Joint Biotechnology Master Program

Palestine Polytechnic University

Deanship of Graduate Studies and Scientific Research

Bethlehem University

Faculty of Science

A novel Splice Site Mutation in ADGVR1 Detected in Palestinian Family with Hearing Loss by Next Generation Sequencing

By

Fida Ahmad Tomizy

In Partial Fulfillment of the Requirements for the Degree

Master of Science

April 2016

The undersigned hereby certify that they have read and recommend to the Faculty of Scientific Research and Higher Studies at the Palestine Polytechnic University and the Faculty of Science at Bethlehem University for acceptance a thesis entitled:

A novel Splice Site Mutation in ADGVR1 Detected in Palestinian Family with Hearing Loss by Next Generation Sequencing

By **(**Fida Ahmad Tomizy)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in biotechnology

Graduate Advisory Committee:

Committee Member (Student's Supervisor)	Date
Dr. Moein Kanaan, Bethlehem University	
Committee Member (Internal Examine)	Date
Dr. Yaqoub Ashhab Palestine Polytechnic University	
Committee Member (External Examine)	Date
Dr. Saad Al- Lahham, Al- Najah University	

Approved for the Faculties

Dean ofGraduate Studies andScientificResearch	Dean of Faculty of Science
Palestine Polytechnic University	Bethlehem University
Date	Date

A novel Splice Site Mutation in ADGRV1 Detected In Palestinian Family With Hearing Loss By Next Generation Sequencing.

by (Fida Ahmad Tomizy)

ABSTRACT

Background: hereditary deafness is a genetically heterogeneous trait. To date, 60 genes that have been identified that cause nonsyndromic hearing loss. The extreme genetic heterogeneity of deafness makes the ability to diagnose and define the mutation by conventional method costly and time-consuming. Therefore we employed next generation sequencing to identify the causative gene and mutation in a Palestinian family with congenital inherited hearing loss.

Methods: Targeted next generation exome sequencing was performed to detect (a) mutation(s) responsible for hearing loss in this family after excluding all known deafness causing mutations in Palestinian population employing standard sanger sequencing method. Sanger sequencing method was also employed to validate next generation exome sequencing data and genotype to phenotype segregation and to analyze100 normal and 263 hearing loss Palestinian controls. Also functional assay was performed to determine the effect of the mutation on mRNA level.

Results: Targeted exome sequencing revealed acceptor splicing mutation in *ADGVR1* (p.G2898A). Sanger sequencing indicates that this mutation segregate with the hearing loss in this family under an autosomal recessive mode of inheritance. cDNA analysis shows skipping of exon 13, 14 completely with deletion of 114 aa from Calx B domain of ADGVR1.

Conclusion: A novel splicing site mutation detected in ADGRV1 gene by next generation sequencing in Palestinian family with hearing loss, this mutation affect the splicing of mRNA resulting in skipping of (exon 13 and 14) of ADGRV1 gene.

Keywords: Next-generation sequencing (NGS), Hearing loss, Heterogeneous, ADGRV1

اكتشاف طفرة جديدة ف الجين (ADGRV1) تسبب فقدان السمع باستخدام الجيل الثاني من تقنية معرفة تسلسل المادة الوراثية

فداء أحمد اطميزي

ملخص الدراسة:

الصمم الوراثي يعتبر من أكثر الأمراض الوراثية تعقيدا حيث أنه هناك أكثر من ستين جينا لها علاقة بفقدان السمع حتى الآن. نظرا للتعقيد في عملية فقدان السمع فان الكشف عن الطفرات المسببة باستخدام الطرق التقليدية يعد أمرا مكلفا ويحتاج لوقت طويل، ولكن باستخدام تقنية Next Generation Sequencing (NGS) تمكن الباحثين من حل مشكلة الفحوصات التقليدية ومن ثم توفير الوقت والمال اللازمين لعمل تلك الفحوصات. الهدف الرئيسي في هذه الدراسة معرفة الجين المسبب لعملية فقدان السمع الوراثي في إحدى العائلات الفلسطينية.

في هذه الدراسة تم فحص جميع الجينات المسؤولة عن فقدان السمع في العائلات الفلسطينية ولم يظهر أي من هذه الجينات له علاقة بفقدان السمع في هذه العائلة الفلسطينية، ولكن باستخدام تقنية (NGS) التي تفحص 246 جينا لها علاقة بفقدان السمع فلقد تم تحديد الجين المسؤول عن هذا الخلل في إحدى العائلات الفلسطينية، وكذلك تم استخدام (Sanger Sequencing) لتأكيد نتائج (NGS) واكتشاف كيفية انتقال الطفرة المسؤولة عن فقدان السمع من الآباء الحاملين لها إلى الابناء المصابين في تلك العائلة.

كشفت تقنية (NGS) عن وجود طفرة جديدة في الجين (ADGRV1). تؤثر هذه الطفرة على عملية معالجة RNA وينتج عن هذا حذف (exon 14,15) من هذا الجين مما يؤدي إلى فقدان114 حمض أميني من البروتين النهائي الناتج عن ترجمة هذا الجين من قواعد نيتروجينية إلى أحماض أمينية.

DECLARATION

I declare that the Master Thesis entitled "dissertation title" is my own original work, and hereby certify that unless stated, all work contained within this thesis is my own independent research and has not been submitted for the award of any other degree at any institution, except where due acknowledgment is made in the text.

Name and signature: Fida Tomizy

Date: 2016

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for the joint master degree in biotechnology at Palestine Polytechnic University and Bethlehem University, I agree that the library shall make it available to borrowers under rules of the library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from, reproduction, or publication of this thesis may be granted by my main supervisor, or in [his/her] absence, by the Dean of Higher Studies when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature : Fida Tomizy

Date: 2016

Dedication

To the best supporters I ever had, whose support knows no bounds, mother, and husband.

To the best kids in my life for their patience with me: Kenan, Yazid and Mohammad.

To My first family; brothers and sisters.

To every person who stood with me, supported me, and encouraged me to go on from amongst my family circle, friends, and others.

To all these, I dedicate this thesis.

Acknowledgment

I would like to thank my supervisor, Prof. Moein Kannan for his guidance, support, patience and encouragement throughout this study.

I would like to thank all the faculty staff in the Biotechnology Master program in Bethlehem University.

Abbreviations:

HL	Hearing loss
ARHL	Age related hearing loss
Hz	Hertz
dB	Decibel
IHC	Inner hair cells
OHC	Outer hair cells
SHL	Syndromic hearing loss
NSHL	Non syndromic hearing loss
DFN	Deafness locus
ARNSHL	Autosomal recessive nonsyndromic hearing loss
NIHL	Nois-induced hearing loss
ADGRV1	adhesion G protein-coupled receptor V1
VLGR1	Very large G protein coupled receptor 1
GPCR	G protein coupled receptor
USH2	Usher syndrome type 2
NGS	Next generation sequencing

List of Figures:

List of Tables:

List of Contents

CHAPTER 1

1. Introduction

Hearing loss (HL) is the most common sensory disorder in humans. It affects 1 in 500 newborn and approximately 278 million people worldwide (Shearer et al., 2011). In Palestinian population the frequency of hearing impairment is approximately 1.7 per 1000. This high frequency reflect the high consanguinity (Shahin et al., 2002). Many types of hearing loss are age related hearing loss (ARHL), also known as presbyacusis. More than 50% of the population older than 80 years old, and more than 25% of 65 years old suffers from different degrees of age-related hearing loss (ARHL) (Dror and Avraham, 2009). If HL starts at birth before speech development, it is called prelingual, and postlingual starts after speech development (Hilgert et al., 2009). Severe hearing loss that occurs during infancy can have dramatic effects on speech acquisition and literacy (Raviv et al., 2010).

Hearing loss in older adults might have a deleterious effect on the quality of life, because of the difficulties of the communication and exchange of information with others. So individuals with hearing loss might have social isolation, depression, and increased of emotional dysfunction (Raviv et al., 2010, Chou et al., 2011, Dalton et al., 2003).

Hearing loss may be conductive, sensorineural, or a combination of both (Kochhar et al., 2007). If a defect is in the external or middle ear, it causes conductive Hl. However, sensorineural HL results from a defect to the inner ear or auditory nerve (Friedman et al., 2007, Smith et al., 2014). Sensorineural hearing loss can be present at birth or acquired at later stages in life (Raviv et al., 2010). It affects 4% of the population before the age of 45 (Idan et al., 2013).

Hearing loss among Palestinians is caused by multiple different mutations

CHAPTER 2

Literature Review

2.1 Degree of hearing loss

The measurement of hearing consists of two parameters: the frequency (HZ) and the intensity or loudness of a sound (dB).

The apparatus used to measure responses to sound is called an audiometer. Responses to the tones given through the audiometer are recorded on a graph called an audiogram. The responses plan on the audiogram illustrate levels of hearing for each ear. Hearing is considered normal if an individual's thresholds are within 15 dB of normal thresholds. HL can be usually described as mild, moderate, severe, moderately severe, or profound (Figure 1) (Smith et al., 2014).

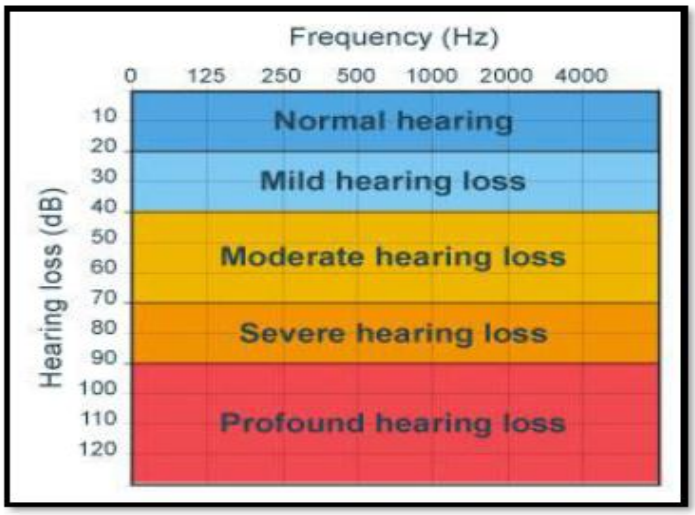


Figure1: degree of hearing loss X-axis of the audiogram represents frequency (HZ), and Y-axis of the audiogram represent the perception of loudness, 0dB is the softest sound that individuals with the best hearing 50% of the time. http://wwwcochlea.org/en/treatments.

2.2 The structure of ear and mechanisms of hearing

The ear is anatomically divided into three parts: the outer, the middle, and the inner ear (Fig 2a). The outer ear consists of auricle, external auditory canal, and the eardrum (tympanic) membrane. The middle ear is composed of three small bones the ossicles, malleus (hammer), the incus (anvil), and the stapes (stirrup).

Sound waves are collected by the auricle, then passes through the external auditory canal to reach the eardrum (Dror and Avraham, 2009). The vibration of the eardrum membrane is transferred by the bones of the middle ear to the inner ear. This vibration pass through the oval window and consequently lead to movement of fluid-filled inner ear.

The inner ear is composed of the vestibular system, that responsible for gravity perception and maintains the balance, and the auditory system, which is composed, of a cochlea that responsible for sound sensation.

The cochlea is a fluid-filled coiled tube in a spiral shape. Cross section in cochlea showing three parts (scalae), the two larger perilymphatic filled compartment, the scala tympani and scala vestibule surrounded the scala media which is filled with endolymph (Fig2B) (Dror and Avraham, 2010). The scala media contains the cochlear sensory epithelium of the auditory system (Brownstein et al., 2013) the organ of corti, which embedded on the basilar membrane that separates the scala media from scala tympani containing specialized cells called hair cells, and supporting cells. The hair cells consist of a single row of inner hair cells (IHC), and three rows of outer hair cells (OHC) (Fig2C).

When sound strikes the tympanic membrane, the vibration transmitted to the cochlea by the bones of the middle ear through the oval window, consequently this generated movement of the fluids through the cochlear duct against the hair cells(Brownstein et al., 2012).

The apical surface of the hair cells contains actin- rich filaments known as stereocilia that form atypical staircase arrangement, connected by network of tip links (Fig 2D). On the top of hair cells embedded by a tectorial membrane which is a collagen rich extracellular matrix (Dror and Avraham, 2010).

Triggering the sensory cells by the vibration caused by sound activates mechanoelectrical transduction and allowing influx of potassium through the transduction channels that are located at stereocilia tips. This process depolarizes the cells.

Development, differentiation and maintenance of this complex machinery contribute in hearing mechanism. Any pathological changes in different parts of hearing apparatus results in HL (Brownstein et al., 2013).

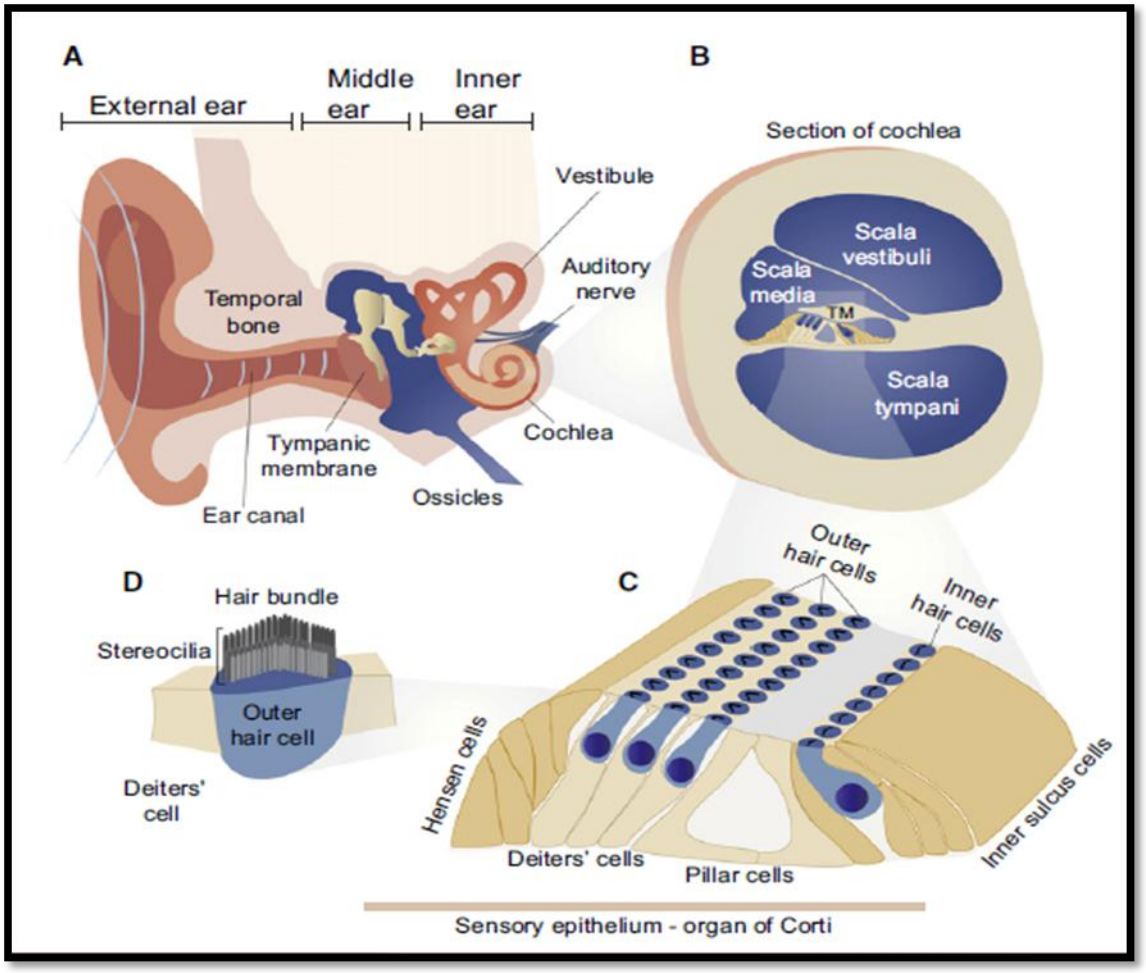


Figure2: Schematic Illustration of the human Inner Ear (Dror and Avraham, 2010).

2.3 Causes of hearing loss

Like other sensory loss, hearing loss can be caused by environmental factors, genetic factors or combination of both (Idan et al., 2013).

2.3.1 Environmental causes:

Acquired hearing loss caused by many factors include exposure to excessive noise, acoustic trauma, effects from different infection and disease and ototoxic drugs (Dror and Avraham, 2009, Dror and Avraham, 2010).

2.3.2 Hereditary causes of hearing loss:

Approximately 60–70% of hearing loss is due to genetic factors and is genetically heterogeneous, with single gene defects in many different genes.

Genetic hearing impairment can present in the form of syndromic hearing loss (SHL), where deafness is associated with other phenotypes; syndromes such as Usher, Pendred, Jervell and Lange-Nielsen syndromes usually include HL (Avraham and Kanaan, 2012). Alternatively, deafness can occur as an isolated disorder in the form of nonsyndromic hearing loss (NSHL), which accounts for about 70% of the cases (Brownstein and Avraham, 2009, Yang et al., 2013).

Hearing loss can be inherited either in a dominant or recessive mode, whether it is syndromic or nonsyndromic(Raviv et al., 2010).Depending on the inheritance mode deafness is classified accordingly: (DFNA) autosomal dominant, (DFNB) autosomal recessive,(DFN) X-linked, Mitochondrial genes, and (DFNM) Modifier genes are genes that modify the expression of other genes in the auditory system and (Goldfarb and Avraham, 2002). ARNSHL is the most prevalent form worldwide, particularly in regions of Consanguinity (Avraham and Kanaan, 2012). It is inherited by 80% of NSHL cases (Friedman et al., 2007). Less abundant forms of NSHL are X-linked and mitochondrial (Avraham and Kanaan, 2012).

The partition between environmental and genetics factors is not always clear. Several studies have shown that mutations in some genes are associated with noise-induced hearing loss (NIHL), increasing the risk of affected individuals to lose their hearing under exposure to high intensity sounds (Dror and Avraham, 2009).

2.4 Genes of hearing loss:

Given the complex process of the hearing mechanism, it should come as no surprise that a large number of genes are implicated in hearing loss. Over 100 loci have been linked to dominant, recessive, and X-linked forms of nonsyndromic HL (NSHL), and over 60 genes have been identified. Some genes are associated with both recessive and dominant forms of deafness and others are associated with both syndromic and nonsyndromic deafness (Figure3) (Idan et al., 2013).

All mapped loci and identified deafness genes are listed and updated regularly at the Hereditary Hearing Loss Homepage (http://hereditaryhearingloss.org) (Goldfarb and Avraham, 2002).

Deafness at the DFNB1locus, which is the first recessive deafness locus identified, can be caused by mutation in *GJB2*gene which encodes the gap-junction protein connexin 26 (Cx26).Today, *GJB2* is the most common deafness-causing gene, since mutation in this genes are responsible for 50% of autosomal recessive nonsyndromic deafness in many parts of the world(Kochhar et al., 2007, Smith et al., 2014).

The genes responsible for nonsyndromic deafness encode many proteins, some of these, gap junctions (GJB2, GJB6), transcription factors (POU4F3, POU3F4, TFCP2L3, PAX3), ion channels (KCNQ1, KCNE1, KCNQ4), molecular motors (MYO6, MYO7A, SLC26A4, Prestin), extracellular proteins (TECTA, OTOA, COLL11A2), and structural proteins (OTOF, DIAPH1).

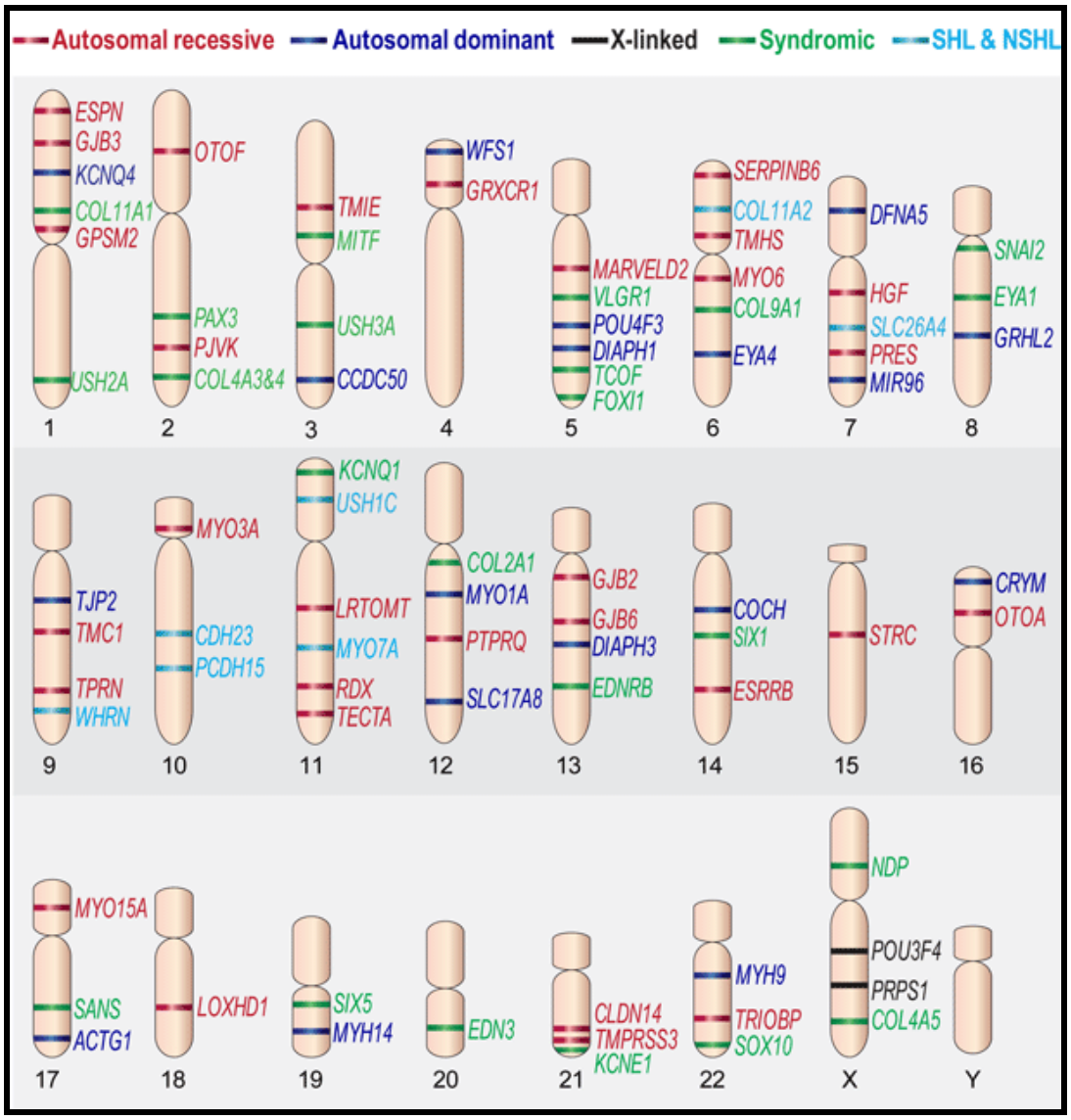


Figure 3:The chromosomal location of selected hearing loss genes (Raviv et al., 2010).

In the Middle East, as in many parts of the world, the most common gene involved in HL is *GJB2*, responsible for 14% of HL cases among Palestinian Arabs (Idan et al., 2013, Brownstein et al., 2011).

2.5 ADGRV1

The *ADGRV1*(adhesion G protein-coupled receptor V1) formerly known as *GPR98* or *MASS1*http://ghr.nlm.nih.gov/gene/ADGRV1, also known as*FEB4;USH2B; USH2C; VLGR1; VLGR1b*(http://www.ncbi.nlm.nih.gov/gene/84059), is one of the largest genes and is found only in vertebrates (Yang et al., 2012).

The *ADGRV1* gene is located on the long (q) arm of chromosome 5 at position 13 (Figure 4). Specifically, it is starts from base pair 90,558,800 to base pair 91,164,223 on chromosome 5,assembly genome reference consortium human build 38 patch release 2 (GRCh38.p2) (http://ghr.nlm.nih.gov/gene/ADGRV1).

It spans 605 Kb, has 90 exons, and encodes a protein with 6306 amino acids(McMillan and White, 2010).

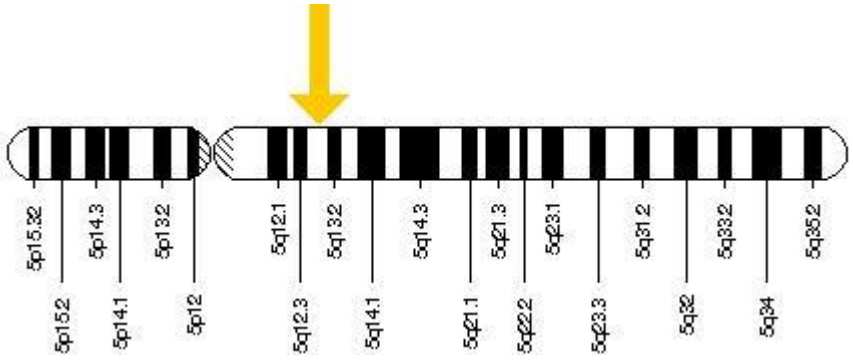


Figure 4: location of *ADGRV1*gene (http://ghr.nlm.nih.gov/gene/ADGRV1).

Several splicing variants are transcribed from the ADGRV1/Vlgr1 locus. Until now eight isoforms of Gpr98/vlgr1 are known, the largest isoform is vlgr1b, with the vlgr1a isoform is restricted to human, partial deletion of exon 31 resulting in vlgr1c Two additional soluble isoform have been identified, vlgr1d is the smallest isoform, and vlgr1e is 1218 amino acids.

Three internal isoforms have been defined, MASS1.1, MASS1.2, MASS1.3, that initiate 5' boundary of the vlgr1 locus and terminate within intron 39 (Figure 5).

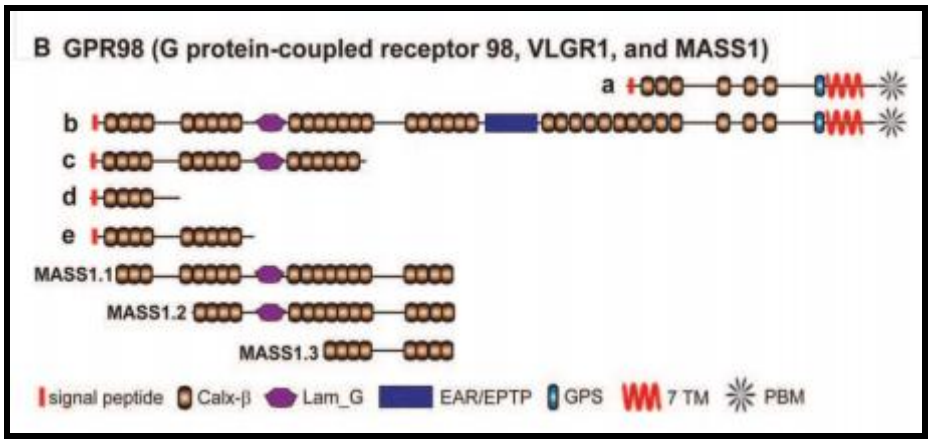


Figure5: Domain structure of VLGR1/MASS1and isoforms(Yang et al., 2012).

2.5.1 Protein structure

The ADGRV1 protein belongs to a member of the secretin family (family 2 or B) of G protein-coupled receptors. It is distinguished by a very large ectodomain consisting mainly of calcium-binding calx-B repeats(McGee et al., 2006).

The full-length isoform Vlgrb is by far the largest GPCR it is approximately 19 kb in size (McMillan and White, 2010), its intracellular domain contains a PBM motif at its terminus along its long extracellular region (Yang et al., 2012)this motif may interact with harmonins PDZ domain (Friedman et al., 2007).

The large ectodomain of vlgr1b contains 35 copies of calx-B motifs, a calcium-binding domain spaced a long the length of its extracellular domain. The function of calx-B was to mediate receptor-ligand interaction, until now a specific ligand for vlgr1b has not been identified. Vlgr1b also contains pentraxin (PTX) domain, the function for this domain in ADGRV1 is suspect as ligand interaction (McMillan and White, 2010).

"the epilepsy associated repeat" (EAR) that occurs as a repeated set of domains that together are resulting to form a seven bladed "beta propeller. This domain termed also as "epitempin"

(EPTP) domain, its function is uncertain but presumed to function as interaction domain(Gibert et al., 2005).

The structural features of the EAR domain and its occurrence in the ectodomain of a receptor suggest a role in ligand binding. Further, the EAR domain is likely to play an important role in the pathogenesis of epilepsy, either by binding to an unknown anti-epileptic ligand, or more likely by interfering with axon guidance or synaptogenesis(Scheel et al., 2002). The carboxyl-terminus of vlgr1 that have a consensus sequence motif recognized by proteins bearing PDZ domain such proteins act as scaffolds for assembling single transduction protein into functional signaling units. C-terminus interacts with two PDZ domain proteins whirlin and harmonin (Gibert et al., 2005)

2.5.2 Speculation on the function of ADGRV1

ADGRV1 is expressed in most adult tissues, with the highest expression found in developing central nervous system during embryogenesis (McMillan and White, 2010).This is considered with another study that show the presence of*ADGRV1* only in chordates and vertebrates(McMillan et al., 2002).

In the cochlea, the expression of*ADGRV1* was identified in the hair bundles of hair cells embedded from the top of the cochlea to the basal turn of it (Yagi et al., 2007). Specifically localized at the base of developing stereocilia in the same places where the ankle links are formed (Figure 6), where it is expressed strongly at the perinatal period and during hair cell development it will decreased (Friedman et al., 2007).

The structural of *ADGRV1* is in agreement with being a major component of the ankle link. The ectodomain of vlgr1b was previously noted to contain a total of 35 calx-B repeats spacing ~120 amino acid. If calx-B repeat unit were globular they would have a diameter of ~ 4 nm. But if these repeats arranged as a string, it would be ~ 180 nm long, so it suggest to span the gap between relative stereocilia in the ankle link region where it is ~ 150 nm (Figure 6)(McMillan et al., 2002, McGee et al., 2006).

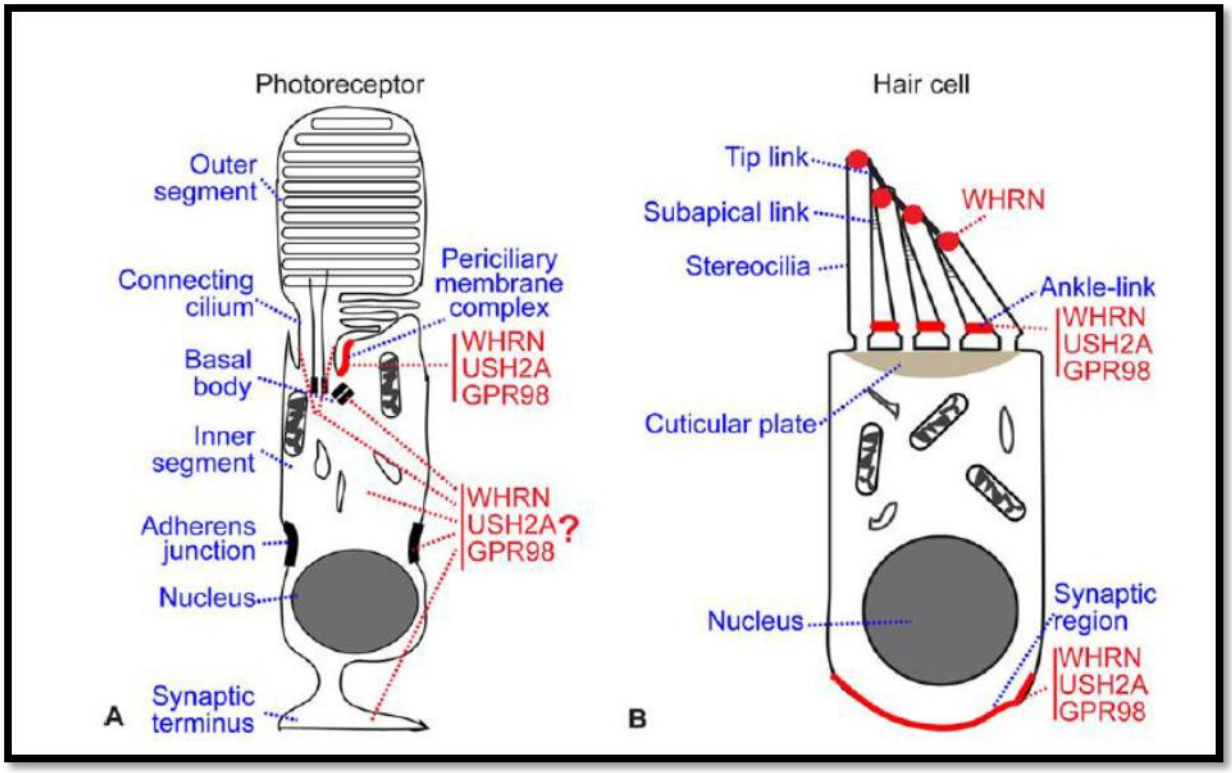


Figure 6: Schematic diagrams show localization of ADGRV1 in photoreceptor and hair cells (Yang et al., 2013)

Mutations in*ADGRV1*lead to many disorders in human like febrile and a febrile seizures, and to Usher syndrome(Gibert et al., 2005).

Usher syndrome is the most common condition of combined blindness and deafness. It is categorized into three clinical subtype differentiated according to severity of hearing loss and presence or absence of vestibular(Weston et al., 2004), Type I (USH1) is manifested as congenital profound deafness as well as vestibular dysfunction; USH2 exhibits congenital moderate hearing loss and normal vestibular function; and USH3 is characterized by progressive hearing impairment and occasional vestibular dysfunction(Yang et al., 2012). Moreover, mutation in ADGRV1 were identified in patients with Usher syndrome type 2(Yang, 2012).Other genes are responsible for Ush2 syndrome, usherin (USH2A), and whirilin (USH2D) (Hilgert et al., 2009, Millán et al., 2010).The products of *USH2A* and*ADGRV1* genes forms ankle links, that interconnect stereocilia at the bases (Jocelyn F.

Krey, 2012) where is the product of whirlin have a role in control of stereocilia actin dynamics. There is also a putative trans-membrane protein called vezatin, which is a part of the ankle link complex (Figure 7) (Hilgert et al., 2009).

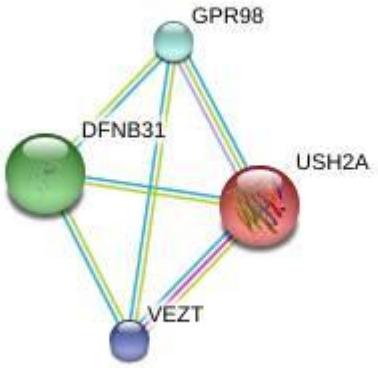


Figure 7: Interacting Proteins for *ADGRV1* Gene (http://www.genecards.org/cgi bin/carddisp.pl?gene=ADGRV1)

Further,*ADGRV1* may have an significant role in development of the mammalian eye(Gibert et al., 2005), where is this gene expressed in photoreceptor cells of the retina (Figure 6) (Millán et al., 2010), and mutations in this gene cause retinitis pigmentosa associated with Usher syndrome(Gibert et al., 2005, Weston et al., 2004).

Mice carrying a targeted mutation in *ADGRV1* (vlgr1/del7TM) revealed that *ADGRV1* is a component of the ankle link complex. In homozygous *vlgr1*/del7TM mice, the hair bundle become disorganized by postnatal day 2 when ankle links are completely formed and ~ 7-8 d after the hair bundle has arise for the first time. This indicate that ADGRV1 is not required for the initial stages of hair bundle morphogenesis, but it has critical role in the development and maintenance of hair bundle structure. Moreover knockout mice in which the transmembrane domain ofADGRV1 was lost there is morphological change of the hair cells(McGee et al., 2006)

Another study shows that the*ADGRV1*mutated mice manifest distortion of the stereocilia, were it is obviously normal at birth and became disturbed at postnatal day 8 (Yagi et al., 2007). Additionally, *ADGRV1*-mutated mice showed abnormal morphology of the inner ear(Yagi et al., 2007).

In conclusion, the loss ofADGRV1 results in morphological defect in stereociliar organization that helps to explain the deafness phenotype.

2.6 Targeted Next Generation Sequencing

Previously, 60 genes were found to be associated with human deafness by using linkage analysis followed by Sanger sequencing. However this technique, has limitation as time- and cost required for gene identification(Idan et al., 2013).

Next-generation sequencing (NGS), also known as massively parallel sequencing (MPS) , has an exceptional upgrade that enables the screening of a huge number of coding genes and also identification of mutation in non-coding region that may affect gene expression.

A lot of challenges in the identification of causative mutation in heterogenic diseases like deafness are decreased by using NGS (Elkan-Miller'r and Avraham, 2013).

While Sanger sequencing gives an output of 120,000 bp (base pairs) per day for approximately $4,000 per mega base (Mb) sequenced. The output of a single MPS machine is more than 30 Gb for the same time and costs approximately $2 per Mb. To conclude NGS can sequence a 3.2 Gb of a single human genome in one day for a far lower cost in comparison to sanger sequencing that would take 73 years and cost $200,000(Shearer et al., 2011, Brownstein et al., 2012).

In NGS process, genomic DNA isolated from the blood, followed by shearing it into fragments, approximately 200bp in size to make a library, next each end of DNA fragment is hybridized to oligonucleotide adapters, then sequencing using one of the massively parallel sequencers such as IIIuminaHiSeq 2000, see figure below (Walsh et al., 2010, Lin et al., 2012).

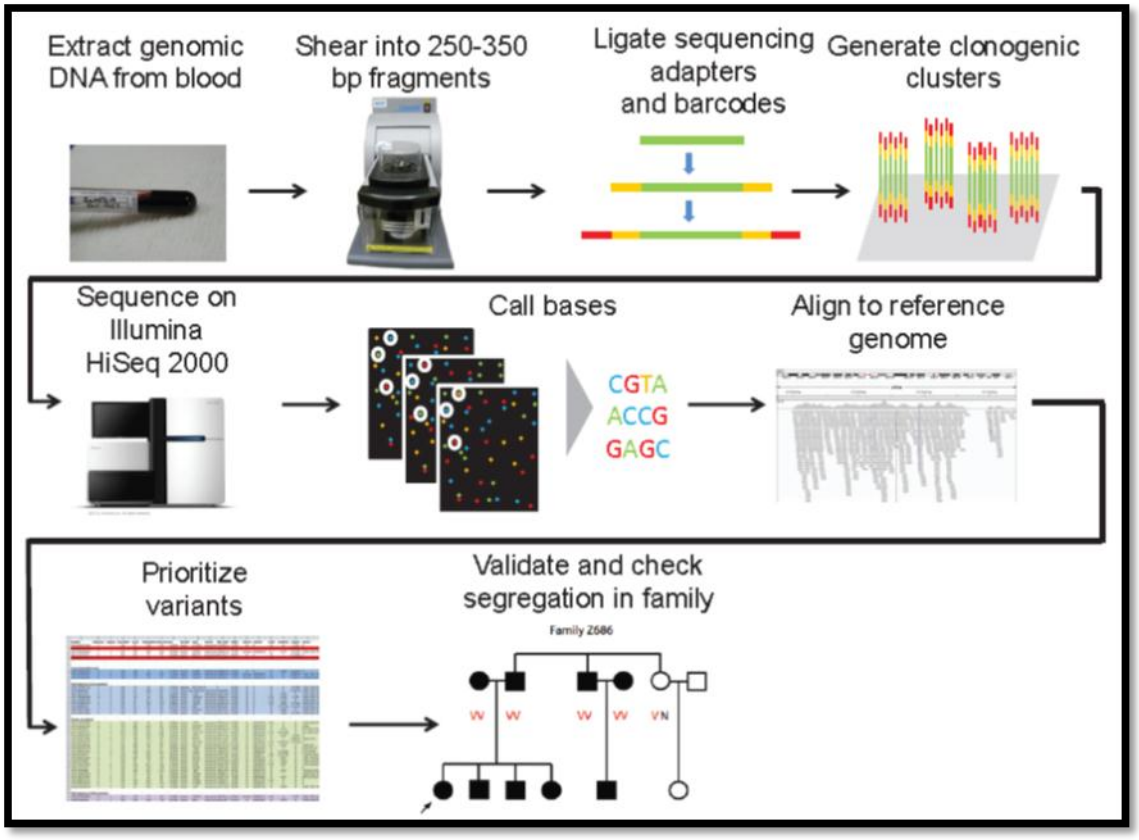


Figure 8:A schematic diagram for identification of the causative mutation for deafness by targeted capture and high-throughput sequencing(Brownstein et al., 2012).

There are three commercial MPS platforms routinely used in the research, including, 454 sequencers (454 Life Sciences, Roche company) using pyrosequencing, the Illumina platform (Illumina, Inc.) using cyclic reversible termination (CRT) technology, and sequence-by-ligation (SBL) technology by Solid platform (Applied Biosystems), the most one used is IIIumina.

The major challenges in the NGS technology are the massive amount of data generated, so there is a need to additional improvement on bioinformatics and analysis tools.

There are two technique narrowing the studied genomic region and can reduce the cost and the needed bioinformatics work this includes targeted genomic capture and whole exome sequencing (Elkan-Miller'r and Avraham, 2013).

In targeted genomic enrichment, also called targeted sequence capture, a portion of the genome is targeted for sequencing(Shearer et al., 2011, Zhang et al., 2011). Targeted genomic enrichment is the most successful MPS clinical platform, this is due to, 85% of pathogenic mutation are found in exon or splice site regions, And thus it makes interpretation easier for the results. Additionally The low amount of sequence that results from decreased amount of data and analysis lead to decrease the cost (Shearer et al., 2011).

By using targeted exomic sequencing eight genes have been discovered including syndromic or nonsyndromic deafness genes. TPRN (Taperin) was the first gene and mutation in it causes nonsyndromic hearing loss at DFNB79 locus. Taperin expressed in a region known as taper, its location exactly at the base of stereocilia of the mouse inner and outer hair cells. When TPRN is mutated stereocilia diameter in this region narrows and weakness and may underlie DFNB79 deafness(Shearer et al., 2011).

Whole exome sequencing (WES) screens the exons of all genes in the human genome, which will lead to discover of novel deafness gene(Brownstein et al., 2013). By using whole exome sequencing combined with homozygosity mapping the huge amount of data are reduced. There is more than one study using this approach, one of these was, a study that discovered a mutation in the gene GPSM2 as the cause of NSHL DFNB82 in the consanguineous Palestinian family (Elkan-Miller'r and Avraham, 2013).

CHAPTER 3

3.1 Problem statements and objectives

Hearing loss (HL), is the most common sensory impairment in human, affects approximately 28 and 22.5 million Americans and Europeans, respectively (Dror and Avraham, 2009). This number rises dramatically in Palestinian population, which has along historical tradition of endogamous marriage (Shahin et al., 2006).

Using linkage analysis more than 60 genes were found to be involved in hearing loss by (Idan et al., 2013).

Despite the great contribution of the conventional methods over the past two decades, many deafness genes remain to be elucidated.

With the development of next generation sequencing (NGS) it is becoming possible to screen for mutation in hundreds of genes for hearing impairment in a relatively short period (Dror and Avraham, 2010).

The main purpose of this study aims to identify the causative gene and mutation in a Palestinian family with congenital recessively inherited HL.

3.2 Specific objectives

1-Whole exome sequencing to identify the mutation responsible for the hearing loss in family CX.

2- Validation of results by checking segregation of the variant/s with the deafness in the affected family.

3-Normal control testing

4- mRNA splicing analysis to determine the pathogenic effect of the mutation discovered.

CHAPTER 4

Materials and methods:

4.1.Clinical evaluation:

A medical history of family (GQ and CX) including hearing loss level and consanguinity was collected and a pedigree was constructed

Figure 9 shows the pedigree of family GQ. The pedigree spans three generations. seven individuals were diagnosed as affected.

Figure 10 shows the pedigree of family CX. The pedigree spans three generations. Two individuals were diagnosed as affected. The study was obtained in accordance with the guidelines of the Bethlehem University Research Committee.

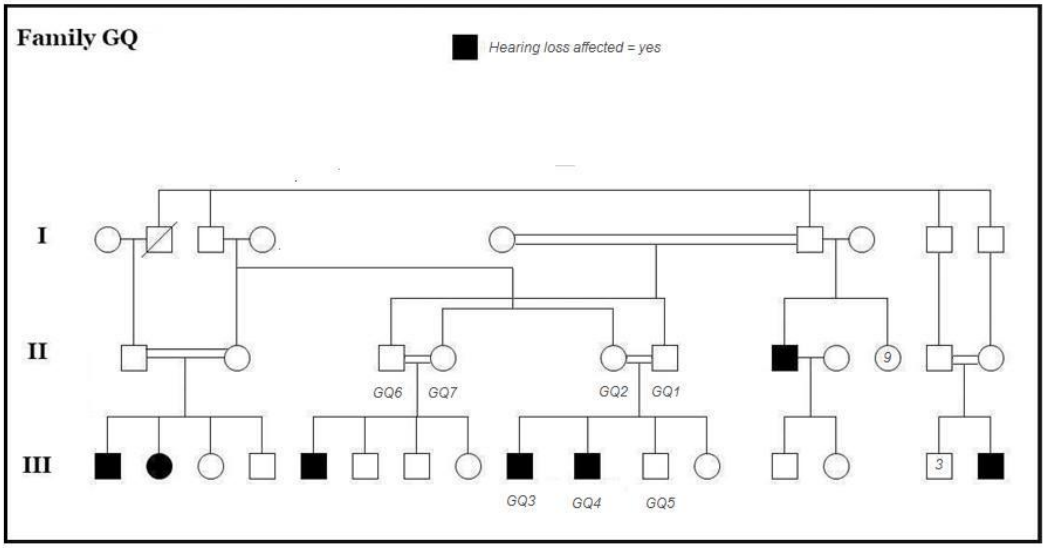


Figure 9: **Pedigree of family GQ**

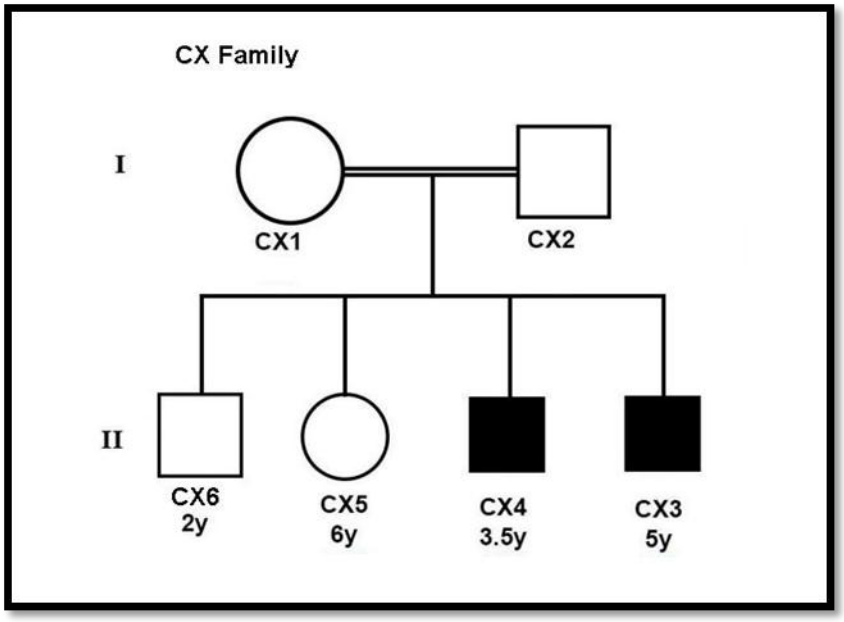


Figure 10: pedigree of family CX.

4.2 Blood collection:

Blood samples were collected from affected and non-affected individuals for family (GQ and CX), in addition to control DNA samples for HL individuals and normal hearing individuals were obtained from the hereditary research laboratory at Bethlehem University.

4.3 Isolation of the DNA by salting-out technique

Approximately 10 ml of the blood was collected in a sterile EDTA vacutainer tube. 20ml of the red blood cell lysis buffer was added, and mixed gently.

Tubes were kept on ice for 30 minutes (min), and then Centrifugation was done at 2000 round per minute (rpm) for 10min at 4˚C. The supernatant was carefully removed and the pellet was re-suspended in 3 ml red blood cell lysis buffer and centrifugation was repeated.

The pellet was then suspended in a mix of 1Xlysis buffer, 100µl of 20%SDS (Amresoc-Cat#M112) and 100µl of 5mg/ml Proteinase K(Amresco-Cat#E115). Incubation was done at 37°Cover night. Then 1 ml of 6M NaCl was added to the lysate and vigorously mixed.

Centrifugation was done at 3000 rpm for 20 min at room temperature, followed by gently transferring the clear upper phase into a 15 ml- tube, avoiding the salt protein deposit.
Two volumes of 100% cold ethanol (EtOH) were added to the upper phase and gently mixed by inverting the tube. DNA was removed with a glass Pasteur pipette, and washed in 70% EtOH (in eppendorf tube) and then air drying for a few minutes on Pasteur pipette. DNA was then dissolved in in 0.02% Sodium Azide(Sigma-Aldrich-Cat# S2002)(200-1000ml depending on the amount of DNA) and left at room temperature overnight

4.4 Mutation screening:

Deaf patient was screened for all known mutations in genes causing HL among the Palestinian population by using sanger sequencing.

4.5 Target genomic capture and massively parallel sequencing

Genomic DNA from (family GQ and CX) only deaf proband (GQ3, CX3) respectively were sheared to approximately 200 bp fragments. Then using Agilent SureSelect Target Enrichment System to design (cRNA) oligonucleotide that cover exons, UTRs, 40bp flanking of 246 genes known to be responsible for either (human or mouse deafness). The final capture design from the 246 genes was 1.43Mb. This 1.43 Mb were sequenced using the Illumina HiSeq 2000. The output data of sequencer was presented as an excel sheet file which contains a list of large number of SNPs and variants.

4.6 Bioinformatics analysis

The data obtained from sequencing of the coding exons of nearly of 246 genes, and 40 bp flanking intronic nucleotides, aligned to the reference human genome (hg19), then SNPs, indels, calls were generated, and variants with read depth less than 10 or score less than 20 were discarded. Rare variants were filtered against their presence in known databases such as dbSNP135,1000 Genomes Project, next using damage prediction web tools that assess the effect of variants on the protein level like **Polyphen2**, **SIFT**. Finally, we had only one mutation that have a damaging effect on the protein level

4.7 Mutation validation

4.7.1 Detection of c.G 2898A by direct sequencing.

4.7.1.1 Polymerase chain reaction (PCR)

920 bp fragments including the c.G2898A mutation was amplified using the following primer pair.

Primer	Primer sequence
FORWARD PRIMER	TCATGGGTGGTTAGTCCAGA
BACKWARD PRIMER	CATGTCTTCAGGGAAACCAAA

Table 4.1: - Standard PCR reaction mix per 25µl of total volume

Reagent	Volume in µl
10X Buffer	2.50
Polymerase (supertherm polymerase)	0.25
Q-Solution	5.00
dNTPs (2.5Mm)	2.00
Forward primer	0.50
Reverse primer	0.50
Nuclease free H2O	1.00
100ng/ µl DNA Template	13.25

Program, (T.D 60 Programs):

94°C	**5 min**	
94°C	30sec	X2
68°C	30sec	
72°C	30sec	
94°C	30sec	X2
66°C	30sec	
72°C	30sec	
94°C	30sec	X2
64°C	30sec	
72°C	30sec	
94°C	30sec	X2
62°C	30sec	
72°C	30sec	
94°C	30sec	X35
60°C	30sec	
72°C	30sec	
94°C	30sec	
4°C	∞	

4.7.1.2 Electrophoresis of PCR product using agarose gels.

Two percent agarose gel containing 0.01% ethidium bromide (Amresco-Cat # E406) was prepared using 1X TAE running buffer. The concentration of agarose dissolved in the TAE was determined by the size of the PCR product. 3µl of PCR product were loaded onto the gel, and run in 1X TAE running buffer at 100V for 30-45 min. DNA fragments were observed using BioRad ultraviolet imaging system. 6µl of 100bp plus DNA molecular weight marker (Thermo Scientific, GeneRuler™, and) were loaded in the first well.

4.7.1.3 Cleaning of PCR Product

Fragments amplified by PCR can be purified directly from the PCR reaction or from the gel after the agarose gel electrophoresis. PCR products were purified directly from PCR reaction using Exonuclease I and Antaractic Phosphatase enzymes.(BioLabs, Cat#M0289L).

Table 4.2: - Master Mix of enzymes per 5 µl of PCR

Reagent	Volume in µl
Exonuclease	I0.25
Antaractic Phosphatase	0.25
Nuclease free H2O	1.5
PCR Product	5

The reaction was completed to a final volume of 7 µl.

4.7.1.4 DNA Sequencing

Direct sequencing of the purified products was done using approximately 10ng of PCR product per 100 bp of the length of PCR fragment. With 10pmol of a single primer was added to 1µlof BigDyeTM Terminators V1.1 Cycle Sequencing Reaction Kit (Applied Biosystems) along with 4 µl of BigDye reaction buffer, and adding ddH_20 to a total volume of 20µl. Sequencing was run using the ABI Prism BigDye Terminator Cycle Sequencing Ready

Reaction Kit (Applied Biosystems, Cat#4337450) and the ABI 3130 DNA Sequencer (Applied Biosystems, S/N: 20355-023).

Table 4.3: - Standard Sanger sequencing reaction mix per 20µl total volume

Reagent	Volume in µl
5x buffer	4
BigDye Terminator V1.1	1
Forward primer	0.5
DNA Template	1
Nuclease Free H2O	13.5

Sequencing PCR Reaction Program:

96 °C	**1min**	
96 °C	10sec	X25
50 °C	5sec	
60 °C	4min	
4 °C	10min	

4.7.1.5 EDTA/Ethanol Precipitation of Cycle-Sequenced Products

For every sequencing reactions 5µl of 125mM EDTA (Amresco-Cat # 0720) and 100µl of absolute ethanol were added and mixed gently. The reaction was incubating at -20°C for 30 min. Then the samples were centrifuged at 3800 rpm for 30 min at 4C°. The supernatant was discarded and 60µl of 70% ethanol were added to each sample and centrifuged again at 3800 rpm for 20 min at 4C°.

The supernatant was discarded and the samples were inverted on tissue paper for, PCR tubes or plate was spin up at 500 rpm for 1 min.

Drying the samples was performed at 95°C for 5min, in order to remove any residual ethanol.16µl of Hi-DiFormamide (Cat # 4311320, Applied Biosystems)were then added, and the samples were dried again at 95°C for 2 min. finally, the reactions were placed on ice for 5min before run on sequencing machine.

4.8 Identification of ADGRV1 splicing variant

4.8.1 RNA extraction from Peripheral Blood

Total RNA was extracted from 5ml of fresh blood using RNeasy® Mini Kit (Qiagen-Cat#74104) according to the manufacturer's protocol.

4.8.2 cDNA synthesis

Total RNA extracted was isolated and was reverse transcribed into first-strand cDNA, using the qScript™ cDNA Synthesis Kit(Quanat Biosciences-Cat#95047-500) according to the manufacturer's protocol.

To determine the effect of the mutation on splicing of mRNA, cDNA from lymphoblast cells of affected individuals (CX4) and a control were amplified and sequenced, using primer pair as seen below:

Primer	Primer sequence
LEFT PRIMER	TTATAAAAGAAGGAGAATCTGTAGAGC
RIGHT PRIMER	CAAGAGGATTATATAAAAGGGCTCA

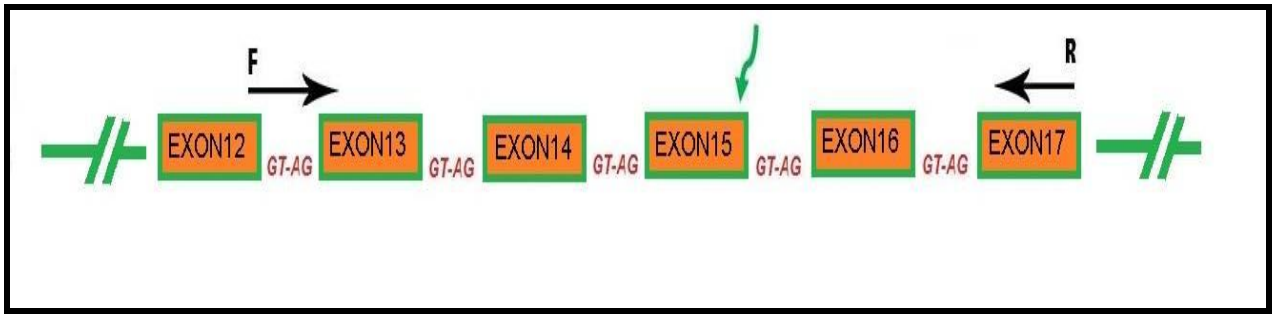


Figure 11: Illustration diagram shows *ADGRV1* splice site mutation (green arrow) and primer positions.

CHAPTER 5

Results:

5.1 Audiogram results:

Audiograms of affected proband (GQ3, GQ4) of family GQ (Figure 12) revealed bilateral mild to severe sensorineural hearing loss (SNHL), and family CX (CX4) (Figure 13), showed moderate to severe right and moderate left ear SNHL .

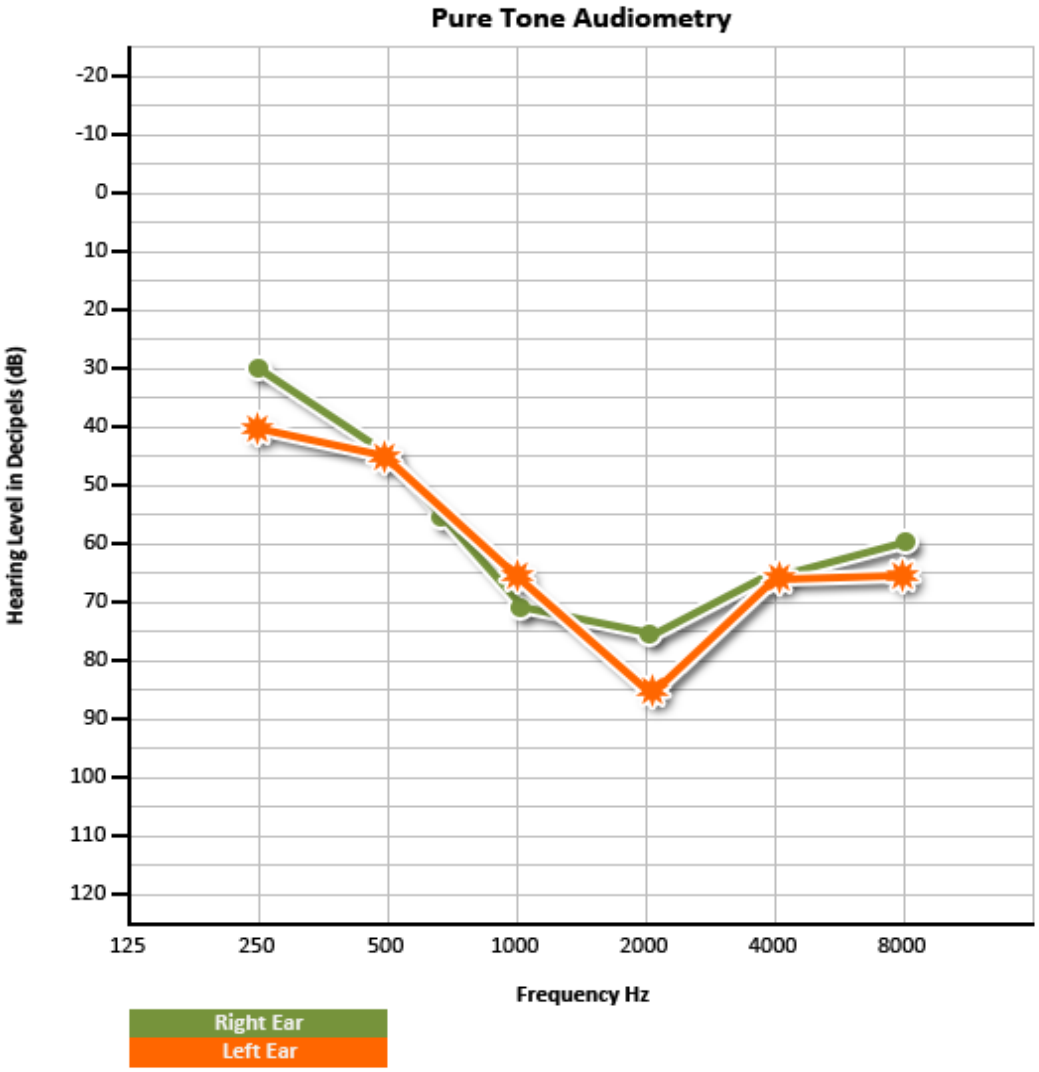


Figure 12: Audiogram of affected individual (GQ3).

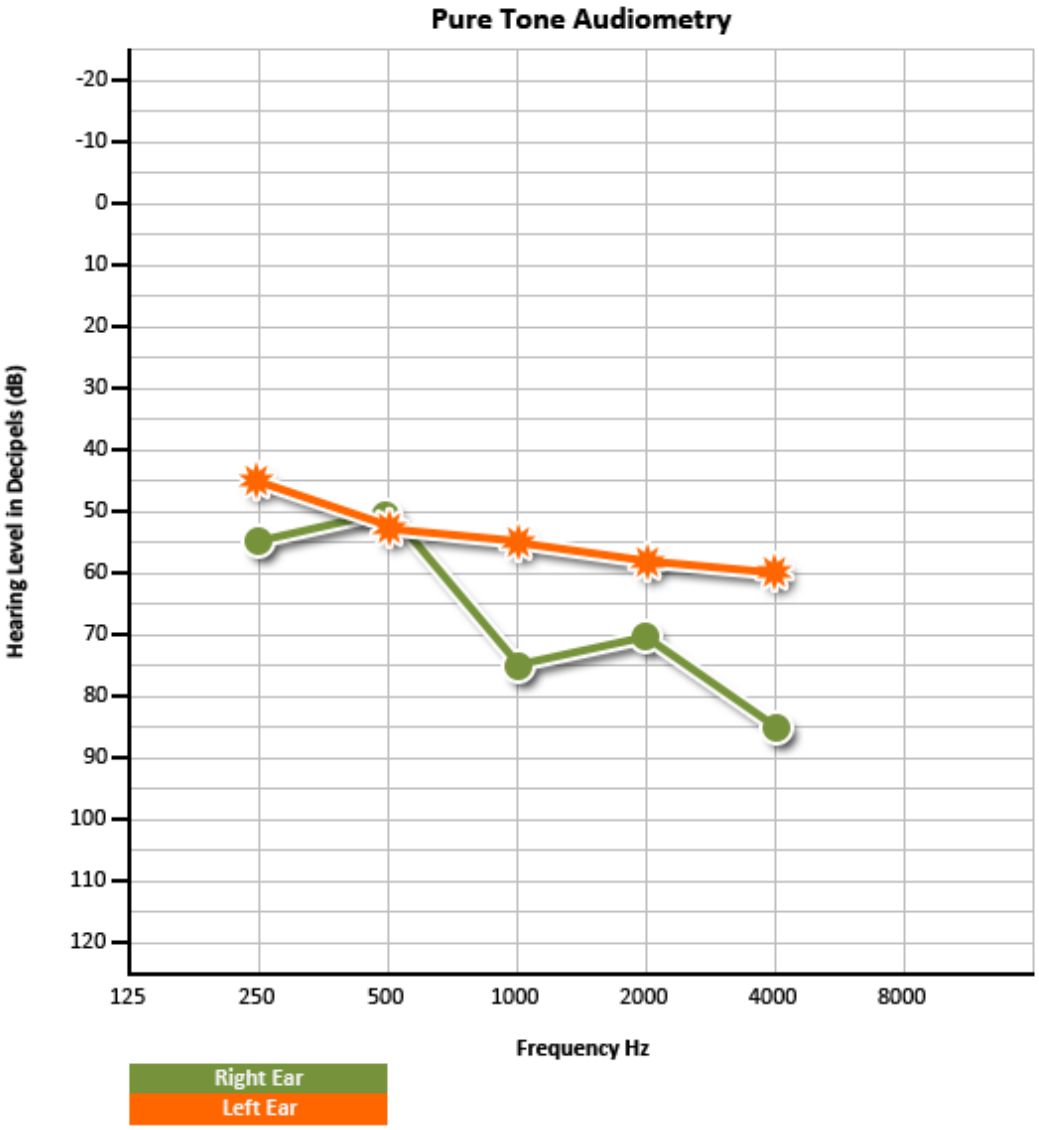


Figure 13: Audiogram of affected individual (CX4).

5.2 Mutation screening result:

The result was negative and none of these mutation that listed in the table that present in the appendices is the causative of HL.

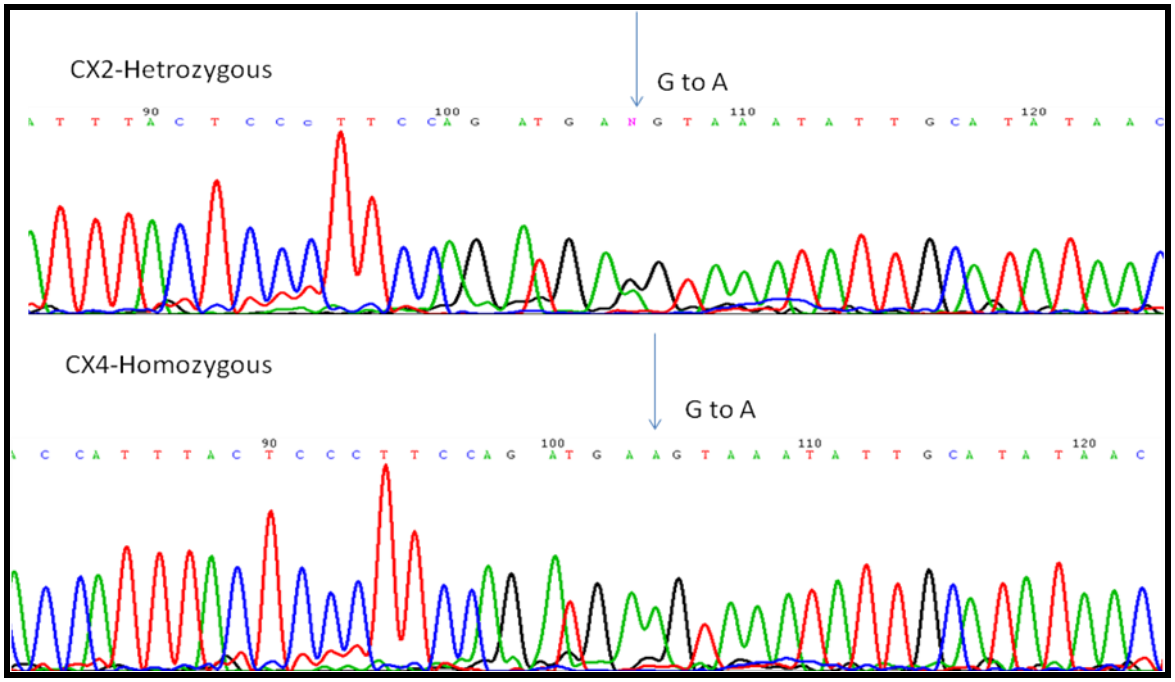
5.3 Targeted genomic capture and MPS results:

Exome sequencing for one affected individuals (GQ3) from GQ family show no result. Whereas in CX family the result of exome sequencing of (CX3) revealed a homozygous donor splice site mutation in *ADGRV1* gene. This mutation c.G 2898A causes the transition of G to A.

5.4 Sanger sequencing results:

Sanger sequencing for the all individuals of the CX family, both affected and unaffected revealed that this mutation c.G2898A segregates with the hearing loss in family CX in a recessive manner. The two affected in this family are homozygous for an c.G2898A allele that lead to a splicing defect, whereas the unaffected parents and siblings are heterozygous for this allele, (Figure 14).

A



B

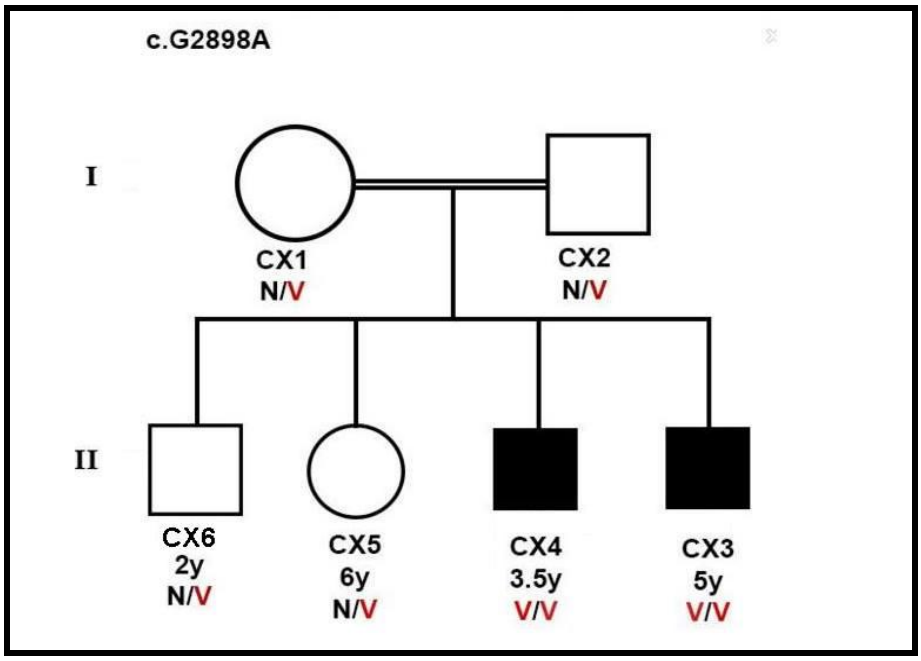


Figure 14: **Verification of splice mutation by Sanger sequencing.** (A) Chromatogram for one affected of the CX family compared with healthy control. (B) Segregation of the mutation with the HL in family CX.

5.5 Sanger sequencing of Palestinian control:

Unaffected Palestinian controls (100 individuals), and 263 unrelated deaf proband were tested wiled type for c.G2898A.E966E mutation.

5.6 cDNA analysis results:

In order to see the effect of the mutation on pre-mRNA splicing the product of cDNA PCR for wild type, affected patient and the carrier visualized on 2 % gel. As shown in figure 15, the product of cDNA PCR for patient CX4 shorter than the wild type, this suggest that the mutation affect the splicing

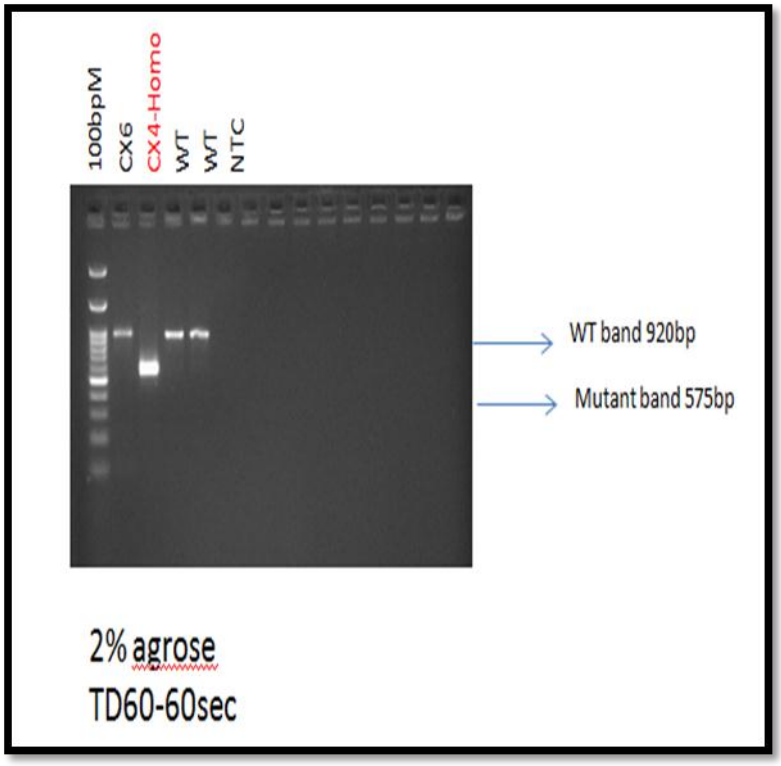


Figure 15: Agarose gel analysis of cDNA-PCR products.

This confirmed by sanger sequencing that revealed abnormal splicing in affected patient (CX4) which is a homozygous for G-A transition at the splice – donor site of exon 15 , which resulting in skipping of exon (14and15) completely, whereas 114 nucleotides are deleted. While in the wild type the splicing occurs normally near the exon 15 (figure 16).

RNA splicing requires conserved sequences which placed at the exon/intron boundaries, start with GT (5') and end with AG (3') of an intron, other nucleotides that found relative to 3' end of the intron, and a sequence that usually placed very close to the end of intron termed branch site. all of these conserved intronic sequence known to be important in splicing. (Figure17a) show that dinucleotide GT-AG at exon 15 will be recognized by splicosome complex and the splicing process occurs normally in the wild type, while in the affected patient, the mutation changes G to A at the splice donor site of exon 15, so it will be undefined by splicosome complex, that is will leading to abnormal splicing of exon 14 and 15 as shown in (figure 17b).

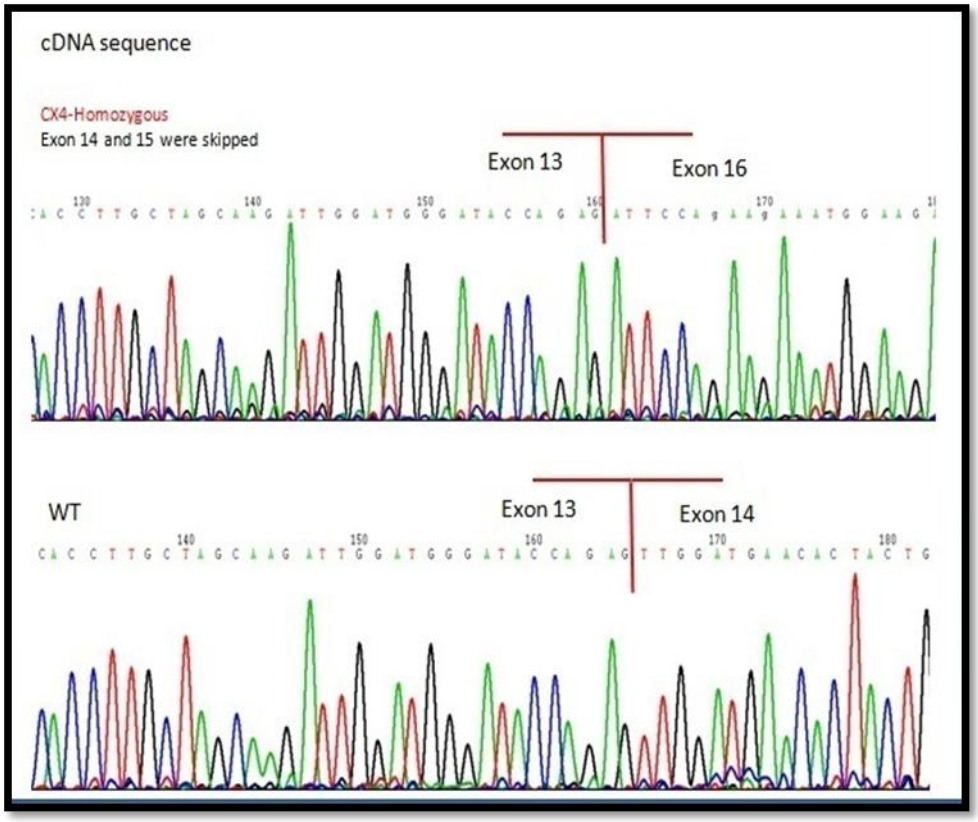


Figure 16: splicing of exon 15 of ADGRV1. Chromatogram of II-I mutant cDNA shows exon 14 and 15 skipping compared to a wild type sample.

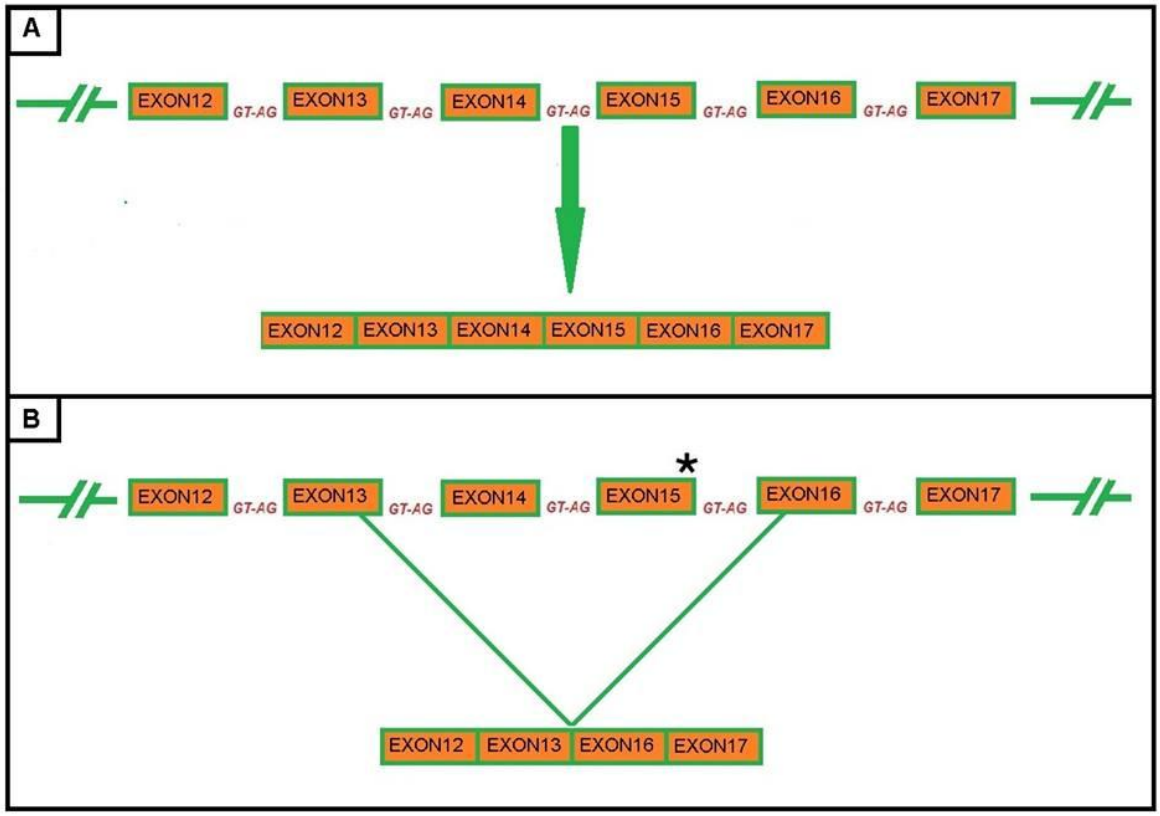


Figure 17. Diagram shows splicing of pre-mRNA in both normal and affected. A. Normal splicing placed between (exon 13 and 14), and between (exon 14 and 15). B. aberrant splicing of pre-mRNA resulting in skipping of exon 14 and 15 completely.

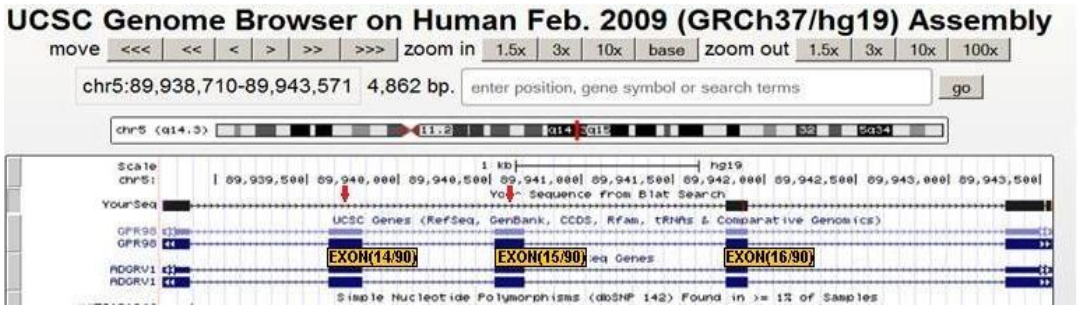


Figure 18: UCSC shows skipping of exon (13and 14) from the sequence.

CHAPTER 6

Discussion

We identified a novel splice site mutation in CX family, that causes the transition of G to A. Two affected individuals in this family are with congenital moderately sever and recessively inherited hearing loss. In addition, the mutation was verified by Sanger sequencing, and showed segregation with hearing loss in CX family. Also 263 unrelated deaf proband and 100 normal controls were screened for *ADGRV1*mutations, and this mutation was not detected among Palestinian probands. These results indicate that the mutation in*ADGRV1* causing deafens in this family.

Beside, using mRNA assay to further verify these finding, PCR amplification was performed around exon 15 using gene specific intron primers flanking exon 15 (Figure11).

The sequence of patient CX3 revealed a G-A transition at the splice-donor site of exon 15, whereas the parents were heterozygous for this base change. Thus, the patient in this family is homozygous for an c.G2898A allele that leads to the skipping of exon (14and15), whereas the healthy parents are heterozygous for the allele.

Several reports of mutation at single-splice junctions that result in the skipping of two or more exon (Takahara et al., 2002, Hori et al., 2013, Fang et al., 2001, Hayashida et al., 1994, Schneider et al., 1993, Yamada et al., 2007, Haire et al., 1997).

(Takahara et al., 2002) propose that the two exon skipping that resulted from mutation at splice site differ depending on the position of mutation at the acceptor site or the donor site. If it was at the acceptor site resulting in deletion of the two downstream exon. It seems that the two exons can be fused as a "single" exon then the next downstream intron and the intron that has a mutation will be stay in the transcript. According to this study the mutation is at a donor site. It will leading to delete the upstream intron that will creates a single exon and if the next upstream intron retained, this resulting in skipping of the two exons upstream from the mutation site (Takahara et al., 2002)

The effect of the mutation on the protein levelwas manifest by deleting 114 residue of ADGRV1 Calx-B domain (Figure18).

Calx-B motif, is a calcium binding domain, which is 35 Calx-B repeats are spaced along the length of vlgrb extracellular domain, which is the longest isoform (McMillan and White, 2010).

Homology to the calx-B repeat is found in the core protein of marine sponge aggregation factor MAFp that form species-specific protein core of large extracellular proteoglycan complex. (McMillan et al., 2002, McMillan and White, 2010).

Calx-B motif also found in integrin bu and in CD97, a related OPCR that have where's the calcium required for interacting the motif with the ligand.

This data suggesting that the function of Calx-B was to mediat receptor-ligand interaction. Mutation in *ADGRV1* are most likely resulted in Usher type 2 that characterized by congenital moderate HL, as well as normal vestibular function and later onset retinitis pigmentosa (RP) (McMillan and White, 2010). Different mutation that found in murine *ADGRV1* gene involved in the same phenotype in vision and hearing. (Yang et al., 2012). Moreover, study was performed on mice shows that no apparent abnormalities in retinas, this suggest that mice with mutation in usher genes do not appear good models to study (RP). This because the phenotype of RP in human with usher syndrome manifest in the teenage years and progress during life, and the mice do not live for long time (McGee et al., 2006)

In this study, there is no indication for RP in affected individuals CX3 and CX4 who are 13 and 11 years old now, respectively. This is could be explained from literature in original gene discovery paper for Ush2C (Weston et al., 2004) that RP onset ranged from ages 17-42 in 4 families' with *ADGRV1* mutation. This suggests that Usher syndrome cannot be ruled out in this family because the HL probands may develop RP in the future.

CHAPTER 7

References

AVRAHAM, K. B. & KANAAN, M. 2012. Genomic advances for gene discovery in hereditary hearing loss.

BROWNSTEIN, Z. & AVRAHAM, K. B. 2009. Deafness genes in Israel: implications for diagnostics in the clinic. *Pediatric research,* 66**,** 128-134.

BROWNSTEIN, Z., BHONKER, Y. & AVRAHAM, K. B. 2012. High-throughput sequencing to decipher the genetic heterogeneity of deafness. *Genome Biol,* 13**,** 245.

BROWNSTEIN, Z., FRIEDMAN, L. M., SHAHIN, H., ORON-KARNI, V., KOL, N., RAYYAN, A. A., PARZEFALL, T., LEV, D., SHALEV, S. & FRYDMAN, M. 2011. Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. *Genome biology,* 12**,** R89.

BROWNSTEIN, Z., SHIVATZKI, S. & AVRAHAM, K. B. 2013. Molecular Etiology of Deafness and Cochlear Consequences. *Deafness.* Springer.

CHOU, R., DANA, T., BOUGATSOS, C., FLEMING, C. & BEIL, T. 2011. Screening for hearing loss in adults ages 50 years and older: a review of the evidence for the US Preventive Services Task Force.

DALTON, D. S., CRUICKSHANKS, K. J., KLEIN, B. E., KLEIN, R., WILEY, T. L. & NONDAHL, D. M. 2003. The impact of hearing loss on quality of life in older adults. *The Gerontologist,* 43**,** 661-668.

DROR, A. A. & AVRAHAM, K. B. 2009. Hearing loss: mechanisms revealed by genetics and cell biology. *Annual review of genetics,* 43**,** 411-437.

DROR, A. A. & AVRAHAM, K. B. 2010. Hearing impairment: a panoply of genes and functions. *Neuron,* 68**,** 293-308.

ELKAN-MILLER'R, T. & AVRAHAM, K. B. 2013. Deafness gene discovery in the genomic era. *ISBN*.

FANG, L. J., SIMARD, M. J., VIDAUD, D., ASSOULINE, B., LEMIEUX, B., VIDAUD, M., CHABOT, B. & THIRION, J.-P. 2001. A novel mutation in the neurofibromatosis type 1 (NF1) gene promotes skipping of two exons by preventing exon definition. *Journal of molecular biology,* 307**,** 1261-1270.

FRIEDMAN, L. M., DROR, A. A. & AVRAHAM, K. B. 2007. Mouse models to study inner ear development and hereditary hearing loss. *International Journal of Developmental Biology,* 51**,** 609.

GIBERT, Y., MCMILLAN, D. R., KAYES-WANDOVER, K., MEYER, A., BEGEMANN, G. & WHITE, P. C. 2005. Analysis of the very large G-protein coupled receptor gene (Vlgr1/Mass1/USH2C) in zebrafish. *Gene,* 353**,** 200-206.

GOLDFARB, A. & AVRAHAM, K. B. 2002. Genetics of deafness: recent advances and clinical implications. *Journal of basic and clinical physiology and pharmacology,* 13**,** 75-88.

HAIRE, R., OHTA, Y., STRONG, S. J., LITMAN, R. T., LIU, Y., PRCHAL, J. T., COOPER, M. D. & LITMAN, G. W. 1997. Unusual patterns of exon skipping in Bruton tyrosine kinase are associated with mutations involving the intron 17 3'splice site. *American journal of human genetics,* 60**,** 798.

HAYASHIDA, Y., MITSUBUCHI, H., INDO, Y., OHTA, K., ENDO, F., WADA, Y. & MATSUDA, I. 1994. Deficiency of the E1β subunit in the branched-chain α-keto acid dehydrogenase complex due to a single base substitution to the intron 5, resulting in two alternatively spliced mRNAs in patient with maple syrup urine disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1225**,** 317-325.

HILGERT, N., SMITH, R. J. & VAN CAMP, G. 2009. Function and expression pattern of nonsyndromic deafness genes. *Current molecular medicine,* 9**,** 546.

HORI, T., FUKAO, T., MURASE, K., SAKAGUCHI, N., HARDING, C. O. & KONDO, N. 2013. Molecular Basis of Two-Exon Skipping (Exons 12 and 13) by c. 1248+ 5g> a in OXCT1 Gene: Study on Intermediates of OXCT1 Transcripts in Fibroblasts. *Human mutation,* 34**,** 473-480.

IDAN, N., BROWNSTEIN, Z., SHIVATZKI, S. & AVRAHAM, K. B. 2013. Advances in genetic diagnostics for hereditary hearing loss. *Journal of basic and clinical physiology and pharmacology,* 24**,** 165-170.

JOCELYN F. KREY, P. G. G. 2012. Deafness Genes. *American Society for Neurochemistry*.

KOCHHAR, A., HILDEBRAND, M. S. & SMITH, R. J. 2007. Clinical aspects of hereditary hearing loss. *Genetics in Medicine,* 9**,** 393-408.

LIN, X., TANG, W., AHMAD, S., LU, J., COLBY, C. C., ZHU, J. & YU, Q. 2012. Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. *Hearing research,* 288**,** 67-76.

MCGEE, J., GOODYEAR, R. J., MCMILLAN, D. R., STAUFFER, E. A., HOLT, J. R., LOCKE, K. G., BIRCH, D. G., LEGAN, P. K., WHITE, P. C. & WALSH, E. J. 2006. The very large G-protein-coupled receptor VLGR1: a component of the ankle link complex required for the normal development of auditory hair bundles. *The Journal of neuroscience,* 26**,** 6543-6553.

MCMILLAN, D. R., KAYES-WANDOVER, K. M., RICHARDSON, J. A. & WHITE, P. C. 2002. Very large G protein-coupled receptor-1, the largest known cell surface protein, is highly expressed in the developing central nervous system. *Journal of Biological Chemistry,* 277**,** 785-792.

MCMILLAN, D. R. & WHITE, P. C. 2010. Studies on the very large G protein-coupled receptor: from initial discovery to determining its role in sensorineural deafness in higher animals. *Adhesion-GPCRs.* Springer.

MILLáN, J. M., ALLER, E., JAIJO, T., BLANCO-KELLY, F., GIMENEZ-PARDO, A. & AYUSO, C. 2010. An update on the genetics of usher syndrome. *Journal of ophthalmology,* 2011.

RAVIV, D., DROR, A. A. & AVRAHAM, K. B. 2010. Hearing loss: a common disorder caused by many rare alleles. *Annals of the New York Academy of Sciences,* 1214**,** 168-179.

SCHEEL, H., TOMIUK, S. & HOFMANN, K. 2002. A common protein interaction domain links two recently identified epilepsy genes. *Human molecular genetics,* 11**,** 1757-1762.

SCHNEIDER, S., WILDHARDT, G., LUDWIG, R. & ROYER-POKORA, B. 1993. Exon skipping due to a mutation in a donor splice site in the WT-1 gene is associated with Wilms' tumor and severe genital malformations. *Human genetics,* 91**,** 599-604.

SHAHIN, H., WALSH, T., SOBE, T., LYNCH, E., KING, M.-C., AVRAHAM, K. B. & KANAAN, M. 2002. Genetics of congenital deafness in the Palestinian population: multiple connexin 26 alleles with shared origins in the Middle East. *Human genetics,* 110**,** 284-289.

SHAHIN, H., WALSH, T., SOBE, T., RAYAN, A. A., LYNCH, E. D., LEE, M. K., AVRAHAM, K. B., KING, M.-C. & KANAAN, M. 2006. Mutations in a novel isoform of TRIOBP that encodes a filamentous-actin binding protein are responsible for DFNB28 recessive nonsyndromic hearing loss. *The American Journal of Human Genetics,* 78**,** 144-152.

SHEARER, A. E., HILDEBRAND, M. S., SLOAN, C. M. & SMITH, R. J. 2011. Deafness in the genomics era. *Hearing research,* 282**,** 1-9.

SMITH, R. J., SHEARER, A. E., HILDEBRAND, M. S. & VAN CAMP, G. 2014. Deafness and hereditary hearing loss overview.

TAKAHARA, K., SCHWARZE, U., IMAMURA, Y., HOFFMAN, G. G., TORIELLO, H., SMITH, L. T., BYERS, P. H. & GREENSPAN, D. S. 2002. Order of intron removal influences multiple splice outcomes, including a two-exon skip, in a COL5A1 acceptor-site mutation that results in abnormal pro-α1 (V) N-propeptides and Ehlers-Danlos syndrome type I. *The American Journal of Human Genetics,* 71**,** 451-465.

WALSH, T., SHAHIN, H., ELKAN-MILLER, T., LEE, M. K., THORNTON, A. M., ROEB, W., RAYYAN, A. A., LOULUS, S., AVRAHAM, K. B. & KING, M.-C. 2010. Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPSM2 as the cause of nonsyndromic hearing loss DFNB82. *The American Journal of Human Genetics,* 87**,** 90-94.

WESTON, M. D., LUIJENDIJK, M. W., HUMPHREY, K. D., MöLLER, C. & KIMBERLING, W. J. 2004. Mutations in the VLGR1 gene implicate G-protein signaling in the pathogenesis of Usher syndrome type II. *The American Journal of Human Genetics,* 74**,** 357-366.

YAGI, H., TOKANO, H., MAEDA, M., TAKABAYASHI, T., NAGANO, T., KIYAMA, H., FUJIEDA, S., KITAMURA, K. & SATO, M. 2007. Vlgr1 is required for proper stereocilia maturation of cochlear hair cells. *Genes to Cells,* 12**,** 235-250.

YAMADA, K., FUKAO, T., ZHANG, G., SAKURAI, S., RUITER, J. P., WANDERS, R. J. & KONDO, N. 2007. Single-base substitution at the last nucleotide of exon 6 (c. 671G> A), resulting in the skipping of exon 6, and exons 6 and 7 in human succinyl-CoA: 3-ketoacid CoA transferase (SCOT) gene. *Molecular genetics and metabolism,* 90**,** 291-297.

YANG, J. 2012. *Usher syndrome: genes, proteins, models, molecular mechanisms, and therapies,* INTECH Open Access Publisher.

YANG, J., WANG, L., SONG, H. & SOKOLOV, M. 2012. Current understanding of usher syndrome type II. *Frontiers in bioscience: a journal and virtual library,* 17**,** 1165.

YANG, T., WEI, X., CHAI, Y., LI, L. & WU, H. 2013. Genetic etiology study of the non-syndromic deafness in Chinese Hans by targeted next-generation sequencing. *Orphanet J Rare Dis,* 8**,** Q110K.

ZHANG, J., CHIODINI, R., BADR, A. & ZHANG, G. 2011. The impact of next-generation sequencing on genomics. *Journal of genetics and genomics,* 38**,** 95-109.

Appendix

No.	Gene	Inheritance	Mutation	Prevalence in deaf population	Experiment	Phenotype	Status
1	**CACNA1D NM_000720**	Recessive	A376V 1127T>C	1 Hom	MPS2	Moderate SNHL	Novel-unpublished
2	**CDH23 NM_022124.2 GI:16507961**	Recessive	P346S 1427C>T	1 Hom	Linkage analysis	Severe to Profound	Shahin et al, 2010[6]
		Recessive	P346L 1428C>T	1Hom	Linkage analysis	Moderate SNHL	Shahin et al, 2010[6]
		Recessive	P559S 2065C>T	1 hom 1 compound	Linkage analysis	Severe to Profound	Shahin et al, 2010[6]
		Recessive	p.D273V c.A818T	1 Hom	MPS3	Severe to Profound	Novel-unpublished
		Recessive	E1059K 3181G>A	1 hom 1 compound	Linkage analysis	Moderate SNHL	
		Recessive	L495P	1 hom	MPS4	congenital profound	
		Recessive	E1917K	1homo	MPS4	HL with RB	
3	**CLDN14 NM_144492**	Recessive	P28L 83T>C	1 Hom	MPS2	Severe to Profound	Novel-unpublished
4	**EPHB2 NM_017449**	Dominant	p.1027_1028del c.3081_3082del	1 Hom	MPS3	Mild to Moderate	Novel-unpublished
5	**ESRRB NM_004452**	Recessive	R182H 545 A>G	1 Hom	MPS2	Severe to Profound	Novel-unpublished
		Recessive	G263S c.787G>A	1 hom	MPS4	profound	
6	**GJB2 NM_004004.5, GI:195539329**	Recessive	IVS1+1 G to A -3172G->A	1 compound	Sanger sequencing	Severe to Profound	Brownstein, 2009[1]
		Recessive	c.35delG	22 Hom 3 compound	Sanger sequencing	Severe to Profound	Sobe, 2000[2]
		Recessive	E120 del 358_360delGAG	1 hom	Sanger sequencing	Severe to Profound	Mani,2009[3]
		Recessive	235DelC	2Hom 3 compound	Sanger sequencing	Severe to Profound	Fuse ,1999
		Recessive	167delT	3 Hom 3compound	Sanger sequencing	Severe to Profound	Sobe, 1999[4]

No.	Gene	Inheritance	Mutation	Prevalence in deaf population	Experiment	Phenotype	Status
		Recessive	W77R 229T->C	1compound	Sanger sequencing	Severe to Profound	Carrasquillo, 1997[5]
7	**GPR98 NM_032119**	Recessive	G1182R 3544G>A	1 hom	MPS2	Moderate to Severe SNHL	Novel-unpublished
8	**GPSM2 NM_013296.4**	Recessive	R127X 875C>T	1 Hom	WES	Severe to Profound	Walsh et al., 2010[11]
		Recessive	W326X 977A>G	1 Hom	MPS2	Severe to Profound	Novel-unpublished
10	**MYH9 NM_002473**	Recessive	p.S1713G c.A5137G	1 Hom	MPS3	Moderate	Novel-unpublished
11	**MYO15A NM_016239.3**	Recessive	Asp2403 fsX2414 7545G>T	1Hom	Linkage analysis	Severe to Profound	Shahin et al , 2010[6]
		Recessive	R3191H c.8183G>A	2 Hom 1compound	Linkage analysis	Severe to Profound	Novel-unpublished
		Recessive	Splice mutation in hypothetical new exon 2A Chr17: 18,026,708, G>A	1 Hom 2compound	Linkage analysis	Moderate	Novel-unpublished
		Recessive	del Glu2769 in exon 45 8309delagg	1 compound	Linkage analysis	Severe to Profound	Novel-unpublished
		Recessive	A408V 1223T>C	1compound	MPS2	no audiograms	Brownstein, 2013[8]
		Recessive	E1414K 4240G>A	2 Hom 1compound	Linkage analysis	Severe to Profound SNHL	Novel-unpublished
12	**MYO6 NM_004999**	Dominant	E299D 897T>G	1 het	MPS2	Late onset (>30), progressive, SNHL	Brownstein, 2013[8]
		Recessive	L926Q 2777A>T	1 Hom	MPS2	Severe to Profound	Brownstein, 2013[8]
13	**MYO7A NM_000260.3**	Recessive	G2123S 6487G>A	1 Hom	Linkage analysis	Severe to Profound	Shahin et al., 2006[9]
		Recessive	Q2066X	1Hom	Linkage analysis	Severe to Profound	Janecke et al., 1999[10]

No.	Gene	Inheritance	Mutation	Prevalence in deaf population	Experiment	Phenotype	Status
		Recessive	Splice mutation in hypothetical new exon 2A Chr17: 18,026,708, G>A 4153-2A>G	2 Hom	MPS2	Severe to Profound	Brownstein, 2013[8]
		Recessive	p.Q2066X c.C6049T	1 Hom	MPS3	Severe to Profound	Novel-unpublished
		Recessive	Q2071X 6211C>T	1 Hom	MPS2	Severe to Profound	Brownstein, 2013[8]
		Recessive	N769fs (c.2307delC)	1 homo	MPS4	severe to profound	
14	**OTOA NM_001161683**	Recessive	D356V 788A>T	1 Hom	Linkage analysis	Moderate	Walsh et al.,2006[7]
		Recessive	Δ 320-550Kb	1 Hom	Linkage analysis	Moderate to severe	Walsh et al.,2006[7]
15	**OTOF NM_194248.2**	Recessive	R577X 4157C>T	1Hom	Linkage analysis		Shahin et al., 2010[6]
		Recessive	p.R1583C c.C4747T	1 Hom	MPS3	Severe to Profound	Novel-unpublished
16	**PAX3**	Dominant	S84F 251T>C	3 Het	MPS2	Waardenburg Syndrome	Zlotogora,1995[12]
17	**POU3F4 NM_000307**	X-linked	R282L 858G>T	1 Het	Sanger Sequencing	Severe to Profound	Novel-unpublished
18	**PTPRQ NM_001145026**	Recessive	Q429X 1285C>T	1Hom	Linkage analysis	Moderate to severe	Walsh et al., 2010[11]
19	**SLC26A4 NM_000441.1 GI:4505696**	Recessive	1001G>T	2 Hom	Linkage analysis	Severe to Profound SNHL	Walsh et al.,2006[7]
		Recessive	ivs11(+1)delG	1 Hom	Linkage analysis	Severe to Profound SNHL	Walsh et al.,2006[7]
		Recessive	F683S 2048C>T	2 hom	MPS2	Moderate	known-unpublished
		Recessive	S399fs 1197delT:S399fs	1 Hom	MPS2	Moderate to Severe	known-unpublished
		Recessive	V239D 716T > A	1 Hom	Linkage analysis	Severe to Profound	Walsh et al.,2006[7]

No.	Gene	Inheritance	Mutation	Prevalence in deaf population	Experiment	Phenotype	Status
		Recessive	K447fs	1 compound het	MPS4	severe to profound w/endolymphatic sac	
		Recessive	Q383fs	1 compound het	MPS4	severe to profound w/endolymphatic sac	
		Recessive	R373C	1 homo	WES-MC_ 2015	severe-profound	
20	**TECTA NM_005422**	Recessive	C1619X 4857C>A	3 Hom	Linkage analysis	Moderate to severe	Shahin et al., 2010
21	**LHFPL5 NM_182548**	Recessive	M1V 1A>G	1 Hom	Linkage analysis		Shahin et al., 2010[6]
22	**TMPRSS3 NM_024022**	Recessive	C194X 783T>A	1 Hom	Linkage analysis	Severe to Profound	Shahin et al., 2010[6]
		Recessive	1190delA	6 Hom	Linkage analysis	Severe to Profound	
		Recessive	Ins(β-sat)+ del	1 Hom	Linkage analysis	Severe to Profound	Walsh et al.,2006[7]
23	**TRIOBP NM_007032**	Recessive	R347X 1039C>T	7 Hom	Linkage analysis	Severe to Profound	Shahin et al., 2006[9]
		Recessive	Q581X 1741C>T	2Hom 1 compound	Linkage analysis	Moderate to severe	Shahin et al., 2006[9]
		Recessive	G1019R 3055G>A	1compound het	Linkage analysis	Moderate to severe	Shahin et al., 2006[9]
24	**USH1G**	Recessive	20-bp deletion	1hom	Linkage analysis	Severe to Profound	
25	**LOXHD1**	Recessive	c.3404insAGCT	1 compound het	MPS4		
		Recessive	G1615R	1 compound het	MPS4		
		Recessive	del18kb (exons 22-25, out of frame)	1 hom	by SG from MPS2		
26	**PTRH2**	Recessive	Q85P	2 hom	WES-MC_ 2015	bilat severe w/mild MR and motor delay	Alazami et al. (2015)
27	**USH1C**	Recessive	USH1C c.103(+5)G>C	2 hom	WES-MC_ 2015	bilat profound	