Bethlehem University

Faculty of Science

Biotechnology Master Program

A Novel Mutation in GPR98 gene detected by Next Generation Sequencing Causes Hearing Loss in a Palestinian Family

By

Abeer H. Anabtawi

In Partial Fulfillment of the Requirements for the Degree of

Master of Science in Biotechnology

April, 2014

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 Mutation in GPR98 Gene Detected by Next Generation Sequencing Causes Hearing Loss in a Palestinian Family"

By

Abeer H. Anabtawi

In partial fulfillment of the requirements for the degree of Master of Science in biotechnology

Graduate Advisory Committee:

Prof. Moien Kanaan, Student Advisor, Bethlehem University — Date: 14/05/2014

Prof. Mazin Qumsieh, Internal Examiner, Bethlehem University — Date: 14/05/2014

Prof. Hisham Darwish, External Examiner, Al Quds University — Date: 19/05/2014

Approved for the Faculties

Dean of Faculty of Scientific Research and Higher studies
Palestine Polytechnic University
22.05.2014
Date

Director of the Master Program
Bethlehem University
19/05/2014
Date

A Novel Mutation in GPR98 gene detected by Next Generation Sequencing Causes Hearing Loss in a Palestinian Family

Abeer H. Anabtawi

Biotechnology Master's Program, Bethlehem University, Bethlehem, Palestine

Abstract

Inherited hearing loss is a genetically heterogeneous disease. Screening patients to detect genes and mutations represents a major challenge. Next generation sequencing is used to overcome this challenge. The primary purpose of this study is to determine the causative gene and mutation in a Palestinian family with congenital recessively inherited hearing loss in three siblings.

The proband was screened for all know mutations causing hearing loss among Palestinian populations. Then, her DNA was subjected to massively parallel sequencing of targeted genomic capture of all suspected hearing loss genes and loci. The data were filtered and analyzed and the mutation was validated via Sanger sequencing and we fully checked its co-segregation with the disease in the affected family. The mutation was further confirmed by testing 100 normal and 263 hearing loss Palestinian controls and functionally validated to determine the effect of the mutation on mRNA level.

We identified a novel exonic splicing mutation in GPR98 gene. It causes the transition of the last nucleotide of exon 49 from G (Guanine) to A (Alanine) at chr5: 90024750. The novel mutation is private to the tested family and it was not detected in tested controls. This mutation leads to abnormal splicing of exon 49. We predicted that it would lead to translation of truncated GPR98 message with a deletion of highly conserved EPTP domain in addition to partial deletion of EAR (3, 4, and 5) repeats.

GPR98 gene is involved in Usher syndrome type 2. Therefore, it is possible that congenital hearing loss in our family is due to Usher syndrome and may develop Retinitis Pigmentosa later in their life. However, symptoms of Retinitis Pigmentosa are still absent in the two elder adolescent sisters and it is likely that this is nonsyndromic condition.

Keywords: Hearing loss, next generation sequencing, heterogeneous, consanguineous, Usher syndrome type 2, Gpr98 gene.

DECLARATION

I declare that the Master Thesis entitled “A Novel Mutation in GPR98 gene detected by Next Generation Sequencing Causes Hearing Loss in a Palestinian Family" 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: Abeer H. Anabtawi

Date:

STATEMENT OF PERMISSION TO USE

Presenting of this thesis is considered as partial fulfillment of the requirements needed to be awarded a Master of Science degree in Biotechnology. 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 absence, by the Dean of Higher Studies. 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.

Name and Signature: Abeer H. Anabtawi

Date:

Dedication

I would to dedicate this thesis to my beloved family

Mom , Dad and Sisters

Thank you for believing in me

Acknowledgment

First and foremost, I would like to thank Almighty Allah most Gracious, who kindly helped me to complete this thesis.

I wish to express my deep gratitude to my advisor Professor Moien Kanaan who offered invaluable supervision, support and guidance.

My sincere thanks to the members of the thesis committee, Professor Mazin Qumsieh and Professor Hisham Darwish; for their remarks, guidelines, careful reviewing and for hard questions.

I would like to thank the staff of heredity research laboratory at Bethlehem university, especially, Amal Abu Rayan who accompanied me during the practical work from the first to the last moment.

I would like to express my love and gratitude to my beloved family; for their understanding, endless love ,continuous support and encouragement through the duration of my study.

Table of Contents

Abbreviations

HL	Hearing loss
IHC	Inner row of hair cells
OHC	Outer rows of hair cells
MET	Mechano-electrical transduction
WHO	World health organization
HOH	Hard of Hearing
dB	Decibel
Hz	Hertz
HHL	Hereditary hearing loss
SHL	Syndromic Hearing Loss
NSHL	Non Syndromic Hearing Loss
OMIM	Online Mendelian inheritance of man
AR	Autosomal recessive
DFN	Deafness locus
GPCR	G protein coupled receptor
GPR98	G protein coupled receptor 98
VLGR1	Very large G protein coupled receptor 1
EAR	Epilepsy associated repeat
EPTP	Epitempin domain
USH2C	Usher syndrome type 2C
RP	Retinitis pigmentosa
NGS	Next generation sequencing
MPS	Massively parallel sequencing
PCR	Polymerase chain reaction
RT-PCR	Reverse transcriptase polymerase chain reaction

List of tables

List of figures

CHAPTER 1: LITERATURE REVIEW

Introduction

"After a lifetime in silence and darkness that to be deaf is a greater affliction than to be blind... Hearing is the soul of knowledge and information of a high order. To be cut off from hearing is to be isolated indeed" (Helen Keller: 1880- 1968)[1]

Sometimes we need to read these words from deaf people to realize that the sense of hearing occupies a prominent center among all other senses due to its extreme importance to the life of every human beings. So how do we hear? and why others are born deaf?. Two questions, scientists ask and apply the scientific methods to reveal the secrets of hearing mechanism, and they are still to that moment discovering a lot of things about this small intricate structure: the human ear.

1.1 Hearing and hearing loss

1.1.1 Hearing Loss: From past to present

"The marriage of deaf people should be avoided" a sentence was said by an Italian physician named Paulus Zacchias (1548-1659). He explained that deaf parents would have deaf children. This reflects early recognition of contribution of hereditary in congenital deafness. It was believed that hearing was the "gate of mind" and therefore congenital deafness would negatively affect child's intelligence (Raviv et al., 2010).

By 1880, it was stated that causes of congenital deafness refers to heredity. Accordingly, in 1883, marriage between deaf people was legally banned. Alexander Graham Bell was one of those who encouraged and defended this decision. Furthermore, he suggested canceling deaf schools, sign language and not to employ disabled people in governmental jobs (Drorr and Avraham, 2009).

Social restrictions on deaf individuals continued even further. The idea of eugenic that calls to improve the quality of human population by discouraging reproduction by

[1] Jean Christie, "Keller, Helen", in the *Gallaudet encyclopedia of deaf people and deafness* (New York: McGraw-Hill, 1987, vol.2, p.125) Retrieved from http://www.audira.org.uk/en/modernising-attitudes-to-hearing-care/where-our-attitudes-to-hearing-come-from/respect-for-hearing/item/16-hearing-our-most-important-sense (date: July 6 ,2013)

persons with genetic defects, made things worse, especially, during the Nazi rule in Germany. 1600 deaf individuals were murdered and 17000 were sterilized (Drorr and Avraham, 2009; Raviv et al., 2010).

Today, the situation has changed. Deaf people are respected and they have rights to live a decent life in most societies. Also, they have specialized schools and associations around the world (Raviv et al., 2010). In Palestine, the deaf are integrated in society with specific laws to protect their rights and with specialized institutions that cater to their educational and vocational needs.

1.1.2 Structure of ear:

Anatomically, the ear is composed of three compartments: the outer, the middle and the inner ear. The outer ear is composed of auricle and external auditory canal and is ended by the ear drum (tympanic) membrane (Petit et al., 2001). The auditory canal is lined with hairy skin that has sweat glands and oily sebaceous gland which together form wax. Hairs and wax in ear canal function as protective barrier and disinfectant that trap foreign objects and prevent them from getting down to the tympanic membrane. The last separates the ear canal from the middle ear and it is the first part of sound transducing mechanism.

The middle ear consists of air cavity that has a chain of three ossicles, the malleus (hammer), the incus (anvil) and the stapes (stirrup). The middle ear is connected to the back of nose by a thin long tube called Eustachian tube. The malleus is attached to the tympanic membrane and stapes are attached to a bony footplate, which is the oval window of the inner ear (Figure1.1a) (Alberti, 2001).

The inner ear is enclosed in the temporal bone. Its location and structure provides protection for this sensory organ and it forms an auditory chamber that helps in catching sound with a very low intensity (Drorr and Avraham, 2009).

The inner ear is composed of two systems: the vestibular system, represented by five vestibular end organs which are accountable for balance and the auditory system, which is the cochlea that is responsible for sound sensation.

The cochlea is a snail shaped fluid filled coiled tube. It is composed of three compartments (scalae). Scala vestibuli and scala tympani are large perilymphatic filled cavities that surround scala media which is an endolymph filled cavity (Figure 1.1b) (Dror et al., 2010).

In the scala media, the corti organ is located, which is the sensory epithelium of the cochlea and it is embedded on the basilar membrane. It has specialized sensory cells known as hair cells, arranged in one inner row of hair cells (IHC) and three outer rows of hair cells (OHC) (Figure 1.1c) (Dror et al., 2010; Drorr and Avraham, 2009).

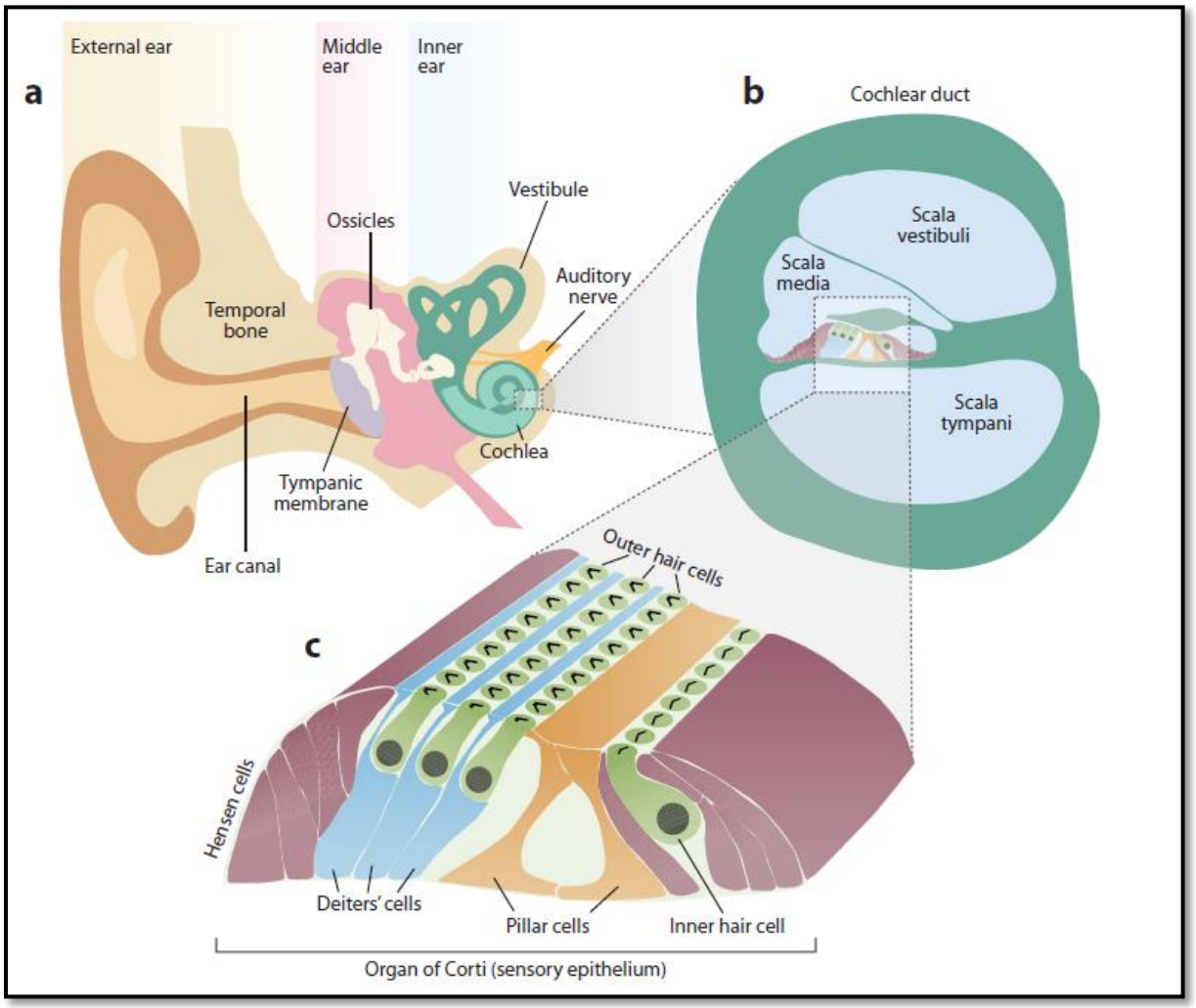


Figure 1.1 Anatomical structure of ear. (a) The ear is composed of the outer, middle and inner ear. (b) Cross section of cochlea duct demonstrating the three fluid filled compartments. (c) Organ of corti that lies on scala media. Hair cells are arranged of inner row of hair cells and three outer rows of hair cells (Drorr et al., 2009).

The apical surface of each hair cell has actin rich projections known as stereocilia (Drorr and Avraham, 2009). They are arranged into bundles of rows of increasing heights forming a stair case arrangements stabilized via lateral, ankle and tip links (Figure 1.2) (Rzadzinska et al., 2004).The hair cells are covered by tectorial membrane, a collagen rich extracellular matrix.

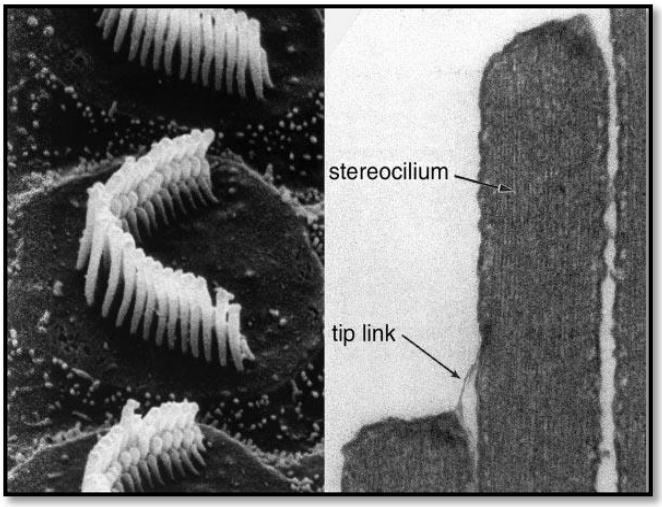


Figure 1.2 Stereocilia. Electron scanning microscopy shows stereocilia arrangement as bundles of rows of increasing height (left side). Within the hair bundle, the apex of each stereocilium is linked to the adjacent taller one via tip links (right side)(Petit et al., 2001).

1.1.3 Mechanisms of hearing:

The auricle of the outer ear collects sound waves that are conducted through the ear canal to reach the ear drum membrane. The vibration is transferred by middle ear bones to the bony footplate of the oval window of the inner ear. Once they move, they cause movement of fluid inside the inner ear (Drorr and Avraham, 2009).

In the cochlea, when sound is induced, mechanical energy of sound causes movement of fluids through the cochlear duct and consequently the basilar membrane with the sensory epithelium vibrate (Drorr and Avraham, 2009). This leads to the movement of stereocilia of the outer and inner hair cells. Bending of stereocilia increases the tension of tip links that stimulate opening of the Mechano-electrical transduction (MET) channels found at stereocilia tips across the bundle. Once they open, an influx of potassium and calcium ions depolarizes the cell, and as a result, it triggers the release of neurotransmitters at the base pole of hair cells, which activate the auditory nerve (Figure1.3) (Dror et al., 2010).

After this action, the endocochlear potential must be reestablished. The potassium ions are recycled and constantly supplied to endolymph, by the action of the sodium/potassium pumps in the stria vascularis cells (Figure 1.4) (Goldfarb et al., 2002).

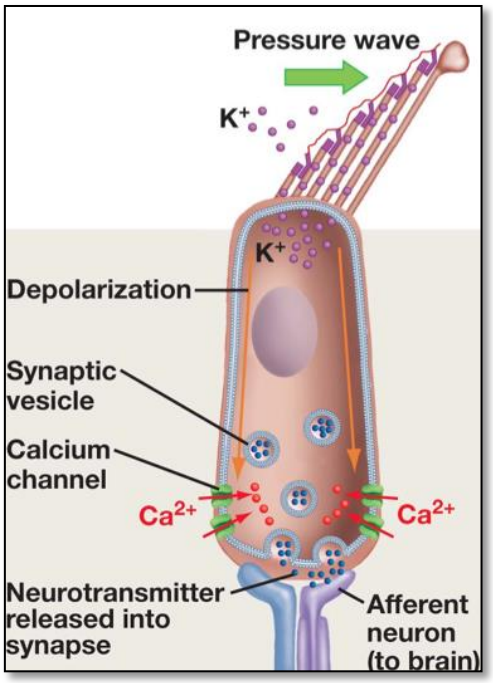


Figure 1.3 Depolarization of hair cells. Bending of stereocilia stimulates opening of MET channels found at the tip of stereocilia. That causes influx of K^+ and Ca^{+2} ions and depolarizes the cells, which stimulates release of neurotransmitter at synapse.Source: http://www.studyblue.com/notes/note/n/ch-46-animal-sensory-systems-and-movement/deck/3080378

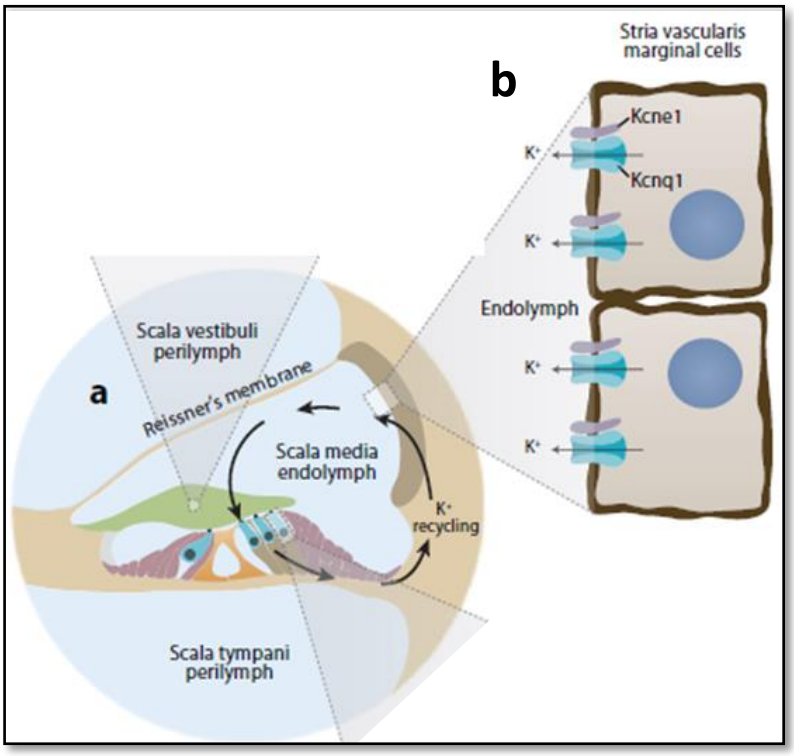


Figure 1.4 The stria vascularis cells. (a) They are located at the lateral wall of scala media. (b) They supply K^+ to the endolymph as well as maintaining endocochlear potential (Drorr et al., 2009).

Proper structure and function of all elements described previously are needed to be able to hear, otherwise hearing loss would occur (Goldfarab et al., 2002).

1.1.4 Hearing loss:

Most references classify hearing loss by type, onset, severity and frequency.

Type: there are four types of hearing loss

- Conductive hearing loss: occurs due to problems in the outer and/or middle ear.
- Sensorineural hearing loss: occurs due to defects in the inner ear organs.
- Mixed hearing loss: occurs due to combination of both conductive and sensorineural hearing loss.
- Central auditory dysfunction: occurs due to dysfunction at the level of eighth cranial nerve, auditory brain stem, or cerebral cortex (Figure 1.5).

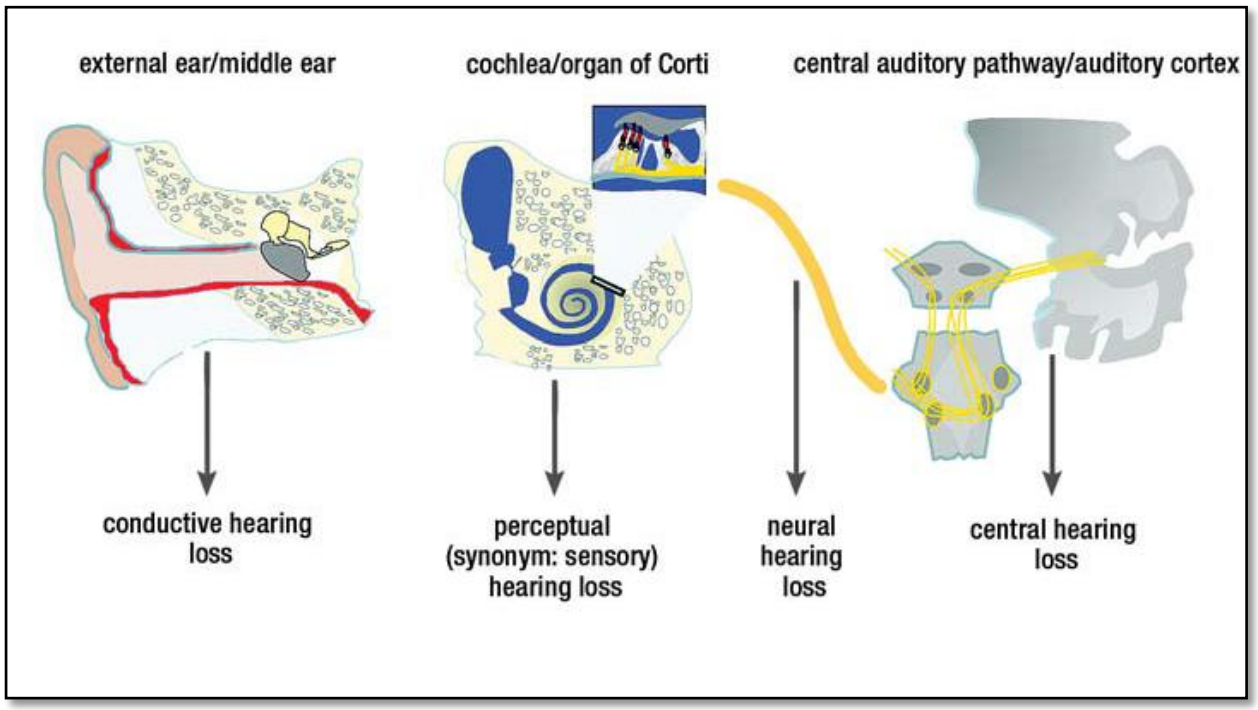


Figure 1.5: Classification of hearing loss according to type. (From left to right: conductive, sensory, neural and central hearing loss) (Zahnert, 2011).

Onset: onset of occurrence of hearing loss can be prelingual or postlingual.

- Prelingual hearing loss: is hearing loss that exists before speech acquisition. And when hearing loss presents at birth, it is termed as congenital hearing loss and it is considered prelingual.
- Postlingual hearing loss occurs after the usual time of speech development.

Severity:

Severity of hearing loss is the degree of hearing loss. It refers to the quietest sound that an individual is able to hear**.** Hearing is measured in decibels unit (dB). The threshold of normal hearing individual occurs within range of 0 -20 dB of normal threshold. Severity of HL of an individual is determined according to the following gradient (Figure 1.6).

- Mild (26–40 dB)
- Moderate (41–55 dB)
- Moderately severe (56–70 dB)
- Severe (71–90 dB)
- Profound (90 dB)

Frequency: hearing loss frequency is determined as follows.

- Low: below 500 Hz.
- Middle: ranges between 501-2000 Hz.
- High: higher than 2000 Hz.

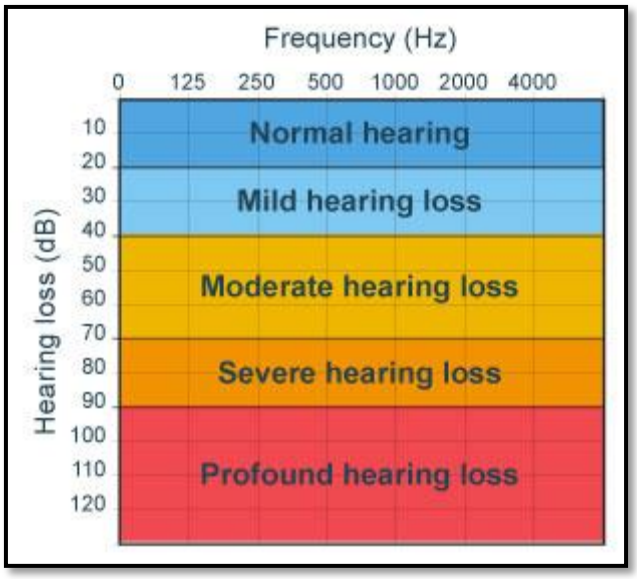


Figure 1.6 Degrees of hearing loss. It is shown as an audiogram which is a plot of an individual's hearing thresholds. The horizontal axis represents the frequency ranges. The vertical axis represents the level of hearing in Decibels. Normal threshold ranges from 0 to 20 dBs. Source: http://www.cochlea.org/en/treatments/

1.1.4.2 WHO Grades of hearing impairment:

WHO classification of hearing impairment according to pure tone audiogram is shown in table 1. It categorized hearing impairment into five sets ranging from no impairment to profound impairment according to the hearing threshold level. It is taken using audiometry for four frequencies 500, 1000, 2000 and 4000 hertz (Mathers et al., 2000).

Table 1.1 The WHO classification of hearing impairment.[2]

Grades of hearing impairment	Audiometric ISO value (average of 500, 1000, 2000 and 4000)	Clinical findings	Recommendations
0 – No impairment	25 dB or better	No or very slight hearing problems. Able to hear whispers.	Counseling, follow-up examination; if conductive hearing loss is present, evaluate indication for surgery
1 – Slight impairment	26–40 dB	Able to hear and repeat words spoken in normal voice at 1 meter.	Counseling, hearing aids may be advisable; if conductive hearing loss or combined hearing loss is present, surgical treatment may be indicated
2 – Moderate impairment	41–60 dB	Able to hear and repeat words spoken in raised voice at 1 meter.	Hearing aids recommended; if conductive hearing loss or combined hearing loss is present, surgical treatment may be indicated
3 – Severe impairment	61–80 dB	Able to hear some words when shouted into better ear.	Hearing aids needed; if an external hearing aid is not possible, consider an implanted hearing aid or cochlear implant; lip-reading and signing for supportive treatment
4 – Profound impairment including deafness	81 dB or higher	Unable to hear and understand even a shouted voice.	Failure of a hearing-aid trial is now usually considered an indication for a cochlear or brainstem implant; lip-reading and signing can be taught in addition

[2] www.who.int/pbd/deafness/hearing_impairment_grades/en/index.html Retrieved on June 26,2013

1.1.4.3 Other terms associated with hearing loss are shown below:

Bilateral versus unilateral: if hearing loss occurs in both ears then it is bilateral hearing loss. And if it occurs in one ear, and the other ear is normal then it is unilateral hearing loss.

Progressive versus sudden hearing loss: if hearing loss is getting worse over time, then it is termed as a progressive. While sudden hearing loss occurs quickly and spontaneously.

Symmetrical versus asymmetrical: if the severity of hearing loss across different frequencies are the same in each ear then it is termed as symmetrical. While, if severity of hearing loss across frequencies is different in each ear then is termed as asymmetrical.

Fluctuating versus stable hearing loss: if hearing loss is unstable, where, sometimes it is getting better while sometimes it is getting worse then it is termed as fluctuating. While, stable hearing loss is not changing over time (American speech language hearing association, 2011).

1.2 Epidemiology and causes of hearing loss

1.2.1 Prevalence of Hearing loss:

WHO has estimated in 2014 that more than 5 % of the world's population has disabling hearing loss, 360 million people, (328 million adults and 32 million are children). Most of these people live in countries with low or middle income. (WHO, 2014)

1.2.2 Causes of Hearing loss

Hearing loss may occur due to environmental factors, genetic factors or combination of both.

1.2.2.1 Environmental hearing loss

Nearly 40% of deafness cases have environmental causes. In children, acquired hearing loss result from prenatal infection of TORCH organisms including toxoplasmosis,

rubella, cytomegalovirus and herpes. Postnatal infection caused by bacteria meningitis, fetal distress, hyperbilirubinemia (high amounts of bilirubin in the blood) or ototoxic drugs also leads to hearing loss (Kochhar et al., 2007).
While, in adults acquired hearing loss occurs due to exposure to loud noise and it is known as noise induced hearing loss (NIHL). Moreover, many cases of acquired hearing loss are age related, known as age related hearing loss (ARHL) or presbycusis (Drorr and Avraham, 2009). Predisposition of individual to hearing loss in the latter two cases reflects an environmental- genetic interaction (Kochhar et al., 2007).

1.2.2.2 Hereditary hearing loss:

More than 60% of hearing loss (HL) cases are inherited (Brownstein and Avraham, 2009). Hereditary hearing loss (HHL) can occur in two forms; syndromic HL (SHL), in which there are other medical problems in addition to HL, and non syndromic HL (NSHL), where HL is the only obvious medical problem (Lenz et al., 2011).

Hearing loss can be inherited in various mode of inheritance, as well as autosomal recessive, autosomal dominant, X-linked and mitochondrial inheritance.

1.2.2.2.1 Syndromic hearing loss (SHL):

Nearly 30% of HHL cases are SHL. There are more than four hundreds syndromes associated with hearing loss described in the Online Mendelian inheritance of Man (OMIM) database. Some of the common syndromes are: Alport, Branchio-oto-renal, Jervell & Lange-Nielsen, Norrie, Pendred, Stickler, Teacher Collins, Usher, Waardenburg and Perrault syndromes (Shen and Deskin, 2004).

The different gene loci for SHL have been named with letters indicating the syndrome. For instance, the genetic locus of Usher syndrome is USH, and it is followed by a number, refers to the order in which the subtype of the disease was discovered (for example: USH1A-F, 2A-D, USH3) (Avraham and Kanaan, 2012).

1.2.2.2.2 Non Syndromic Hearing Loss (NSHL)

NSHL accounts for approximately 70% of HHL cases (Farpon and Ba˜nales, 2011).

It is always designated to single gene defect and the altered genes could be inherited in several pattern mentioned previously (Mahbobi et al., 2012).

Autosomal recessive (AR) mode is the most common mode, and it represents 80% of NSHL cases. While, nearly 18% follows autosomal dominant (AD) pattern, and 2% correspond with X-linked and mitochondrial patterns (Brownstein and Avraham, 2009). The different gene loci for NSHL have been named DFN (derived from deafness) and followed by a number refer to chronological order for their discovery. DFNA refers to AD loci, while DFNB refers to AR loci, and DFNX stands for X-linked loci (Farpon and Ba˜nales, 2011). DFNM stands for modifiers of other genes (Mahbobi et al., 2012). Other definite symbols are used for different forms of hearing loss as well as (AUNA) auditory neuropathy, (OTSC) otosclerosis and mitochondrial (MRTNR, MTTS) genes (Dror et al., 2010).

1.2.3 Deafness genes:

Hearing loss (HL) is genetically highly heterogeneous, whereas, many different genes are involved in HL cases (Figure 1.7) (Lenz et al., 2011). A number of genes responsible for syndromic and non-syndromic deafness have been identified in humans. And they are found in a specified database specialized for hereditary hearing loss. (http://hereditaryhearingloss.org/).

In SHL, 32 genes at least have been identified (Raviv et al., 2010). While In non-syndromic deafness cases, more than 130 loci have been mapped and 46 genes have been discovered to date. They include twenty seven *DFNB* genes, twenty two *DFNA* genes; seven genes can cause dominant and recessive forms of deafness and only one X-linked gene and two mitochondrial genes (Browenstein and Avraham, 2009).

There is also number of genes that are associated with both SHL and NSHL (Shalit and Avraham, 2008). For example, SLC26A4 gene is associated with autosomal recessive NSHL (DFN4) and Pendred syndrome that causes congenital deafness and thyroid goiter. Also WFS1 gene is associated with autosomal dominant NSHL (DFNA 6/ 14/38) and autosomal recessive Wolfram syndrome or DIDMOAD that causes diabetes insipidus, diabetes mellitus, optic atrophy and deafness (Bitner-Glindzick, 2002).

Most of genes that cause NSHL encode proteins accountable for proper development, structure and function of the cochlea (Farpon and Ba˜nales, 2011). These genes are classified into four functional groups: transcription factors, extracellular matrix composition, ion homeostatis and hair bundle morphogenesis. However, the proper function of some genes is yet to be identified (Hilgert et al., 2009).

DFNB1 locus is responsible for the most common form of autosomal recessive NSHL. It is related with connexin 26 GJB2 mutations. Whereas, more than 50% of hereditary hearing loss cases are genetically diagnosed due to mutations in the connexin 26 (GJB2) gene. The last encodes the protein connexin which is vital for the accurate function of gap junctions between cells (Shalit and Avraham, 2008).

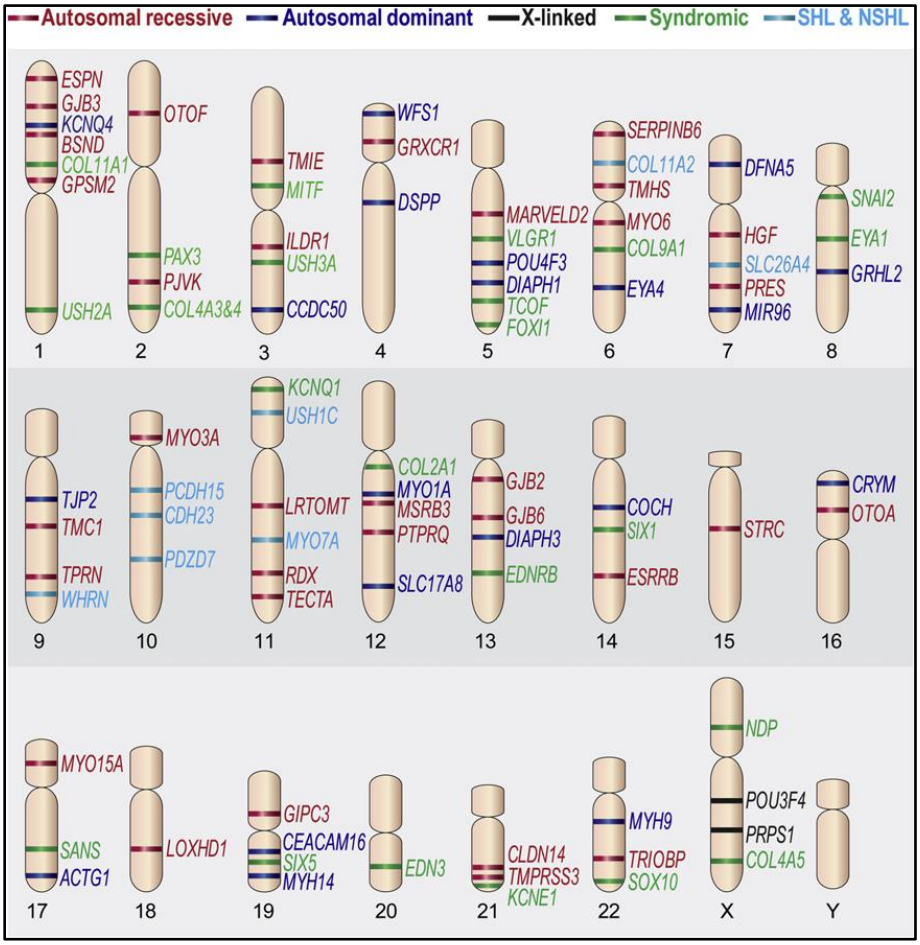


Figure 1.7 Genetic heterogeneity of hereditary hearing loss. Genes are classified as autosomal recessive (red), autosomal dominant (blue), x-linked (black), syndromic (green) and genes that causes both SHL and NSHL (light blue) (Lenz et al., 2011).

Most of the identified mutations in HL genes affect the inner ear and cause sensorineural HL, while only few numbers of them are responsible for middle ear conductive HL (Drorr et al., 2009). Figure 1.8 shows number of these genes and the location of proteins encoded in the cochlea.

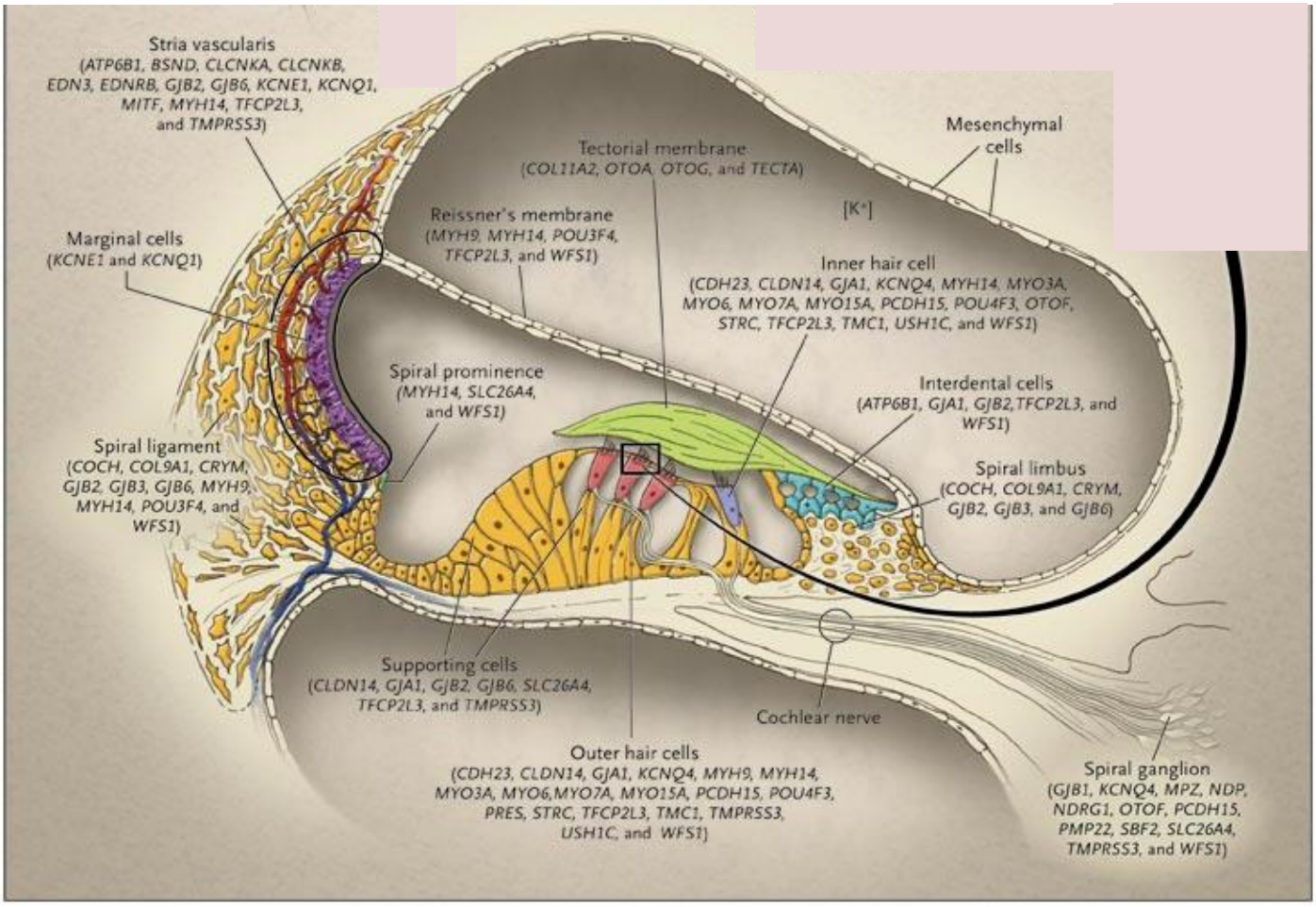


Figure 1.8 Cross section of cochlea indicating responsible genes and proteins. Number of genes involved in deafness and location of protein they encode are shown (Morton and Nance, 2006).

To conclude, the genetics of hearing loss is complex. In the following sections, I have concentrated on GPR98 gene and the discovery of a novel mutation in this gene underlying hearing loss in a Palestinian family.

1.3 Gpr98 gene and its role in hearing loss

The *Gpr98* (G protein coupled receptor 98), *VLGR1* (Very large G protein- coupled receptor I) and *MASS1* (Monogenic Audiogenic Seizure Susceptibility 1) are three different names used for this gene.

1.3.1 Gene structure:

According to GeneCards database; *Gpr98* is located in the q arm of chromosome 5 at position 14.3 (figure 1.9). More specifically, it starts from base pair 89,854,617 to base pair 90,460,033, and it is about 605,417 bp ("Gpr98 Gene-Genecards", 2013).

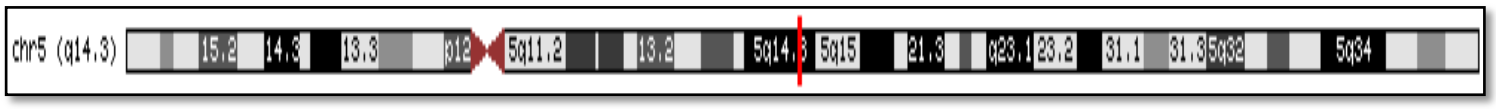


Figure 1.9 GPR98 genomic location. Source: http://genome.ucsc.edu/

GPR98 is one of the largest genes and it is found only in vertebrates. (Yang, 2012) It spans 605 Kb of chromosome 5 and its primary RNA transcript has 90 exons and encodes 6307 amino acids. This gene is highly conserved between human and mouse genomes. In mice, it also has 90 exons, spans 538 kb of chromosome 13 and encodes about 6299 amino acids residues (McMillan and White, 2010).

1.3.2 Splice variants and protein isoforms:

Organization of exons and introns allows alternative m-RNA splicing that result in varied protein isoforms. To date, eight isoforms of Gpr98/Vlgr1 are known and they are labeled as Gpr98-a to Gpr98-e, and Mass 1.1 to Mass 1.3. Vlgr1b is the largest isoform and it encodes about 6309 amino acids residues (Figure 1.10).

1.3.3 Protein structure:

GPR98 is the largest known G protein coupled receptors (GPCR). These receptors have seven transmembrane (7TM) helical domains. They are involved in signal transduction process.

According to phylogenetic studies, GPCRs are classified into five classes as follows: class A is Rhodopsin- like receptors, class B is Secretin receptors, class C is metabotropic glutamate receptors, class D is adhesion receptors, and class E is frizzled/smoothened receptors (Sun et al., 2012).

GPR98 belongs to class D adhesion receptors, and it has shared structural features with other adhesion GPCRs classes; including large extracellular domain, seven transmembrane domain and GPCR proteolysis site (GPS).

VLGR1b isoform was determined to be the full length protein. (McMillan and White, 2010) It has a large extracellular N- terminal domain, 7 transmembrane protein helices, and a small intracellular C- terminal domain. The N terminus of VLGR1b has 35 calx-beta motifs, pentraxin (PTX) domain inserted between the 9 and 10 calx motifs, and epilepsy associated repeats (EAR) inserted between the 22 and 23 Calx beta (β) motifs. GPS is located next to the last N- terminal Calx- β motif (Figure 1.10) (Sun et al., 2012).

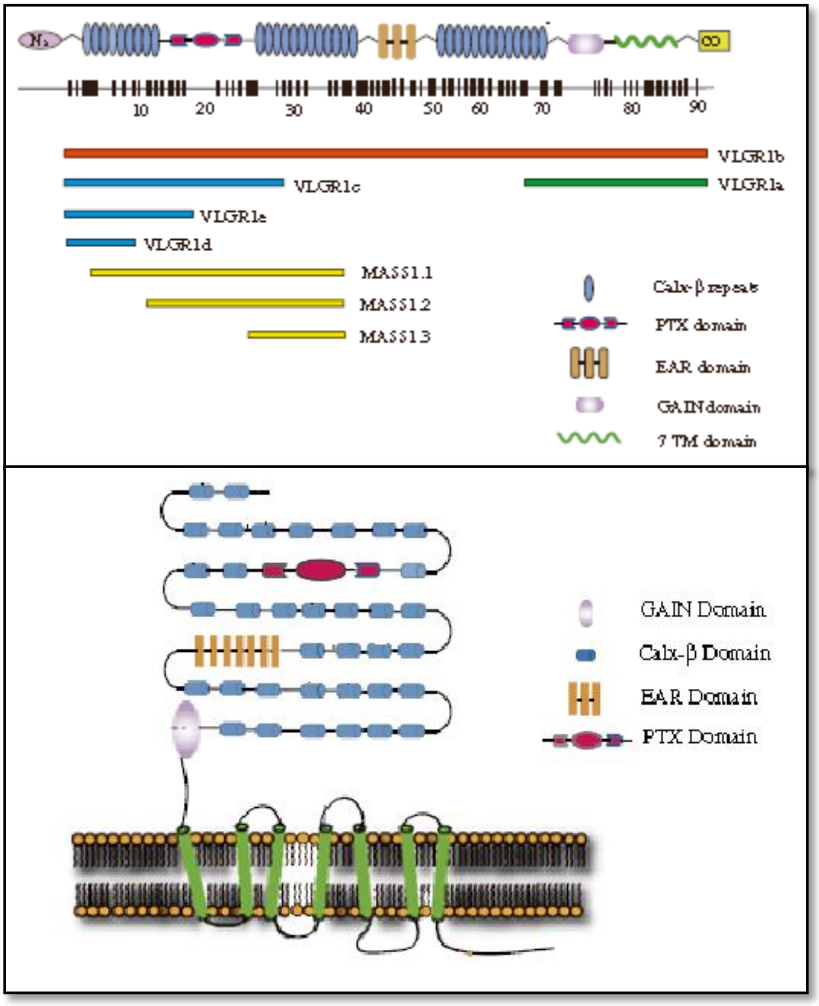


Figure 1.10 Isoforms of GPR98 protein and domains architecture. Domains are shown in the upper side as shaded shapes with different colors. Isoforms are illustrated as lines with different colors. In the lower side, architecture of domains is illustrated. Domains are indicated too, and the seven green transmembrane bars represent the 7 transmembrane domains (Sun et al., 2012).

1.3.3.1 Function of each domain in GPR98:

- **Calx- β motif**: it is a calcium binding domain, its function is unknown but it may act as extracellular calcium receiver and plays a role in signal transduction or regulation cell to cell conjunction (Sun et al., 2012). White and McMillan ,(2010), proposed that the Calx β motif mediate receptor ligand interaction.
- **PTX domain:** its role in VLGR1 has not been demonstrated yet. But it might mediate extracellular protein interaction and it may be involved in wide range of cellular functions (McMillan and White, 2010).

- **EAR domain:** Seven copies of nearly 50 amino acids repeat, termed as Epitempin (EPTP) or epilepsy associated repeats (EAR) are located near the center of the ectodomain. It is assumed to be made of β sheets folded into seven bladed β- propeller structures. Although its role is still unknown, but it is thought to mediate protein interaction (McMillan and White, 2010). Similar EAR repeats domain are found in other proteins as Lgi1-4 and Tspear, they are thought to be implicated in different types of seizures or hearing loss (Sun et al., 2012).
- **GPS:** the large ectodomain is linked to the 7TM domain by GPS, which might separate VLGR1 into two parts after proteolytic cleavage. GPS motif is embedded in GPCR autoproteolysis–inducing (GAIN) domain. It is hypothesized that GAIN domain could regulate receptor signaling through transmembrane helices (Araç et al., 2012).
- **The C- terminus**

The C-terminal residues of VLGR1 have a consensus sequence motif that is recognized as a ligand for PDZ domain containing protein (Weston et al., 2004). PDZ containing proteins are cytoplasmic scaffolding proteins that play a role in several cellular processes, such as receptor clustering and trafficking, and connecting receptors with their downstream signaling proteins (McMillan and White, 2010).

Because of PDZ– domain binding motif at the VLGR1 C-terminus, it may bind and interact with first PDZ domain of Whirlin (USH2D), second PDZ domain of PDZD7, or first PDZ domain of Harmonin (USH1C). Interaction between VLGR1 and Myosin VIIa

is facilitated by the cytoplasmic region of VLGR1 and the C-end MyTH4-FERM fragment of Myosin VIIa (Sun et al., 2012) (See figure 1.11). To date, the ligand of Vlgr1 protein is still unknown (Araç, et al., 2012; Sun et al., 2012; Weston et al., 2004).

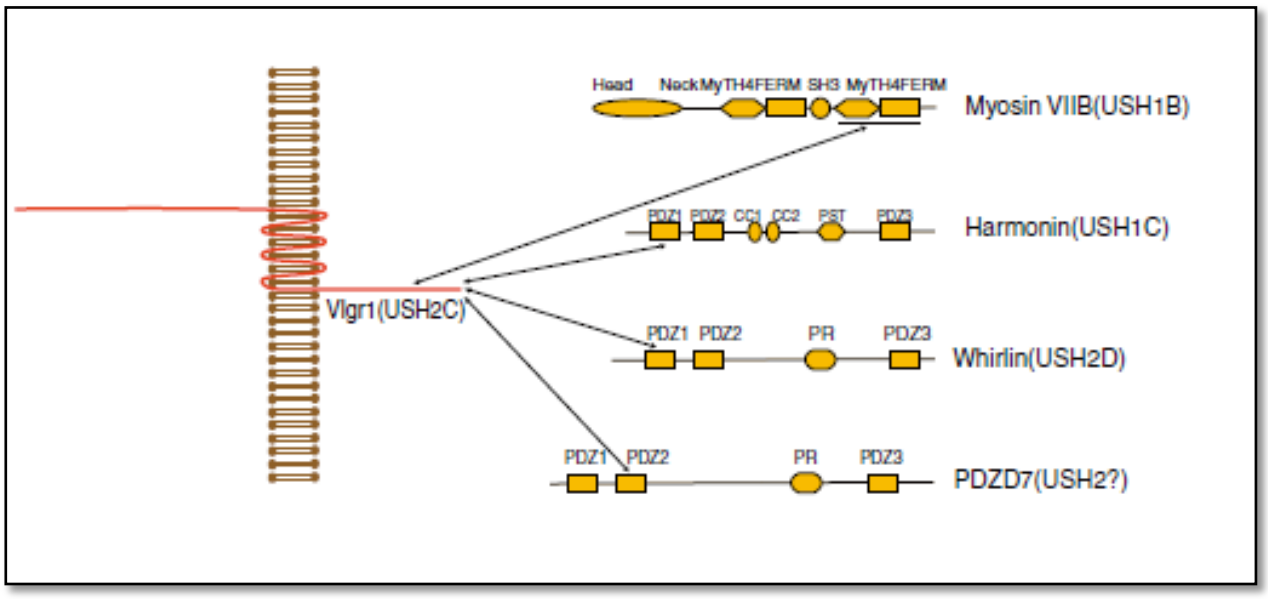


Figure 1.11 Possible interacting proteins with VLGR1 (Sun et al., 2012).

1.3.4 Expression of GPR98:

GPR98 is expressed in cochlea and other tissues as well as brain, lung, eye and kidney. In cochlea GPR98 is expressed in synaptic region and stereocilia of hair cells (Yang et al., 2012b). As explained previously, hair cells are receptor cells that play a main role in transforming mechanical stimuli into electrical signals in Mechano-electrical transduction.

GPR98 are confined at the core of stereociliary ankle links (See Figure 1.12), which connect stereocilia to their nearest neighbors. Besides, three other proteins, Usherin, Vezatin and Whirlin are found in the ankle link region of stereocilia in developing cochlear hair cells. These four proteins interact with each other and form ankle link complex. And When GPR98's expression is disturbed all other three proteins are not found in these regions (Sun et al., 2012).

The expression of GPR98 in eye is localized in photoreceptor cells of retina. Specifically, it is a component of the precilliary complex which is a fibrous structure that links connecting cilium with the inner segment of photoreceptor cells in the retina. And it is assumed that there is homology between precilliary complexes with stereociliary ankle links. Moreover, it is expressed in the synaptic terminus of photoreceptor cells (McMillan and White, 2010) (See figure 1.12).

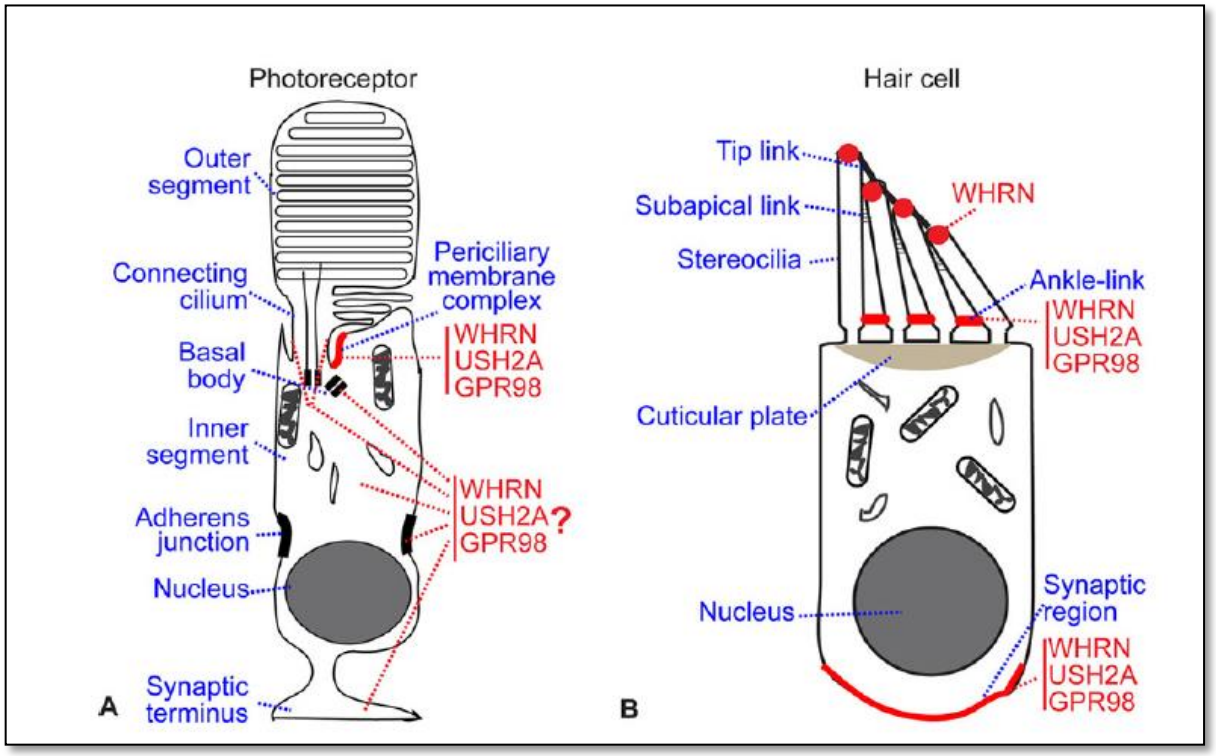


Figure 1.12 Cellular localization of VLGR1. Schematic diagram of photoreceptor and hair cells. (A) GPr98 is expressed in synaptic terminus and it is a component of percilliary complex which links connecting cilium with inner segment of photoreceptor. (B) GPR98 is localized in stereocilliary ankle links and synaptic region of hair cells (Yang et al., 2012b).

1.3.5 Pathology Action:

GPR98 may have a significant role in development of vertebrate nervous system. Mutations in Vlgr1 gene are associated with several disorders in humans including febrile and afebrile seizures, as well as Usher syndrome (combined sensorineural deafness and Retinitis Pigmentosa). Moreover, studies on mice revealed that it has negative consequences on nervous system including audiogenic seizures (Gibert et al., 2005).

However, Nakayam et al.,(2002), proposed that the isoform b does not contribute significantly to any seizure disorder (Weston et al., 2004).

Moreover, mutations in VLGR1 b isoform were identified in patients with usher syndrome type 2 (USH2C) (Yang et al., 2012b). Furthermore, diverse mutations along the murine Gpr98 gene share common phenotypes in vision and hearing. These outcomes propose too that isoform b is the major isoform in both the retina and the inner ear and it is crucial for vision and hearing (Yang, 2012a).

1.3.6 VLGR1 in usher syndrome:

Usher syndrome is a genetically heterogeneous autosomal recessive disease described by combined deafness and blindness. It is categorized into three clinical subtypes based on severity of hearing loss and vestibular symptoms. The following table illustrates Usher syndrome types (Weston et al., 2004).

Table 1.2 Clinical features of usher syndrome types. (Millan et al., 2010)

	USH1	USH2	USH3
Hearing loss	Severe to profound Congenital Stable	Moderate to severe Congenital stable	Moderate to severe Progressive
Vestibular Function	Altered	Normal	Variable
Retinitis pigmentosa onset	Usually pre pubertal	Around puberty or post pubertal	Around puberty or post pubertal

It is noticeable from the table that usher syndrome types 1, 2 and 3 share common phenotypes and that refer to the fact that most of usher proteins are expressed in the hair cell stereocilia and form complexes with each other (McMillan and White, 2010; Sun et al., 2012). This confirms what Arac et al.,(2012), indicated that GPR98 is a component of USH protein networks in inner ear hair cells and retinal photoreceptor cells (Araç et al., 2012).

Different mutations in GPR98 gene cause type IIC Usher Syndrome (USH2C). Patients with USH2C have congenital deafness and progressive retinitis pigmentosa (Sun et al, 2012). They usually start to lose their vision at their late teenage years and progress during their life (Weston et al., 2004).

A study showed that only 6.4% of USH2 mutations are caused by mutations in GPR98. Other genes involved with USH2 are Usherin (USH2A) and Whirlin (USH2D) (Sun et al., 2012).

1.3.7 Mice models with mutation in GPR98

Till now four mice models of GPR98 are known. *Frings* mice and BUB/BnJ mice are two naturally occurring mice strains. They share a single base deletion in exon 31 of GPR98 gene. It consequently leads to substitution of valine at position 2250 into stop codon, which leads to premature termination of GPR98 proteins.

The other two mutant mice models were engineered. The first one is Vlgr1/del7TM, a knock in mutation, which was developed by insertion of cDNAs encoding antigenic tags into GPR98 gene leading to removal of the expression of GPR98 trans membrane and cytoplasmic domains. The second model is Knock out mutation developed by deletion of exons two, three and four. This would delete isoforms b, c, d and e, but it will keep the internal MASS1 isoforms safe (McMillan and White, 2010; Sun et al., 2012).

Studies on mice revealed that all four mice models are subjected to develop audiogenic seizure. And relating to usher syndrome symptoms they are partially phenocopied in mice models particularly the hearing loss phenotype. Also, minor retinal defect was also stated in matured GPR98 knock in mice (Sun et al., 2012).

McMillan and White,(2010), in their study reported that the positions of mutations give essential information on the functional information of the different GPR98 isoforms. They concluded that as Vlgr1b is the only isoform affected by a mutation in all four mutant mice models, and then it is clear that VLGR1b plays an important role in the development and conservation of normal hearing. They added that other isoforms expression cannot substitute for VLGR1b in hearing function (McMillan and White, 2010).

Study on cochlear structure of mutant mice models showed that ankle links that connect stereocilia together is not found and stereociliary development is corrupted. Moreover, a study done by Michalski and his colleagues in 2007, showed that in the knockout mice, the other ankle link components Usherin, Vezatin and Whirlin are not found any more in ankle link region, from this he inferred that this protein complex depends on existence of functional GPR98 protein (Sun et al., 2012).

Another study on mice done by Yagi et al.,(2006), showed that expression of vlgr1 in cochlea was localized at the base of stereocilia and between the stereocilia where ankle link is found. They found in homozygous GPR98 mutated mice, stereocilia are not arranged in row and it did not form a staircase like shape. When they studied the eyes of 5-10 month old mice they did not find clear defects of retinas of GPR98 mutated mouse. Therefore, they concluded that GPR98 plays important role in the maturation of stereocilia, particularly in the arranging of stereocilia and the formation of staircase structures by stereocilia. Also, they showed that deletion of GPR98 causes hearing loss and malformed stereocilia structure (Yagi et al., 2006).

And this conclusion is consistent with the one drawn by White and McMillan ,(2010), which mentioned loss of GPR98, causes morphological defects in stereocilia organization, and that clarify deafness phenotype. And the function of ankle link may be crucial in stereocilia maturation (McMillan and White, 2010).

To conclude, GPR98 is one of the largest receptor proteins, it is localized in ankle links that connect the stereocilia together. Any mutations would affect the proper development of hair cells and stereocilia and would lead to hearing impairment. To detect these mutations we may use conventional sequencing technologies and take months to screen all exons. But with the introduction of next generation technologies, we will save time and effort and increase the pace of mutations identifications and gene discovery.

1.4 Genetic diagnosis of HL and NGS

Over the past two decades, the effort to identify genes of hearing loss by conventional approaches has made a great progress. Mapping chromosomal location of deafness genes has been achieved by linkage analysis using genetic markers such as SNPs (Single nucleotide polymorphism) and microsatellite markers. And once the linkage region is identified, Sanger sequencing is used to determine the causative mutations (Dror and Avraham, 2010).

But, when the linkage regions include a large number of genes, identification the causative mutations can be challenging using the conventional method (Rehman et al., 2010). The challenge is to decide which gene is responsible and which is the causative mutation in the patient. Besides, screening with Sanger sequencing of all genes is expensive and time consuming. Hence, the use of advanced approach, as next generation sequencing (NGS) has a prospective to be more cost effective and time saving (De Keulenaer et al., 2012).

1.4.1 Overview on Next Generation Sequencing:

Next generation sequencing (NGS), also called Massively parallel sequencing (MPS) and second generation sequencing, has been developed recently to overcome the challenges that face researchers and scientists in diagnosis mendelian disorder (Shearer et al., 2011).

Next generation DNA sequencing is currently applied, specifically in research. High throughput DNA sequencers have been developed by companies such as Roche (454 Genome Sequencer FLX), Applied Biosystems (SoLid System) and Illumina (Genome Analyzer) (De Keulenaer et al., 2012; Tucker et al., 2009). Each one has its own weaknesses and strengths, but they have common features including, high sequencing depth, high output and short sequence read- length (Shearer et al., 2011). Besides, a common technological feature shared by all NGS platforms is their ability to MPS of clonally amplified or single DNA molecules that are spatially isolated in a flow cell (Majewski et al., 2011).

The massive ability of NGS permits the sequencing of many DNA segments; consequently, each nucleotide in targeted regions may be incorporated in many reads. This allows repeated analysis and gives depth of coverage, which is defined as a measure

of numbers of time each nucleotide is sequenced during a single sequencing run. Accordingly, increase depth of coverage usually increases sequencing accurateness (Lin et al., 2012).

In the NGS process, DNA library is prepared by fragmentation of genomic DNA (mechanically or chemically) into small fragments, ranging in size from 300 to 500 bps. Next, each end of DNA fragments is ligated to oligonucleotide adaptors. Then, these ligated sequences are attached to a solid surface to perform the sequencing process. (Illumina, 2012; Lin et al., 2012). The used solid surface and sequencing of these small DNA fragments are varied according to the used platforms.

The sequencing process in Illumina platform depends on cyclic reversible termination (CRT) technology. In CRT, all four nucleotides with fluorescent reversible terminators are added and once a single nucleotide is added to the strand by DNA polymerase, fluorescence imaging follows. Then, the fluorescence is chopped and the cycle is repeated when another nucleotide joins the growing strand (Figure 1.13) (Shearer et al., 2011).

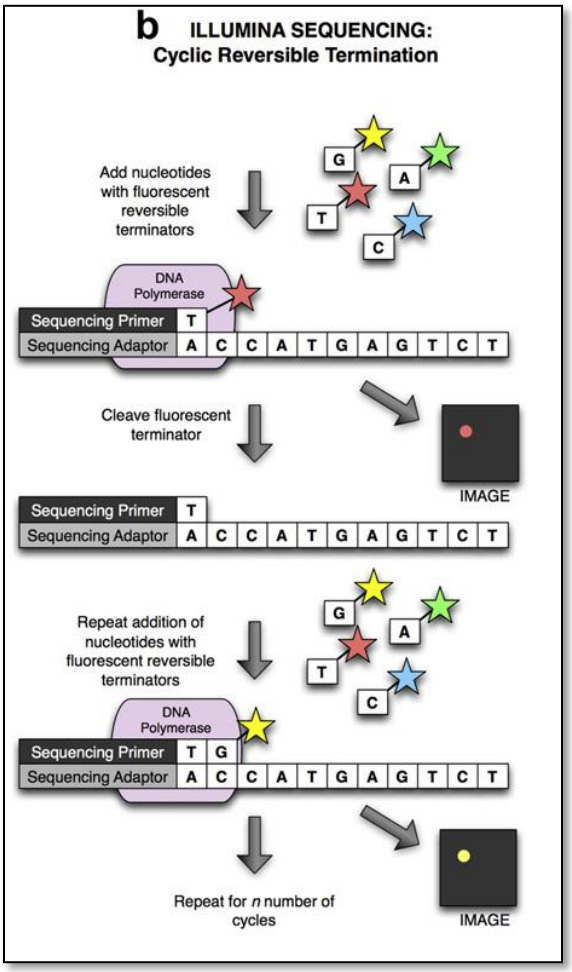


Figure 1.13 Overview of illumina platform. In illumina sequencing (cyclic reversible technology), all four nucleotides with fluorescent reversible terminators are added, and when one nucleotide is integrated into the growing strand, it is detected by imager, then fluorescent terminator is cleaved and the whole process is repeated for *n* number of cycles (Shearer et al., 2011).

Depending on the used platform, MPS produces several hundred megabases to gigabases of short sequenced segments that are called "reads", which could range in size from 25 to 100 bps from one or both ends in a single sequencing run (Majewski et al., 2011; Lin et al., 2012). Any genomic region has an equal chance of being sequenced, unless a genomic region is targeted (Lin et al., 2012). Therefore, scientists develop targeted genomic enrichment technologies, to isolate and amplify targeted regions of the genome (Shearer et al., 2011).

1.4.2 Targeted genomic enrichment:

In targeted genomic capture, also called targeted sequence capture, a portion of DNA is captured for sequencing, and thus the proportion of DNA segments enclosing or near targeted regions is greatly amplified, and sequence coverage of interest is increased (Lin et al., 2012). Rather than amplifying a single genomic site, hundreds to millions genomic regions are isolated and amplified to make a DNA library, then they are sequenced by one of the MPS platforms (Shearer et al., 2011).

Any genomic region can be targeted for MPS, including exons, highly conserved sequences in genome, promoters regions, non-coding RNA, or other region of interest (Majewski et al., 2011). Targeted exome sequencing means that all exons in each gene are targeted for enrichment and MPS. It examines only known disease causing genes for pathogenic alleles.

In genetic diagnosis sequencing a portion of genome is preferable to sequencing the whole genome. That refers to two reasons, first, 85% of mutations are supposed to be found in the coding regions (exons) or splice site regions, which makes results obtained from exon targeting sequencing are easier to interpret. Second, the amount of sequence is low, and that consequently decrease the cost of data generation and analysis. Currently, the most effective MPS clinical platforms depend on targeted genomic enrichment (Shearer et al., 2011).

1.4.3 NGS and deafness (MPS and novel deafness gene discovery)

NGS technologies accelerate the pace of gene discovery and mutation identifications (See figure 1.14). Many publications shows that a number of deafness genes and mutations have been identified using this technology. Probably, the identification of more deafness genes will follow (De Keulenaer et al., 2012).

Taperin (TPRN) is a gene identified by next generation sequencing technology. Rehman et al., (2010), combined targeted genomic capture with MPS to identify a gene responsible for NSHL in consanguineous Pakistani family. At the DFNB79 locus, 2.9Mb was mapped on chromosome 9q34.3, and having 108 candidate genes. They identified one nonsense mutation in a predicted gene, C9orf75, renamed TPRN. Later, they identified three additional frame-shift mutations after sequencing the same gene in other three DFNB79- linked Pakistani families. Immunolocalization study on mouse revealed that TPRN gene is expressed in the cochlea, particularly in the taper region of hair cell stereocilia (Rehman et al., 2010).

In a Turkish study done by Sirmaci et al.,(2012), these researchers combined both autozygosity mapping and whole exome sequencing in a family with 3 children having NSHL and born to consanguineous parents. They identified two novel missense mutations, c.508C>A (p.H170N) in GIPC3 and c.1328C >T (p.T443M) in ZNF57. They were found in the same autozygous region on chromosome 19 in affected members of the family. Both variants co segregated with the family and were absent in 335 ethnically matched controls (Sirmaci et al., 2012).

A novel nonsyndromic deafness gene was identified in a Palestinian family. Walsh et al., (2010), combined homozygosity mapping and exome sequencing. They mapped DFNB2 locus at chromosome 1p13.3. After filtering variants yielded from whole exome sequencing, only a single deleterious mutation remained and linked to DFNB82. They found that this mutation causes early truncation of the G protein-signaling modulator GPSM2. Subsequently, they found that GPSM2 is localized to apical surfaces of hair cells and supporting cells and is most highly expressed during embryonic development (Walsh et al., 2010).

Additional study done by De Keulenaer et al.,(2012), estimated whether NGS allows detection of mutations or not. They screened fifteen deafness genes using Roche 454 Genome sequencer for five patients with familial congenital deafness. In three of them, two new mutations were identified in CDH23 and OTOF, respectively. In the fifth patient, they confirmed a known mutation (c.236 + 1 G>A) in TMC1 and they were unable to identify mutation in the other patient (De Keulenaer et al., 2012).

It is obvious from previous publications that they share a common strategy to identify mutations via NGS. This strategy is illustrated in a diagram presented by Brownstein et al., (2012), (Figure 1.14).

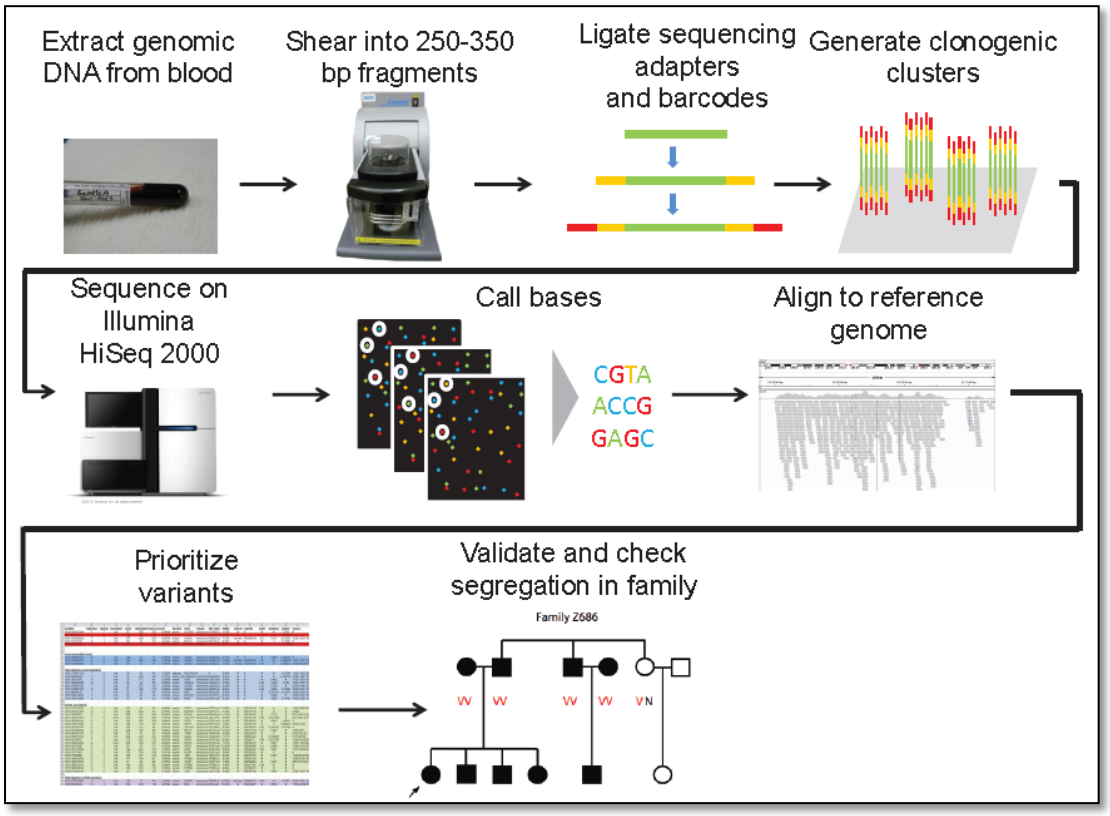


Figure 1.14 procedure flowchart for identification mutations in HL patients by targeted genomic capture and massively parallel sequencing. It starts with extracting genomic DNA, followed by shearing it into fragments, then ligation of sequencing adapters and barcodes, next cluster generation is performed and then followed by sequencing using illumina or other platforms mentioned previously. Bioinformatics analysis is done including aligning the sequence to reference genome to help prioritize variants. And once a mutation is identified, it has to be validated and checked for segregation in family (Brownstein et al., 2012).

To conclude, identifying genes and mutations responsible for HL by conventional methods is time consuming, labor intensive and expensive. Introducing the NGS technologies has come to overcome challenges that face researchers and scientists. Recently, many researchers start to apply this approach to identify genes and even mutations within known genes. Thus, using these technologies is expected to add more genes to list of human genes for hearing loss in a short period of time.

CHAPTER 2: Problem statement and objectives of the study

2.1 Introduction:

Hearing loss is the most common sensory defect in humans, affecting nearly one in 1000 infants and over 50% of the populations by the age of 80 (Raviv et al., 2010). Most cases of hearing loss are inherited (Lenz and Avraham, 2011; Avraham and Kanaan, 2012).

In Palestine itself, the frequency of inherited deafness among Palestinian population is nearly 1.7 per 1000, and it is thought to be higher in some villages. This high frequency reflects high rate of consanguineous marriages in Palestine (Shahin et al., 2002).

In Palestinian communities, consanguineous marriages are common. Whereas, nearly 50% of it are between first cousins and nearly 40% of it are between relatives. In these families the rate of having recessively inherited phenotypes is high.

Moreover, deafness in these families is highly heterogeneous. Hearing loss among Palestinians is caused by multiple different mutations which could be either allelic or in different genes that are responsible for a similar phenotype (Zoltogora, 2010).

The highly heterogeneity of inherited deafness in Palestinian populations is posing a major challenge to diagnosis. However, with the advent of next generation sequencing (NGS) approaches, this challenge is vanishing. With NGS, screening of large number of genes can be performed in single test. Thus, identifying causative mutations of hearing loss in the Palestinian population will allow proper genetic counseling. Besides, it will add to the scientific knowledge we need to comprehend the underlying mechanism of hearing loss.

2.2 Problem statement

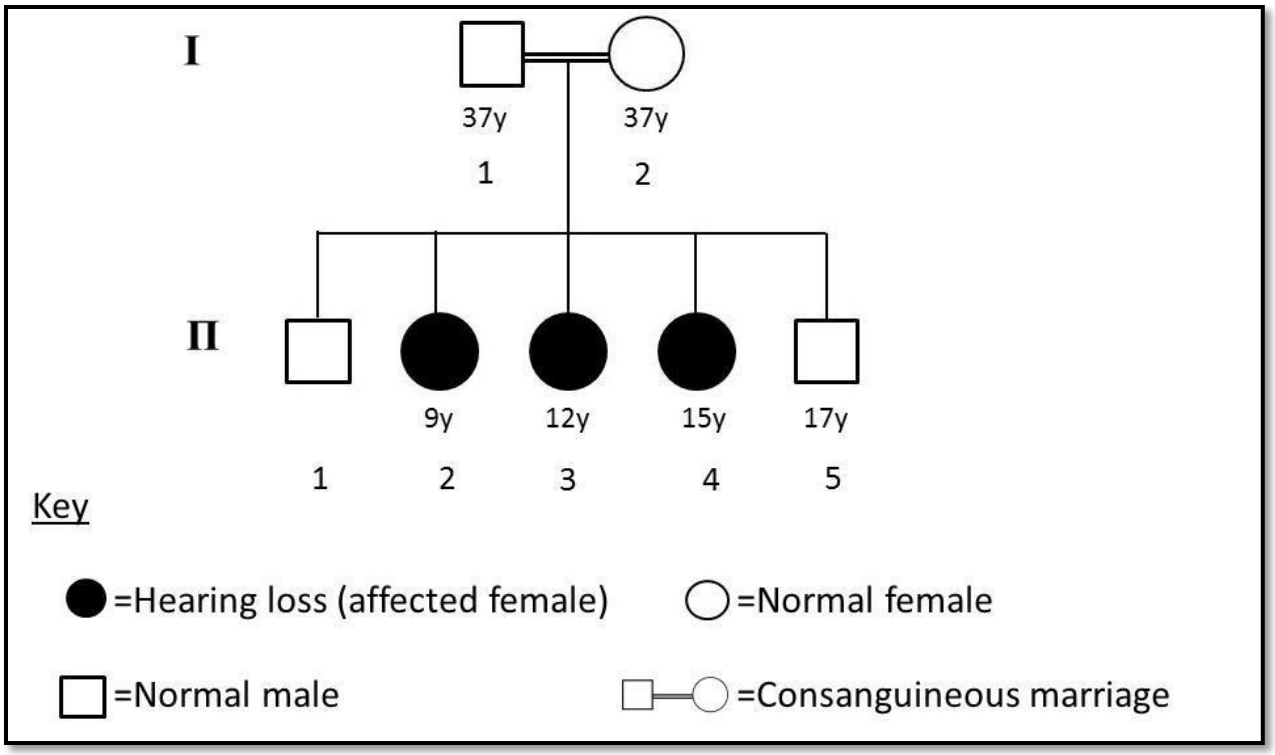


Figure 2.1 Pedigree of Family EI.

The patient II-2 is born to consanguineous parents in family EI of Palestinian origin. She suffers with her two elder sisters from congenital, moderatly severe and receissvely inherited hearing loss (HL) (Figure 2.1). In order to determine the causative mutation of HL, she was screened for all known mutations in genes causing HL among the Palestenian polulation. But the result was negative and none of these mutation is the causative of HL . Hence, her DNA was subjected for diagnoses by targeted genomic capture using masively parallel sequencing.

Next generation sequencing data provided us with large number of variants and the data had to be filtered in order to determine which gene and which mutation is responsible for hearing loss in II-2 member.

We were able to identify a novel mutation. We validated it by Sanger sequencing, checked segregation in the family and filtered it among deaf and normal Palestenian

controls. As the mutation is in the exonic splice site, we did functional assay to determine the effect of mutation on pre- mRNA splicing.

2.3 Objectives:

1. To identify the causitive gene and mutation responsible for congenital deafness in family member Π-2.
2. To validate the mutation generated from NGS data via Sanger sequencing.
3. To check segregtion of the mutation in the family.
4. To valditate the mutation by its exclusion in 100 individuals with normal hearing and 263 individuals with hearing loss (HL).
5. To determine the mutation effect on pre-mRNA splicing.

CHAPTER 3: MATERIALS AND METHODS

3.1 Materials:

All materials and instruments used in this study are listed in the following table.

Table 3.1 list of used materials and instruments.

Name of material	Company	Cat number
Nuclease free water (H_2O)	Ambion – The RNA company	AM9932
SeaKem® LE Agarose	Lonza	50004
PCR Master mix	Thermo Scientific	AB-0575/DC/LD
Exonuclease I	New Engnland Biolabs	M0293L
Antractic Phosphatase	New Engnland Biolabs	M0289L
Big dye ® Terminator v1.1 Cycle sequencing kit- (Include 5x buffer)	Applied Biosystem	4337451-100
(BigDye enhancing buffer) BDX64	MCLAB	BDX-100
Super-DI Formamide	MCLAB	SDI-100
Ethanol Absolute	SIGMA-ALDRICH	UN 1170
EDTA, DISODIUM, DIHYDRATE	aMResco	638-29-6
Tris Acetate EDTA (TAE) buffer	aMResco	77-86-1
Acetic Acid Glacial	FRUTAROM LTD	2355500300
Ethidium Bromide	United States Biochemical Corporation	23813
qScript™ cDNA Synthesis Kit	Quanta Biosciences	95047100
Custom Tagman SNP genotyping assays	Applied Biosystem	4334431
Instruments used		
Instrument	**company**	Notes
Thermocycler- (gene Amp PCR system 9700)	Applied Biosystem	
Sequencer: 3130 Genetic Analyzer	Applied Biosystem	
7300 Real time PCR system	Applied Biosystem	
Heraeus Megafuge 1.0	Kendro	
Gel Doc XR+ imaging system	BIO-RAD	
Mettler AE 50 digital analytical balance	-	

3.2 Methodology:

3.2.1 Subjects:

This study was accepted by the Human Subjects Committees of Bethlehem University. As the criteria for our study are limited to hearing loss, a medical history of family EI including hearing loss level and consanguinity was collected and a pedigree was constructed. Also, an informed consent was obtained from parents.

Genomic DNA and RNA from white blood cells for family EI, in addition to control DNA samples for HL individuals and normal hearing individuals were obtained from the hereditary research laboratory at Bethlehem University.

3.2.2 Gene Exclusion:

The deaf patient (II-2) was tested for GJB2 gene by standard sequencing. Besides she was screened for all known mutations known in causing hearing loss in the Palestinian population, including mutations in: MYO7A, TMHS, TECTA, SLC26A4, CDH23, MYO15A, PJVK, GPMSM2, PTPRQ, OTOA and TMPRSS3.

3.2.3 Targeted genomic capture and massively parallel sequencing:

Exons and the flanking 40bp into introns of 246 genes of both human and mouse (known to cause SHL and NSHL) were selected for capture and underwent MPS using illumina paltform. In family EI only (II-2) member was subjected to targeted genomic capture and MPS. (**Note**: this step was performed in specialized laboratory in Technion University.)

3.2.3.1 Variants Analysis:

The sequences were aligned to the reference human genome (current version is hg19) to identify SNPs and mutations. Then, the variants were identified and aligned. After that we got the output which presented as an excel sheet file which contains a list of large number of SNPs and variants that contains all the following information:

- Chromosome number
- Position
- Reference nucleotide
- Variant nucleotide

- Score – this score accounts for both quality of the sequencing process and the alignment to hg19.
- Read depth – the number of times the nucleotide was covered (includes both variant and reference)
- Genomic context – whether the SNP is in an intron, exon, splice site, non-coding RNA.
- Gene – the name of the gene the SNP is in, or the nearest gene if it's intergenic
- Synonymy – whether the mutation is synonymous or not (synonymous = no change in amino acid; non-synonymous = change in amino acid).
- SNV Change – the affected exon, the change in the coding region, the change in the protein sequence,
- 1000 Genomes – the frequency of the SNP in the 1000 genomes project (which has 1000 genomes).
- Clinical – whether the SNP has ever been clinically associated with hearing loss.
- dbSNP134 – whether the SNP appears in dbSNP134, a database covering most known SNPs.
- avSIFT and PolyPhen2 – web tools prediction programs to assess protein damaging.
- PhyloP – the evolutionary conservation score for the nucleotide.
- ESP – the frequency of the SNP in the Exome Sequencing Project

3.2.3.2 Prioritization of variants

To identify the causative mutations we followed the following steps:

1. We sorted out all SNPs that are not exonic, splicing or exonic splicing SNPs.
2. Then, we sorted out all synonymous variants.
3. After that, we discarded all SNPs that are seen more than 15% in 1000 Genomes project.
4. Next, we discarded all known variants that appears in dbSNP134 and is not clinically associated.

5. Subsequently, we neglected all variants with low evolutionary conservation score across 46 vertebrates (PhyloP score lower than 0.9) as they are considered as non-conserved.
6. Then, we sorted out all non-synonymous SNPs seen in more than 5 deaf probands.
7. Now, we had a small number of variants, so we looked at the PolyPhen 2 (Prediction of functional effects of human nsSNPs) and SIFT (Sorting Tolerant Form Intolerant) scores. These are damage prediction web tools that assess the effect of non- synonymous SNPs on the protein level.
8. Finally, we had only one mutation and to ensure that it is novel we searched the ESP (Exome sequencing project) database. This variant has not seen before. So, we have a novel mutation.

To confirm the mutation in Π-2 patient and check segregation in Family EI, we performed PCR amplification followed by Sanger sequencing. Steps are clarified below.

3.2.4 PCR amplification.

To target the region flanking the mutation we designed primers using Primer 3 (http://bioinfo.ut.ee/primer3-0.4.0/), as shown in table 3.2 below.

Table 3.2 Sequences and conditions of primers used to target the novel mutation.

Name of Primer	Sequence of Primers 5'-----3'	Primer length	GC % content	Tm°C
GPR98-Ex-49-F	TGGGACAACAGAAGTTGAGG (Forward)	20	50.00	58.69
GPR98-Ex-49-R	TCAAAAATCCATTTACAGCTGAC (Reverse)	23	34.78	58.31

Note: the expected product size = 236bp

Steps of PCR reactions

1. We prepared the PCR mix tube using the following reagents and volumes as illustrated in table 3.3. (**Note**: Six samples to be tested: I1, I2, Π2, Π3, Π4, Π5 as shown in the pedigree (Figure 2.1), in addition to non-template control (NTC). Extra volume was added to account for pipetting loss, so we had 7 PCR tubes.)

Table 3.3 Reagents and volumes used to set up 8 PCR reactions

reagents	**1X**	**8 X**
PCR master mix	12.5 µL	100 µL
Forward primer (GPR98-Ex-49-F)	0.5 µL	4 µL
Reverse primer (GPR98-Ex-49-R)	0.5 µL	4 µL
Nuclease free H_2O	9.5 µL	76 µL
DNA	1 µL	Distribute 1µl
Total volume	25µL	184 µL

2. We pipetted 24 µL of PCR mix to each PCR tube.
3. Then, we added 1 µL of DNA (100ng/ µL) samples to each PCR tube for tubes I-1 to Π-5, and 1µL of nuclease free H_2O to NTC tube.
4. Finally, we spun them down and put the PCR tubes in the thermocycler (Gene Amp PCR system 9700 – applied biosystem) in the program named touch down 60 (td. 60) using the following conditions.

1 cycle	4 min	95°C	(initial deanaturation)
2 cycles	30 sec	95°C	(denaturation)
	45 sec	66°C	(annealing)
	30 sec	72°C	(extension)
2 cycles	30 sec	95°C	(denaturation)
	45 sec	64°C	(annealing)
	30 sec	72°C	(extension)
2 cycles	30 sec	95°C	(denaturation)
	45 sec	62°C	(annealing)
	30 sec	72°C	(extension)
35 cycles	30 sec	95°C	(denaturation)
	45 sec	60°C	(annealing)
	30 sec	72°C	(extension)
1 cycle	10 min	72°C	(final extension)

- Hold at 4°C until reactions are removed.

Visualize PCR product.

To see the result of PCR product, 1.5% agrose gel is prepared as follows.

1. We weighted 3 gm of agrose gel powder to 200ml of 1X TAE buffer.
2. Then, we put it on microwave for 3 minutes.
3. Next, we added 2 drops of ethidium bromide and then poured the visicous gel in the loading tank and wait until it is soldified.
4. When the gel is soldified, we added to the loading tray 1X TAE buffer.
5. Then we laoded 3µl of each PCR product, and perform electrophoresis with 100 volts for 15 minutes.
6. After that, we examined each lane for the presence of band with 263bp product, except for NTC lane (no band should appear).
7. Finally, we took a photo of gel using molecular imager (Gel Doc XR+ imaging system).

3.2.5 Sanger sequencing

3.2.5.1 PCR clean

This step is necessary to remove excess primers and deoxynucleotides (dNTPs). So we did the following steps.

1. We prepared the PCR clean reaction mix tube using reagents and volumes as illustrated in the table below. (**Note**: PCR samples to be cleaned is I-1to II-5, extra volume was added to account for pipetting loss so we have 6 PCR clean tubes.)

Table 3.4 Reagents and volumes used to set up 7 reactions of PCR clean

Reagent	1x	7x
Exonuclease I	0.25 µL	1.7 µL
Antractic Phosphatase	0.25 µL	1.7 µL
Nuclease free H_2O	1.5 µL	10.5 µL
Total volume	7µL	13.9 µL

2. Then, we pipetted 5µL of PCR clean mix into each sample tube.
3. Next, we added 2µL of PCR product to each sample tube.

4. Finally, we spun them down and incubated the reactions tubes in thermocycler, in program named PCR clear under the following conditions.

1 cycle	30 min	37° C
1 cycle	20 min	80° C

Hold at 4°C, until the tubes are removed.

3.2.5.2 Big dye fast kit sequencing

1. We prepared the Big dye fast kit sequencing reaction mix tube using components and volumes as shown below. (**Note**: two sequencing mixes were prepeared, in one of them we used forward primer and in the second one we used reverse primer)

Table 3.5 Reagents and volumes used to set up 7 reactions of Big dye fast kit sequencing.

components	1x	7x
Big dye terminator	0.1µL	0.7 µL
64X buffer	0.75 µL	5.25 µL
5X buffer	1.5 µL	10.5 µL
Cleaned PCR	2 µL	-
Primer (GPR98-Ex-49-F)/(GPR98-Ex-49-R)	0.5 µL	3.5 µL
Nuclease free H_2O	10.5 µL	73.5 µL
Total volume	15.4 µL	93.45 µL

2. Second, we pipetted 13 µL of big dye fast kit mix into each sample tube.
3. Then, we added 2µL of cleaned PCR to each sample tube.
4. Finally, we spun them down and put the reaction tubes in the thermo-cycler (Gene Amp PCR system 9700), in the program named seq-buf64x under the following conditions.

1 cycle	3 min	96°C	(initial denaturation)
30 cycles	10 sec	96°C	(denaturation)
	0.5 sec	50°C	(annealing)
	2min	60°C	(extension)

- Hold at 4°C, until the reaction is removed.

3.2.5.3 Cleaning via ethanol:

This step is essential prior to electrophoresis to get rid of unincorporated dye terminators in the sequencing reactions. (BigDye Terminator v3.1 Cycle Sequencing Kit protocol, 2010)

1. First, we added to each sample tube 100µL of 100% ethanol.

2. Second, we added 5 µL of 125M EDTA.
3. Third, we incubated samples at 20°C for 20 minutes.
4. Fourth, we centrifuged for 30 minutes at 4°C, speed 3800 RPM.
5. Fifth, we discarded the supernatant, and then added 60 µl of 70% ethanol.
6. Sixth, we centrifuged at 4°C for 20 minutes, speed 3800 RPM.
7. Seventh, we discarded the supernatant, turned tubes upside down, and then centrifuged for 1 minute speed 500 rpm.
8. Eighth, we left tubes open and made them dry completely by leaving them 2 minutes on heater (96°C).
9. Ninth, we added to each tube 10µL of Hi-Di formamide, and vortex for 30 seconds.
10. Tenth, we incubated at 96°C for 2 minutes and then incubated directly on ice for 5 minutes.
11. Next, we performed spinning down and transferred them to 96 well plate to put them on sequencer (3130 genetic analyzer, Applied Biosystem).
12. Finally, we analyzed the sequence to detect the mutation using Finch TV software version 1.4.0.(http://www.geospiza.com/).

3.2.6 Taqman genotyping assay (Real time –PCR) (concept is explained in the appendix)

We filtered the mutation in 100 normal hearing samples and 263 Deaf samples using Taqman genotyping assay. We followed the steps illustrated in Allelic Discrimination starting guide and they are as follow:

1. We calculated number of reactions to be performed in each assay.
a. For normal hearing controls, we had 100 DNA samples, in addition to 4 NTCs and extra volume was added to account for pipetting loss, so we had 106 reactions.
b. For HL controls, we had 263 DNA samples, in addition to 4 NTCs and extra volume was added to account for pipetting loss, so we had 270 reactions.
2. Then we prepared the reaction mixes tubes using components and volumes as shown below. (**Note**: we had two reaction mix tubes one for normal hearing controls and the other is for HL controls.)

Table 3.6 Volume of components needed for all wells in the real time reaction plate

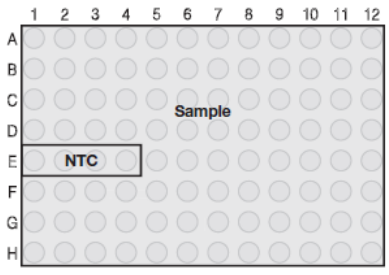
component	Volume 1X (µL/ reaction)	Volume of 106 reactions (normal hearing)	Volume of 270 reactions (deaf probands)
2x TaqMan Universal PCR Master Mix, No AmpErase UNG	0.5 µL	53 µL	135 µL
20x SNP Genotyping Assay Mix	5.0 µL	530 µL	1431 µl
Nuclease free H_2O	4.0 µL	424 µL	1080 µL
DNA	0.7 µL	Distribute 0.7 µL	Distribute 0.7 µL
Total	10.2 µL	1.007 ml	2.646 ml

3. Then we centrifuged each tube briefly to spin down the contents and to eliminate air bubbles.

4. Next, we pipetted 0.7µL of DNA (100ng/ µL) samples into each well in a 96-well reaction plate (shown below), and 0.7µL NF H_2O to NTCs to wells E1 through E4.

5. After that, we pipetted 9.5µL of reaction mix into each well.

6. Next, we covered the reaction plate with an optical adhesive cover.



7. Finally, we kept the reaction plate on ice until loading in the 7300 Real time PCR system. The output was analyzed via SDS software v1.4.

3.2.7 Functional assay: Study the effect of mutation on pre-mRNA splicing

We performed nested RT–PCR to determine the effect of exonic splicing mutation on pre-mRNA splicing. So we performed reverse transcription to synthesize c-DNA, followed by nested RT- PCR, visualized the PCR product on gel electrophoresis and finally we analyzed PCR product by sanger sequencing.

We chose three members of EI family to perform the functional assay:

1. I-1 carrier (Heterozygous)
2. I-2 carrier (Heterozygous)
3. Π-2 affected (Homozygous)

- AV5 (Wild type) was used as negative control.

3.2.7.1 c-DNA synthesis:

1. We Prepared the reaction mix tube using the reagents and quantities illustrated in the table below. (**Note**: 4 samples to be tested: I-1, I-2, Π-2 and AV5, in addition to NTC. Extra volume was added to account for pipetting loss, so we had 5 PCR tubes.)

Table 3.7 Reagents and volumes used to prepare 6 reactions of cDNA synthesis

Components	**1X**	**6X**
5 X buffer	4µL	24 µL
RT- Enzyme	1 µL	6 µL
Nuclease free (nf) H_2O	15µL	90 µL
Total volume	20 µL	120µL

2. Then, we pipetted 20 µL of reaction mix into each sample tube and NTC sample tube
3. After that, we added 5µL of RNA samples to each sample tube. While, to NTC sample we added 5µl of nfH_2O. (Note: total volume is 25 µL.)
4. Next, we spun them down and incubated the complete reaction mix in thermocycler (Gene Amp PCR system- 9700 Applied Biosystem) in the program named cDNA using the following conditions:

5 min	22°C
30 min	42°C
5 min	85°C

- Hold at 4°C until the reactions are removed.

3.2.7.2 Designing Primers:

As the mutation is found in the last base of exon 49, we designed primers to target exonic splice site mutation based on the flanking sequences of exon 49. As shown below, the flanking sequences are exon 48 and exon 50.

Forward Primer

GTACAAACAATCATTATTCTGGAAAGTTCTCAAGTAAGATATTTTACTTCAGACAGCCAA GATTATTTAATCATTGCAAGTCAAAGAGATGATTCCGAATTAACTCAG — Exon 48

GTCTTCAGGTGGAATGGAGGAAGCTTCGTGTTGCATCAAAAACTCCCTGTCCGAGGTGTG CTGACCGTGGCCTTGTTCAACAAGGGAGGCTCTGTGTTCTTAGCCATTTCCCAGGCTAAT GCCAGGCTAAACTCCCTTTTATTCAGATGGTCTGGCAGTGGGTTTATTAACTTTCAAGAG GTGCCTGTCAGTGGGACAACAGAAGTTGAGGCTTTGTCTTCAGCCAATGATATTTACCTA ATATTTGCCGAAAATGTCTTTCTAG — Exon 49

Reverse Primer

GAGATCAGAATTCAATTGATATTTTCATCTGGGAGATGGGACAGTCTTCCTTCAGGTATT TTCAGTCTGTAGATTTTGCTGCTGTTAACAGAATCCACTCCTTCACACCAGCCTCAGGAA TAG — Exon 50

The labeled nucleotides in yellow represent the forward primer and the reverse primer, respectively. The start and the end of each exon is the nucleotide colored with orange. And the last nucleotide of exon 49 represents the site of mutation.

- Transcripts of exons 48,49,50 respectively were copied from Ensemble database (ENST00000405460)
- We designed primers using primer 3 (http://bioinfo.ut.ee/primer3-0.4.0/) and they have the following conditions illustrated in the following table:

Table 3.8 Sequences and conditions of primers used to target the novel mutation for nested RT-PCR first round.

Name of Primer	Sequence of Primers 5'-----3'	Primer length	GC % content	Tm°C
GPR98-Ex48-F	CAATCATTATTCTGGAAAGTTCTCAA (Forward)	26	30.77	59.91
GPR98-Ex50-R	AGACTGTCCCATCTCCCAGA (Reverse)	20	55.00	59.64

Note : the expected product size= 413bp

3.2.7.3 First round of Nested RT- PCR:

1. We prepared the reaction mix tube using reagent and quantities illustrated in the following table.

Table 3.9 Reagents and volumes used to prepare 6 reactions of first round of nested Rt-PCR

Reagents	**1X**	**6X**
PCR master mix	12.5 µL	75 µL
Forward primer (GPR98-Ex48-F)	0.5 µL	3 µL
Reverse primer (GPR98-Ex50-R)	0.5 µL	3 µL
Nuclease free (nf) H_2O	9.5 µL	57 µL
DNA	2 µL	Distribute 2µl
Total volume	25µL	138 µL

2. After that, we pipetted 23 µL of nested Rt-PCR mix into each sample tube and NTC tube.
3. Then, we added 2 µL of DNA to each sample tube, and 2µL of nfH_2O to NTC.
4. Next, we put the reactions in the thermocycler (Gene Amp PCR system 9700 – applied biosystem) in the program named touch down 60-45. (**Note: the same conditions as mentioned previously in section 3.2.4 point 4, but the extension time is 45 seconds.)**
5. Then we visualized PCR products by performing gel electrophoresis. **(Note: the same steps as mentioned previously in section 3.2.4)**

3.2.7.5 Second round of Nested RT-PCR:

1. We prepared the reaction mix by using reagents and quantites as shown in table 3.10.
2. Then, we pippeted 23 µL of nested Rt-PCR mix into each sample tube and NTC tube.
3. Then we added 2 µL of nested Rt-PCR product of first PCR round to each sample tubes, and 2µL of nfH_2O to NTC tube.
4. Afterwards, we put the reactions in the thermocycler (Gene Amp PCR system 9700 – applied biosystem) in the program named touch down 60 using the same conditions used previously in first round nested RT-PCR,and as shown in section 3.2.4 step 4.

5. Finaly, we visualized PCR products by performing gel electrophoresis. (**Note**: the same steps as mentioned previously in section 3.2.4.3, but we load 5µL instead of 3µL.)

Table 3.10 Reagents and volumes used to prepare 6 sets of nested Rt-PCR reactions second round

Reagnets	**1X**	**6X**
PCR master mix	12.5 µL	75 µL
Forward primer	0.5 µL	3 µL
Reverse primer	0.5 µL	3 µL
Nuclease free (nf)H_2O	9.5 µL	57 µL
DNA (PCR product of first PCR round)	2 µL	(distribute to each tube)
Total volume	25µL	138 µL

3.2.7.6 Sanger sequencing

1. We sequenced only the normal (AV5) sample and the mutant (II-2) sample to see the effect of mutation. Steps of PCR clean and big dye fast kit sequencing and cleaning via ethanol are the same as illustrated previously in sections 3.2.5.1, 3.2.5.2 and 3.2.5.3, respectively.

2. We analyzed sequences using Finch TV software version 1.4.0.(http://www.geospiza.com/)

3.2.8 Bioinformatics:

1. We determined the exact location of protein sequence predicted to be translated by mRNA by blasting it using Blastx on Blast NCBI database.
2. We used Interproscan bioinformatics tool to identify conserved domain on the targeted protein sequence.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Identification a novel mutation in GPR98 gene.

Family EI is of Palestinian Arab origin, three daughters in the family are with congenital, moderately severe and recessively inherited hearing loss (Figure 2.1). The genomic DNA of EI3 was subjected to targeted genomic capture and MPS. Consequently, she was identified to be homozygous for one novel variant in GPR98 gene (Table 4.1).

Table 4.1: Information about the novel mutation identified in GPR98 gene.

Gene	GPR98
chromosome	Chr5:90024750 ,
Type of mutation	Exonic splicing
Novelty of mutation	Novel
Nucleotide change	G to A
Protein change	Glycine (G) 3476 Arginine (R)
Depth coverage	237/240 (98.75%)
Phylop score	2.60829
PolyPhen score	0.976 (probably damaging)
SIFT score	Not provided
Sanger sequencing	Validated and confirmed
Segregation in family	Confirmed (segregate)

In this study we identified a novel exonic splicing mutation in GPR98 gene. It causes the transition of the last nucleotide of exon 49 from G (Guanine) to A (Alanine) at chr5: 90024750. Prediction program PolyPhen classified the mutation as probably damaging. Besides, it was not reported before nor as disease causing mutation nor as a SNP. Moreover, the PhyloP Score (evolutionary conservation across 49 vertebrates) indicates that the site of mutation is highly conserved. Therefore, we considered c.G10426A mutation in GPR98 a strong candidate for causing deafness in this patient.

As this mutation was identified by applying targeted exome capture and MPS, we verified it via PCR amplification and Sanger sequencing for the same patient sample. This step is essential, because NGS is still not approved by Clinical Laboratory Improvement Amendments (CLIA) (Lin et al., 2012).

Subsequently, the mutation was analyzed in the whole family and showed complete segregation with hearing loss in three members (II-2, II-3 and II-4), while three members are carriers for the mutation (I-1, I-2 and II-5). This step is important too to determine the parental origin of the identified mutation. See figure 4.1.

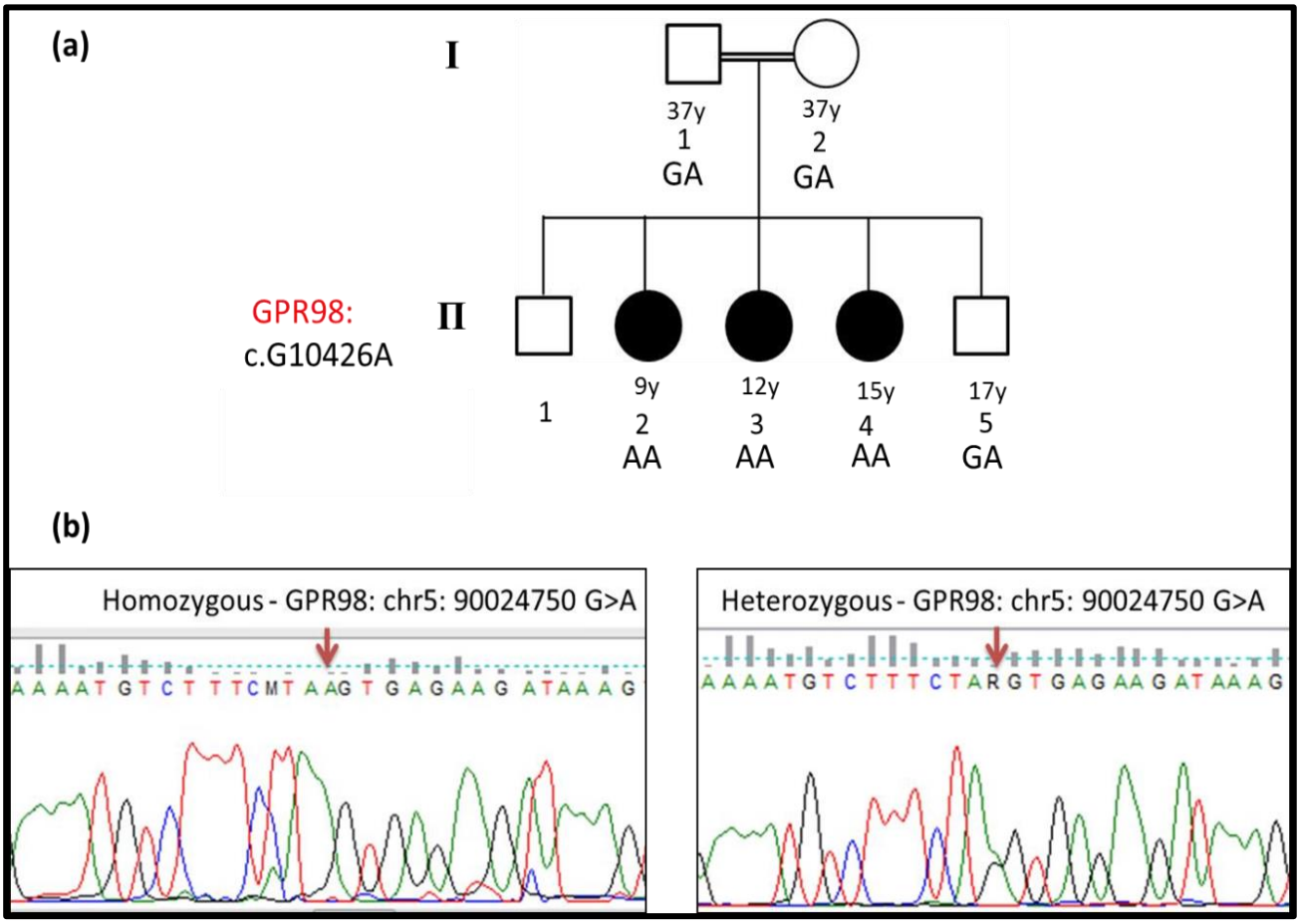


Figure 4.1 Pedigree of family EI with GPR98 mutation. (a)GPR98 chr5:90024750 G>A was discovered by targeted genomic capture and MPS. It was identified in II-2, who is homozygous (AA) for mutation, and it segregates in her two elder sisters. The parents and one brother (I-1, I-2 and II-5) respectively, are heterozygous for mutation. **(b)** Sanger sequences of variant for representative homozygous and heterozygous individuals. The mutation is indicated by the red arrow.

4.2 Filtration of the novel mutation in Palestinian controls.

In order to ensure that the found mutation is not an unrevealed common polymorphisms in the Palestinian population, we genotyped it among hearing controls and deaf controls and all of them share the same ethnicity. For this purpose we used Taqman genotyping assay, to save time and effort. (See table 4.2)

Table 4.2 Results of mutation filtration among deaf and normal controls.

Controls	Mutation presence
100 normal controls	negative
263 deaf controls	negative

263 unrelated deaf probands were screened for the new GPR98 mutation. They were all wild type for the mutation. 100 normal controls were screened for the same mutation, and found to be wild type. Moreover, no carrier of mutation was detected in either of them. In total, the approximate probability of the GPR98: c.G10426A mutation among deaf population in Palestine is nearly 0.4% (1 in 264), which is considered low.

So, we detected the mutation only in three members of family EI, and no carrier was found in both deaf and normal controls. This led us to conclude that this mutation in GPR98 gene is private to family EI.

This case reflects the high rate of consanguinity marriage among Palestinian population. Also, discovering a novel mutation in a single family proposes happening of de novo event. Whereas, the mutation would occur de novo to one of grandparents, those would marry between themselves from generations to another, and when families become large and extended, the probability of marrying two carriers increases, so the first affected child would be born in the single family or extended family (Zlotogora, 2010; Zlotogora, 2002).

4.3 Effect of novel mutation of pre-mRNA splicing of GPR98

In this study we present a patient with congenital deafness having a novel exonic splicing mutation in exon 49 c.G10426A of the GPR98 gene.

In this context we proved that this exonic splice site mutation leads to abnormal splicing of pre- mRNA of GPR98 and provides confirmation for the disease causing effect of the newly identified mutation.

To see the effect on m-RNA we performed nested RT-PCR, since the RNA transcript of GPR98 is found in low concentration in blood leukocytes, nested PCR would improve sensitivity and specificity of target as mentioned by Goode et al., (2002), in their research paper.

So, nested RT-PCR was performed on three family members in addition to anonymous wild type control. After two rounds we were able to see and analyze the product on 1.5% agarose gel. See figure 4.2.

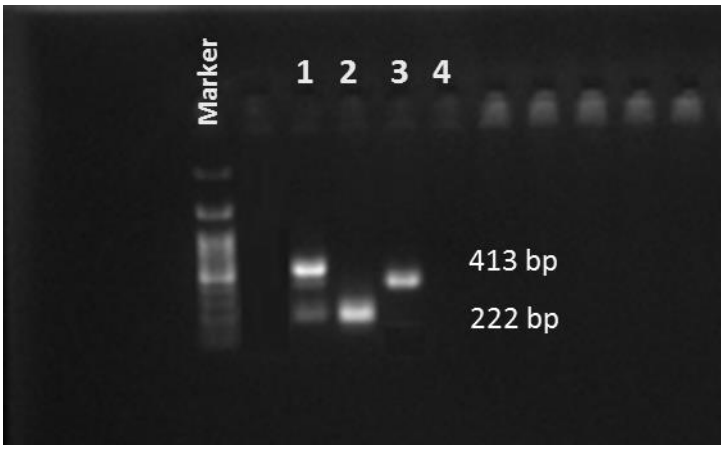


Figure 4.2 Analysis of nested RT-PCR product of GPR98: 90024750 (G>A). lane 1 shows cDNA of I-1 who is heterozygous; lane 2 shows cDNA of II-2 which is homozygous; lane 3 represents cDNA of AV5(wild type control), and lane 4 is NTC. Sizes of bands are indicated. (Note cDNA for I-2 didn't work)

As shown in the previous figure, the product of nested RT-PCR visualized on 1.5% gel has differentiated between the normal, affected patient and the carrier. As shown in figure 4.2, the father (I-1) is heterozygous to the mutation and has 2 bands with size 413 and 222 bp, respectively, and II-2 is mutant and has only one small band with size 222bp. While, the wild type control, has only a large band with size 413bp. This indicates that the abnormal splicing occur in the affected patient due to the mutation which leads to partial deletion of exon 49.

And this result is confirmed when we did Sanger sequencing, see figure 4.3. When we compare the wild type (AV5) with the mutant (II-2), we can notice in (II-2) that mutation results in activation of cryptic splice site, causing splicing to occur at earlier position between exon 49 and 50 and causes a partial skip of exon 49. Whereas, 210 nucleotides are deleted from exon 49. While, in AV5, splicing occurs normally between exon 49 and 50.

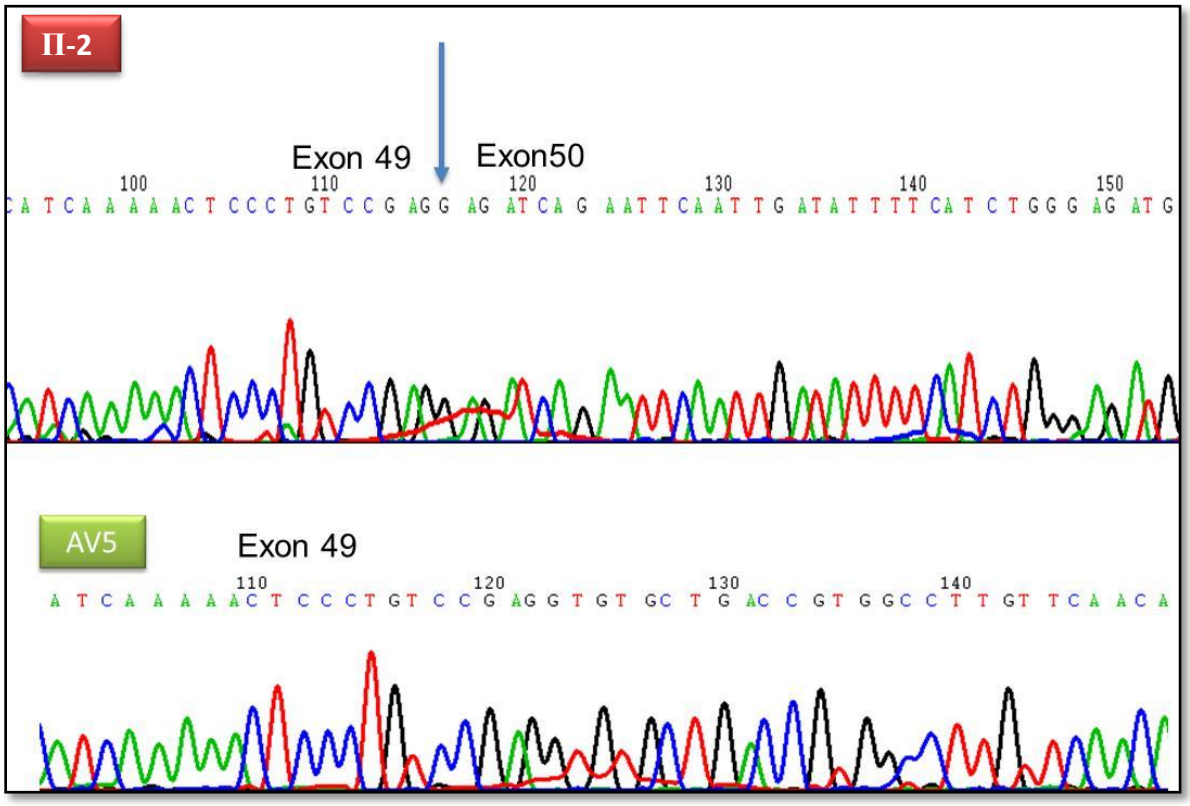


Figure 4.3 Electropherograms of nested RT-PCR product for GPR98 Chr5: 90024750 G>A mutation. Sequence of II-2 which is homozygous to mutation shows partial deletion of exon 49 at the position of arrow. While in AV5 which is wild type no deletion occurs to exon 49.

The following diagram illustrates how splicing occurs in both normal and mutant case. In normal case, the highly conserved consensus sequences, which are essential for normal splicing are: GT and AG dinucleotides, placed respectively at the 5' and 3'ends of an intron, besides directly adjacent nucleotides preceding ends of an intron and splice branch site in intron (Strachan and Andrew, 2004). These consensus sequences are recognized by

splicesome machinery, and subsequently splicing of pre-mRNA occurs normally for exons 49 and 50. See figure 4.4 (a).

When the mutation occurs at the last base of exon 49, which is an authentic splice site, then the 5' splice donor site is not recognized by splicing machinery. And then in its role, it searches for possible alternatives cryptic splice site to perform splicing. In case of mutant pre-mRNA, it is activated in earlier place within the exon as shown in figure 4.4 (b). Therefore, abnormal splicing occurs and causes partial deletion of exon 49.

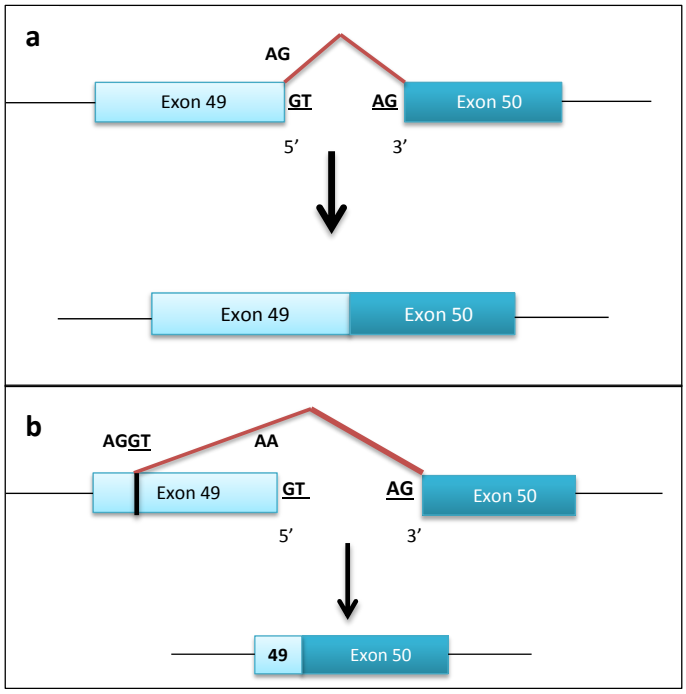


Figure 4.4 Splicing of pre-mRNA for exons 49 and 50 in both normal and mutant cases. (a) Normal splicing occurs between exons 49 and 50. (b) Aberrant splicing of pre-mRNA due to mutation occur at last nucleotide of exon 49, thus, alternative 5'splice site were recognized and causes partial skip of exon 49.

Partial skip of exon 49 may affect mRNA stability which would lead to its possible degradation or it may produce unstable protein.

4.4 Prediction the effect of novel mutation on protein sequence of GPR98

To see the effect on protein sequence and to locate exactly which part exon 49 translates in GPR98 protein, we blasted the translated nucleotides (transcript) to BlastX in Blast NCBI database. The normal protein sequence locates at position 3388 to 3476 of GPR98

total length 6307 amino acids. The normal sequence has 88 amino acids while the mutated sequence has only 18 amino acids. This means that 70 amino acids are missed in the mutant sequence. So, we conclude that the aberrant splicing on mRNA produces truncated protein missing 70 amino acids, which may affect protein stability. See figure 4.5.

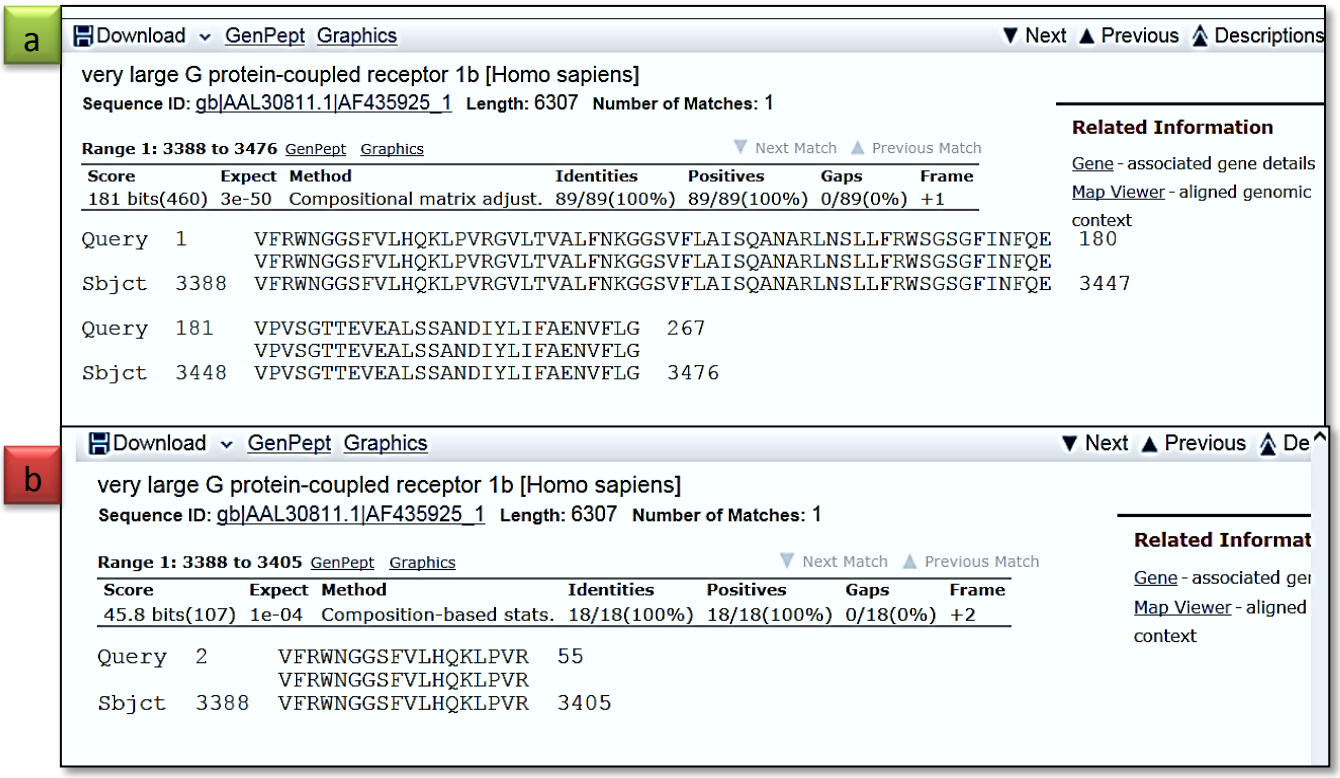


Figure 4.5 Blastx of exon 49 of GPR98. Part of GPR98 protein sequence resulted from blastX of translated nucleotide sequence of exon 49. (a) Normal protein. (b) Abnormal protein with 70 deleted amino acids.

In order to see which domains are found in residue (3388- 3476) of GPR98 protein, we put the normal sequence in Interproscan bioinformatics tool, which indicates active sites and domains in the protein sequence. We found that it has a highly conserved EPTP (Epitempin) domain besides part of epilepsy associated repeats (part EAR3, EAR4, of EAR5). EPTP domain and part of EARs repeats are not found in the mutated protein sequence. See figure 4.6.

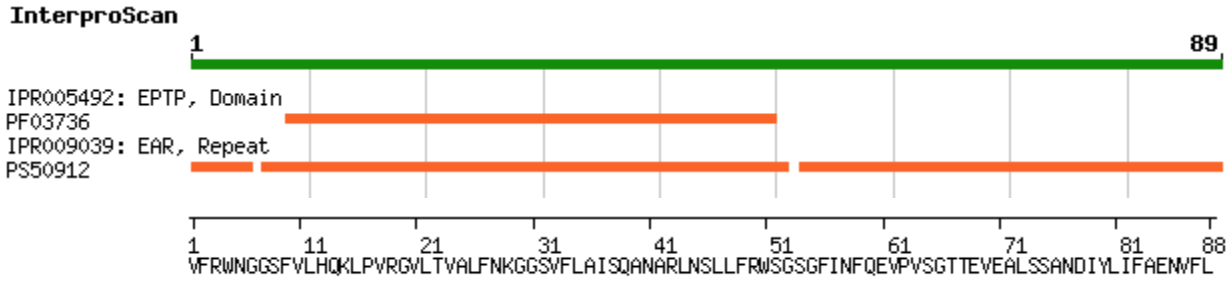


Figure 4.6 Interproscan of GPR98 protein sequence residues (3388-3347). The length of this part of protein is 89 amino acids. It indicates domains as follows; EPTP domain from residues 9 to 51, part of EAR3 from residues 1 to 6, EAR4 from residues 7 to 52 and part of EAR5 from residues 53 to 89.

Epitempin (EPTP) domain or epilepsy associated repeats (EAR) are seven copies of nearly 50 amino acids repeat, are located near the center of the ectodomain of VLGR1 specifically at residues 3198 to 3532 (Giber et al., 2005). It is assumed to be made of β sheets folded into seven bladed β- propeller structures. Although its role is still unknown, but it is thought to mediate protein interaction and ligand recognition by the receptor (McMillan and White, 2010; Scheel et al., 2002).

Similar EAR repeats domain are found in other proteins as Lgi1-4 and Tspear, they are thought to be implicated in different types of seizures and/or hearing loss (Sun et al, 2012). Besides Delmaghani et al.,(2012), showed in their study, proteins that contain EAR repeats are vital for development and function of auditory system and this explain deafness phenotype in our patient II-2 and her sisters (Delmaghani et al., 2012).

In our study the protein sequence translated by normal exon 49 occupies residues 3388 to 3476 of GPR98 protein. The novel identified mutation causes deletion of EPTP domain and partial deletion of EAR repeats, particularly EAR3, EAR4 and EAR5 in GPR98/VLGR1. Since the mutation occurs in a conserved location in exon 49, it causes deletion of highly conserved EPTP domain. According to Scheel et al.,(2002), mutations at conserved position of EAR repeats in GPR98 probably make the domain functionally inactive or unstable (Scheel et al., 2002).

In human, expression of GPR98 is basically localized in the stereocilary ankle links of cochlea and in photoreceptor cells of retina (McMillan and White, 2010). Mutations in Gpr98 gene are involved in causing seizure disorders and/or usher syndrome type 2 (USH2C) (Satub et al., 2002). However, it is proposed that the isoform b of GPR98 does not contribute significantly to any seizure disorder (Weston et al., 2004). Moreover, mutations in VLGR1b isoform were identified in patients with USH2C who have congenital deafness and progressive development of Retinitis Pigmentosa at the second decade of life or even later. (Yang et al., 2012b). Besides, it plays an important role in the development and conservation of normal hearing (McMillan and White, 2010).

This is confirmed in our study where that patient II-2 has mutation in GPR98, and isoform b is the one that was affected by the mutation. That explains why she expresses deafness phenotype and has no seizure disorder. Although mutations in GPr98 are involved in causing USH2C, novel mutations in GPR98 were identified in Chinese patients with nonsyndromic hearing loss (Yang et al., 2013). This paradox put us in front of a question yet to be answered: is hearing loss in this family syndromic or non syndromic?. Symptoms of Retinitis Pigmentosa are absent in our patient and her two elder adolescent sisters and it is possible that HL in this case is nonsyndromic. However, symptoms of Retinitis Pigmentosa does not develop until the second decade of life or even at later age. So, Usher syndrome cannot be ruled out in this family because the HL probands may develop Retinitis Pigmentosa in the future. Therefore, if they are suspected to have usher syndrome, we recommend to include PDZD7 gene in the molecular diagnostic process. Based on the information that PDZD7 is localized in ankle link complex and it localizes and stabilizes GPR98 in ankle link complex of cochlear hair cells (Garti et al., 2012). Besides, monoallelic mutations in both PDZD7 and GPR98 gene were found to cause digenic USH2 phenotype (Ebermann et al., 2010).

In our study, the EI pedigree shows that all affected patients are females. This seems consistent with what was thought that all found mutations in Gpr98 gene which cause USH2C affect only females. GPR98 mutations would cause a lethal phenotype in males (Wenston et al., 2004). However, Ebermann et al.,(2009), in their study identified novel GPR98 mutations in two male patients with USH2C, one of the patient shares the same

mutation and phenotype with his sister. So they concluded that gpr98 mutations can cause usher syndrome in both males and females (Ebermann et al., 2009).

This contraction leads us to conclude that there is probability that some mutations in GPR98 could be lethal to male, while others could affect both males and females. It might depend on the type and position of mutation in GPR98. But, in our case of study, we cannot infer that the novel mutation we found affects only females. Because, mutation is found in single family and it is likely not associated with usher syndrome type 2, so we have to monitor the track of the mutation when it passes from generation to another in the same family so we can draw a conclusion.

Studies on mice cochlea showed that different mutations in VLGR1 may lead to absence of ankle link complex and consequently lead to corruption of stereocilia development and maturation. Besides, the unique staircase arrangement of stereocilia would be destructed. (Yagi et al., 2006; Sun et al., 2012; McMillan and White, 2010). Moreover, it would lead to abnormal retinal function in aged mutants (McGee et al., 2006).

The same effect on stereocilia may be done by mutation GPR98:c.G10426A that is responsible for having congenital deafness in our patient. But in order to understand the specific cause of deafness due to a novel GPR98 mutation, we recommend performing mouse model studies to investigate mutation's effect on hair cells of cochlea and photoreceptor cells of retina.

CONCLUDING REMARKS

By applying targeted exome sequencing and MPS we were able to identify the causative gene and a novel mutation (GPR98 c.G10426A) responsible for having congenital deafness in patient II-2 in family EI. Also, it is found so far only in this single family in Palestine.

The GPR98 gene is implicated in causing usher syndrome type 2, which is characterized by congenital, moderately severe hearing loss and progressive Retinitis Pigmentosa. However, symptoms of Retinitis Pigmentosa are absent in our patient and her two adolescent sisters and it is possible that HL in this case is nonsyndromic .

The position of mutation leads to deletion of EPTP domain and partial deletion of Ear (3, 4, and 5) repeats in GPR98 protein. That is believed to play a role in ligand recognition by the receptor. The novel mutation may make the domain unstable or inactive.

To understand the effect of mutation on stereocilia in cochlea, we recommend performing studies on mouse model to determine the cause of deafness due to Gpr98 novel mutation.

As HL is highly heterogeneous in our Palestinian community we recommend using NGS approach to be able to identify the causative mutation with reduced time and effort. In addition, it is a good tool to discover new genes and mutations responsible for deafness yet to be identified.

Knowing causative genes and mutations is essential for genetic counseling; also it will contribute in understanding the mechanisms underlying hearing and hearing loss.

REFERENCES

Alberti, P. W. (2001). The anatomy and physiology of the ear and hearing. Occupational exposure to noise: Evaluation, prevention, and control: 53-62.

Araç, D., Aust, G., Calebrio, D., Engel, F., B., Formstone, C., Goffinet, A. et al. (2012). Dissecting signaling and functions of adhesion G protein–coupled receptors. Annals of the New York Academy of Sciences. 1276(1): 1-25.

American speech language hearing association. (2011). Type, degree and configuration of hearing loss. Audiology information series.7976-16.

Avraham, K., B and Kanaan, M. (2012). Genomic advances for gene discovery in hereditary hearing loss. Journal of basic clinical physiology and pharmacology. 23(3):93-97.

BigDye Terminator v3.1 Cycle Sequencing Kit. Protocol. (2006- 2010). Part number 4337035 Rev. B. Applied Biosystems. Foster City, CA.

Bitner-Glindzicz, M. (2002). Hereditary deafness and phenotyping in humans. British medical bulletin. 63: 73-94.

Brownstein, Z. and Avraham, K., B. (2009). Deafness genes in Israel: Implications for diagnostics in the clinics. Pediatric Research.66 (2):128-134.

De Keulenaer, S., Hellemans, J., Lefever, S., Renard, JP, De Schrijver, J., Van de Voorde, H. et al. (2012). Molecular diagnostics for congenital hearing loss including 15 deafness genes using a next generation sequencing platform. BMC Medical Genomics 5(1): 17.

Delmaghani, S., Aghaie, A., Michalski, N., Bonnet, C., Weil, D. and Petit, C. (2012). Defect in the gene encoding the EAR/EPTP domain-containing protein TSPEAR causes DFNB98 profound deafness. Human molecular genetics, 21(17), 3835-3844.

Dror, A. A. and K. B. Avraham (2010). "Hearing Impairment: a panoply of Genes and Functions." Neuron 68(2): 293-308.

Dror, A. A. and Avraham, K., B. (2009). Hearing Loss: mechanisms revealed by genetics and cell Biology. Annual Review of Genetics .43(1): 411-437.

Ebermann, I., Phillips, J. B., Liebau, M. C., Koenekoop, R. K., Schermer, B., Lopez, I., ... & Bolz, H. J. (2010). PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. The Journal of clinical investigation, 120(6), 1812.

Ebermann, I, Wiesen, M., HJ, Zrenner, E., Lopez, I., Pigeon, R. Kohl, S., et al. (2009). GPR98 mutations cause Usher syndrome type 2 in males. Journal of Medical Genetics 46(4): 277-280.

Farpon, R., C. and Ba˜nales, J, C. (2011). Hereditary hearing loss: Genetic counseling. Acta Otorrinolaringol Espanola. 63(3):218-229.

Gibert, Y., McMillan D., R., Kayes-Wandover, K., Meyer, A., Begemann, G. and 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. and Avraham K., B. (2002). Genetics of deafness: recent advances and clinical applications. Journal of basic and clinical physiology and pharmacology.13 (2):1-13

Goode, T., Ho, W-Z, O'Connor, T., Bustee, S., Doughlas, S., D., Shanahan, F. and O'Connell, J. (2002). Nested RT PCR, In O'Connell J. pp. 65-79. RT-PCR protocols. Springer

Grati, M. H., Shin, J. B., Weston, M. D., Green, J., Bhat, M. A., Gillespie, P. G., & Kachar, B. (2012). Localization of PDZD7 to the stereocilia ankle-link associates this scaffolding protein with the Usher syndrome protein network. The Journal of Neuroscience, 32(41), 14288-14293.

Hilgert, N., Smith, R., JH and Van Camp, G. (2009).Function and expression pattern of nonsyndromic deafness genes. Curr Mol Med. 9(5): 546–564

Illumina. (2012). An introduction to next generation sequencing technology Retrieved on April 5, 2013. From: www.illumina.com/NGS

Kochhar, A., Hildebrand M., S. and Smith, R., JH. (2007). Clinical aspects of hereditary hearing loss. Genetics in Medicine. 9(7): 393-408.

Lenz, D., R. and Avraham, K., B. Hereditary hearing loss: From human mutations to mechanism. 281(1–2): 3-10.

Lin, X., Tang, W., Ahmad, S., Lu, J., Colby, C., C., Zhu, J. et al. (2012). Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. Hearing Research 288(1–2): 67-76.

Mahboubi, H., Dwabe, S., Fradkin. M., Kimonis, V. and Dajlilian, H., R. (2012): Genetis of hearing loss: where are we standing now? Eur Arch Otorhinolaryngol .269:1733–1745

Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A. and Jabado, N. (2011). What can exome sequencing do for you? Journal of Medical Genetics. 48(9): 580-589.

Mathers, C., Smith, A. and Concha. (2000). Global burden of hearing loss in the year 2000. Global burden of disease, 18.

McGee, J., Goodyear, R., J., McMillan, D., R., Stauffer, E., A., Holt, J., R, Locke, K., G., et al., (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. Journal of Neuroscience. 26(24): 6543–6553.

McMillan D., R. and 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. In: Yona, S. and Stacey, M. (Eds). Adhesion GPCRS. (pp 76-86). Springer Us. ISBN: 978-1-4419-7913-1.

Millan, J., Aller, E., Jaijo, T., Blanco-Kelly, F., Gimenez-Pardo, A. and Ayuso, C. (2010). An update on the genetics of usher syndrome. Journal of ophthalmology.2011

Morton, C. C. and Nance, W.E. (2006). Newborn hearing screening: A silent revolution. New England Journal of Medicine. 354(20): 2151-2164.

Petit, C., Levilliers, J. and Hardelin, JP. (2001). Molecular genetics of hearing loss. Annu. Rev. Genet. 35:589–646

Raviv, D., Drorr A., A. and Avraham, K., B. (2010) Hearing loss: a common disorder caused by many rare alleles. Annals of the New York Academy of Sciences. 1214(1): 168-179

Rehman, A., U., Morell R., J., Belyantseva, I., A., Khan, S., Y., Boger, E., T., Shahzad, M.et al. (2010). Targeted capture and next-generation sequencing identifies C9orf75, encoding Taperin, as the mutated gene in nonsyndromic deafness DFNB79.The American Journal of Human Genetics 86(3): 378-388.

Rzadzinska, K., A., Schneider, M., E., Davies, C., Riordan, G., P. and Kachar, B. (2004). An actin molecular treadmill and myosins maintain stereocilia functional architecture and self-renewal. The journal of cell biology. 164 (6):887-897.

Scheel, H., Tomiuk, S. and Hofmann K. (2002). A common protein interaction domains links two recently identified epilepsy genes. Human molecular genetics. 11(15):1757-1762.

Shalit, E. and Avraham, K., B. (2008).Genetics of hearing loss. In: Schact, J. and Fay, R., R (Eds.). Auditory Trauma, Protection, and Repair. (Pp9-47). Springer. ISBN: 978-0-387-72561-1

Shearer, A. E., Hildebrand, M., S., Sloan, C., M. and Smith R., JH. (2011). Deafness in the genomics era. Hearing Research 282(1): 1-9.

Shen, J. and Deskin, R., W. (2004). Genetics hearing loss. In: Quinn, F., B. and Rayan, M., W. (Eds). Grand Rounds Presentation, UTMB, Dept. of Otolaryngology. Retrieved on January 5, 2013 from: http://www.utmb.edu/otoref/grnds/Genetic-HL-0004/Genetic-HL-0004.htm

Sirmaci, A., Edwards, Y., J., Akay, H. and Tekin, M. (2012). Challenges in whole exome sequencing: an example from hereditary deafness. PloS one. 7(2): e32000.

Smith, R., JH, Shearer, A., E., Hildebrand, M., S.and Camp, G., V. Deafness and Hereditary Hearing Loss Overview. 1999 Feb 14 [Updated 2013 Jan 3]. In: Pagon RA, Adam MP, Bird TD, et al., editors. GeneReviews™ [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2013. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1434

Strachan, T. and Read, A., P. (2004). Human molecular genetics. 3rd ed. Garland Science: Taylor and Francis group, London and New York.

Staub, E., Perez-Tur, J., Siebert, R., Nobile, C., Moschonas, N., K., Deloukas, P., and Hinzmann, B. (2002). The novel EPTP repeat defines a superfamily of proteins implicated in epileptic disorder. Trends in biochemical sciences. 27(9): 441-444.

Sun, J.-P, Li, R., Ren, HZ, Xu, AT, Yu, X. and Xu, ZG. (2012). the very large G protein coupled receptor (Vlgr1) in hair cells. Journal of Molecular Neuroscience: 1-11.

Tucker, T., Marra, M. and Friedman J., M. (2009). Massively parallel sequencing: the next big thing in genetic medicine. The American Journal of Human Genetics .85:142–154.

Walsh, T., Shahin, H., Elkan-Miller, T., Lee, M., K., Thornton, A., M., Roeb, W.et al. (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(1): 90-94.

Weston, M.,D., Luijendijk, M.,WJ, Humphrey, K.,D., Mo¨ller, C. and Kimberling, W.,J. (2004).Mutations in the VLGR1 gene implicate G- protein signaling in the pathogenesis of usher syndrome type II. American journal of human genetics. 74:357–366

World Health Organization. (February, 2013). Deafness and Hearing loss. [Fact sheet]. Retrieved from http://www.who.int/mediacentre/factsheets/fs300/en/

Yagi, H., Tokano, H., Maedo, M., Takabayashi, T., Nagano, T., Kiyama, H., et al. (2007). Vlgr1 is required for proper stereocilia maturation of cochlear hair cells. Genes to cells.12: 235-250.

Yang, J., Wang, L., Song, H. and Sokolov, M. (2012). Current understanding of usher syndrome type II.

Yang, J. (2012). Usher Syndrome: Genes, Proteins, Models, Molecular Mechanisms, and Therapies. In Sadaf Naz (Ed.), Hearing Loss (pp. 293-328). Rijeka: Intech, ISBN: 978-953-51-0366-0Frontiers in Bioscience .17: 1165-1183

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 journal of rare diseases, 8(1), 85.

Zahnert, T. (2011). The differential diagnosis of hearing loss. Deutsches Ärzteblatt International. 108(25): 433-442

Zlotogora, J. (2010). Molecular basis of autosomal recessive diseases among the Arabs and Druze in Israel. Hum Genet.128:473–479

Zlotogora, J. (2002). Molecular basis of autosomal recessive diseases among Palestinian Arabs. American journal of medical genetics. 109: 176-182

Bioinformatics tools:

- Blast NCBI: web page: http://blast.ncbi.nlm.nih.gov/Blast.cgi

- G-protein coupled receptor 98. (2013,March,18) Gene cards, Retrieved, March 25,2013 from http://www.genecards.org/cgi-bin/carddisp.pl?gene=GPR98

- Interproscan: webpage on Swiss model: http://swissmodel.expasy.org

Zdobnov E.M. and Apweiler R. (2001). InterProScan - an integration platform for the signature-recognition methods in InterPro Bioinformatics 17: 847-848

- Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402

Internet websites.

http://www.studyblue.com/notes/note/n/ch-46-animal-sensory-systems-and-movement/deck/3080378 Retrieved on June 25, 2013

http://www.who.int/topics/deafness/en/index.html Retrieved on April 5, 2013

http://www.who.int/mediacentre/factsheets/fs300/en/ Retrieved on April 23, 2014

http://www.cochlea.org/en/treatments/, Retrieved on July 4, 2013.

http://genome.ucsc.edu/ , Retrieved on March 25, 2013.

http://dictionary.reference.com/browse/eugenics Retrieved: 11.4.2013

http://news.xinhuanet.com/english/world/2013-02/19/c_132178891.htm Retrieved: 11.4.2013

Jean Christie, "Keller, Helen", in the *Gallaudet encyclopedia of deaf people and deafness* (New York: McGraw-Hill, 1987, vol.2, p.125) Retrieved from http://www.audira.org.uk/en/modernising-attitudes-to-hearing-care/where-our-attitudes-to-hearing-come-from/respect-for-hearing/item/16-hearing-our-most-important-sense (date: July 6 ,2013)

APPENDICES

Taqman assay (Allele discrimination (AD) assay)

This assay relies on the use of fluorogenic 5’ chemistry recognized also as Taqman probe based chemistry. Whereas, two pairs of primer and probe in each reaction are used to permit genotyping of two possible variants at the single-nucleic polymorphism (SNP) site in a target template sequence.

Each probe is attached to fluorescent dye detectors which are VIC dye and FAM dye. These two probes differ in sequence only at the SNP site. For example to target an SNP site in the DNA sequence. One fluorescent dye detector is a perfect match to the wild type (allele 1) and the other fluorescent dye detector is a perfect match to the mutation (allele 2).

So AD assay can classify unknown samples as:

1. Homozygous (Have only allele 1 or have only allele 2)
2. Heterozygous (Have both alleles 1 and 2)

   The process of real time PCR reaction is illustrated in Figure A-1.

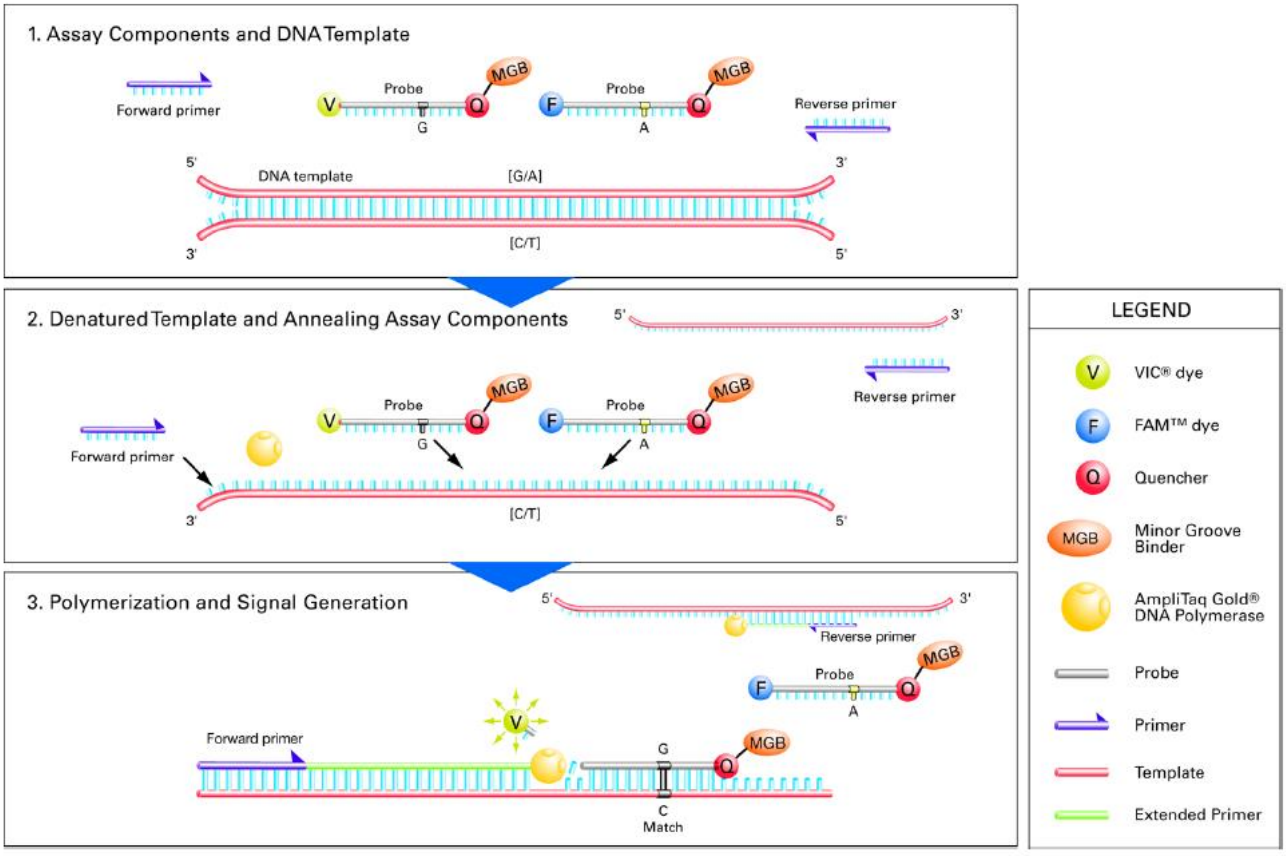


Figure A-1. Real time PCR reaction. 1. Assay component all available two primers and two probes, whereas, a 5' reporter dye (VIC or FAM) and a 3' quencher dye are covalently linked to the wild-type and mutant allele probes. 2. When denaturation occurs, primers anneal to the template DNA and Taqman probes hybridize to the targeted SNP site. 3. Once matched probe bind, extension begins, the reporter and quencher dyes are released due to the 5' nuclease activity of the Taq polymerase, resulting in fluorescence of the reporter dye.

Experiment flowchart:

The experiment flowchart is illustrated below. After preparing the reaction plate and load it into instrument. Three important steps are followed:

1. Pre-read run: it creates AD plate document that records information about the run such as sample names, detectors and markers. Besides it stores baseline fluorescence associated with primers and probes for each well before PCR amplification starts.
2. Amplification run: it makes real time PCR data by creating an AQ (Absolute quantitation) plate document.
3. Post read run: it uses AD plate document to assign allele calls using the amplified data, by subtracting the baseline fluorescence recorded during the pre-read run. (Allelic discrimination getting started guide, 2006-2010)

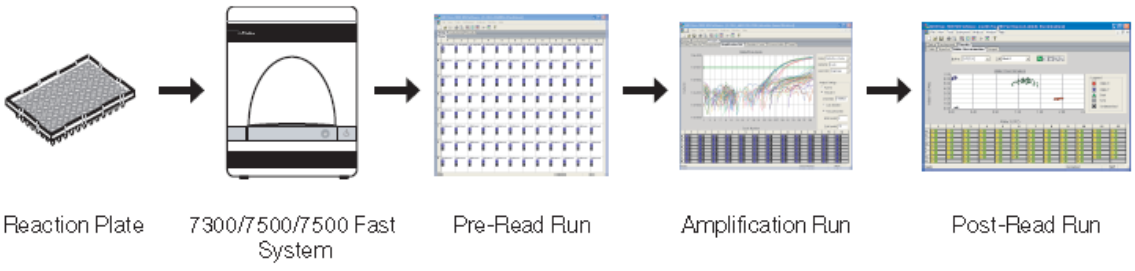


Figure A-2. Procedure flowchart of Taqman based genotyping assay.

Nested RT-PCR:

Nested Reverse Transcriptase (RT) PCR is a technique used to increase sensitivity of the RT-PCR product. It reduces the presence of primer dimers and non-specific binding in PCR products that results due to binding primers to unexpected site in template. It consists of two rounds and they are like normal PCR amplification. But, in the second round the product of primary RT-PCR is used as a template instead of DNA template. Whereas, only the amplified products are re-amplified and no amplification of nonspecific band would occur. And thus enabling product specificity to be maintained over the high number of amplification cycles combined in the primary and secondary PCRs. (Goode et al, 2002)

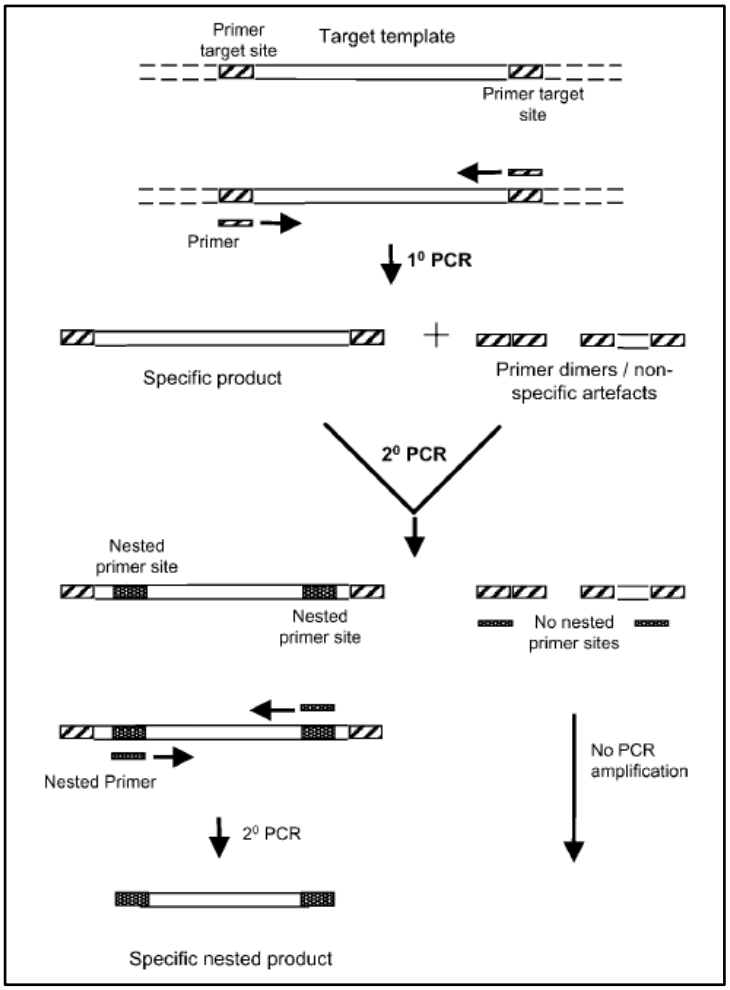


Figure A-3. Nested RT-PCR. It consists of two rounds of PCR. in second round instead of using template DNA, the primary RT_PCR product is used template .Two sets of primers are used and thus enabling product sensitivity by getting rid of non -specific bands that may result in the first round of PCR.