genetic and physical map data, it is the most accurate map available. This genetic map has information on the genetic and physical position for 2,889 marker loci from Build 34 of the human reference sequence, of which ~28% are STR markers. Such a map with a high degree of accuracy in terms of the order of markers allows for more reliable linkage analysis.

IDENTIFICATION OF HEARING IMPAIRMENT GENES

The positional candidate gene approach that is used for gene identification involves choosing a gene within a locus interval based on the most currently annotated physical map of known genes (e.g. recent build of the human reference sequence as viewed by the UCSC Genome Browser). Positional candidate genes are prioritized based on the following criteria: (1) genes that were identified to cause syndromes with a deafness component; (2) genes that belong to the same family as other NSHI or syndromic HI genes; (3) genes that are expressed in the inner ear [24-25]; (4) genes that were naturally mutated or knocked out in animals that had hearing or vestibular defects; and lastly, (5) genes that may have a neurosensory function.

After the selection of a candidate gene for NSHI, direct sequencing is the preferred method among numerous methods of detecting mutations because it has the best sensitivity and specificity rates. Also due to the capabilities and cost ($0.20 per reaction) of sequencing within the Human Genome Sequencing Center at Baylor College of Medicine, it is definitely more cost-effective. Only the exons and the promoter region are sequenced, so a regulatory intragenic or an intronic functional variant may be missed, although such an occurrence is probably rare. One unaffected and two affected family members are first tested for variants in a candidate gene within the linked genetic interval. Once a potentially functional sequence variant is detected and found to segregate with HI status of the tested family members, the DNA from the other pedigree members is sequenced with the same primer pairs with which the variant was first identified. If the

sequence variant segregates in those family members who have HI, then 100-150 unaffected controls from an ethno-linguistically matched population are screened with the same primer pairs in order to help determine if this variant is responsible for the NSHI. Aside from comparison of allele frequencies, bioinformatics tools from publicly available databases and software are also used to assess the functional significance of a variant. The different methods of determining functionality and prevalence of functional variants are described in the next section. If through these steps there are evidences that a genetic variant is responsible for NSHI within a family, collaboration with other laboratories is in place in order to test the expression of the corresponding protein within animal inner ear systems.

Through ascertainment of more families, further fine mapping and direct sequencing, it is hoped that the HI genes in the reported loci will be identified in the near future.

STUDY OF NOVEL VARIANTS IN HEARING IMPAIRMENT GENES AND THEIR FUNCTIONAL SIGNIFICANCE

After the identification of a sequence variant which segregates with NSHI within a family, two further steps in the course of the study are performed: one is the determination of the functionality of identified sequence variants; the other is the study of prevalence rates of functional NSHI gene variants within hearing-impaired populations. Both prevalence and functionality are important for prioritization of sequence variants for screening in specific populations. Before any reliable prevalence estimate for HI due to a specific gene can be made, the pathogenicity of variants must be determined, otherwise, the inclusion of benign variants can falsely inflate prevalence rates. In this thesis, only six of 10 variants found in *GJB2* were labeled as functional through the use of prediction software and allele frequency comparisons [Section 3.1]. Using a similar strategy, five novel functional variants were identified in *TMC1*, and two in *TMIE* [Sections 3.2-3.3]. Though performance of functional experiments are beyond the scope of this study, the strong evidence that functional studies provide to support pathogenicity must always be included in examining specific sequence variants. On the reverse side, when studying the effects of variants in inner ear proteins, it is the more prevalent variants that are subject to scrutiny, and it is in this indirect way that prevalence studies contribute to knowledge of molecular biology. Discussed in this section are the different ways by which investigators currently judge whether an identified sequence variant is pathogenic or not, with the *GJB2* gene used as an example for each strategy.

COMPARISON OF ALLELE FREQUENCIES OF VARIANTS IN CASES AND IN CONTROLS

It is common practice that the allele frequency of an identified sequence variant is checked not only in cases but also in controls (at least 100 individuals with normal hearing or 200 control chromosomes). Since most NSHI genes known to date are fully penetrant, it is expected that variants of autosomal dominant inheritance will not be found among control chromosomes. On the other hand, variants in ARNSHI genes would be found rarely in the heterozygous state and never in the homozygous state. The expectation of a higher allele frequency of a functional variant among cases may be tested statistically by chi-square tests, Fisher exact tests, likelihood ratio tests, or by comparison of 95% confidence intervals on the estimates of allele frequencies among cases versus that of controls.

There are some caveats to using comparison of allele frequencies among cases and controls. Firstly, the selection of the control population may not be correct, rendering any comparison between cases and controls invalid. This is obvious if the cases come from a different ethnic population as the controls (e.g. East Asian vs. European Caucasian). However more subtle errors occur when ethnic homogeneity is assumed to occur due to non-awareness of the sample population's ethnic diversity. To illustrate, Pakistan is divided into four provinces – Punjab, Sindh, Northwestern Frontier Province (NWFP) and Federally Administered Tribal Areas (FATA), and Balochistan (http://www.infopak.gov.pk/public/country_profile_index.htm) – which roughly represent the country's major ethnic groups. More closely related to ethnic origin, though, is the division of the population in terms of linguistic group: although Urdu is the official language, most people in Punjab province speak either Punjabi or Sairiki; in Sindh, Sindi; in Balochistan, Pushto; and in Azad Jammu and Kashmir, Kashmiri. A smaller percentage of the population speaks any of the minority languages, for example, Hindko in NWFP. Thus if a sample from the Sairiki-speaking population is compared to a control population labeled generically as "Pakistani", a benign variant within the Sairiki population may not be identified simply because that variant is rarer in the other subpopulations and will thus have a lower frequency among the screened controls. Therefore it is recommended that the controls be as closely matched to the cases as possible, e.g. by using both ethnic and linguistic information if available.

Another problem is the phenotypic variability of the NSHI trait. In most cases of ADNSHI, the HI is postlingual in onset, and may occur even late in adulthood. Thus depending on the age at which controls are screened, a

carrier may be misclassified as a control individual. For ARNSHI, which is mostly prelingual, the chance of this occurrence is minimal if adult control individuals are selected.

To date, more than 100 sequence variants have been identified in the *GJB2* gene, which is the most studied genetic determinant of NSHI in humans. Upon review of *GJB2* studies in various populations that had both patient and control frequencies published [26-33], in the case of missense substitutions, the decision if a substitution is benign may be based on non-significant differences between case and control allele frequencies (Table 5.2). However, the delineation of functional variants based on allele frequencies may not be as informative. Sometimes it is difficult to tell whether the allele frequencies among controls are low because the variant is functional, or rather the variant is simply rare in the study population. As can be seen in Table 5.2, the 35delG variant, which has been established as a common cause of prelingual HI among Caucasians and Ashkenazi Jews [34-35], has extremely low prevalence rates among patient alleles in the Chinese population and, given the sample size used in the study, a significant difference in the allele frequency between cases and controls could not be detected. On the other hand, in the French population where the 35delG mutation has a high prevalence, the difference in allele frequencies between patients and controls was significant. From this example, it can be shown that comparison of allele frequencies between cases and controls helps to determine whether a certain sequence variant is a benign polymorphism, but may require a large sample for both patients and controls to produce reliable statistical inference. Also results may be equivocal for functional variants, particularly when the allele frequency is low in the population.

FUNCTIONAL ANALYSES USING ANIMAL MODELS OR CELL SYSTEMS

In most reports of gene isolation, the functionality of the encoded protein is usually tested by using cell expression systems (i.e. to determine if the mRNA of the isolated gene is expressed in the inner ear of an animal model or in human tissue) or studying the phenotype of animal mutants / knockouts (e.g. mouse, rat, gerbil), either by histology or audiology (Preyer reflex, auditory brainstem response). When a novel variant is identified in a known gene, investigators have to be more creative in designing experiments that are specific to the gene or protein function. For example, for transmembrane transporters or ion channels, in vitro expression studies may be coupled with intercellular dye transfers or voltage measurements.

Table 5.2 Comparison of allele frequencies of selected *GJB2* variants among patients vs. controls

GJB2 variant	Known functionality	*N* cases	Allele frequencies among cases (95% CI)	*N* controls	Allele frequencies among controls (95% CI)	Reference population
35delG	Functional	262	0.008 (0.0009-0.03)	200	0 (0-0.01)	Chinese [26]
35delG	Functional	318	0.12 (0.09-0.16)	12586	0.008 (0.006-0.01)	French [27]
V27I	Benign	332	0.10 (0.07-0.13)	410	0.13 (0.10-0.17)	Thai [28]
I30V	Benign	648	0.002 (0.00004-0.01)	864	0 (0-0.004)	Taiwanese [29]
235delC	Functional	262	0.08 (0.05-0.12)	200	0.005 (0.0001-0.03)	Chinese [26]
V153I	Benign	430	0.04 (0.02-0.06)	120	0.06 (0.02-0.12)	Indian [30]
F191L	Benign	648	0.003 (0.00008-0.02)	864	0.001 (0.00003-0.01)	Taiwanese [29]
I203T	Benign	332	0.02 (0.01-0.05)	410	0.05 (0.03-0.07)	Thai [28]

For both cases and controls, *N* denotes the number of chromosomes tested.

Table 5.3 Effects of selected *GJB2* variants based on functional studies

GJB2 variant	Functional effect	Reference
M1V	No membranous localization and absent intercellular coupling	[36]
35delG	No protein expression detected	[37]
235delC	Loss of targeting activity to plasma membrane with defective oligomerization and assembly	[38]
L90P	Absent intercellular coupling and partly disturbed hemichannel assembly	[36-37]
E114G	No effect	[38]
R127H	Variable: may have no effect or may cause low protein expression and very low intercellular coupling	[36-37, 39]
F161S	Very weak membranous localization and absent intercellular coupling	[36]
P173R	No membranous localization and absent intercellular coupling	[36]
R184P	No membranous localization, absent intercellular coupling and no hemichannel formation	[36-38]

Because the *GJB2* (*CX26*) protein functions as a building block for a transmembrane channel, immunolocalization and dye transfer experiments have been done to test specific variants [36-39]. These tests were done to determine which of the following requirements for proper *GJB2* function are affected by *GJB2* sequence variants: (1) protein stability and/or accumulation; (2) proper targeting or trafficking from the endoplasmic reticulum or Golgi apparatus to the plasma membrane; and (3) competent hemichannel formation and intercellular coupling to permit transfer of ions, metabolites and second messengers [37]. Table 5.3 enumerates the effect of several *GJB2* variants on function based on functional studies.

The assumption in these studies is the ability to transpose results in animal or recombinant expression models to human systems, which might not always be true. Also such studies may be time-consuming, labor-intensive and costly. Nevertheless, when performed well, they provide good insights into protein mechanisms which may include protein-protein interactions, domain roles, or cellular pathways. On the molecular biologic and tissue-specific levels, functional analyses add the most to knowledge of inner ear pathophysiology.

STUDY OF PROTEIN STRUCTURE THROUGH BIOINFORMATICS

One major stumbling block in the analysis of functionality of sequence variants is the scarcity of functional studies that support the roles of protein products within the inner ear. At present, there are many databases and software that are readily available to predict the possible functional effect of a specific variant. Homologous sequences from a growing number of completed genome projects for different species, known protein structures from databanks, and prediction software all contribute to the assessment of the importance of specific amino acid residues to function. This is based on the premise that sequences are conserved through evolution because they are functionally important, and so if a specific amino acid residue or genomic sequence is minimally changed across divergent species or within protein families, it is predicted that the conserved sequence or motif has a bearing on function. Prediction software include: multiple sequence alignment for evolutionary conservation; prediction of location of transmembrane domains, transcription factor binding sites, alternative splice sites, and known motifs or patterns; and secondary and tertiary protein structure analysis. It should be stated that the predictions are limited to published information, and accuracy may vary according to software design, thus the results should always be taken in context.

To illustrate, the human *CX26* (*GJB2*) protein sequence, which is composed of cytoplasmic amino- and carboxyl-terminals, and four transmembrane (TM), one intracellular and two extracellular (EC) domains, was aligned with the sequences of seven other human connexin proteins. It was shown in this analysis that transmembrane domains TM1 and TM2 were most conserved, and extracellular segments EC1 and EC2 and the amino terminal were more conserved than TM3 and TM4 (data not shown). The alignment was then repeated using connexin protein sequences from 23 species that were most similar to human *CX26*, which then resulted in a combination of different connexins (mostly sequences homologous to *CX26*) in different species. The latter analysis demonstrated that the degree of conservation was higher in the two extracellular regions than in the transmembrane domains (data not shown). This is supported by results with the TMHMM program [40], in which the four transmembrane regions predicted for *CX26* was relatively stable across species, but TM3 and TM4 were variable when compared to other connexins. These results taken together may imply that TM1 and TM2, being most conserved in the alignment of human connexins, are probably basic structural components for connexins in general. On the other hand the extracellular segments are most conserved in multiple alignments across species, which may suggest that they are defining regions for *CX26* as compared to other connexins, and might have a role in substrate specificity. The occurrence of variants in the conserved regions (TM1, TM2, EC1, EC2) provides clues to their functional effect, but given the existence of multiple variants in several specific amino acid residues within *GJB2* (e.g. Q80K-Q80L-Q80R; L90P-L90V), the type of amino acid change must also be taken into account. Also for each variant, the allele frequencies in patients and controls, the results of functional experiments and the audiologic phenotype must be reviewed, and may require a decision paradigm that can settle conflicting results.

ESTIMATION OF NSHI GENE-SPECIFIC PREVALENCE RATES

Validation of the existence of functional variants within populations is important for an effective genetic screening paradigm for NSHI is to be developed. In addition there are probably many sequence variants in any NSHI gene, and each variant has its own prevalence within each study population, such that each population carries a different spectrum of variants for each gene. Gene-specific prevalence estimation among individuals with ARNSHI has been possible with the larger subset of families from Pakistan, and was done for three genes, *GJB2*, *TMC1* and *TMIE* [Chapter 3].

Mutations in the *GJB2* gene are said to be the most common cause of ARNSHI, with as much as 50% of ARNSHI attributable to *GJB2* in certain populations [34-35]. Thus it is common to first screen for functional variants in the *GJB2* gene as a first tier for genetic analysis, both in the clinical setting and in genetic studies prior to carrying out a genome scan. In the Pakistani population, however, the *GJB2* prevalence is lower at 6.1% [Section 3.1]. After the publication of our *GJB2* article, two out of 64 additional Pakistani families were found to be *GJB2*-positive, which further brings down the prevalence to 5.4% (95%CI: 3.0, 8.9). This is lower than the *GJB2* prevalence rate of 9.5% (95%CI: 6.0, 14.1) among Dutch patients who are homozygous or compound heterozygous for functional alleles in *GJB2* but not *GJB6* [Section 4.1]. There is overlap of the 95% confidence intervals between the Pakistani and Dutch populations, which might be due to the small sample sizes and/or difference in ascertainment (i.e. an expectation of higher prevalence rates in mostly consanguineous multiplex families with ARNSHI among Pakistanis than in Dutch patients which were mostly sporadic). Four additional Dutch patients were compound heterozygous for a *GJB2* variant plus del(*GJB6*-D13S1830) [Section 4.1]. Therefore the Dutch *GJB6* allele frequency of 0.9% (95%CI: 0.2, 2.3) is closer to the rates found in Belgium (1.4%) and is lower than the *GJB6* prevalence in Mediterranean Europe [41-42].

Two other genes, *TMC1* and *TMIE*, also had low prevalence rates in the Pakistani population, with 4.4% and 1.7% respectively [Sections 3.2-3.3]. Functional variants in *GJB2* were identified in this population through direct sequencing of at least two hearing-impaired individuals from each family. On the other hand, for *TMC1* and *TMIE*, the families first underwent a genome scan, then for those families for which there was either established or suggestive linkage, two family members with NSHI were screened for potentially functional variants via sequencing. Of 184 Pakistani families in the current registry that underwent genome scan, 66 (36%) are linked (LOD $\geq$3.0) to regions containing one of 15 previously identified NSHI genes.

Several hypotheses have been put forward to explain the high prevalence of HI in populations, such as:

- consanguinity [43-44];
- higher rates of intermarriage among the deaf, which is also related to group formation (e.g. schools for the deaf) and relaxation of negative selection against hearing-impaired people through time [45];
- positive selection through biological phenomena, such as better skin phe-

notype (e.g. increased thickness of epidermis) among *GJB2* carriers [46];
- random drift [47]; and
- population substructure (e.g. based on ethnicity and not due to deafness) [48].

Several studies have shown an association between consanguinity and increased incidence of HI within the population [43-44]. Contrary to expectations, consanguinity was associated with decreased *GJB2* prevalence in Turkish probands [49] and was therefore considered a possibility also within the Pakistani population [Chapter 3.1]. To test this hypothesis, the consanguinity rates in different countries were plotted against *GJB2* allele frequencies (with all potentially functional variants combined) among patient populations (Figure 5.1). The *GJB2* gene is by far the most extensively researched cause of NSHI worldwide, hence the availability of prevalence data across populations. Using data from 25 countries (8 European, 6 Middle Eastern, 3 East Asian, 3 African, 2 South American, plus India, Pakistan and Australia) [27, 30, 32-33, 49-69], the correlation coefficient (r) between consanguinity and *GJB2* allele frequencies was -0.42 (p <0.05). This may imply that there is indeed a trend in which countries with higher consanguinity rates have lower *GJB2* prevalence rates. However when the analysis was repeated after transforming the values to a logarithmic scale (Figure 5.2), the association became non-significant (r = -0.37, p = 0.08). One limitation of this analysis is the sampling of the patient populations, which are not uniform across the various countries in terms of ascertainment, method of screening and common gene variants. Similarly the consanguinity rates were derived from different studies conducted at different time periods and within populations that may not exactly match the patient group. Thus there is currently no strong evidence to support the relation between increased consanguinity rates and decreased *GJB2* prevalence. Similarly, there are not enough prevalence studies on other non-*GJB2* genes that cause NSHI, therefore any statement that relates consanguinity rates to the prevalence of HI, specifically if based solely on *GJB2* prevalence rates, cannot be relied upon. Additionally there may be other important population-specific factors (e.g. population history) that must be considered when examining trends in gene-specific prevalence rates.

Intermarriage among the deaf is unlikely as a cause of an increase in HI prevalence in Pakistan because of the low occurrence of deaf-by-deaf matings in the country. In our subset of families, in the rare instances when deaf-by-deaf matings occur, the assortative mating occurred not due to deafness but primarily due to being from the same family (in four of five

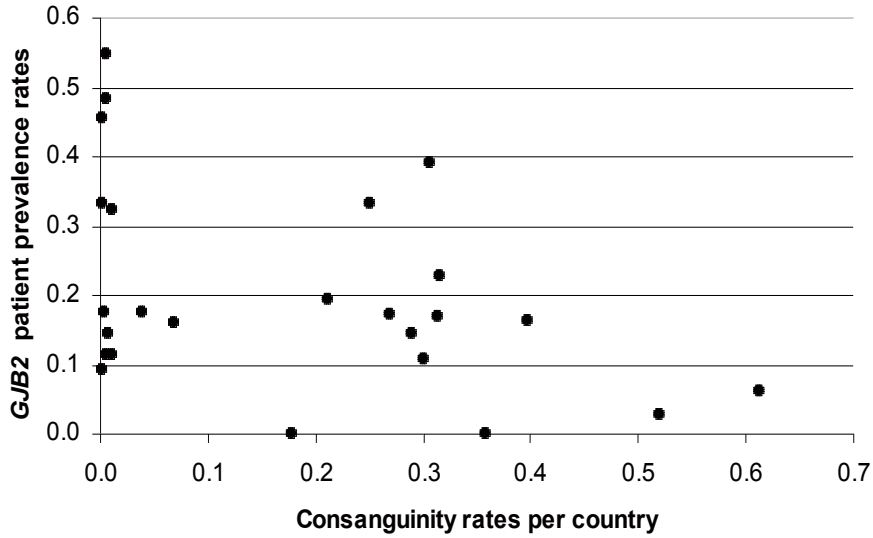


Figure 5.1 Consanguinity and GJB2 prevalence in patient populations

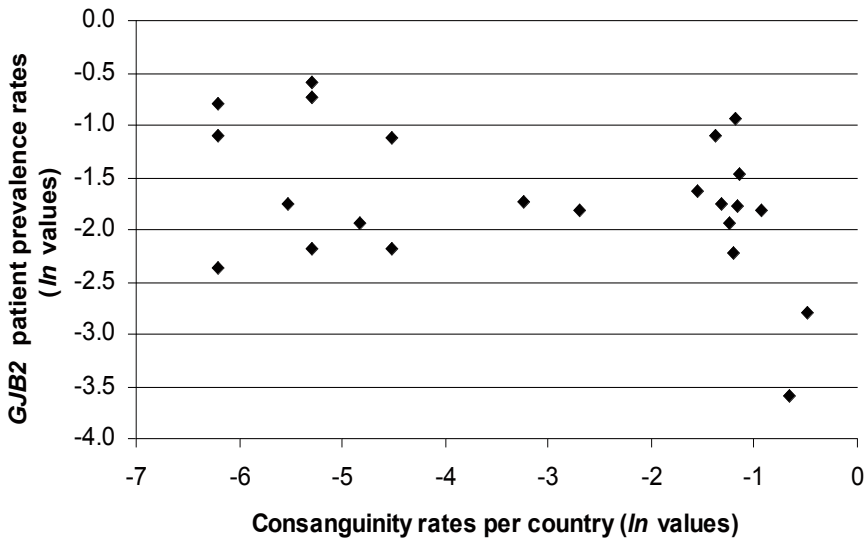


Figure 5.2 Consanguinity and GJB2 prevalence in patient populations after logarithmic (ln) transformation

families, deaf-by-deaf matings were between closely related individuals). Preliminary analyses using *F*-statistics in this collection of families showed that the contribution of population substructure based on provincial region and language group, if ever present, would be minimal ($F(ST) < 0.05$), and it is expected that the excess homozygosity in this sample is largely due to inbreeding. Between positive selection and random drift, drift seems to be a more plausible explanation for the lower prevalence rates in Pakistan.

ANALYSIS OF HEARING IMPAIRMENT PHENOTYPES

In general, ARNSHI is associated with prelingual, sensorineural, non-progressive, moderate to profound HI affecting all frequencies, with a flat or gently downsloping audiometric shape [70]. On the other hand ADNSHI has more locus variability, although in most cases it is progressive and post-lingual in onset. Due to interlocus and even intralocus variability, detailed phenotypic description of HI is required to facilitate definition of locus- or gene-specific audiometric profiles. In one study, pooling families with the same audiometric profile facilitated gene mapping and identification [71].

For HI, the audiogram is the most widely used measure that guides the clinician for diagnosis and management. Because the standard test is performed for two ears per individual, is repeated over several frequencies and may be done at different time points, several HI parameters may be studied, which includes severity, audiometric shape, symmetry, age at onset, and progression over time. Of these parameters, severity is of immediate clinical interest, and may be correlated with genotype. In a study among Dutch patients, it was shown that the ARNSHI due to *GJB2* was commonly bilateral, flat or sloping in audiometric shape, variable in severity depending on genotype, and non-progressive, although there was one case of progression (0.8-2.2 dB/year) at the lower frequencies for a patient with the [c.299A>C + c.416G>A] genotype [Section 3.1]. On the other hand, in a Swiss-German family with ADNSHI due to *DFNA24*, the phenotype was bilateral, sensorineural, prelingual and progressive. The down-sloping audiometric profile showed moderate high-frequency HI at birth, which steadily progresses to moderate HI at the lower frequencies, severe HI at mid-frequencies, and profound HI at high frequencies [Section 3.2].

Depending on the number of individuals and families that carry a particular genotype for an NSHI gene, genotype-phenotype correlation may be done across individuals in a specific large family consisting of several branches and generations [Section 3.2] or across families [Section 3.1]. If done in single large families segregating a specific NSHI locus, there is

some worry that the phenotype may be a product of a gene-gene or gene-environment interaction that is family-specific, or that there is genetic or phenotypic heterogeneity within the family. Should genetic/phenotypic heterogeneity exist among different branches or generations within one family, deeper statistical analysis and further investigation of family history and clinical features are required. This is particularly true in cases when the gene identified as causative of the HI also causes a syndrome. There have been several occasions in which previous reports of NSHI cases were later proven to have syndromes [72], especially in cases where the other syndromic features are expressed much later in life (e.g. retinitis pigmentosa in Usher syndrome), therefore comprehensive phenotype characterization cannot be overemphasized. Moreover, genetic/phenotypic heterogeneity within a sample can decrease power to detect linkage, mislead as to the true location of the causative HI gene, and underestimate prevalence estimates; in such cases, individuals with unclear phenotypes must be removed from analysis.

Often there is a very limited number of individuals or families with one particular genotype, so the genotypes may be grouped or classified according to type of mutation, e.g. by specific combinations of "inactivating" or "radical" mutations (deletions, insertions, nonsense and splice site mutations) and/or "non-inactivating" mutations or missense substitutions. Studies have shown that the presence of at least one allele with a missense substitution in individuals with biallelic *GJB2* mutations had milder HI compared to those with inactivating/radical mutations [Section 3.1; ref. 73].

Recently a large multi-center study was completed with a roster of *GJB2* genotypes associated with a specific audiometric profile [74]; this study includes the Dutch sample described in Section 3.1. One limitation of this study, however, is the inadvertent inclusion of patients who were homozygous or compound heterozygous for *GJB2* variants with questionable functionality (e.g. c.107G>A (p.V37I)). Because *GJB2* screening is the first step for genetic HI diagnosis in most centers, given the genetic heterogeneity of HI, there is a danger that a clinician may conclude that a *GJB2* variant is the culprit for the HI and stop further testing, while the HI is actually caused by another variant or another gene. Aside from genetic heterogeneity, there may also be phenotypic heterogeneity, such that individuals who were homozygous for 35delG (N = 889), though having a median audiometric threshold of 102 dB, have a wide range of HI (~25-120dB) [74]. This further opens the question of other possible genetic or environmental influences for the *GJB2* genotype.

FUTURE TRENDS IN GENETIC HEARING IMPAIRMENT RESEARCH

Gene-gene and Gene-environment Interaction

With multiple NSHI genes, it is not hard to imagine that there may be some interaction between these genes that could produce specific phenotypes. At present two modifier genes have been published [72]. These modifier genes were shown to prevent the development of HI or cause a milder phenotype despite the occurrence of the disease haplotype/variant in an individual. Several genes have also been identified to have digenic inheritance, including: *GJB2* and *GJB6* [75], and *PCDH15* and *CDH23* [76]. As to gene-environment interaction, the 1555A>G mitochondrial mutation has been associated with aminoglycoside-induced HI, which might also be modulated by *GJB2* mutations [77-78]. Other exposures that cause HI such as noise and viral infections are being actively investigated for susceptibility-inducing polymorphisms [79]. Though it is commonly assumed that NSHI is a single gene disorder with full penetrance, it is not surprising if it is later found that the trait is much more complex than previously thought.

Decision Paradigms for Genetic HI Screening and Functionality of Variants

In the course of this study, the first line of screening occurs with *GJB2* sequencing for both AD and AR NSHI families. This is a strategy that has been adopted in many clinical and research institutes, and has been helpful because the diagnosis of *GJB2* results in less laboratory work-up for syndromes. It has been demonstrated that *GJB2* patients do very well with cochlear implantation because the defect lies only within the cochlea and does not affect the auditory nerve [80-81]. However, if a patient has NSHI and is negative for *GJB2* screening, how to proceed with the genetic diagnosis is not yet established. Some genes have as much as 50 exons, with more than 100 variants spread out over the exons, so it may not be cost-effective to sequence the whole gene. In a research lab with the capability for linkage analysis, if the family is large enough it is possible to carry out a genome scan and fine-mapping to exclude the known NSHI genes as the cause of HI. Based on the experience of our lab with *TMC1* and *TMIE*, and given that 83% of the families at hand can establish linkage (LOD >3.0), criteria were introduced in order to decide if re-sequencing of an NSHI gene should be pursued (Table 5.4). Our lab is fortunate enough to have fast analytical capabilities and very low sequencing costs, but in the clinical setting and in other labs that have less resources, an effective paradigm for

NSHI genetic screening that includes as many genes as possible is yet to be built.

Another decision paradigm that has to be designed is that which shows if a variant is functional or benign. As discussed in the section above, there are several ways by which researchers settle this question, but if there are conflicting results, the interpretation will be subject to much debate. The Polyphen website has shown much promise to tackle this problem [82]. However as we have shown with the *TMC1* study [Section 3.2], the strategy employed by Polyphen is not as effective if the protein has not been well-studied.

In tandem with studies of variants, genotype-phenotype correlations will contribute much to the development of genetic screening algorithms. This is of particular interest when faced with positive results for genes that cause postlingual HI, and thus the clinical interventions are more preventative in nature. For both congenital and postlingual HI, genetic counseling will depend heavily on known audiometric profiles for the gene in question. Among neonates, Universal Newborn Hearing Screening is conducted via testing of otoacoustic emissions (OAE), auditory brainstem response (ABR), or a combination of both, after which the family might be counseled on the possibility of genetic causes of HI. As of yet, most genotype-phenotype correlation studies utilize audiometric data. The effects of genetic variants on OAEs and ABRs are largely unexplored.

Table 5.4 Criteria used to determine which families will be re-sequenced for known NSHI genes

LOD at NSHI gene region	LOD at other genomic regions	Re-sequence NSHI gene?
< 1.2	Any	No
≥ 1.2	<3.0	Yes
1.2 ≤ LOD < 2.0	3.0 ≤ LOD < 4.0	No
≥ 2.0	3.0 ≤ LOD < 4.0	Yes
< 3.0	≥ 4.0	No
≥ 3.0	Any	Yes

LOD scores are based on maximum LOD scores obtained from linkage analyses. Note that only families with a minimum maximum LOD score of ≥1.2 based on simulation studies were submitted for genome scan. After linkage analysis of genome scan genotypes, if a two-point LOD score of ≥1.2 (nominal p-value ≤0.01) is obtained from a known NSHI gene region, but the multipoint LOD score is <1.0 across the region, sequencing will not be performed. Although it has been established that for an infinitely dense map of marker loci, LOD ≥3.3 denotes a genome-wide *p*-value of 0.05 [83], the criterion has been relaxed to LOD ≥3.0 [84]. This is to prevent accidental exclusion of known NSHI genes for sequencing, particularly in families with reduced power to detect linkage due to pedigree structure and limited DNA availability.

Mass Screening of HI Variants through Gene Chips and Microarrays

One possible way to simplify genetic HI screening is to develop HI gene chips and microarrays, particularly for newborns, in which a single blood spot can be used to screen thousands of HI variants across multiple genes [85]. The technology for this is already under development. The main question then is which genes and variants to screen for. If this is not addressed, then the sensitivity and specificity of the screen will be low. This goes back to the need for prevalence studies across populations and functional studies for variants. For genes in which there are so many mutations that have been found within the whole length of the sequence (e.g. *GJB2*), direct sequencing is still the most cost-effective screening method. Eventually technology might allow the screening of all variants in all NSHI genes with just one gene chip. Until then screening for specific functional variants in genes with high prevalence rates may be more efficient. Because various populations tend to have different allelic frequencies and spectra for each gene, there may be a need to design screening methods that are population-specific.

Gene Therapy

As the molecular mechanisms for the inner ear are unraveled with the continuing discovery of NSHI genes, genetic therapy, through the introduction of gene products into the inner ear, will be eventually available. Because the inner ear is a closed system, injection of gene products through the round window is theoretically possible. If mRNA products or growth factors will be effective in restoring certain inner ear functions, an effective vector must be developed. Many viruses have detrimental effects to the sensory epithelium and the auditory nerve, so viral vectors that combine targeted long-term expression of transgenes and minimal cytotoxicity have been created [86]. The use of stem cells is also being studied [87], and the identification of growth factors that stimulate these stem cells and also of genes that contribute to hair cell development is an exciting field of research.

Genetic Counseling

The increase in the sophistication of knowledge regarding NSHI has a tremendous impact on the clinical management of patients and genetic counseling of families. Genetic testing for NSHI has the following known benefits: (1) the avoidance of unnecessary and costly testing; (2) the accurate estimation of risk recurrence of NSHI within the family; (3) the dismissal of incorrect notions of the true cause of HI; and (4) prognostication in terms of expected degree and progression of HI and habilitation outcomes [88].

Currently many clinics in developed countries provide routine testing for *GJB2*, and in some, for *WFS1* and *SLC26A4* [89]. One possible difficulty is that, given a negative test result, it is still not established that the HI is non-genetic. This is mainly due to the current limitations of knowledge about NSHI and the trait's extensive genotypic and phenotypic heterogeneity. As the number of known NSHI genes and gene variants increase and the audioprofile for each gene and/or variant is described, genetic counseling becomes more case-specific and serviceable for both families and clinical personnel involved. Through collaborative efforts throughout the HI research and clinical community, it is hoped that a comprehensive genetic testing program for HI that incorporates advances in both knowledge and technology can be designed for implementation across different ethnic populations around the globe.

REFERENCES

1. Ansar M, Amin ud Din M, Arshad M, Sohail M, Faiyaz-Ul-Haque M, Haque S, Ahmad W, Leal SM. Autosomal recessive non-syndromic deafness locus (*DFNB35*) maps to 14q24.1-14q24.3 in large consanguineous kindred from Pakistan. Eur J Hum Genet 2003; 11: 77-80.
2. Ansar M, Ramzan M, Pham TL, Jamal SM, Ahmad MM, Haque S, Ahmad W, Leal SM. Localization of a novel autosomal recessive non-syndromic deafness locus (*DFNB38*) to 6q26-6q27 in consanguineous kindred from Pakistan. Hum Hered 2003; 55: 71-74.
3. Wajid M, Al Abbasi A, Ansar M, Pham TL, Haque S, Ahmad W, Leal SM. *DFNB39*, a recessive form of sensorineural hearing impairment, maps to chromosome 7q11.22-q21.12. Eur J Hum Genet 2002; 11: 812-815.
4. Aslam M, Wajid M, Chahrour MH, Ansar M, Haque S, Pham TL, Santos RP, Yan K, Ahmad W, Leal SM. A novel autosomal recessive nonsyndromic hearing impairment locus (*DFNB42*) maps to chromosome 3q13.31-q22.3. Am J Med Genet A 2005; 133: 18-22.
5. Ansar M, Chahrour MH, Amin Ud Din M, Arshad M, Haque S, Pham TL, Yan K, Ahmad W, Leal SM. *DFNB44*, a novel autosomal recessive non-syndromic hearing impairment locus, maps to chromosome 7p14.1-q11.22. Hum Hered 2004; 57: 195-199.
6. Mir A, Ansar M, Chahrour MH, Pham TL, Wajid M, Haque S, Yan K, Ahmad W, Leal SM. Mapping of a novel autosomal recessive nonsyndromic deafness locus (*DFNB46*) to chromosome 18p11.32-p11.31. Am J Med Genet A 2005; 133: 23-26.
7. Xu J, Meyers DA, Pericak-Vance MA. Lod score analysis. p. 253-272. In: Haines JL, Pericak-Vance MA, ed. Approaches to gene mapping in complex human diseases. New York: Wiley-Liss, 1998.
8. Terwilliger JD, Ott J. Handbook of human genetic linkage. Baltimore: Johns Hopkins University Press, 1994.
9. Freimer NB, Sandkuijl LA, Blower S. Incorrect specification of marker allele frequencies: effects on linkage analysis. Am J Hum Genet 1993; 52: 1102-1110.

10. Cottingham R, Indury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993; 53: 252-263.
11. O'Connell JR. Rapid multipoint linkage analysis via inheritance vectors in the Elston-Stewart algorithm. Hum Hered 2001; 51: 226-240.
12. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996; 58: 1347-1363.
13. Abecasis GR, Wigginton JE. Handling marker-marker linkage disequilibrium: pedigree analysis with clustered markers. Am J Hum Genet 2005; 77: 754-767.
14. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A. Allegro, a new computer program for multipoint linkage analysis. Nat Genet 2002; 25: 12-13.
15. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996; 58: 1323-1337.
16. Gudbjartsson DF, Thorvaldsson T, Kong A, Gunnarsson G, Ingolfsdottir A. Allegro version 2. Nat Genet 2005; 37: 1015-1016.
17. Liu F, Elefante S, van Duijn CM, Aulchenko YS. Ignoring distant genealogic loops leads to false-positives in homozygosity mapping. Ann Hum Genet (in press).
18. Leutenegger AL, Prum B, Genin E, Verny C, Lemainque A, Clerget-Darpoux F, Thompson EA. Estimation of the inbreeding coefficient through use of genomic data. Am J Hum Genet 2003; 73: 516-523.
19. Boyles AL, Scott WK, Martin ER, Schmidt S, Li YJ, Ashley-Koch A, Bass MP, Schmidt M, Pericak-Vance MA, Speer MC, Hauser ER. Linkage disequilibrium inflates type I error rates in multipoint linkage analysis when parental genotypes are missing. Hum Hered 2005; 59: 220-227.
20. Lindholm E, Zhang J, Hodge SE, Greenberg DA. The reliability of haplotyping inference in nuclear families: misassignment rates for SNPs and microsatellites. Hum Hered 2004; 57: 117-127.
21. Evans DM, Cardon LR. Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. Am J Hum Genet 2004; 75: 687-692.
22. International Human Genome Sequence Consortium. Initial sequence and analysis of the human genome. Nature 2001; 409: 860-921.
23. Kong X, Murphy K, Raj T, He C, White PS, Matise TC. A combined linkage-physical map of the human genome. Am J Hum Genet 2004; 75: 1143-1148.
24. The Hearing Research Group at Brigham & Women's Hospital. Human Cochlear cDNA library and EST database. 2002. URL: http://hearing.bwh.harvard.edu/estinfo.HTM
25. Holme RH, Bussoli TJ, Steel KP. Table of gene expression in the developing ear. 2002. URL: http://www.ihr.mrc.ac.uk/Hereditary/ genetable/index.shtml
26. Xiao ZA, Xie DH. *GJB2* (*CX26*) gene mutations in Chinese patients with congenital sensorineural deafness and a report of one novel mutation. Chin Med J (Engl) 2004; 117: 1797-1801.
27. Roux AF, Pallares-Ruiz N, Vielle A, Faugere V, Templin C, Leprevost D, Artieres F, Lina G, Molinari N, Blanchet P, Mondain M, Claustres M. Molecular epidemiology of DFNB1 deafness in France. BMC Med Genet 2004; 5: 5.
28. Wattanasirichaigoon D, Limwongse C, Jariengprasert C, Yenchitsomanus PT, Tocharoenthanaphol C, Thongnoppakhun W, Thawil C, Charoenpipop D, Pho-iam T, Thongpradit S, Duggal P. High prevalence of V37I genetic variant in the connexin-26

(*GJB2*) gene among non-syndromic hearing-impaired and control Thai individuals. Clin Genet 2004; 66: 452-460.
29. Hwa HL, Ko TM, Hsu CJ, Huang CH, Chiang YL, Oong JL, Chen CC, Hsu CK. Mutation spectrum of the connexin 26 (*GJB2*) gene in Taiwanese patients with prelingual deafness. Genet Med 2003; 5: 161-165.
30. RamShankar M, Girirajan S, Dagan O, Ravi Shankar HM, Jalvi R, Rangasayee R, Avraham KB, Anand A. Contribution of connexin26 (*GJB2*) mutations and founder effect to non-syndromic hearing loss in India. J Med Genet 2003; 40: e68.
31. Posukh O, Pallares-Ruiz N, Tadinova V, Osipova L, Claustres M, Roux AF. First molecular screening of deafness in the Altai Republic population. BMC Med Genet 2005; 6: 12.
32. Snoeckx RL, Djelantik B, Van Laer L, Van de Heyning P, Van Camp G. *GJB2* (connexin 26) mutations are not a major cause of hearing loss in the Indonesian population. Am J Med Genet A 2005; 135: 126-129.
33. Snoeckx RL, Hassan DM, Kamal NM, Van Den Bogaert K, Van Camp G. Mutation analysis of the *GJB2* (connexin 26) gene in Egypt. Hum Mutat 2005; 26: 60-61.
34. Zelante L, Gasparini P, Estivill X, Melchionda S, D'Agruma L, Govea N, Mila M, Monica MD, Lutfi J, Shohat M, Mansfield E, Delgrosso K, Rappaport E, Surrey S, Fortina P. Connexin26 mutations associated with the most common form of non-syndromic neurosensory autosomal recessive deafness (DFNB1) in Mediterraneans. Hum Mol Genet 1997; 6: 1605-1609.
35. Morell RJ, Kim HJ, Hood LJ, Goforth L, Friderici K, Fisher R, Van Camp G, Berlin CI, Oddoux C, Ostrer H, Keats B, Friedman TB. Mutations in the connexin 26 gene (*GJB2*) among Ashkenazi Jews with nonsyndromic recessive deafness. N Engl J Med 1998; 339: 1500-1505.
36. Thonnissen E, Rabionet R, Arbones ML, Estivill X, Willecke K, Ott T. Human connexin26 (*GJB2*) deafness mutations affect the function of gap junction channels at different levels of protein expression. Hum Genet 2002; 111: 190-197.
37. D'Andrea P, Veronesi V, Bicego M, Melchionda S, Zelante L, Di Iorio E, Bruzzone R, Gasparini P. Hearing loss: frequency and functional studies of the most common connexin26 alleles. Biochem Biophys Res Commun 2002; 296: 685-691.
38. Choung YH, Moon SK, Park HJ. Functional study of *GJB2* in hereditary hearing loss. Laryngoscope 2002; 112: 1667-1671.
39. Wang HL, Chang WT, Li AH, Yeh TH, Wu CY, Chen MS, Huang PC. Functional analysis of connexin-26 mutants associated with hereditary recessive deafness. J Neurochem 2003; 84: 735-742.
40. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol 2001; 305: 567-580.
41. Del Castillo I, Moreno-Pelayo MA, Del Castillo FJ, Brownstein Z, Marlin S, Adina Q, Cockburn DJ, Pandya A, Siemering KR, Chamberlin GP, Ballana E, Wuyts W, Maciel-Guerra AT, Alvarez A, Villamar M, Shohat M, Abeliovich D, Dahl HH, Estivill X, Gasparini P, Hutchin T, Nance WE, Sartorato EL, Smith RJ, Van Camp G, Avraham KB, Petit C, Moreno F. Prevalence and evolutionary origins of the del(*GJB6*-D13S1830) mutation in the DFNB1 locus in hearing-impaired subjects: a multicenter study. Am J Hum Genet 2003; 73: 1452-1458.
42. del Castillo FJ, Rodriguez-Ballesteros M, Alvarez A, Hutchin T, Leonardi E, de Oliveira CA, Azaiez H, Brownstein Z, Avenarius MR, Marlin S, Pandya A, Shahin H, Siemering KR, Weil D, Wuyts W, Aguirre LA, Martin Y, Moreno-Pelayo MA, Villamar M, Avraham KB,

Dahl HH, Kanaan M, Nance WE, Petit C, Smith RJ, Van Camp G, Sartorato EL, Murgia A, Moreno F, del Castillo I. A novel deletion involving the connexin-30 gene, del(*GJB6*-d13s1854), found in trans with mutations in the *GJB2* gene (connexin-26) in subjects with DFNB1 non-syndromic hearing impairment. J Med Genet 2005; 42: 588-594.

43. Ben Arab S, Masmoudi S, Beltaief N, Hachicha S, Ayadi H. Consanguinity and endogamy in Northern Tunisia and its impact on non-syndromic deafness. Genet Epidemiol 2004; 27: 74-79.
44. Bener A, ElHakeem AAM, Abdulhadi K. Is there any association between consanguinity and hearing loss. Int J Pediatr Otorhinolaryngol 2005; 69: 327-333.
45. Nance WE, Kearsey MJ. Relevance of connexin deafness (DFNB1) to human evolution. Am J Hum Genet 2004; 74: 1081-1087.
46. Common JE, Di WL, Davies D, Kelsell DP. Further evidence for heterozygote advantage of *GJB2* deafness mutations: a link with cell survival. J Med Genet 2004; 41: 573-575.
47. Falconer DS. Introduction to quantitative genetics, 2nd ed. New York: Longman Inc., 1981.
48. Overall AD, Ahmad M, Thomas MG, Nichols RA. An analysis of consanguinity and social structure within the UK Asian population using microsatellite data. Ann Hum Genet 2003; 67: 525-537.
49. Tekin M, Duman T, Bogoclu G, Incesulu A, Comak E, Ilhan I, Akar N. Spectrum of *GJB2* mutations in Turkey comprises both Caucasian and Oriental variants: roles of parental consanguinity and assortative mating. Hum Mutat 2003; 21: 552-553.
50. Bittles A. consang.net. 2004. URL: http://www.consang.net/
51. Estivill X, Fortina P, Surrey S, Rabionet R, Melchionda S, D'Agruma L, Mansfield E, Rappaport E, Govea N, Mila M, Zelante L, Gasparini P. Connexin-26 mutations in sporadic and inherited sensorineural deafness. Lancet 1998; 351: 394-398.
52. Abe S, Usami S, Shinkawa H, Kelley PM, Kimberling WJ. Prevalent connexin 26 gene (*GJB2*) mutations in Japanese. J Med Genet 2000; 37: 41-43.
53. Kudo T, Ikeda K, Kure S, Matsubara Y, Oshima T, Watanabe K, Kawase T, Narisawa K, Takasaka T. Novel mutations in the connexin 26 gene (*GJB2*) responsible for childhood deafness in the Japanese population. Am J Med Genet 2000; 90: 141-145.
54. Masmoudi S, Elgaied-Boulila A, Kassab I, Ben Arab S, Blanchard S, Bouzouita JE, Drira M, Kassab A, Hachicha S, Petit C, Ayadi H. Determination of the frequency of connexin26 mutations in inherited sensorineural deafness and carrier rates in the Tunisian population using DGGE. J Med Genet 2000; 37: E39.
55. Sobe T, Vreugde S, Shahin H, Berlin M, Davis N, Kanaan M, Yaron Y, Orr-Urtreger A, Frydman M, Shohat M, Avraham KB. The prevalence and expression of inherited connexin 26 mutations associated with nonsyndromic hearing loss in the Israeli population. Hum Genet 2000; 106: 50-57.
56. Wilcox SA, Saunders K, Osborn AH, Arnold A, Wunderlich J, Kelly T, Collins V, Wilcox LJ, McKinlay Gardner RJ, Kamarinos M, Cone-Wesson B, Williamson R, Dahl HH. High frequency hearing loss correlated with mutations in the *GJB2* gene. Hum Genet 2000; 106: 399-405.
57. Mustapha M, Salem N, Delague V, Chouery E, Ghassibeh M, Rai M, Loiselet J, Petit C, Megarbane A. Autosomal recessive non-syndromic hearing loss in the Lebanese population: prevalence of the 30delG mutation and report of two novel mutations in the connexin 26 (*GJB2*) gene. J Med Genet 2001; 38: E36.
58. Simsek M, Al-Wardy N, Al-Khayat A, Shanmugakonar M, Al-Bulushi T, Al-Khabory M, Al-

Mujeni S, Al-Harthi S. Absence of deafness-associated connexin-26 (*GJB2*) gene mutations in the Omani population. Hum Mutat 2001; 18: 545-546.
59. Liu XZ, Xia XJ, Ke XM, Ouyang XM, Du LL, Liu YH, Angeli S, Telischi FF, Nance WE, Balkany T, Xu LR. The prevalence of connexin 26 (*GJB2*) mutations in the Chinese population. Hum Genet 2002; 111: 394-397.
60. Medlej-Hashim M, Mustapha M, Chouery E, Weil D, Parronaud J, Salem N, Delague V, Loiselet J, Lathrop M, Petit C, Megarbane A. Non-syndromic recessive deafness in Jordan: mapping of a new locus to chromosome 9q34.3 and prevalence of DFNB1 mutations. Eur J Hum Genet 2002; 10: 391-394.
61. Najmabadi H, Cucci RA, Sahebjam S, Kouchakian N, Farhadi M, Kahrizi K, Arzhangi S, Daneshmandan N, Javan K, Smith RJ. *GJB2* mutations in Iranians with autosomal recessive non-syndromic sensorineural hearing loss. Hum Mutat 2002; 19: 572.
62. Oliveira CA, Maciel-Guerra AT, Sartorato EL. Deafness resulting from mutations in the *GJB2* (connexin 26) gene in Brazilian patients. Clin Genet 2002; 61: 354-358.
63. Shahin H, Walsh T, Sobe T, Lynch E, King MC, Avraham KB, Kanaan M. Genetics of congenital deafness in the Palestinian population: multiple connexin 26 alleles with shared origins in the Middle East. Hum Genet 2002; 110: 284-289.
64. Maheshwari M, Vijaya R, Ghosh M, Shastri S, Kabra M, Menon PS. Screening of families with autosomal recessive non-syndromic hearing impairment (ARNSHI) for mutations in *GJB2* gene: Indian scenario. Am J Med Genet 2003; 120A: 180-184.
65. Gasmelseed NM, Schmidt M, Magzoub MM, Macharia M, Elmustafa OM, Ototo B, Winkler E, Ruge G, Horstmann RD, Meyer CG. Low frequency of deafness-associated *GJB2* variants in Kenya and Sudan and novel *GJB2* variants. Hum Mutat 2004; 23: 206-207.
66. Toth T, Kupka S, Haack B, Riemann K, Braun S, Fazakas F, Zenner HP, Muszbek L, Blin N, Pfister M, Sziklai I. *GJB2* mutations in patients with non-syndromic hearing loss from Northeastern Hungary. Hum Mutat 2004; 23: 631-632.
67. Dalamon V, Beheran A, Diamante F, Pallares N, Diamante V, Elgoyhen AB. Prevalence of *GJB2* mutations and the del(*GJB6*-D13S1830) in Argentinean non-syndromic deaf patients. Hear Res 2005; 207: 43-49.
68. Hutchin T, Coy NN, Conlon H, Telford E, Bromelow K, Blaydon D, Taylor G, Coghill E, Brown S, Trembath R, Liu XZ, Bitner-Glindzicz M, Mueller R. Assessment of the genetic causes of recessive childhood non-syndromic deafness in the UK - implications for genetic testing. Clin Genet 2005; 68: 506-512.
69. Minarik G, Ferakova E, Ficek A, Polakova H, Kadasi L. *GJB2* gene mutations in Slovak hearing-impaired patients of Caucasian origin: spectrum, frequencies and SNP analysis. Clin Genet 2005; 68: 554-557.
70. Huygen PLM, Pauw RJ, Cremers CWRJ. Audiometric profiles associated with genetic non-syndromic hearing impairment, a review and phenotype analysis. Proceedings from the Caserta GENDEAF meeting, 2005.
71. Topsakal V, Pennings RJE, te Brinke H, Hamel B, Huygen PLM, Kremer H, Cremers CWRJ. Phenotype determination guides swift genotyping of a DFNA2/KCNQ4 family with a hot spot mutation (W276S). Otol Neurotol 2005; 26: 52-58.
72. Van Camp G, Smith RJH. Hereditary Hearing Loss Homepage. 2006. URL: http://dnalab-www.uia.ac.be/dnalab/hhh.html
73. Cryns K, Orzan E, Murgia A, Huygen PL, Moreno F, del Castillo I, Chamberlin GP, Azaiez H, Prasad S, Cucci RA, Leonardi E, Snoeckx RL, Govaerts PJ, Van de Heyning PH, Van de

Heyning CM, Smith RJ, Van Camp G. A genotype-phenotype correlation for *GJB2* (connexin 26) deafness. J Med Genet 2004; 41: 147-154.

74. Snoeckx RL, Huygen PL, Feldmann D, Marlin S, Denoyelle F, Waligora J, Mueller-Malesinska M, Pollak A, Ploski R, Murgia A, Orzan E, Castorina P, Ambrosetti U, Nowakowska-Szyrwinska E, Bal J, Wiszniewski W, Janecke AR, Nekahm-Heis D, Seeman P, Bendova O, Kenna MA, Frangulov A, Rehm HL, Tekin M, Incesulu A, Dahl HH, du Sart D, Jenkins L, Lucas D, Bitner-Glindzicz M, Avraham KB, Brownstein Z, Del Castillo I, Moreno F, Blin N, Pfister M, Sziklai I, Toth T, Kelley PM, Cohn ES, Van Maldergem L, Hilbert P, Roux AF, Mondain M, Hoefsloot LH, Cremers CW, Lopponen T, Lopponen H, Parving A, Gronskov K, Schrijver I, Roberson J, Gualandi F, Martini A, Lina-Granade G, Pallares-Ruiz N, Correia C, Fialho G, Cryns K, Hilgert N, Van de Heyning P, Nishimura CJ, Smith RJ, Van Camp G. *GJB2* Mutations and Degree of Hearing Loss: A Multicenter Study. Am J Hum Genet 2005; 77: 945-957.
75. Pallares-Ruiz N, Blanchet P, Mondain M, Claustres M, Roux AF. A large deletion including most of *GJB6* in recessive non syndromic deafness: a digenic effect? Eur J Hum Genet 2002; 10: 72-76.
76. Zheng QY, Yan D, Ouyang XM, Du LL, Yu H, Chang B, Johnson KR, Liu XZ. Digenic inheritance of deafness caused by mutations in genes encoding cadherin 23 and protocadherin 15 in mice and humans. Hum Mol Genet 2005; 14: 103-111.
77. Hutchin T, Haworth I, Higashi K, Fischel-Ghodsian N, Stoneking M, Saha N, Arnos C, Cortopassi G. A molecular basis for human hypersensitivity to aminoglycoside antibiotics. Nucleic Acids Res 1993; 21: 4174-4179.
78. Abe S, Kelley PM, Kimberling WJ, Usami SI. Connexin 26 gene (*GJB2*) mutation modulates the severity of hearing loss associated with the 1555A-->G mitochondrial mutation. Am J Med Genet 2001; 103: 334-338.
79. Davis RR, Kozel P, Erway LC. Genetic influences in individual susceptibility to noise: a review. Noise Health 2003; 5: 19-28.
80. Green GE, Scott DA, McDonald JM, Teagle HF, Tomblin BJ, Spencer LJ, Woodworth GG, Knutson JF, Gantz BJ, Sheffield VC, Smith RJ. Performance of cochlear implant recipients with *GJB2*-related deafness. Am J Med Genet 2002; 109: 167-170.
81. Bauer PW, Geers AE, Brenner C, Moog JS, Smith RJ. The effect of *GJB2* allele variants on performance after cochlear implantation. Laryngoscope 2003; 113: 2135-2140.
82. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002; 30: 3894-3900.
83. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 1995; 11: 241-247.
84. Morton NE. Sequential tests for the detection of linkage. Am J Hum Genet 1955; 7: 277-318.
85. Green NS, Pass KA. Neonatal screening by DNA microarray: spots and chips. Nat Rev Genet 2005; 6: 147-151.
86. Van de Water TR, Staecker H, Halterman MW, Federoff HJ. Gene therapy in the inner ear. Mechanisms and clinical implications. Ann N Y Acad Sci 1999; 884: 345-360.
87. Li H, Roblin G, Liu H, Heller S. Generation of hair cells by stepwise differentiation of embryonic stem cells. Proc Natl Acad Sci U S A 2003; 100: 13495-13500.
88. Robin NH, Prucka SK, Woolley AL, Smith RJ. The use of genetic testing in the evaluation of hearing impairment in a child. Curr Opin Pediatr 2005; 17: 709-712.
89. Smith RJH. Clinical application of genetic testing for deafness. Am J Med Genet A 2004; 130: 8-12.

Chapter 6
Summary and Conclusions

This thesis presents studies on the genetic determinants of non-syndromic hearing impairment (NSHI) with the following categorization: (1) the mapping of novel NSHI loci using linkage analyses, and the exclusion of several candidate genes; (2) the identification of novel variants in previously isolated NSHI genes, with prediction of functionality of variants and estimation of gene-specific prevalence rates; and (3) the correlation of genotype with audiometric phenotype.

In Chapter 2, five novel autosomal recessive (AR) NSHI loci are described. Table 6.1 shows the different loci that were mapped, with the corresponding chromosomal region, the STR markers which flank the region of homozygosity, and the interval size in cM and Mb. Two of these loci had overlap with other previously known loci: *DFNB55* lies within the genetic interval of autosomal dominant (AD) NSHI locus *DFNA27*, thus it is possible that different mutations within the same gene can cause both AD and AR forms of NSHI. *DFNB68* was mapped to the 19p region of *DFNB15*, but since the *DFNB15* family has non-significant LOD scores at both 19p and 3q, the mapping of a small interval in two families with highly significant LOD scores shows better evidence of a novel NSHI gene within the *DFNB68* interval.

Table 6.1 Novel ARNSHI loci that were presented in this thesis

Locus	Chromosomal region	Flanking markers	Interval size (cM)	Interval size (Mb)
DFNB47	2p25.1-p24.3	D2S2952, D2S131	13.2	5.3
DFNB55	4q12-q13.2	D4S2978, D4S2367	8.2	11.5
DFNB62	12p13.2-p11.23	D12S358, D12S1042	22.4	15.0
DFNB65	20q13.2-q13.32	D20S480, D20S430	10.5	4.3
DFNB68	19p13.2	D19S916, D19S584	1.9	1.4

The NSHI transmitted within the pedigrees that were ascertained had primarily autosomal recessive inheritance. As discussed in this thesis, the high consanguinity rate in Pakistani society has resulted in the availability for study of multigenerational extended family structures with multiple consanguinity loops, which are thus very useful for homozygosity mapping.

The analyses for loci mapping have employed recent advances in genetic epidemiology, including more accurate physical and genetic maps, error-checking programs (MERLIN), faster multipoint linkage analysis (ALLEGRO), and fine-mapping with dense maps of SNP markers combined with STR markers.

For the novel loci which are described in this thesis, an average of 37 known genes (range 14-75) occurs within each genetic/physical interval, and often predicted genes also lie within these intervals. A number of genes were selected as candidate genes based on inner ear expression and association with syndromic forms of HI and were sequenced. This endeavor has resulted in the exclusion of several genes as the genetic cause of HI in these families. Currently for the five loci, a total of 19 genes and 6 expressed sequence tags (average 4 per locus) have been screened for functional variants, but the causative NSHI gene remains unknown. Much laboratory and analytical work remains to be done for the identification of novel NSHI genes.

Studies of putatively functional variants in known NSHI genes are included in Chapter 3. Three genes, *GJB2*, *TMC1* and *TMIE*, were screened among Pakistani families with ARNSHI through direct sequencing. *GJB2* was screened for all families prior to submission for genome scan. For *TMC1* and *TMIE*, marker genotyping and linkage analysis were performed prior to selection of families for sequencing, which is part of a streamlined approach to the genetic cause of NSHI in these families. These two genes were selected based on the number

of families that were linked to the NSHI gene region, the ease of sequencing based on the number of reactions required and G-C content of the genetic region, and the paucity of known gene variants. This has resulted in the discovery of five new *TMC1* variants and two for *TMIE* (Table 6.2). One of the *TMIE* variants is a splice mutation that was identified in a family from Jordan. No new variants were found for *GJB2*. Discrimination between functional variants and benign polymorphisms was determined based on bioinformatics tools such as multiple sequence alignment and transmembrane domain prediction, and allele frequency comparison with control populations.

Because of the large number of Pakistani families screened, gene-specific prevalence estimates were derived, as follows: *GJB2*, 5.4%; *TMC1*, 4.4%; and *TMIE*, 1.7%. The *GJB2* gene has been studied across populations, thus allowing comparison of prevalence estimates among patients and their corresponding confidence intervals. It was shown that the spectrum of alleles within the Indian subcontinent is unique as compared to other world regions, which may imply the occurrence of common founders within the region. Additionally, although similar in allelic spectrum, the prevalence rate in Pakistan is much lower than in India. The hypothesis is that the lower gene-specific prevalence rates within Pakistan may be related to the more frequent occurrence of consanguinity within families, thus resulting in multiple mutated genes across the population that are preserved within highly inbred lineages.

Chapter 4 presents two studies on audiometric phenotype. The *GJB2* phenotype in unrelated Dutch patients was described in Section 4.1. Hearing impairment was significantly worse in patients with biallelic mutations as compared to wild type and heterozygote individuals. The presence of at least one missense variant tends to lead to better hearing thresholds as compared to biallelic radical mutations. Non-progressive hearing impairment was demonstrated in some patients, but the occurrence of progression in the lower frequencies (with the genotype [299A>C + 416G>A]) was also demonstrated.

In a large multigenerational Swiss-German family segregating NSHI due to *DFNA24* [Section 4.2], the audiometric phenotype was shown to be sensorineural, bilateral, prelingual, steeply down-sloping and progressive. The age-related typical audiogram showed that moderate HI starts at the high frequencies at birth, then steadily progresses to profound HI by age 70, and eventually affects low-to-mid frequencies with moderate-to-severe HI. Annual threshold deterioration was ~0.5

Table 6.2 Novel variants in the *TMC1* and *TMIE* genes

Gene	Nucleotide Change	Amino Acid Change	Domain[a]	Evolutionary Conservation[b]	Allele Frequencies among Controls[c]
TMC1	c.830A>G	p.Y277C	TM2	Identical	0/234
TMC1	c.1114G>A	p.V372M	TM3	Conserved	0/234
TMC1	c.1334G>A	p.R445H	TM4	Identical	0/234
TMC1	c.2004T>G	p.S668R	EC	Conserved in 11 of 14 species	0/234
TMC1	c.2035G>A	p.E679K	EC	Conserved in 8 of 13 species	0/234
TMIE	c.92A>G	p.E31G	EC	Conserved in mouse only	0/110
TMIE	IVS2 -2A>C	Splice acceptor mutation	N/A	N/A	0/176

[a] Transmembrane domains for *TMC1* and *TMIE* were determined via the TMHMM v.2.0 program. *TM*, transmembrane; *EC*, extracellular; *N/A*, not applicable.
[b] Similarity search was done using most similar proteins, one from each species, from the ExPaSy/UniProt database and the NCBI BLASTP 1.5.4-Paracel program. Additional sequences were obtained from Ensembl (www.ensemble.org).Multiple protein sequence alignment was then performed with ClustalW.
[c] Allele frequencies were determined using ethnically-matched hearing controls.

dB/year at 1-2 kHz after correction for presbycusis. These studies on genotype-phenotype correlation are expected to contribute to genetic counseling and clinical management. In the case of *DFNA24* in which the gene is unknown, it may also lead to the identification of other families that may aid in fine mapping and fine-tuning of the phenotype.

In the overall discussion in Chapter 5, how the different studies in the previous chapters fall within the realms of NSHI genetic epidemiology are discussed. The different methods and assumptions that are used for locus mapping are examined, and the positional candidate gene approach is briefly outlined. For the screening of variants in NSHI genes, the advantages and disadvantages of different ways in which a variant can be identified as putatively functional are reviewed in detail. The problems associated with phenotypic description are also scrutinized. Lastly, future areas of NSHI research are enumerated, including complex interactions, decision paradigms for genetic screening and functionality of variants, mass screening of variants, gene therapy and genetic counseling.

In conclusion, this thesis has attempted to compile studies that extract genotypic and/or phenotypic information from NSHI families and individuals from different countries. Whether it is the discovery of a novel NSHI gene or locus, variant or phenotype, it is hoped that each finding, however small, may contribute to the large knowledge base that is required for one goal: a comprehensive and effective genetic testing program for NSHI that would benefit populations of various ethnic backgrounds, newborns and adults alike.

Samenvatting

In dit proefschrift worden studies naar de genetische oorzaken van niet-syndromale slechthorendheid (NSHI) gepresenteerd, meer in het bijzonder: (1) het in kaart brengen van nieuwe NSHI genlokaties via genetische koppelingsanalyse, en het screenen van kandidaat-genen; (2) de identificatie van nieuwe varianten binnen eerder geïsoleerde NSHI genen, met voorspelling van hun functionaliteit en schattingen van de gen-speficieke prevalentiecijfers; en (3) de correlatie tussen genotype en audiometrisch fenotype.

In hoofdstuk 2 worden vijf nieuwe autosomaal recessieve (AR) loci voor NSHI beschreven. Tabel 6.1 toont de in kaart gebrachte loci, met de overeenkomstige chromosomale regio, de STR merkers die de homozygote regio flankeren, en de intervalgrootte in cM en Mb. Twee van deze loci overlapten met andere, al bekende genlokaties: *DFNB55* ligt binnen het genetische interval van de autosomaal dominante (AD) NSHI locus *DFNA27.* Het is dus mogelijk dat verschillende mutaties binnen hetzelfde gen zowel AR als AD vormen van NSHI veroorzaken. *DFNB68* werd gelokaliseerd op de 19p regio van *DFNB15*, maar aangezien de *DFNB15* familie niet-significante LOD scores heeft op zowel 19p als 3q, levert het afgebakende kleine interval in twee families met zeer significante LOD scores een sterker bewijs voor een nieuw gen binnen het *DFNB68* interval.

De door ons geverifieerde niet-syndromale slechthorendheid die binnen de stambomen werd overgeërft, had voornamelijk een autosomaal recessief overervingspatroon. Zoals in dit proefschrift werd bediscussieerd, heeft de Pakistaanse samenleving een hoge graad van bloedverwantschap. De vele uitgebreide multi-generatie families met meerdere consanguiene huwelijken, leverden zeer geschikt familiemateriaal op om homozygositeit in kaart te brengen.

We gebruikten recente ontwikkelingen binnen de genetische epidemiologie voor het lokaliseren van de verantwoordelijke genen, zoals meer nauwkeurige fysieke en genetische kaarten, programmas voor het opsporen van fouten (MERLIN), snellere multipoint analyse (ALLEGRO) en fijn-mapping met een dens raster van SNP een STR merkers.

Van de nieuwe loci die in dit proefschrift beschreven worden, bevinden er zich gemiddeld 37 bekende genen (bereik 14-75) binnen elk genetisch/fysieke interval, en meestal liggen de voorspelde genen ook binnen deze intervallen. Een aantal genen werd, gebaseerd op hun expressie in het binnenoor en hun rol in andere vormen van niet-syndromale slechthorendheid, geselecteerd als kandidaatgenen en gesequenced. Zo werden verscheidene genen als mogelijke oorzaak van slechthorendheid binnen deze families uitgesloten.

Tabel 6.1 Nieuwe ARNSHI loci die in dit proefschrift beschreven worden

Locus	Chromosomale regio	Flankerende merkers	Interval grootte (cM)	Interval grootte (Mb)
DFNB47	2p25.1-p24.3	D2S2952, D2S131	13.2	5.3
DFNB55	4q12-q13.2	D4S2978, D4S2367	8.2	11.5
DFNB62	12p13.2-p11.23	D12S358, D12S1042	22.4	15.0
DFNB65	20q13.2-q13.32	D20S480, D20S430	10.5	4.3
DFNB68	19p13.2	D19S916, D19S584	1.9	1.4

Tabel 6.2 Nieuwe varianten in de *TMC1* en *TMIE* genen

Gen	Nucleotide Verandering	Aminozuur verandering	Domein[a]	Evolutionaire Conservatie[b]	Allel Frequenties bij Controles[c]
TMC1	830A>G	Y277C	TM2	Identiek	0/234
TMC1	1114G>A	V372M	TM3	Geconserveerd	0/234
TMC1	1334G>A	R445H	TM4	Identiek	0/234
TMC1	2004T>G	S668R	EC	Geconserveerd in 4/8 species	0/234
TMC1	2035G>A	E679K	EC	Geconserveerd in 2/8 species	0/234
TMIE	92A>G	E31G	EC	Enkel geconserveerdd in muis	0/110
TMIE	IVS2 -2A>C	Splice acceptor mutation	NVT	NVT	0/176

[a] Transmembraan domeinen voor *TMC1* en *TMIE* werden bepaald met behulp van het TMHMM v.2.0 programma. TM=transmembraan; EC=extracellulair; NVT=niet van toepassing.
[b] Similarity search werd uitgevoerd, gebruik makend van de meest soortgelijke eiwitten, één van elke soort, uit de ExPaSy/UniProt database en het NCBI BLASTP 1.5.4-Paracel programma. Multiple protein sequence alignment werd uitgevoerd met ClustalW.
[c] Allel frequenties werden bepaald in etnisch gematchte, goed horende controlepersonen.

Momenteel zijn er voor de 5 loci in totaal 19 genen en 6 'expressed sequence tags' (gemiddeld 4 per locatie) gescreend op functionele varianten, maar het ziekteveroorzakende NSHI gen blijft vooralsnog onbekend. Voor de identificatie van nieuwe NSHI genen zal nog veel laboratorium en analytisch werk gedaan moeten worden.

In hoofdstuk 3 staan studies naar mogelijke functionele varianten van bekende NSHI genen beschreven. In Pakistaanse families met ARNSHI werden drie genen, *GJB2*, *TMC1* en *TMIE*, gescreend via rechtstreekse sequentiebepaling. Als onderdeel van een gestroomlijnde benadering voor de genetische oorzaak van NSHI in deze families, werden voorafgaand aan de genoom scan alle families gescreened op *GJB2* mutaties. Voorafgaand aan de selectie van de families voor sequentiebepaling werden flankerende merkers voor *TMC1* en *TMIE* gegenotypeerd en koppelingsanalyses uitgevoerd. Deze twee genen werden geselecteerd op basis van het aantal families dat gekoppeld was met de desbetreffende regio's, het gemak van sequencen in termen van het aantal vereiste reacties en G-C inhoud van de regio, en het gebrek aan bekende genvarianten. Dit resulteerde in de ontdekking van vijf nieuwe *TMC1* en twee nieuwe *TMIE* varianten (tabel 6.2). Eén van de *TMIE* varianten is een splice site mutatie die werd geïdentificeerd in een familie uit Jordanië. Er werden geen nieuwe varianten gevonden voor *GJB2*. Onderscheid tussen functionele varianten en neutrale polymorfismen werd gemaakt met behulp van bio-informatische hulpmiddelen zoals multiple sequence alignment en transmembraan domein voorspelling, en een vergelijking van allelfrequenties met controle populaties.

Vanwege het grote aantal gescreende Pakistaanse families, werden de volgende gen-specifieke prevalenties geschat: *GJB2* 5,4%; *TMC1* 4,4%; en *TMIE* 1,7%. Het *GJB2* gen is in meerdere populaties gescreend, wat vergelijkingen van prevalentieschattingen met betrouwbaarheidsintervallen mogelijk maakt. Het allelische spectrum in het Indiase subcontinent bleek uniek te zijn in vergelijking met andere regio's in de wereld, hetgeen zou kunnen wijzen op het voorkomen van gemeenschappelijke voorouders in deze regio. Bovendien was het prevalentiecijfer, in Pakistan veel lager dan in India, hoewel het allelisch spectrum vergelijkbaar was. De hypothese is dat de lagere gen-specifieke prevalentiecijfers binnen Pakistan gerelateerd zijn met het vaker voorkomen van bloedverwantschap binnen families, waardoor meerdere gemuteerde genen binnen de populatie behouden blijven binnen afstammingslijnen met veel inteelt.

Hoofdstuk 4 beschrijft twee studies naar audiometrisch fenotype. Het *GJB2* fenotype in niet-gerelateerde Nederlandse patiënten werd beschreven

in hoofdstuk 4.1. Slechthorendheid was significant erger bij patiënten met bi-allelische mutaties in vergelijking met wildtype en heterozygote individuen. De aanwezigheid van tenminste één missense variant lijkt te leiden tot betere gehoorsdrempels dan bi-allelische trunkerende mutaties. Niet-progressieve slechthorendheid kwam bij enkele patiënten voor, maar progressie kwam ook voor bij de lagere frequenties (met het genotype [299A>C + 416G>A]).

In een grotere Zwitsers-Duitse familie van meerdere generaties waarin NSHI overerft door *DFNA24* (hoofdstuk 4.2), was het audiometrische fenotype sensorineuraal, bilateraal, prelinguaal, steil aflopend en progressief. Het leeftijd-specifieke typische audiogram toonde aan dan matige slechthorendheid begon bij de geboorte in de hoge frequenties, met een voortdurende progressie tot ernstige slechthorendheid op de leeftijd van 70 jaar. In de lage en midden-frequenties kwam uiteindelijk matige tot ernstige slechthorendheid voor. De jaarlijkse verslechtering in drempelwaarde was ~0.5 dB per jaar bij 1-2 kHz na correctie voor presbyacusis. Deze genotype-fenotype correlatie studies kunnen een bijdrage leveren aan de genetische counseling en klinische behandeling. In het geval van *DFNA24*, waar het gen onbekend is, kan het ook leiden tot de identificatie van andere families, hetgeen fijnmapping van het gen en fine-tuning van het fenotype mogelijk maakt.

In de algemene discussie in hoofdstuk 5 wordt uiteengezet hoe de verschillende studies uit de voorgaande hoofdstukken kaderen in de genetische epidemiologie van NSHI. De verschillende methoden en aannames die gebruikt worden voor het in kaart brengen van ziekteloci worden onderzocht en de positionele kandidaatgen strategie wordt kort uitgelegd. In het kader van de screening van varianten in NSHI genen, worden de voor- en nadelen van de beschikbare methoden om een genvariant te identificeren als mogelijk functioneel, in detail behandeld. De problemen die samenhangen met de fenotypische beschrijving worden ook onderzocht. En tenslotte worden toekomstige onderzoeksterreinen voor NSHI onderzoek aangegeven, zoals complexe interacties, beslissingparadigmas voor genetische screening en functionaliteit van varianten, massa-screening van varianten, gentherapie en genetische voorlichting.

Samenvattend brengt dit proefschrift studies samen waarin genotypische en fenotypische informatie van NSHI families en individuen uit verschillende landen beschreven worden. Of het nu gaat om de ontdekking van een nieuw NSHI gen of genlocatie, variant of genotype, we hopen dat elke bevinding, hoe klein dan ook, een bijdrage kan leveren aan de grote kennisbasis die nodig is voor maar één doel: een samenhangend en effectief genetisch onderzoeksprogramma voor NSHI waarvan populaties van verschillende etnische achtergronden en leeftijden, zowel pasgeborenen als volwassenen, profijt zullen hebben.

Words of Appreciation

A classmate in medical school once said that she did not want to remember her school days with just books and tests. I have been in school for most of my life, and if I should look back at my days spent, the thrill of learning new things, finding new places and meeting new people supersedes most memories. After all, how do you justify yanking oneself out of a comfortable life in one's homeland, and meeting the challenges of living alone in strange cities? Putting aside bike accident, petty theft (and classmate), cut-off phone line, endless car troubles, paranoid roommate, etc: my experiences, both profound and mundane, are unique and irreplaceable, so much so that the PhD becomes a hefty excuse to live among the good. For example, these weird associations: tobacco puffs + green tea + biking shorts = three topnotch scientists, Cornelia, Suzanne and Yurii -- way above scholarly existence, so that if I live to 60, I'll be happy if I absorb at least 20% of their knowledge, both in science and in life. Top that with a few months with Lodewijk Sandkuijl; I surely got the bases loaded. Only through their zealousness and inspiration was this four-year effort on either side of the Atlantic fully realized. Through Cornelia's and Suzanne's sponsorship, I was able to secure scholarships from Nuffic and Nihes at Erasmus Medical Centre, and then fellowship at Baylor College of Medicine and the Human Genome Sequencing Center (HGSC); for these chances of a lifetime, I will always be grateful.

I am also thankful to the members of the doctoral committee – Professors Uitterlinden, Borst, van Camp, Cremers and Dr.Vincenzo – for taking time out to read my manuscript and be there at my defense. My utmost gratitude to my co-authors who belong to five countries, including wonderful co-workers in the Leal lab: Thanh, Kai, Danny, Maria, Nat, Wambie and Jun; from Nijmegen, Prof. Cremers, Patrick, Lies, Hannie and Martijn; and half of whom I have never even met personally, the Swiss doctors – Drs. Häfner, Linder, Schinzel and Spillmann -- and most notably Dr. Wasim Ahmad and his staff in Pakistan.

So what was initially a sabbatical from clinics became a personal quest for balance, surviving on international phone cards, free home-cooked meals and emails from good friends. My paranimfs, Sandra and Esther, with their superb organizational skills and unfailing friendship, allowed this thesis to reach its culmination even when I was offshore. Esther and

Erik Fransen also helped me in the Dutch translation of my summary and conclusions. Their benevolent examples abound in my universe, friends that are home away from home: Atsuko, the Martin's, Jenny B, the Francisco's, Pineda's, Tita Chol, the Haberstein's, Heijboer's, Goldsack's, Barcelon's, Boni's, Arquiza's, the Calumpang clan, plus the Silva's, de Guzman's, Angie & Ed M, Christine, the Guirao's, de Slegte's and Jansen's. Cheers for scrumptious dinners with Drs. E Cruz and Hardillo, the Cababarroses, my peers from the MSc and DSc programs, the HGSC lunch group, and the Iglesia ni Cristo locales of Dordrecht and Southwest Houston. Here's to sound advice from Omer, Oliver R, Behrooz, Dra. Chiong and other clinicians who allowed me to observe their work. To loyal friends Tita Tess, Michelle, Rina, Cindy, Peter and Prapapan, I salute you. Kuya Dennis, thanks for helping me through crunch time and being patient with me through the layout and printing process. Chris, who would have thought that I could love again? There are many more names and faces to remember, people I have dined, cooked, traveled, conversed with, learned from, but more important to recognize would be the hundreds unnamed, including a small group of unknown editors and reviewers, and a huge conglomerate of families from hidden places, whom I know only as pedigree drawings. To all of you, maraming salamat po.

Papa and Ate Chelah, your devotion is overwhelming, and I can only pray that someday I can make it all worth your while. Finally, to Mama, I will always regret that I was not by your side when you needed me most, so all my sacrifices are yours too. I thank God that you are still here, and my only consolation is that upon closing this book, we are together once more.

Author's Description

Regie Lyn P. Santos was born on October 8, 1973 in Caloocan City, Philippines. In 1997, she graduated from the University of the Philippines (UP) College of Medicine and passed the Philippine Medical Board Examination. She underwent residency training in otorhinolaryngology – head and neck surgery at the UP-Philippine General Hospital (PGH) from 1998 to 2001, thereafter passing the specialty board in otolaryngology. In 2002, she received a one-year scholarship from the Netherlands organization for international cooperation in higher education (Nuffic) via the University Fellowships Program to earn her Master of Science in Genetic Epidemiology from the Netherlands institute of health sciences (Nihes) in cooperation with Erasmus University Rotterdam. For her Doctor of Science in Genetic Epidemiology, she spent another eight months with the Genetic Epidemiology Unit, Department of Epidemiology and Biostatistics, Erasmus Medical Center Rotterdam as a Nihes scholar. She then moved to Houston, Texas to join the Laboratory of Statistical Genetics of the Department of Molecular and Human Genetics, which is situated within the Human Genome Sequencing Center at Baylor College of Medicine. She is now awaiting formal appointment as assistant research professor at the Philippine National Ear Institute, an affiliate of the National Institutes of Health – Philippines, and clinical associate professor at the UP-PGH Department of Otorhinolaryngology.

Other international publications by the author:

Aslam M, Wajid M, Chahrour MH, Ansar M, Haque S, Pham TL, Santos RLP, Yan K, Ahmad W, Leal SM. A novel autosomal recessive nonsyndromic hearing impairment locus (DFNB42) maps to chromosome 3q13.31-q22.3. Am J Med Genet A 2005 Feb; 133(1): 18-22.

Santos RLP, Zillikens MC, Rivadeneira FR, Pols HA, Oostra BA, van Duijn CM, Aulchenko YS. Heritability of fasting glucose levels in a young genetically isolated population. Diabetologia 2006 Apr; 49(4): 667-672.

Stellingen behorend bij het proefschrift

Genetic Determinants of Non-syndromic Hearing Impairment

1. Genetic hearing impairment is extremely heterogeneous, and the number of genes involved may total up to 300 or 1% of all human genes. – *After: Nance WE. Ment Retard Dev Disabil Res Rev 2003; 9: 109-119.*

2. The Pakistani population is a valuable resource for mapping non-syndromic hearing impairment genes, with ~29 loci identified in Pakistani families, including five loci that are described in this thesis.

3. The gene-specific prevalence rates among Pakistani families for functional variants in the genes *GJB2, TMC1* and *TMIE* are low, and must be explained by random genetic drift (this thesis).

4. Bioinformatics tools, such as multiple sequence alignment and transmembrane domain prediction, should be used in order to predict if a variant is functional (this thesis).

5. For new gene discovery in hearing impairment to be clinically useful, studies on gene-specific prevalence rates, spectra of sequence variants, phenotypic description and functional analyses are required (this thesis).

6. Phenotypic description for non-syndromic hearing impairment loci aids not only gene mapping but also clinical management (this thesis).

7. DNA sequence analysis to discover population polymorphisms that may predispose to regionally specific diseases is one of the most promising biotechnologies for improving health in developing countries. – *Daar AS et al. Nat Genet 2002; 32: 229-232*

8. The study of head and neck tumors through microarray data has great potential in the elucidation of progression of molecular disease, the prediction of the occurrence of tumors and morbid outcomes, and the identification of altered biologic pathways that are novel therapeutic targets. – *Choi P and Chen C. Cancer 2005; 104: 1113-1128.*

9. The majority of studies find that racial and ethnic disparities [in health care] remain even after adjustment for socioeconomic differences and other healthcare access-related factors. – *Institute of Medicine. "Unequal treatment: confronting racial and ethnic disparities in health care" Washington DC: National Academies Press, 2003.*

10. The clinical academicians, the physicians who care for patients and also spend time directing their own laboratory effort, represent the critical elements to facilitate clinical-basic collaborations and promote translational studies. – *After: Holcombe RF. Acad Med 2005; 80; 905-907.*

11. Philippine mango is the best mango in the world.

Regie Lyn Pastor Santos, 2006

Genetic Determinants of Non-syndromic Hearing Impairment

Regie Lyn Pastor Santos