UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di: Salute della donna e del bambino

CORSO DI DOTTORATO DI RICERCA IN: MEDICINA DELLO SVILUPPO E SCIENZE DELLA PROGRAMMAZIONE SANITARIA
CURRICULUM: EMATO-ONCOLOGIA, GENETICA, MALATTIE RARE E MEDICINA PREDITTIVA
CICLO XXX

HEREDITARY HEARING LOSS: FROM MOLECULAR BASES TO PHENOTYPIC CARACTHERIZATION

Coordinatore: Ch.mo Prof. Carlo Giaquinto
Supervisore: Ch.mo Prof. Alessandra Murgia
Co-Supervisore: Dr. Elisa Bettella

Dottoranda: Federica Cesca

To my father

INDEX

List of figures

List of Tables

Abbreviations

ABR: Auditory Brain Stem Responses

bp: base-pair

CDS: coding sequence

EAC: External Auditory Canal

ESS: exonic splicing silencer

gDNA: genomic DNA

HL: Hearing loss

IGV: Integrative Genome Viewer

IHC: inner hair cells

INDEL: insertion/deletion

ISP: Ion Sphere™ Particles

MAF: Minor Allele Frequency

NGS: Next Generation Sequencing

NSHL: non-syndromic hearing loss

OAE: Otoacoustic Emissions

OHC: outer hair cells

PGM: Personal Genome Machine

PTA: pure tone average

SD: standard deviation

SHL: syndromic hearing loss

SNHL: Sensorineural Hearing Loss

SNV: Single Nucleotide Variants

ssDNA: single strand DNA

TMAP: Torrent Mapping Alignment Program

UTR: untraslated region

WES: Whole Exome Sequencing

WGS: Whole Genome Sequencing

WS: Waardenburg syndrome

Abstract

Hearing loss is one of the most common birth defects in developed countries. Approximately one/two in 1000 newborns are diagnosed with bilateral permanent sensorineural hearing loss.

Hereditary hearing loss can be syndromic (about 25%), in which deafness is associated with other signs and/or symptoms, and non-syndromic (about 75%), in which no other clinical features are present.

Non-syndromic hearing loss (NSHL) is characterized by a vast genetic heterogeneity with more than 160 loci described in humans and 100 genes so far identified. NSHL generally follows simple Mendelian inheritance and is predominantly transmitted as an autosomal recessive trait (75-80%), although other modes of inheritance are possible: autosomal dominant (20%), X-linked (2-5%) and mitochondrial (1%).

Given the high genetic heterogeneity of HL, tests based on NGS technologies are rapidly replacing many single-gene Sanger sequencing tests, due to their technical limits and higher costs.

The aim of this work was translational, with the goal to develop advanced molecular tools, with high diagnostic rate, and to investigate the genetic bases of NSHL in a population of Caucasian individuals, mainly of pediatric age.

A customized NGS targeted panel of 59 genes, strongly associated, in Caucasians, with NSHL or with SHL, which onset is usually characterized by isolated deafness (i.e. Pendred and Usher syndrome), was designed and developed.

The Ion Torrent PGM™ platform and a customized bioinformatics pipeline have been used for the analysis of DNA samples collected from clinically highly selected subjects with a previous negative test for the frequently mutated *GJB2* gene.

A series of 78 cases has undergone a complete study; an etiological diagnosis has been established for 34 of these subjects, with an overall diagnostic yield of 43.6%.

For each tested subject, an average depth of coverage of 249 X in the analyzed sequences and a mean of 499 variants were obtained.

Likely causative identified variants were located in the following 20 genes: *CDH23, GJB2, COCH, MYO7A, ADGRV1, EYA4, OTOG, SLC17A8, TMPRSS3, ACTG1, CEACAM16, COL11A2, GJB3, KCNQ4, MYH9, MYO6, PTPRQ, SLC26A4, STRC, TMC1.*

The most frequently mutated gene in our cohort was *CDH23,* which, even in our cases, accounted both for NSHL and Usher syndrome type 1 phenotypes.

A novel *EYA4* mutation, identified in two related subjects with post-lingual progressive deafness, was found to co-segregate in two individuals of the same family, with a Waardenburg syndrome phenotype, due to a novel *PAX3* gene mutation.

All the identified variants were collected in an *in-house* database that proved an invaluable tool for the identification of recurrent variants or possible alignment errors, and for further stratification and correlation between genotype and phenotype.

In conclusion, the targeted gene-panel we have developed, in combination with the *in-house* bioinformatics pipeline, is a proven sensitive diagnostic tool capable of providing an extremely competitive diagnostic yield.

Our work further demonstrates the importance of integrating the power of NGS technology and data process with the fundamental role of a strong clinical evaluation in keeping with what is expected from modern medicine.

Riassunto

Nei paesi sviluppati, l'ipoacusia rappresenta uno dei principali deficit presenti alla nascita; circa 1/2 su 1000 neonati vengono diagnosticati con una perdita uditiva neurosensoriale bilaterale permanente. L'ipoacusia ereditaria può essere sindromica, in circa il 25% dei casi, in cui la sordità è associata ad altri segni e/o sintomi e non sindromica (circa il 75%), in cui non sono presenti altre caratteristiche cliniche.

L'ipoacusia neurosensoriale non sindromica è caratterizzata da una vasta eterogeneità genetica con più di 160 loci descritti nell'uomo e circa 100 geni finora identificati. Tale ipoacusia viene prevalentemente ereditata in modo autosomico recessivo (75-80%), poi autosomico dominante (20%) e in rari casi è ereditata per via X-linked (2-5%) o mitocondriale (1%).

Considerando l'elevata eterogeneità dei geni associati ad ipoacusia, le tecnologie NGS stanno rapidamente sostituendo il sequenziamento Sanger su singolo gene, a causa dei limiti tecnici e dei costi di quest'ultimo.

Lo scopo di questo lavoro è stato traslazionale, ovvero sviluppare uno strumento molecolare avanzato, con un elevato *rate* diagnostico, per indagare le basi genetiche dell'ipoacusia non sindromica in una popolazione di individui caucasici, soprattutto di età pediatrica.

Allo scopo, è stato quindi disegnato un pannello genico comprendente 59 geni fortemente associati, nei Caucasici, ad ipoacusia non sindromica o ad ipoacusia sindromica, in particolare includendo quelle condizioni in cui la sordità insorge prima degli altri sintomi, come apparentemente isolata (sindrome di Pendred e sindrome di Usher).

L'analisi NGS è stata condotta sulla piattaforma Ion Torrent PGM™ ed è stata sviluppata una pipeline di analisi dati per ottimizzare la chiamata delle varianti, utilizzando criteri di filtraggio *ad hoc*.

In totale sono stati analizzati 78 soggetti, caratterizzati clinicamente e testati per le mutazioni nel gene *GJB2* e le delezioni del gene *GJB6,* risultando negativi.

Per 34 soggetti sui 78 analizzati è stato possibile ottenere una diagnosi molecolare, consentendo una resa diagnostica complessiva del 43,6%.

Per ciascun soggetto testato è stata ottenuta una copertura media di 249X e una media di 499 varianti.

Le varianti possibilmente causative identificate sono state localizzate nei geni: *CDH23, GJB2, COCH, MYO7A, ADGRV1, EYA4, OTOG, SLC17A8, TMPRSS3, ACTG1, CEACAM16, COL11A2, GJB3, KCNQ4, MYH9, MYO6, PTPRQ, SLC26A4, STRC, TMC1.*

Nella coorte di soggetti analizzati il gene più frequentemente mutato è stato *CDH23*, identificato sia in soggetti con ipoacusia non sindromica che in un bambino con sindrome di Usher di tipo 1.

E' stata inoltre identificata una nuova mutazione nel gene *EYA4*, presente in due soggetti correlati con sordità progressiva post-linguale; tale variante è stata trovata anche in altri due soggetti della stessa famiglia, i quali presentavano un fenotipo di sindrome di Waardenburg e nei quali è stata identificata una nuova variante nel gene *PAX3*.

Tutte le varianti identificate sono state raccolte in un database interno, che costituisce uno strumento prezioso per identificare le varianti ricorrenti, eventuali errori di allineamento e per un'ulteriore stratificazione e correlazione tra genotipo e fenotipo.

In conclusione, il pannello genico sviluppato, in combinazione con una pipeline bioinformatica *ad hoc*, si è dimostrato uno strumento diagnostico molto sensibile ed in grado di fornire diagnosi eziologica con una resa diagnostica estremamente competitiva.

Questo lavoro dimostra l'importanza di coniugare le nuove potenti tecnologie di analisi genomica NGS e di processamento dei dati con un'accurata caratterizzazione clinica, in linea con le richieste della moderna medicina.

1. INTRODUCTION

1.1 Anatomy and physiology of the ear

The auditory system is composed by three parts, illustrated Figure 1.1, that are necessary to proper transduction of the sounds: i) The external ear that comprises the auricle and the external auditory canal (EAC). The EAC directs sounds from the auricle to the tympanic membrane. The external ear alters sound wave amplitudes and provides a mechanism for amplification of different sounds within the frequencies of human speech. The sound waves are translated into mechanical stimuli in the form of vibrations, which are transmitted through the middle ear. ii) The middle ear is an air-filled space that consists of the tympanic membrane, tympanic cavity, ossicles and associated muscles. It is connected to the back of the nose by the Eustachian tube. The three ossicles (malleus, incus and stapes), located between the tympanic membrane and the cochlea, conduct sound from the tympanic membrane towards the cochlea.

iii) The inner ear comprises the cochlea, responsible for hearing, and the vestibular system, responsible for balance [1].

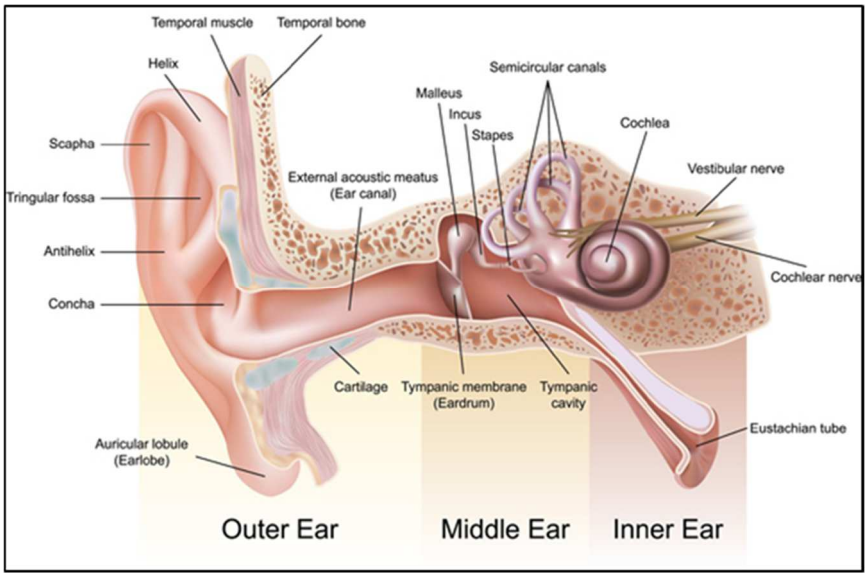


Figure 1.1 Anatomy of the ear; it is divided in three part: the outer ear, middle ear, inner ear.

1.1.1 The cochlea

The cochlea is a bony coiled tube shaped like a snail shell, as shown in Figure 1.2, filled with fluid called perilymph. This shell is divided in three compartments, called *scalae*: the scala vestibuli, the scala media, filled with endolymph, and the scala tympani; the basilar membrane divided the scala timpani from scala media, while Reissner's membrane divide scala media from scala vestibuli.

The scala media contains the cochlear sensory epithelium, the organ of Corti, which sits on top of the basilar membrane and is ultimately responsible for the mechano-electrical transduction of sound. The organ of Corti is composed of two classes of hair cell, inner and outer, with distinct functions and supporting cells.

Information about the acoustic environment (speech, music or other sounds in the outside world) is relayed primarily by the electrical signals of inner hair cells (IHCs), whereas the main task of outer hair cells (OHCs) is to boost the stimulus by electromechanical feedback [2].

Sound-induced mechanical vibration of the middle ear is transmitted to the cochlea, generating movements of its associated fluids. As a consequence, deflection of the basilar membrane activates the sensory cells that transduce the mechanical stimulations into electrical signals, in particular three rows of OHCs amplify the vibrations; then the mechanical signals are transferred onto IHCs, which transmit the information to afferent neurons. Hair cells at the base of the cochlea respond to the highest frequencies and those at the apex to the lowest. Sound frequencies are therefore relayed to the nervous system as a tonotopic map [3].

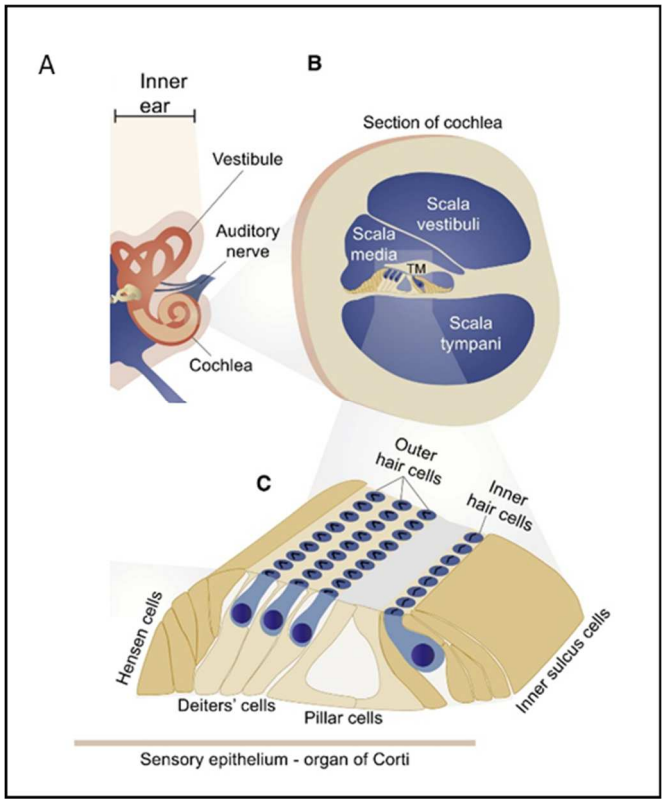


Figure 1.2 A) Schematic representation of the cochlea and the vestibule. B) A cross-section of the cochlear duct presents the scala media, scala tympani and scala vestibule, as well as the tectorial membrane (TM) over the organ of Corti. C) An enlargement of the organ of Corti, showing three rows of OHCs and one row of IHCs, flanked by various types of supporting cells. Figure modified from Dror et al, 2010 [4].

1.1.2 The Vestibular system

The vestibular organ consists of three semicircular canals, and two satellite organs: the utricle and saccule, as illustrated in Figure 1.3. The vestibular system is the sensory system that has a primary role in the sense of balance and is responsible of movement coordination in the spatial orientation. Receptor organs in the semicircular canals respond to angular acceleration while receptor organs in the utricle and saccule respond to linear acceleration and static tilt.

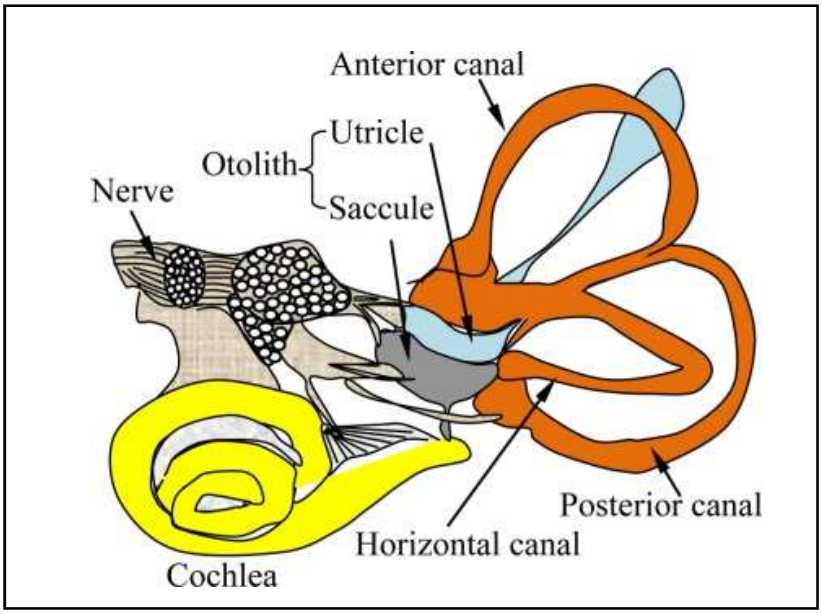


Figure 1.3 Schematic representation of the vestibular system. Figure modified from Zeng et al, 2011 [5]

1.2 Hearing and hearing loss

Normal mammalian auditory functions rely on two categories of function: mechanical and electrochemical. Malfunction of either of these auditory processing components results in hearing loss (HL), with considerably different molecular etiologies, clinical presentations, and possible appropriate management.

Hearing loss is clinically classified as conductive, definition which refers to a malfunction of the mechanical components of the auditory processing, sensorineural referring to malfunctions of the electrochemical or neurological components, and mixed when different aspects are present at the same time. Conductive hearing loss typically results from middle ear pathology, as tympanic membrane perforation, ossicular discontinuity or fixation, or middle ear infections, that are often amenable to improvement with either amplification (i.e. hearing aids) or surgical procedures. Sensorineural hearing loss (SNHL), on the other hand, can result from malfunction anywhere along the auditory pathway, from the hair cell to higher-order central auditory processing loci [6]. SNHL can be congenital or develop at a later stage in life.

Hearing loss is the most common birth defect and sensory disorder in the industrialized countries, with a prevalence of two to four in 1,000 newborns [7]. It can be caused by genetic factors (60-70% of cases), or by environmental causes as trauma, medications, medical problems, environmental exposure and Cytomegalovirus infection, that is the most common non-genetic cause of HL in children and can manifest in early childhood, can be unilateral or bilateral and is often progressive [8].

Newborn screening for hearing defects has an important role in treatment and rehabilitation strategies, such as cochlear implantation. Early etiological diagnosis in newborns and infants allows the effective monitoring of possible complications and can indicate which therapy is most suitable and appropriate. It is also necessary for a more accurate genetic counseling for parents and other family members.

1.2.1 Measuring the hearing threshold

Audiological diagnosis consists of three phases: identification of subjects at risk, definition of hearing loss and/or subjects' features, verification of appropriateness of diagnosis itself and rehabilitation

program. Strategies and methods of audiological diagnosis include an integration of data coming from objective methods with clinical and behavioral data [9].

The hearing ability is tested by stimulating the auditory system with a sound. This can be obtained through air conduction, which tests all the components of the auditory system, or through bone conduction via a transmitter placed on the skull, which bypasses the outer and middle ear.

A conductive hearing loss is suspected when the auditory response to air conduction is reduced but bone conduction is normal. However, SNHL is most likely if both air and bone conductions are reduced.

The subjective test consists on response from the individual that must actively indicates which tones at fixed levels, that are presented, can listen. The test consists of presenting beeps as varying intensities for different test frequencies, recording at each frequency the lowest intensity at which there are responses from the listener.

The response is represented on an audiogram that shows the hearing level in deciBels (dB) at multiple frequencies (Hz). The threshold audiogram produces a picture of how a person hears air-conducted signals, such as pure tones.

Normal hearing frequencies range between 20 and 20,000 Hz, with speech perception that is around 2,000 Hz. The normal speech perception audiogram has the shape of a banana, in which all the phonemes, or sounds of audible human language are represented; it is called “speech banana audiogram”, as reported in Figure 1.4. Those with a normal hearing can also hear sounds above this range, as well as high-frequency and low-frequency sounds.

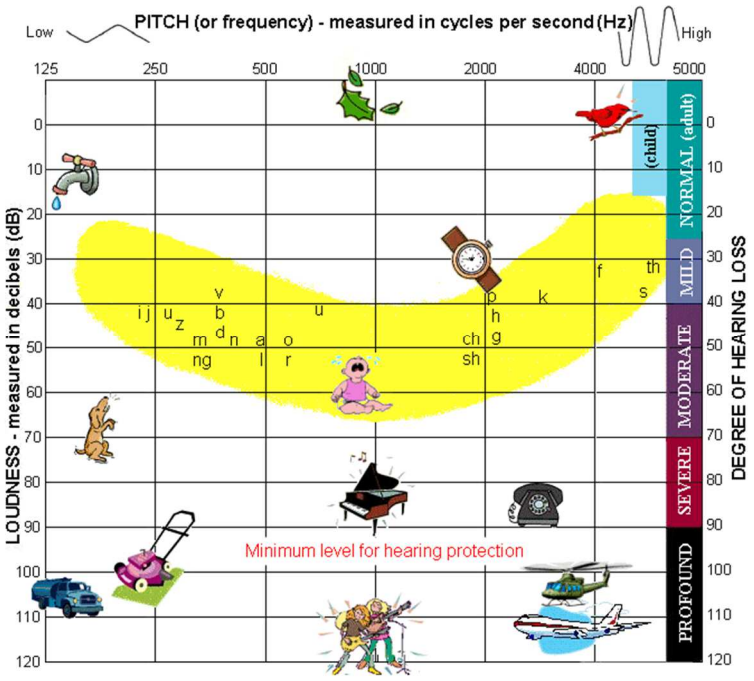


Figure 1.4 Speech banana audiogram. In yellow is represented the audible phonemes and sound. On the X-Axis are plotted the frequencies, measured in Hz. On Y-axis is reported the loudness, expressed in db. On the right is reported the degree of hearing loss.

Hearing loss can be described according to severity, depending on the auditory threshold: 0-20 dB is within normal limits, 21-40 dB: mild; 41-70 dB: moderate, 71-95 dB: severe , and ≥96 dB, profound [10]. Depending on the type of HL, a characteristic audiogram shape appears: high-frequency HL is described as “down-sloping”; low-frequency as “rising” and mid-frequency is “cookie-bite.”

To obtain an objective audiogram profile, especially in young children or during the neonatal screening, different tests are available that should be used in combination. In particular: i) otoacoustic emissions (OAEs) measure the returned sound from the cochlear cells after a sound stimulation is sent to the cochlea. ii) Auditory brain stem responses (ABRs) measure the evoked potential generated by sound after it has been transduced into an electrical signal and is transmitted from the VIII cranial nerve to the brain.

In the context of permanent childhood hearing loss, early audiological diagnosis is a prerequisite for activation of an adequate rehabilitation program to prevent or limit the known effects that auditory deprivation determines on language development and cognitive skills in neonates.

1.3 Hereditary Hearing Loss

Hereditary hearing loss (HHL) is characterized by a vast phenotypic heterogeneity. The first human gene associated with deafness was *GJB2* (Cx26, OMIM # 121011) in 1997 [11]. Even before the identification of the mutations in this gene, it was clear that a significant proportion of HL was hereditary and therefore due to gene mutations. Since the first gene discovered, great scientific advancements have been made and many genes now considered responsible of HL have been identified.

So far 160 loci have been described in humans and 100 genes have been identified, distributed throughout the chromosomes, as reported in Figure 1.5 (Hereditary Hearing Loss Homepage; http://hereditaryhearingloss.org/ last accessed September 2017).

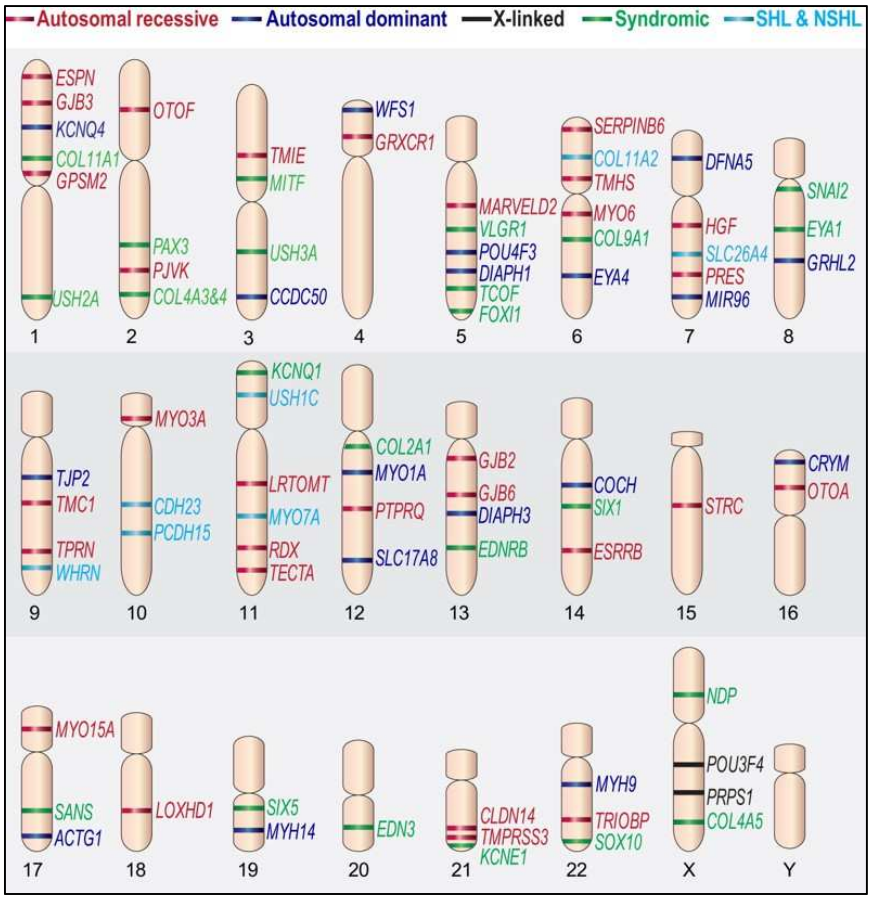


Figure 1.5 Chromosomal location of genes associated with hearing loss. Figure modified from Dror et al, 2010 [3].

The genes so far identified, encode a large variety of proteins with many functions in the inner ear, such as genes responsible for gene-regulation, fluid homeostasis, synaptic transmission and hair cell bundle morphology and development [12].

HHL can be syndromic hearing loss (SHL) in about 25% of cases, in which deafness is associated with other signs and/or symptoms, and non-syndromic (NSHL), in 75% of cases, in which there are no additional clinical abnormalities.

1.4 Syndromic Hearing Loss

SHL is defined by the presence of HL associated with defects in many different organs, such as anomalies of the eyes, kidneys, the musculoskeletal and the nervous systems, as well as pigmentary disorders and other alterations, as shown in Figure 1.6 [13].

Among the well-known complex conditions involving HL, Pendred, Usher and Waardenburg syndromes are the most common ones [14]. In these conditions (mainly Pendred and Usher) the hearing loss often is present before the onset of the other clinical features, giving the impression of a non-syndromic form of deafness. Several genes associated with SHL are also involved in NSHL, as in the case of SLC26A4, whose mutations are responsible both of Pendred syndrome and DFNB4 and hearing loss with Enlarged Vestibular Acqueduct (EVA syndrome).

The complete list of known genes associated with SHL is reported in Appendix 1.

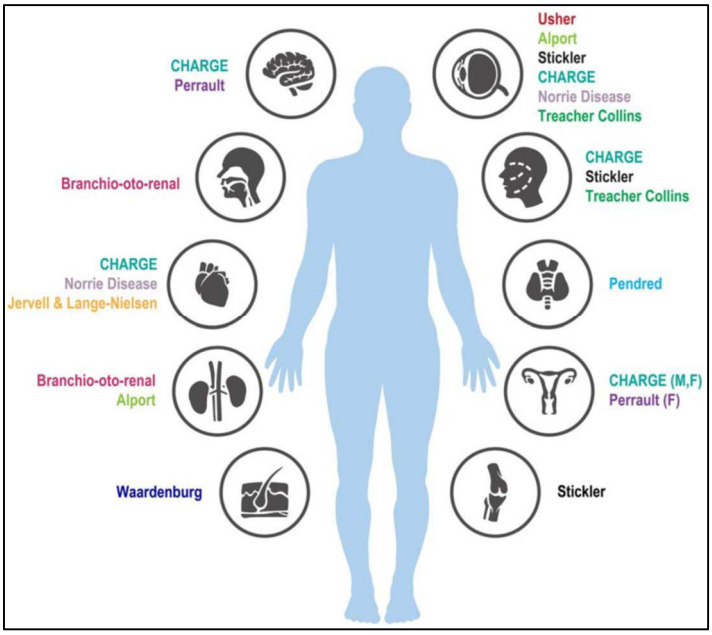


Figure 1.6 Schematic representation of the different organs involved in SHL. F, female genitals; M, male genitals. Figure modified from Koffler et al. 2015 [13].

1.4.1 Pendred Syndrome

The estimated prevalence of Pendred syndrome is 7.5 per 100,000 newborns and it accounts for approximately 1 to 8% of the cases of congenital deafness; it is inherited as autosomal recessive.
The audiological phenotype varies from mild to profound hearing loss with pre or post-verbal onset and a typically fluctuant and progressive course. A common feature among patients is an enlarged vestibular aqueduct (EVA). Pendred syndrome also features thyroid dysfunction, leading to hypothyroidism with euthyroid goiter and vestibular dysfunctions, demonstrated in approximately 65% of affected individuals [13].

1.4.2 Usher Syndrome

The prevalence of Usher syndrome ranges from 1/6,000 to 1/10,000. It is an autosomal recessive genetic disease with clinically and genetically heterogeneous characteristics. It is defined by congenital bilateral sensorineural deafness and a later onset of vision impairment, caused by retinitis pigmentosa. This syndrome is sub-classified into three clinical types, USH1, USH2, and USH3, based on the severity of the SNHL, the presence or absence of vestibular dysfunctions and the age at onset of retinitis pigmentosa [15].

1.4.3 Waardenburg syndrome

Waardenburg syndrome (WS) is estimated to have a prevalence of 1/42,000 and is responsible for 1–3% of all congenital HL cases [16].
It is characterized mostly by SNHL and pigmentation abnormalities that can occur in the eyes, hair, skin and the cochlear stria vascularis. Four clinical types of WS have been described (WS1–WS4).

WS type 1 and type 2 are the most frequent phenotypes distinguished by dystopia canthorum, which is present only in WS1.

1.5 Genetics of non-syndromic hearing loss

NSHL is genetically very heterogeneous, with more than associated 160 loci and 106 identified disease genes. It generally follows simple Mendelian inheritance and is more frequently autosomal recessive (75-80%), than autosomal dominant (20%), X-linked (2-5%) and in few cases (1%) NSHL is due to mitochondrial mutations [17]. The forms of NSHL inherited as autosomal dominant, also referred to as DFNA, are usually post-lingual and progressive, while autosomal recessive forms, referred to as DFNB, are more typically pre-lingual, severe to profound [18]. Until now 36 genes have been reported as responsible of autosomal dominant hearing loss, 65 of autosomal recessive and 5 of X-linked hearing deficit (Hereditary Hearing Loss Homepage. http://hereditaryhearingloss.org/; last accessed October 2017). The complete gene-list is reported in Appendix 2.

Despite the heterogeneity, defects at the DFNB1 locus, which contains the *GJB2* gene (Cx26 OMIM #121011), and the structurally related *GJB6* gene (Cx30 OMIM #204418), play a major role worldwide, although with a frequency of mutant alleles that varies between countries and even between regions of the same country [19].

In Caucasian subjects with pre-lingual non-syndromic autosomal recessive hearing loss, it was estimated an average frequency of *GJB2* mutations of 31.5% [20]. The most frequent *GJB2* mutation in Caucasians is c.35delG, with a carrier rate of 1/30-1/35 in the general population of the Mediterranean area [21].

It was demonstrated that individuals homozygote for c.35delG have a more severe hearing loss with respect to those in which c.35delG co-segregates with other *GJB2* mutations: c.[35delG];

[358_360delGAG], c.[35delG]; [-23+1G>A], c.[35delG]; [269T>C], c.[35delG]; [109G>A] and that the specific genotype has a major impact on the degree of hearing impairment [22].

In up to 50% of subjects with HL and heterozygotes for *GJB2* mutation, a deletion of 309 kb del(GJB6-D13S1830) truncating the neighboring *GJB6* gene was identified [23]. A rarer deletion of 232kb del(GJB6-D13S1854), at the DFNB1 locus, was also found *in trans* with pathogenic *GJB2* mutations in affected subjects [24]. It was demonstrated that these deletions remove putative *cis*-regulatory element upstream of *GJB6* and down-regulate the region of location that impact on *GJB2* expression [25].

Currently there is a general consensus on molecular analysis of DFNB1-related hearing loss (due to mutations in *GJB2* and adjacent deletions in *GJB6*) as the first step in genetic testing for non-syndromic hearing loss [26].

1.6 Genetic testing for deafness

HL is defined based on clinical presentation and age of onset; it is necessary to distinguish syndromic from non-syndromic forms, and performing the appropriate genetic test is of paramount importance. A genetic diagnosis in fact provides essential information on the nature of the hearing loss, on the prognosis as well as and on the best rehabilitation options for affected subject; it also offers the possibility of a precise genetic counseling for the parents and other family members.

Conventional technologies for genetic testing, such as Sanger sequencing, are highly accurate and sensitive but are burdened by an extremely low throughput, which has represented a main obstacle to the expansion of testing. Sanger sequencing is useful to test one gene at a time and it is proven effective in cases in which a single gene, or a very limited number of genes, are held responsible of a certain phenotype such as a HL subtype.

In these instances, a screening by Sanger sequencing can be cost-effective. For example, sequencing *GJB2* can identify the underlying etiology in 50% of subjects with congenital or pre-lingual, likely autosomal recessive, non-syndromic hearing loss.

Modern NGS technologies, allowing massive parallel sequencing, enable simultaneous analysis of a large number of specific disease-genes, in case of a targeted multigene panel strategy, or even the entire coding sequences, in case of whole exome sequencing (WES). Analysis of the whole genome (WGS) has not yet entered the clinical arena.

Genetic tests based on Next Generation Sequencing (NGS) technologies are currently rapidly replacing single gene-sequencing in the diagnostic setting.

According to the ACMG practice guideline, the correct diagnostic workflow for Hearing Loss is to search for *GJB2/GJB6* alterations as a first-tier test and then proceed to testing large targeted gene-panels and then go to WES and eventually WGS analysis [26].

1.7 Next-generation sequencing in hearing loss

The crucial advantage of NGS technologies is the ability to address the problem of genetic heterogeneity in conditions that cannot be easily distinguished clinically [27]. This is very much the case of NSHL, in which many different genes are causally implicated and where it is very difficult to correlate pre-diagnostic hypotheses through clinical history and audiological data. A number of papers have indeed reported very encouraging data on targeted gene panels and WES applied to the etiological diagnosis of hearing loss [28].

Targeted NGS tests restrict sequencing to specific sets of genes, known as disease-causing or highly candidate as associated with hearing loss. Such tests, although as mentioned, limited to a certain set of genes can provide an excellent depth of coverage of the selected sequences assuring therefore the best analysis of these sequences.

There are different targeted NGS platform commercially available, that differ mainly on the library preparation method and on the sequencing chemistry.

Library preparation can be performed with amplification method or solution capturing method. The main advantages of the first method, compared to the second one are the high multiplexing capacity and flexibility as it is easy to design and add more specific primers. However, the main disadvantages are the chemistry issues associated with PCR primers. For examples, PCR primers may not amplified the targeted region because of underlying SNPs [29], causing allele-drop-out, PCR primers may fail to amplify GC-rich region and often is difficult to design primers that can discriminate gene from pseudogene when they are highly homologous.

A number of Hearing Loss NGS studies been performed with targeted panels of different quality both in terms of analysis and number of examined genes have reported highly variable diagnostic yields [30] [31].

Obviously Targeted NGS does not address the issue of detection of new genes which is instead a question that can be answered by WES, and finally by WGS.

WES is also based one exon capture but does not depend on a list of genes involved in a specific disease process. Instead, WES seeks to evaluate all exons in the genome for variations. This approach can identify variants in known hearing loss–related genes and genes that have yet to be associated with hearing loss [26].

WGS otherwise is not limited to screening exons and therefore has the potential to identify changes outside of exons that may be related to hearing loss.

Depth of coverage and data interpretation are very important limits of WES and WGS due to the spreading of the sequencing capacity and the large amount of sequencing data to be analyzed. Furthermore, not all the genomic regions are efficiently captured and analyzed by current exon-capture in WES or WGS approach.

Irrespective of the NGS strategy adopted of course, large deletions and duplications, in addition to copy-number and structural variations, are not be efficiently detected yet [32].

A recent study, performed in a large cohort of individuals from The Netherlands, revealed a diagnostic yield of WES targeting only HL related genes, of 33.5%; a comparable result to other studies performed with other massively parallel sequencing technologies in different populations [33].

1.8 Using the Ion Torrent Platform

1.8.1 Sequencing technology and workflow

The Ion Torrent sequencing method, the only one not employing Sanger sequencing, represents an innovation within the innovative NGS technology. It in fact uses a new semiconductor chip, capable of directly translating chemical signals into digital information, without the use of light, which results in unprecedented speed, scalability and lower costs of sequencing.

The Ion Torrent sequencing chemistry itself is remarkably simple. Naturally, a proton (H^+) is released when a nucleotide is incorporated by the polymerase in the synthetizing DNA molecule, resulting in a detectable local change of pH. Each micro-well of the Ion Torrent semiconductor-sequencing chip contains approximately one million copies of a DNA molecule. The Ion Personal Genome Machine (PGM™) sequencer sequentially floods the chip with one nucleotide after another (solutions of native dATP, dCTP, dGTP, dTTP, respectively). If a nucleotide complements the sequence of the DNA molecule in a particular micro-well, it will be incorporated and H^+ are released. The change of pH is detected by the ion sensor and converted to a voltage and to a digital information. If there are two identical bases on the DNA strand, the voltage is double, and the chip records two identical bases. If the next nucleotide that floods the chip is not a match, no voltage change is recorded and no base is called (Figure 1.7) [34]. After each dNTP flow, a wash flow is performed to fully eliminate the

previous nucleotide prior to attempting the next. The Ion Torrent PGMTM (ThermoFisher SCIENTIFIC) workflow consists on the steps illustrated in Figure 1.8.

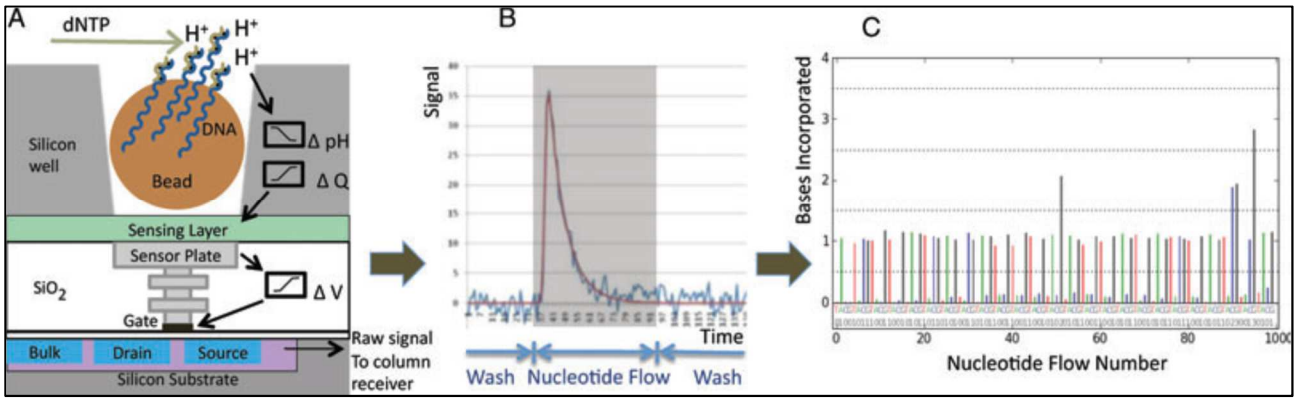


Figure 1.7 Overview of Ion Torrent sequencing process. A) The sequencing chemistry configuration, with templated DNA bead (brown sphere) deposited in a well above the sensor. As a specific dNTP flows into the well, H+ is released if it is incorporated as the next base the resulting H+ release is converted into a voltage change. B) These pH signals are sampled in a plot. (C) The incorporation signals for each flow is converted in a plot called ionogram. Figure modified from Merriman et al. 2012 [34].

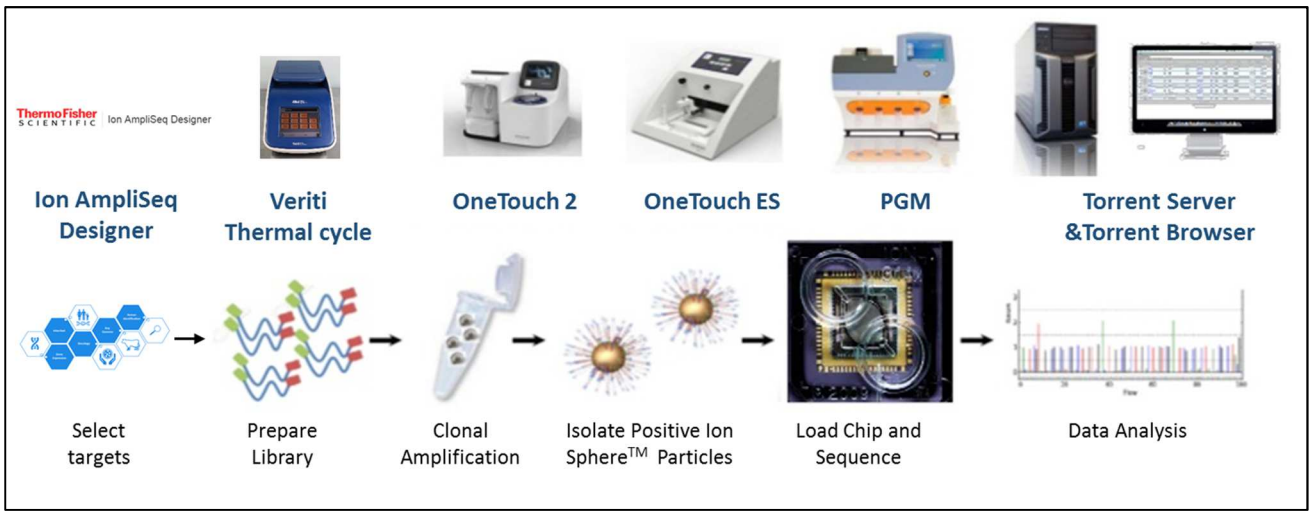


Figure 1.8 Ion torrent process overview.

1.8.2 Library preparation

Genomic DNA or FFPE DNA, or reverse-transcribed RNA are PCR amplified with Ion AmpliSeq™ Library Kit 2.0, using specific primer-pairs designed with Ion Ampliseq Designer or commercially available panels.

The amplicons are partially digested and phosphorylated with FuPa reagents, and then flanked with the Ion Torrent Barcode adapters at 5' and with P1 adapters at 3' end. The libraries are quantified and then combined together according to the manufacturer's instruction. Figure 1.9 represents the entire process of library preparation.

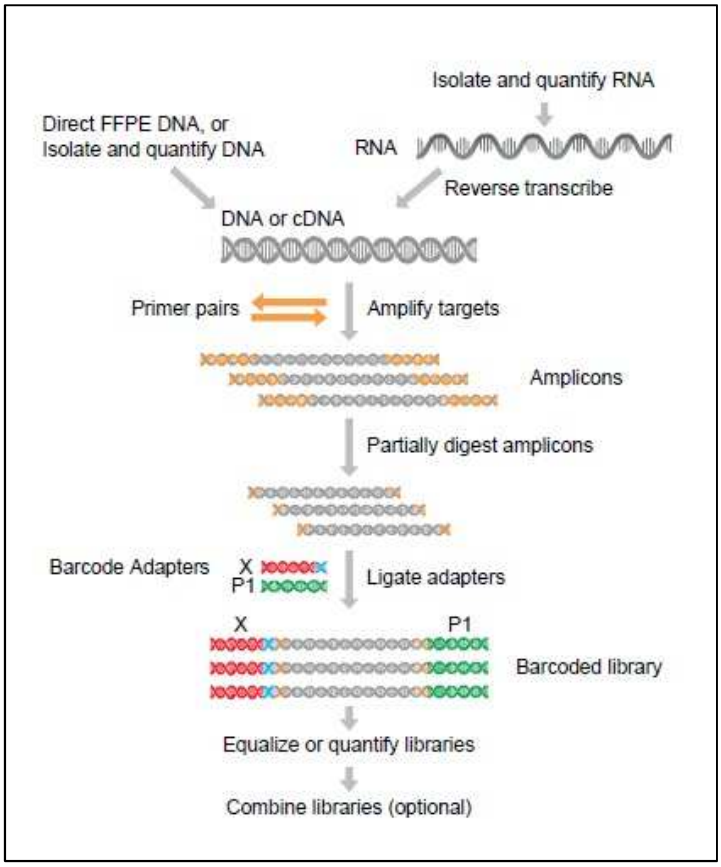


Figure 1.9 Schematic representation of library preparation process. Figure modified from Ion AmpliSeq™ Library Kit 2.0 User Guide.

1.8.3 Template preparation/amplification

The library amplicons are then clonally amplified onto the proprietary Ion Sphere™ particles (ISP) according to the manufacturer's instruction. Clonal amplification is accomplished by emulsion PCR (emPCR) performed in the Ion PGM™ One Touch 2 Instrument and illustrated in Figure 1.10. The ISP-beads coated with complementary primers are mixed with a dilute aqueous solution, containing the amplicons to be sequenced along with the necessary PCR reagents. This solution is then mixed with oil to form an emulsion of microdroplets. The ratio of beads and amplicons is kept low enough so that each microdroplet ideally contains only one of each.

P1 adapters ligated to amplicons are complementary to the sequence coated to the beads, so after a denaturation step, the ssDNA fragment is formed, and then the annealing and extension phase are necessary to synthesize the complementary fragment, which will be further denatured and so on until the PCR reaction ends. At the end, each bead contains ssDNA amplicons.

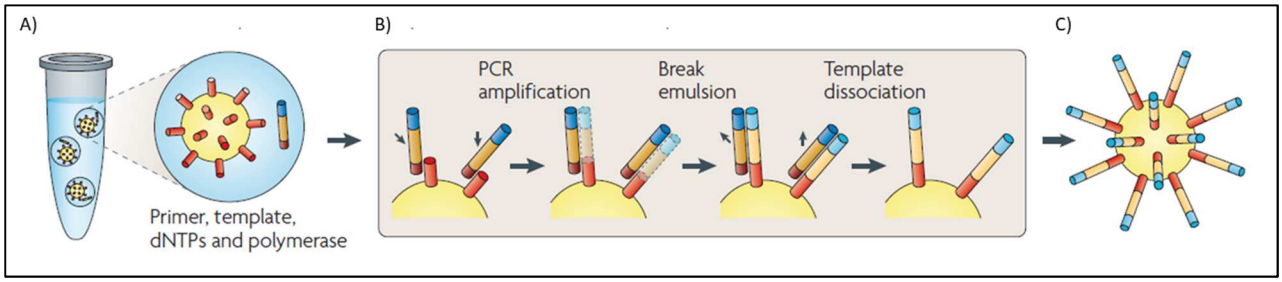


Figure 1.10 Schematic representation of emPCR. A) The tube contains a mixture of beads, primer, template, dNTPs and polymerase. B) The amplicon binds to the bead via the P1 adapter, and then is clonally amplified. The double strand amplicon is then denaturated to form a single strand amplicon. C) Each bead contains single strand amplicon. Figure modified from Metzker 2010 [35].

The Ion Sphere™ Quality Control assay with the Qubit® 2.0 Fluorometer is then performed to evaluate the quality of the templated-ISP and to calculate the percentage of ISPs that correctly bound

to the DNA amplicons. The assay measures the fluorescence of template-positive ISPs labeled with two fluorophores: Alexa Fluor® 488 and Alexa Fluor® 647. The probe labeled with Alexa Fluor® 488 anneals to primer B sites, while the probe labeled with Alexa Fluor® 647 anneals to primer A sites, as shown in Figure 1.11.

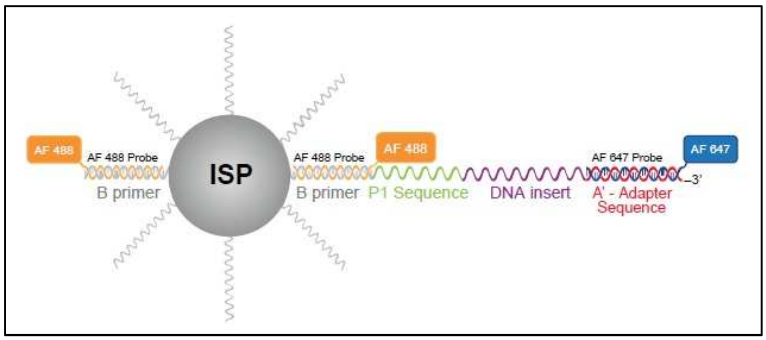
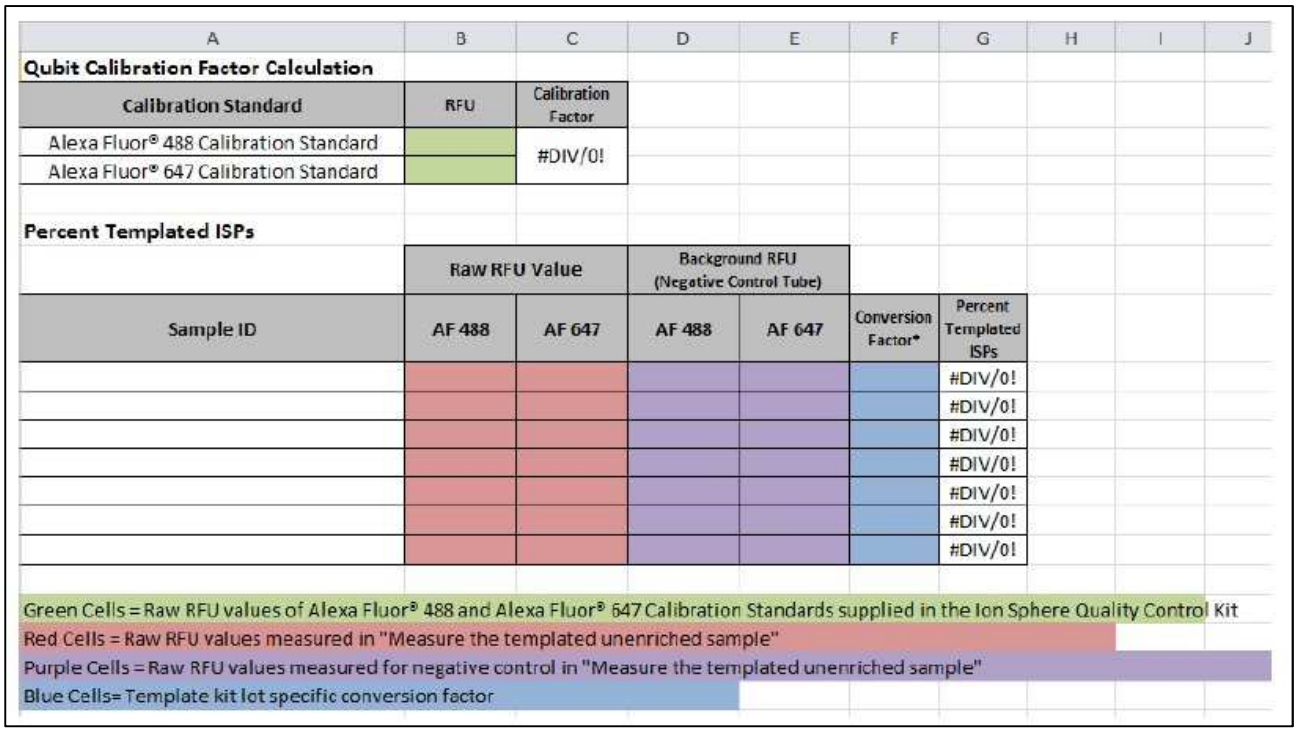


Figure 1.11 Schematic illustration of the Alexa Fluor® 488 and Alexa Fluor® 647, which anneal to the amplicon to primer B sites and to primer A sites, respectively. Figure modified from Ion PGM™ Hi-Q™ OT2 Kit User Guide.

The ratio of the Alexa Fluor® 647 fluorescence (templated ISPs) to the Alexa Fluor® 488 fluorescence (all ISPs present) yields the percentage of templated ISPs and it is calculated with the Qubit® Easy Calculator, reported in Figure 1.12.

The templated-ISPs were then enriched using the Ion OneTouch™ ES instrument, using Dynabeads® MyOne™ Streptavidin C1, that ligate the biotinylated templated-ISP.



A	B	C	D	E	F	G	H	I	J
Qubit Calibration Factor Calculation									
Calibration Standard	**RFU**	**Calibration Factor**							
Alexa Fluor® 488 Calibration Standard		#DIV/0!							
Alexa Fluor® 647 Calibration Standard									
Percent Templated ISPs									
	Raw RFU Value		**Background RFU (Negative Control Tube)**						
Sample ID	**AF 488**	**AF 647**	**AF 488**	**AF 647**	**Conversion Factor***	**Percent Templated ISPs**			
						#DIV/0!			
						#DIV/0!			
						#DIV/0!			
						#DIV/0!			
						#DIV/0!			
						#DIV/0!			
						#DIV/0!			
Green Cells = Raw RFU values of Alexa Fluor® 488 and Alexa Fluor® 647 Calibration Standards supplied in the Ion Sphere Quality Control Kit									
Red Cells = Raw RFU values measured in "Measure the templated unenriched sample"									
Purple Cells = Raw RFU values measured for negative control in "Measure the templated unenriched sample"									
Blue Cells= Template kit lot specific conversion factor									

Figure 1.12 Excel file template used for calculation of the percentage of template ISPs. RFU: relative fluorescence unit. The value of percentage of template ISP appears in column G (#DIV/0!) after entering the RFU values, using a trademark-masked formula. Figure modified Ion PGM™ Hi-Q™ OT2 Kit User Guide.

1.8.4 Sequencing

The enriched templated-ISPs are loaded on the Ion chip, which is placed on the PGM instrument, to perform the sequencing run. The chip capacity varies between different types of chips and amplicons read-length determines the sequence run-time. Each chip produce different amount of output in term of bp, and consequently numbers of reads, as reported in Table 1.1.

		Ion 314™ Chip v2 or Ion 314 Chip v2 BC	Ion 316™ Chip v2 or Ion 316 Chip v2 BC	Ion 318™ Chip v2 or Ion 318 Chip v2 BC
Output*	200 base 400 base†	30–50 Mb 60–100 Mb	300–600 Mb 600 Mb–1Gb	600 Mb–1Gb 1.2–2 Gb
Reads		400–550 thousand	2–3 million	4–5.5 million
Run time	200 base 400 base	2.3 hr 3.7 hr	3.0 hr 4.9 hr	4.4 hr 7.3 hr

Table 1.1 Ion Torrent chip capacity. Mb: Megabases; Gb: Gigabases; hr: hours.

1.8.5 Data analysis

The PGM generates DAT files of electrical signals' raw traces that are the conversion of the raw pH value in a well, to a digital representation of the voltage. This measure is then converted in a graph (ionogram) that represents the base per position image. These raw data are transferred to the Torrent Server for analysis pipeline processing, that is performed by the Ion Torrent Suite Software.

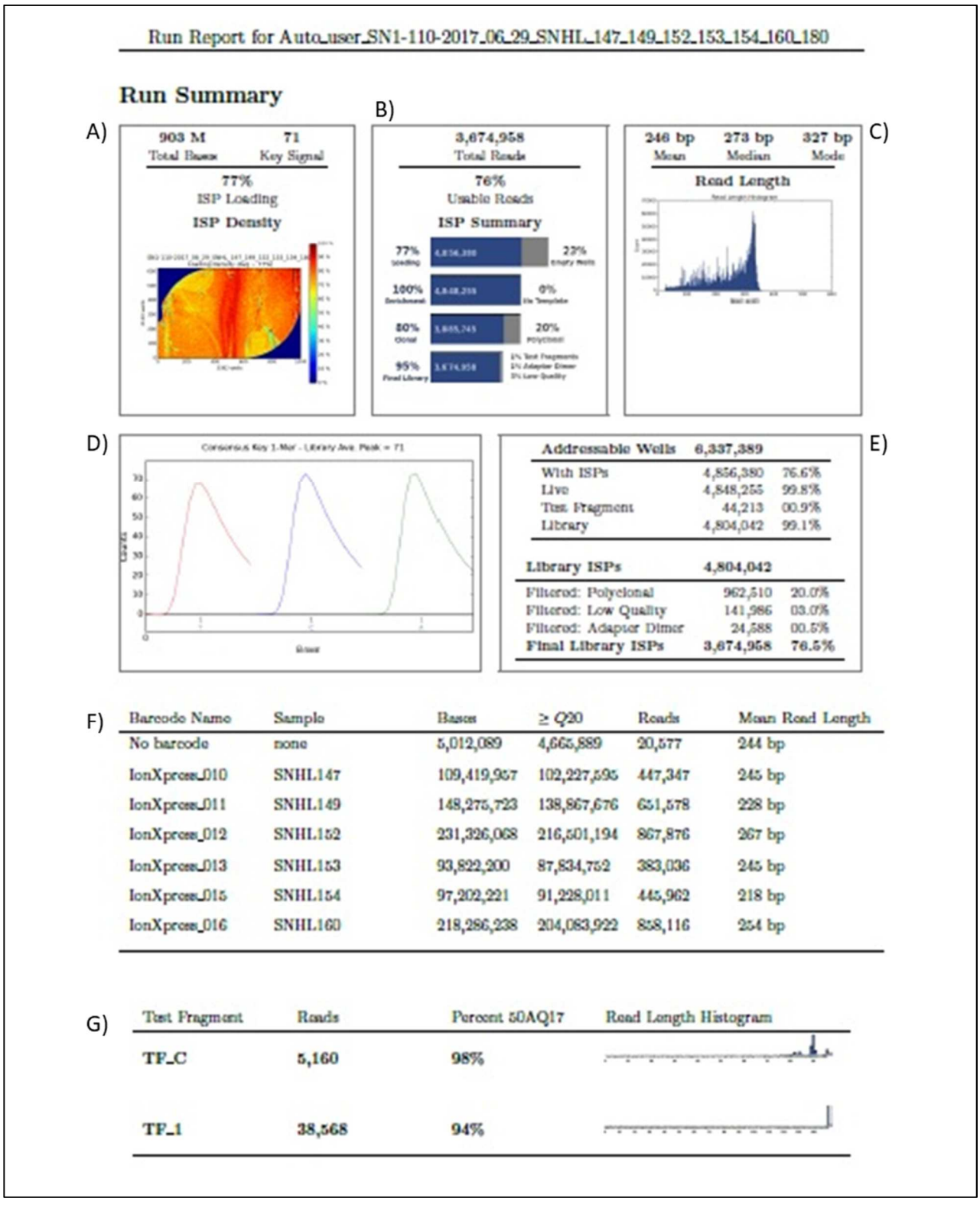
The main steps are the follow: i) the signal processing-step converts the raw traces into a single number per flow per well; ii) The raw Ion signals are then converted to base calls and stored as unaligned BAM file, indexed with a BAI file; iv) The BAM file is passed to TMAP (Torrent Mapping Alignment Program) for alignment.

The Torrent Suite software produce a summary run report that contains the statistics and quality metrics of the run, as illustrated in Figure 1.13. In particular, it contains information on the chip loading, on the total bases and on the key-signal that is a run-quality parameter. The ISP summary table reports the quality of the ISP loaded. It is also reported the metric of the reads length. For each barcoded-subjects the following information are displayed: the number of total bases, the called bases with a predicted quality of Q20, that is reported on the Phred scale, defined as -10×log10 (error probability) and corresponds to a predicted error rate of one percent; the number of total reads and the mean fragment read length.

Test fragments (TFs) provide information about the performance of the experiment; in particular, TF_C evaluates the sequencing run, while the TF_1 evaluates the OT2 process.

Other information contained in the run report are regarding reads' alignments to reference sequence of Homo sapiens (hg19/GRCh37) such the aligned and unaligned bases, the aligned and unaligned reads.

The Raw Accuracy graph plots percent accuracy for each position in an aligned sequence. The information regarding the alignment quality (AQ17, AQ20 and perfect) are represented in table and in a color-plot that illustrate different Phred quality scores (Figure 1.13).



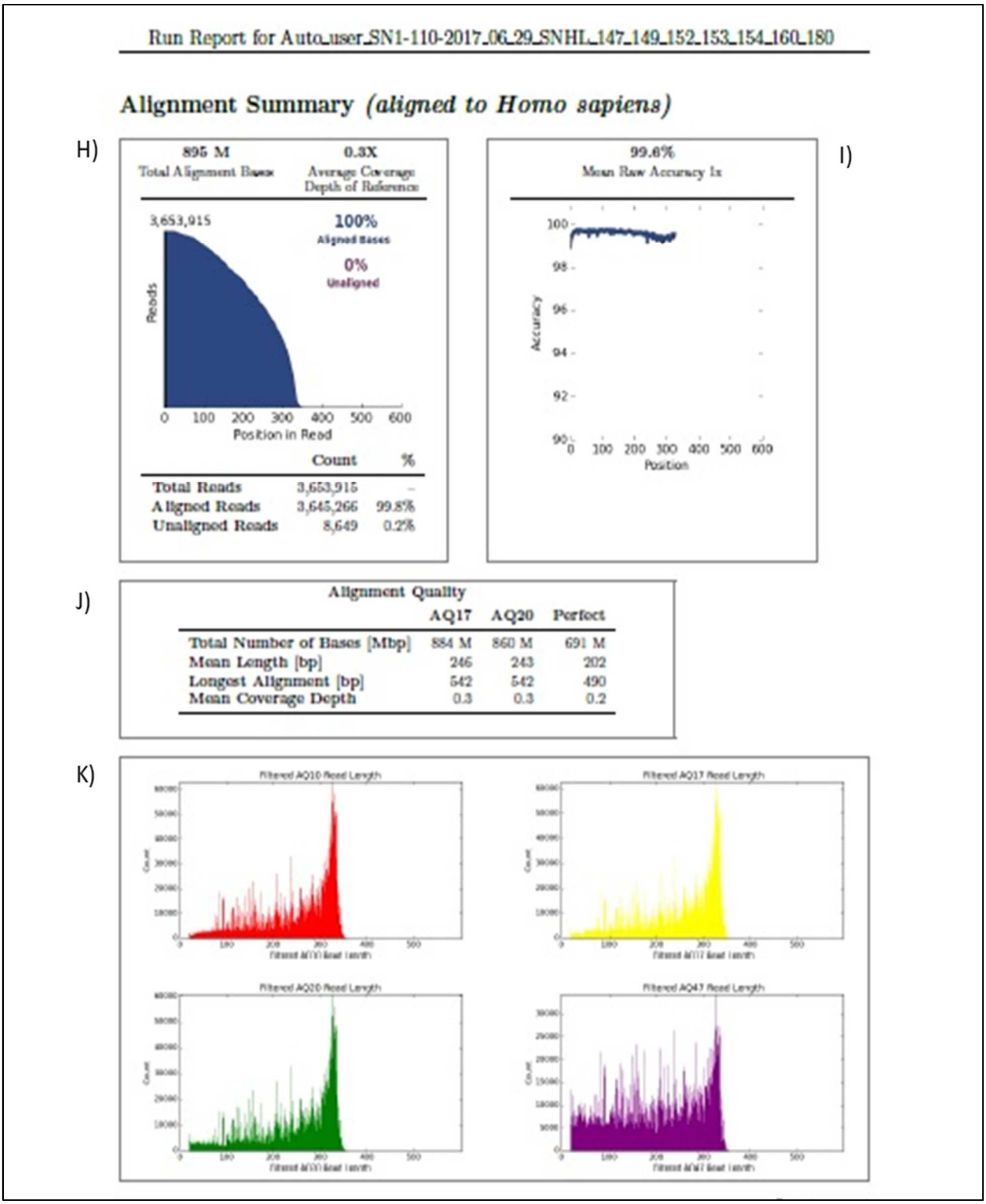
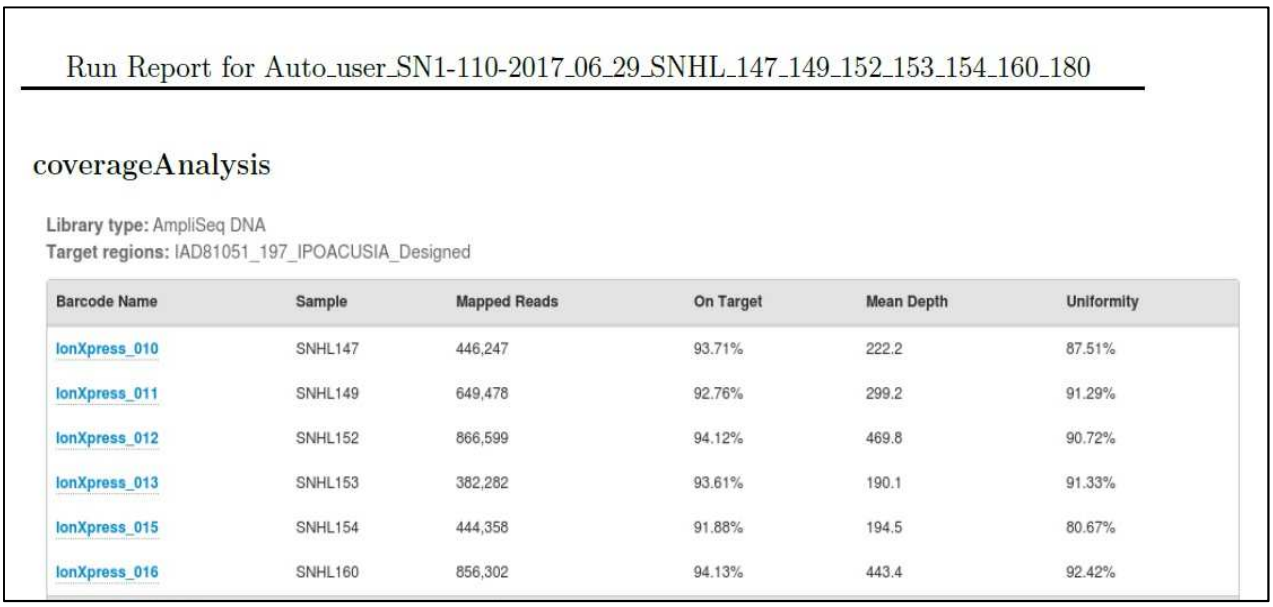


Figure 1.13 Example of summary run report. A) ISP density plot and information about the chip loading. B) ISP summary. C) Read-length with mean, median, mode values. D) Key-signal plot. E) Chip-well details and library ISP details. F) Barcoded-samples information on total reads and reads-length. G) Test-fragments data. H) Alignment summary on hg19. I) Bases alignment accuracy plot. J) Alignment quality table. K) Plot of different Phred Quality score for read length.

The Coverage Analysis Report, produced by the specific plug-in, consists of a collection of summary statistics and graphical representations of the reads coverage for each barcoded sample. In Figure 1.14 is reported the summary of the plug-in results that contains the number of mapped reads, the percentage of the reads mapped on target, the mean depth and the uniformity of coverage expresses as percentage.



Run Report for Auto_user_SN1-110-2017_06_29_SNHL_147_149_152_153_154_160_180

coverageAnalysis

Library type: AmpliSeq DNA
Target regions: IAD81051_197_IPOACUSIA_Designed

Barcode Name	Sample	Mapped Reads	On Target	Mean Depth	Uniformity
IonXpress_010	SNHL147	446,247	93.71%	222.2	87.51%
IonXpress_011	SNHL149	649,478	92.76%	299.2	91.29%
IonXpress_012	SNHL152	866,599	94.12%	469.8	90.72%
IonXpress_013	SNHL153	382,282	93.61%	190.1	91.33%
IonXpress_015	SNHL154	444,358	91.88%	194.5	80.67%
IonXpress_016	SNHL160	856,302	94.13%	443.4	92.42%

Figure 1.14 Example of coverage analysis plug-in summary report.

The Variant Caller plug-in, Figure 1.15, produces the files necessary to further analysis of the variants. In particular the Variant Call Format (.vcf), contains the list of the variants identified for each subject (Single Nucleotide Variants and small insertions/deletions), and is associated with its index file (.vcf.tbi.).

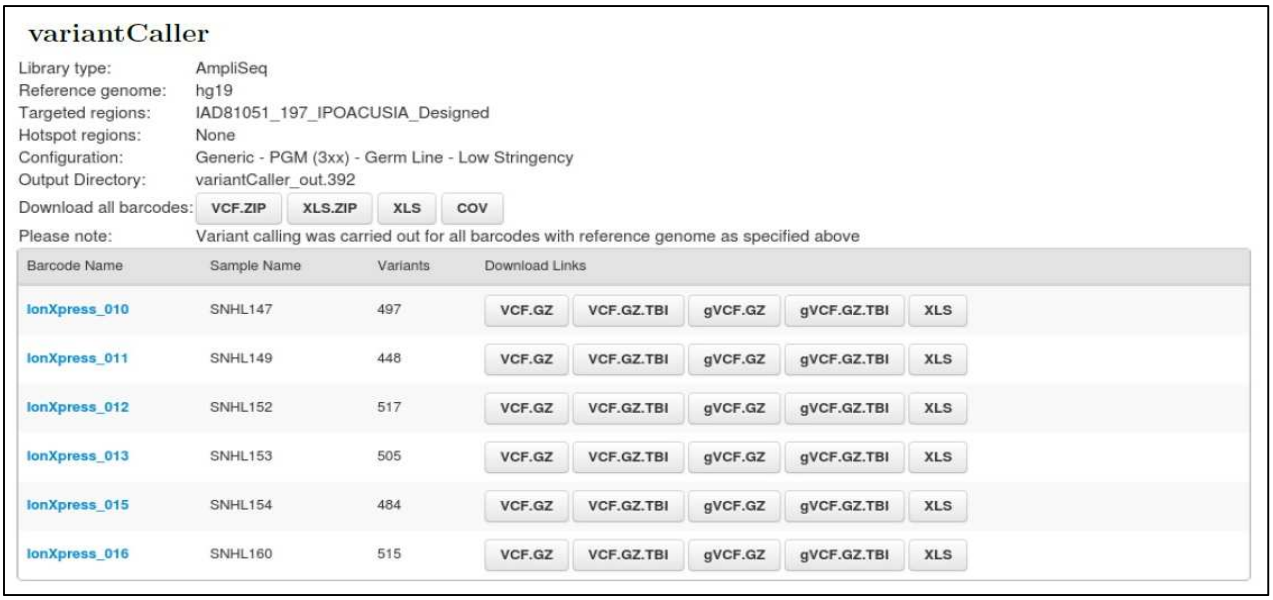


Figure 1.15 Example of Variant caller plug-in summary report.

1.9 Variants annotation

The list of variants obtained from Variant Caller plug-in must be annotated. Different tools are available and the most frequently used is ANNOVAR [36] that is a downloadable command-line tool. It is also available on web a free graphical user interface called wANNOVAR [37].

wANNOVAR utilizes update information to annotate functionally genetic mutations, starting from a list of variants (vcf files) that contains chromosome localization, start position, end position, reference nucleotide and observed nucleotides. The reference genome used to annotate variants is hg19 (GRChg37 genome assembly).

The wANNOVAR output is a table that contains several information on gene-based annotation that identifies whether SNPs or CNVs is located in an exon or in an intron, the amino acids that are mutated and the aminoacids change. The exonic variants are classified as non-synonymous SNV, synonymous SNV, frameshift insertion, frameshift deletion or unknown when different isoforms are available, and it is not possible to assign single annotation.

For each variant identified, wANNOVAR classifies if is reported in specific databases (i.e. dbSNP, 1000 Genome Project, NHLBI-ESP 6500 exomes, Exome Aggregation Consortium) and shows the frequency in different populations. When a dbSNP identifier (rs number) is available, wANNOVAR annotate it. wANNOVAR also reports the scores and the prediction for the aminoacids substitutions using different *in silico* predictor tools [38].

1.10 *In silico* predictor tools

In silico predictor tools are useful to evaluate the pathogenicity of the variants, considering several parameters, such as phylogenetic data, protein structure, and they compute an indicative score of the probability of the variant to be disease-related.

The most frequently used and reported by wAnnovar output are:

- SIFT prediction is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences, collected through PSI-BLAST. The Scores range from 0 to 1; the smaller the score the more likely the SNP has damaging effect [39].
- PolyPhen-2 (Polymorphism Phenotyping v2) predicts the functional significance of an amino acid substitution by Naïve Bayes classifier, using sequence-based and structure-based predictive features. HDIV, or HumDiv, identifies human damaging mutations by assuming differences between human proteins and their closely related mammalian homologus as non-damaging. HVAR, or HumVar, identifies human disease-causing mutations by assuming common human nsSNPs as non-damaging. The score ranges from 0 to 1 [40].
- MutationTaster evaluates the disease-causing potential of DNA sequence alterations by Naïve Bayes classifier, integrating information of evolutionary conservation, splice-site changes,

loss of protein features and changes that might affect the amount of mRNA from different biomedical databases and uses established analysis tools [41].

- Mutation Assessor predicts the functional impact of amino-acid substitutions in proteins based on evolutionary conservation of the affected amino acid in protein homologous. The score ranges from -5.545 to 5.975 in dbNSFP [42].
- LRT (likelihood ratio test) identifies conserved amino acid positions and deleterious mutations using a comparative genomics data set of multiple vertebrate species. The original LRT two-sided p-value (LRTori), ranges from 0 to 1 [43].
- FATHMM (Functional Analysis through Hidden Markov Models) predicts the functional consequences of cancer-associated amino acid substitutions using a model weighted for inherited disease mutations. Scores range from -16.13 to 10.64. The smaller the score the more likely the SNP has damaging effect [44].
- RadialSVM score is an ensemble-based approach integrating multiple scoring systems (function prediction and conservation Score) by radial support vector machine.
- LR score is an ensemble-based approach integrating multiple scoring systems (function prediction and conservation Score) by logistic regression (LR).
- CADD (Combined Annotation Dependent Depletion) score is a framework that integrates multiple annotations into one metric for functional prediction of a SNP. Scores range from -7.535037 to 35.788538 in dbNSFP; the larger the score the more likely the SNP has damaging effect. [45].
- phyloP calculate a conservation score based on the multiple alignments of 7 vertebrate genomes (including human). The larger the score, the more conserved the site. Scores range from -5.172 to 1.062 in dbNSFP [46].
- GERP++ calculates site-specific “rejected substitutions” (RS) scores and to discover evolutionarily constrained elements based on maximum likelihood evolutionary rate

estimation. Scores range from-12.3 to 6.17; the larger the score, the more conserved the site [47].

- SiPhy detects bases under selection from a multiple alignment data using a hidden Markov model. It evaluate the estimated stationary distribution of A, C, G and T at the site, using an algorithm based on 29 mammals genomes [48].

2. AIMS OF THE PROJECT

- Recruitment of a cohort of non-syndromic sensorineural hearing loss subjects negative for *GJB2/GJB6* molecular test.

- Identification of the most comprehensive collection of genes associated with non-syndromic hearing loss, in order to develop a targeted gene panel that could be the translational delivery of the project.

- Identification and molecular characterization of pathogenic variants; characterization of the molecular pathology of selected genes.

- Development of an *in-house* database of variants that could represent an epidemiologic map of the studied population.

- Genotype-phenotype correlation and better stratification of the different phenotypes.

3. MATERIAL AND METHODS

3.1 Patients recruitment and sample collection

A cohort of 78 subjects (54 females and 24 males), mainly of Caucasian origin, with sensorineural hearing loss of variable age of onset and clinical course were recruited in collaboration with many centers of Medical Genetics and departments of Audiology and Otorhinolaryngology, prevalently located in North-East of Italy.

Appropriate genetic counseling was offered to each enrolled subject, provided regular written informed consent for molecular genetic testing was obtained. The work was conducted according to the ethical standards as defined by the Helsinki Declaration and according to indications from the local institutional ethics committee for current molecular diagnosis.

3.2 Clinical information and audiometry evaluation

Family history and clinical information about the hearing loss of the subjects enrolled were collected in an *ad hoc* form, which reported age of onset (congenital, peri-lingual, post-lingual, unknown) and type of hearing loss (neurosensorial, conductive, mixed), as well as laterality (bilateral, symmetric, asymmetric, monolateral), progression (stable, progressive, fluctuating, unknown), severity (mild, moderate, severe, profound) and audiogram shape (flat, U-shaped cookie-bite, down-sloping, rising, unknown). Presence or absence of vestibular dysfunction or auditory neuropathy was also investigated. The previously performed single-gene tests were recorded.

Hearing thresholds for each individual were tested at frequencies from 250 to 8000 Hz (Grason-Stadler GSI 61 audiometer) in a sound-attenuating room. The degree of hearing impairment was defined by the pure tone average (PTA) threshold levels at 0.5, 1, 2 and 4 kHz, and was classified as

mild (PTA 21-40 dB HL), moderate (PTA 41-70 dB HL), severe (PTA 71-95 dB HL) and profound (PTA > 95 dB HL) [10].

3.3 DNA extraction and molecular characterization of *GJB2/GJB6*

Genomic DNA (gDNA) was extracted with Maxwell automatic extractor (PROMEGA), according to the manufacturer's instructions from 400 µl of peripheral blood samples anti-coagulated with EDTA. gDNA was quantified with NanoVue™ Plus Spectrophotometer (GE healthcare).

GJB2 gene (NCBI reference sequence RefSeqNM_004004) was amplified using AmpliTaq Gold® (ThermoFisher SCIENTIFIC) with primers listed in Table 3.1 and PCR-thermal-cycling conditions listed in Table 3.2 and Table 3.3.

2 µl of Exonuclease I and 2 µl of Shrimp Alkaline Phosphatase (IllustraExoProStar GE Healthcare Life Sciences) were added to 10 µl of PCR reaction, to clean-up the PCR products before sequencing; the mixture was incubated at 37°C for 30', followed by incubation at 80°C for 15'.

The sequencing reaction was performed in a total volume of 10 µl, of which 1 µl of BigDye® Terminator v3.1, 1 µl of primer (3.2 µM), PCR products (3-10 ng) and RNase/DNase free water.

The sequencing products were purified with IllustraAutoSeq G-50 columns (GE Healthcare) and then analyzed with ABI PRISM 3130 Genetic Analyzer capillary sequencer (ThermoFisher SCIENTIFIC).

Gene	Name	Sequence	Length (bp)	Tm (C°)	GC%	Ampliconsize	exon
GJB2	cx26_R30	CGTAACTTTCCCAGTCTCCGAGGGAAGAGG	30	70,9	56,7	348 bp	1
GJB2	cx26_L31	GCCCAAGGACGTGTGTTGGTCCAGCCC	27	72,6	66,7		
GJB2	cx26_F1	CATTCGTCTTTTCCAGAGCA	20	55,3	55,3	770 bp	2
GJB2	cx26_R2	CCTCATCCCTCTCATGCTGT	20	59,3	55		

Table 3.1 GJB2 primer sequences for PCR amplification. Tm: melting temperature GC%: percentage of GC bases.

Step	Temperature	Time	Cycle
Initial denaturation	94° C	12′	-
Denaturation	94° C	20′′	3X - 1°C/cycle
Annealing	70° C	20″	
Extension	72° C	30′′	
Denaturation	94° C	30′′	33X
Annealing	68° C	30″	
Extension	72° C	30′′	
Final extension	72° C	7′	-
Cooling	10° C	∞	-

Table 3.2 Thermal cycling conditions for exon 1 of GJB2.

Step	Temperature	Time	Cycle
Initial denaturation	94° C	12′	-
Denaturation	94° C	20′′	3X - 1°C/cycle
Annealing	62° C	20″	
Extension	72° C	30′′	
Denaturation	94° C	30′′	35X
Annealing	60° C	30″	
Extension	72° C	30′′	
Final extension	72° C	10′	-
Cooling	10° C	∞	-

Table 3.3 Thermal cycling conditions for exon 2 of GJB2.

The two *GJB6* deletions (GJB6-D13S1830 and GJB6-D13S1854) were analyzed in a single PCR assay previously reported [24]. PCR products were separated by electrophoresis in a 1.5% agarose gel in TAE 0.5X buffer, stained with GelRed™ 10,000X (Biotium). The PCR products were visually compared with two positive control-sample for the deletions.

3.4 Hearing loss targeted NGS panel

The hearing loss targeted gene-panel was constructed with the aim of developing a molecular tool with a high diagnostic rate for non-syndromic hereditary hearing loss, specifically in the pediatric Caucasian population.

The genes-selection was performed after an accurate review of the literature, after consulting the most comprehensive public databases, such as OMIM (https://www.omim.org/); hereditary hearing loss homepage (http://hereditaryhearingloss.org/); deafness variation database (http://deafnessvariationdatabase.org/) and after the evaluation of the targeted gene-panel already available in the market.

The selection originally included about 100 genes, known to be hearing loss disease-genes or candidate genes; then the filtration process was performed, ending with a panel including 59 genes plus non-coding exon 1 of *GJB2*.

The criteria used for genes inclusion were: i) only genes known to be mutated at least once in Caucasians; ii) genes related to non-syndromic hearing loss, except Pendred and Usher Syndrome genes, because the onset of hearing loss is earlier than other clinical manifestations. This means that these two disorders, at onset, may be clinically indistinguishable from isolated hearing loss, but absolutely need an early diagnosis to optimize clinical follow-up and rehabilitation strategies; iii) genes that were found to be mutated in humans and not only in experimental animal; iv) genes with proven pathogenic role.

3.5 Ion Torrent process overview

3.5.1 Target selection

Ion AmpliSeq Designer v.4.4.1 software (https://www.ampliseq.com) was used to select primers for preparation of DNA libraries. The designer required as input the reference genome, the gene-list, the region of interest (CDS only or CDS and UTR), the padding around exons (10 bp or 25 bp) and the amplicons-size range (125-225 bp or 125-375bp for standard DNA).

GRCh37/hg19 was selected as reference genome to create the amplicons for the 59 genes; it was requested to consider the CDS and UTR with 10 bp of exon-padding. It was also manually added non-coding exon 1 of *GJB2,* by insert genomic coordinate.

The system designed a total of 1,646 amplicons divided in two primer-pools of 827 and 819 amplicons respectively. The amplicon size-range was 125-375 bp, with a mean of 228 bp; the designed gene-panel ensured overall targeted region coverage of 98.6% spanning through 456.8 Kb of sequence, with 3,153 bp missed.

3.5.2 Library preparation

gDNA was quantified with Qubit 2.0 Fluorometer using the Qubit® dsDNA High Sensitivity Assay Kit (ThermoFisher SCIENTIFIC).

The DNA libraries were constructed starting from 10 ng of gDNA and each primer-pool was amplified in two separate reaction with Ion AmpliSeq DNA Library Kit 2.0 (ThermoFisher SCIENTIFIC), using the thermal cycling condition reported in Table 3.4.

Step	Temperature	Time	Cycle
Activate the enzyme	99° C	2′	-
Denaturation and Annealing Extension	99° C 60° C	15′′ 4′	15X
Cooling	10° C	∞	-

Table 3.4 Thermal cycling conditions for library amplification.

After the target-amplification reaction, the tubes containing each amplified pool were combined together and the primer sequences were partially digested with 2 µl of FuPa Reagent using the conditions listed in Table 3.5.

Temperature	Time
50° C	10′
55° C	10′
60° C	20′
10° C	Hold (for up to 1 hour)

Table 3.5 Thermal cycling conditions for FuPa reaction.

The libraries were barcoded with Ion Xpress Barcode Adapters Kit, that were ligated to the amplicons with DNA ligase according to the manufacturer's instructions and with the thermal cycling condition listed in Table 3.6.

Temperature	Time
22° C	30′
68° C	5′
72° C	5′
10° C	Hold (for up to 1 hour)

Table 3.6 Thermal cycling conditions for the ligation of the barcode.

The unamplified libraries were purified with Agencourt AMPure® XP purification system (Beckman Coulter), using a final concentration of 1.5X of beads to sample volume-ratio.150 µl of freshly prepared 70% ethanol was used two times to wash the beads that contained the desired libraries, which were then eluted and amplified with Platinum® PCR SuperMix High Fidelity and Library Amplification Primer Mix, with condition listed in Table 3.7. This step permitted to enrich amplifiable material and obtain sufficient library for accurate quantification.

Temperature	Time	Cycle
98° C	2′	-
98° C 64° C	15′′ 1′	5X
10° C	∞	-

Table 3.7 Thermal cycling conditions for library amplification.

The amplified libraries were purified with Agencourt AMPure® XP reagent using two rounds of purification: first round at 0.5X bead-to-sample-volume ratio to elute the amplicons and primers, and a second round at 1.2X bead-to-sample-volume ratio to bound the amplicons to the beads, that were then eluted with 50 µl of low TE.

10 µl of purified libraries were analyzed with Qubit dsDNA HS Assay Kit and then was determined the dilution factor that results in a concentration of ~100 pM (or ~15-22 ng/mL).

3.5.3 Template preparation

Six libraries at 100 pM were combined in equal volume and then diluted to a final concentration of 16 pM in a total volume of 25 µl.

The diluted libraries were added to 800 µl of Ion PGM™ Hi-Q™ Reagent Mix, in which were also added 25 µl of Nuclease Free Water, 50 µl of Ion PGM™ Hi-Q™ Enzyme Mix and 100 µl of Ion

PGM™ Hi□Q™ ISPs. The solution was loaded in the Ion OneTouch™ Reaction Filter with 1.7 ml of Ion OneTouch™ Reaction Oil. The prepared filter was loaded in the Ion PGM™ One Touch 2 Instrument.

After the run completed the template-positive Ion PGM™ Hi-Q™ ISPs (ThermoFisher SCIENTIFIC) were recovered according to the manufacturer's protocol and then enriched with the Ion OneTouch™ ES instrument (ThermoFisher SCIENTIFIC) using Dynabeads® MyOne™ Streptavidin C1 Beads (ThermoFisher SCIENTIFIC).

2 µl of unenriched template-positive Ion PGM™ Hi-Q™ ISPs were used to perform the quality control by adding 19 µl of Annealing Buffer and 1 µl of Ion Probes. The mix was incubated at 95°C for 2' followed by 37°C for 2'. After three washing steps, the fluorescence of Alexa Fluor® 488 and Alexa Fluor® 647 was measured and then, using the Qubit® Easy Calculator Microsoft® Excel® Spreadsheet file, the percent of templated-ISPs was calculated; according to manufacturer's instruction the optimal values range between 10-30%.

3.5.4 Sequencing

The Ion PGM machine (ThermoFisher SCIENTIFIC) was used to carry out the sequencing process. The instrument was firstly cleaned with 18 MΩ water and then with chlorite solution. After the cleaning procedure, the initialization process was performed according to the manufacturer's instructions. The Ion PGM™ Hi-Q™ Sequencing Kit (ThermoFisher SCIENTIFIC) was used to prepare and sequence the enriched template positive ISP; in particular 5 µl of Control ISPs were added directly to the entire volume of ISPs, followed by a centrifugation step. 12 µl of Sequencing Primer were added to the 15 µl of ISP collected after centrifugation; the mix was incubated at 95°C for 2' and then 37°C for 2'. After the incubation step, 3 µl of Ion PGM™ Hi□Q™ Sequencing Polymerase were added to the ISPs.

The mixture was loaded into the Ion 316™ Chip v2 (ThermoFisher SCIENTIFIC) as recommended by the manufacturer. The prepared chip was then placed into PGM machine for sequencing. 850 flows of sequence were used for the 316 Chip, as suggested by the manufacturer's protocol.

The run was planned in the Torrent Suite Software, giving the information about the setting used in the sequencing, the number of flows, the kit type, the barcode, the run type, the reference file (genome) and the target region (BED file).

3.5.5 Data analysis

Sequencing data were stored in the Torrent Server and then processed by the Ion Torrent Suite™ Softwarev4.4 and later. Single Nucleotide Variants (SNV) and small insertion/deletions (INDELs), generated by the variant caller plug-in, were then annotated using the free web tool wANNOVAR (http://wannovar.wglab.org/) [37].

The reads alignment was visualized using the Integrative Genome Viewer (IGV) software (http://software.broadinstitute.org/software/igv/).

The variants were filtered based on: i) minor allele frequency (MAF) in healthy individuals (<1%) as reported in public database (dbSNP, 1000Genomes, ESP6500, ExAC, gnomAD) to exclude possible polymorphisms. ii) Depth coverage higher than 20X. iii) Localization of the variant (exonic, intronic, 3' UTR or 5'UTR). iv) Type of molecular alteration (frameshift variants, non-synonymous SNV, stop-gain mutation, splice-site mutations).

The pathogenicity of missense mutation was evaluated with a cluster of *in silico* predictor tools: SIFT [39], Polyphen-2 [40], MutationTaster [41], MutationAssessor [42], LRT [43], FATHMM [44], RadialSVM, LR, CADD_phred [45].

The alteration at protein level and the aminoacid conservation was evaluated with PhyloP [46] GERP++ [47], SiPhy [48], Proviz (http://proviz.ucd.ie/) [49]. Information on the structural domain involved and protein structures were obtained consulting UniProt (http://www.uniprot.org/).

The prediction of splice-site mutations was performed with Human Splicing Finder (http://www.umd.be/HSF3/) [50].

The disease-specific database and clinical-associated database as Deafness Variation Database (http://deafnessvariationdatabase.org/), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) and The Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php) were also used to determine the pathogenic role of the variants.

ACMG guidelines [51] and InterVar tool (http://wintervar.wglab.org/) [52] were used to classify the mutations, which were named according to HGVS recommended nomenclature [53].

.

3.6 Sanger sequencing and family segregation analysis

For diagnostic purposes, we decided to adopt a cut-off depth of coverage of 100 X.

Likely pathogenic variants, either new or known in literature, with a depth of coverage of less than 100 X, were systematically validated by conventional Sanger sequencing, using primers designed in-house with Primer3Plus (https://primer3plus.com/) or designed by the AmpliSeq designer. The PCR reactions were performed using AmpliTaq Gold® (ThermoFisher SCIENTIFIC) with standard PCR-thermal-cycling conditions, as reported by the manufacturer's instructions. PCR products were purified using ExoSAP-IT^{TM} (ThermoFisher SCIENTIFIC Waltham, MA USA) and sequencing reactions were performed with BigDye® Terminator v3.1 kit according to manufacturer's instruction. The sequencing products were analyzed with 3130 Genetic Analyzer capillary sequencer (ThermoFisher SCIENTIFIC Waltham, MA USA).

Family segregation analysis was performed, whenever possible, to evaluate the *de novo* or inherited nature of the variant.

3.7 RNA extraction and cDNA analysis

PAXgene Blood RNA Tube was used, according to manufacturer's instruction, to collect blood for subsequent RNA extraction, which was performed with the PAX gene Blood RNA Kit, as reported by the manufacturers's protocol. Total RNA was quantified with NanoVue™ Plus Spectrophotometer (GE healthcare).

100 ng of total RNA was reverse-transcribed using *SuperScript II Reverse Transcriptase* (ThermoFisher SCIENTIFIC) and random hexaprimers (Invitrogen), according to the manufacturer's instructions. The RNA plus random hexaprimers was rapidly denatured at 65°C for 5' and then quick chilled on ice. 5X First-Strand Buffer, 0.1 M DTT and RNAase OUT (40U/µl) were added to the mixture which was incubated at 25° for 2', before adding SuperScript II RT (200 units).

The RT-PCR protocol was carried out as follows: 25°C for 10', 42°C for 50' and 70°C for 15'.

2µl of the obtained cDNA were used to amplify exons 59, 60, 61 of *CDH23* gene, to evaluate the molecular effect of the novel splice-site mutation identified by NGS.

The PCR was performed with AmpliTaq Gold® (ThermoFisher SCIENTIFIC) with two different primers–pair, listed in Table 3.8, that amplify simultaneously the three exons. The PCR-thermal-cycling conditions are listed in Table 3.9.

Gene	Name	Sequence	Length (bp)	Tm (C°)	GC%	Amplicon size	exon
CDH23	CDH23_ex59_a_F	CGGCAACGAAGAGAAGAACT	20	59,6	50	455bp	59,60,61
CHD23	CDH23_ex61_a_F	CCATGAAGAGGTCGAAGGTG	20	60,6	55		
CDH23	CDH23_ex59_b_F	TTGTGCTAGAGGACATCAACG	21	58,9	47,6	445bp	59,60,61
CDH23	CDH23_ex61_b_F	ATGTTGGAGAGCAGGTGGAT	20	59,5	50		

Table 3.8 CDH23 primer sequences for PCR amplification. Tm: melting temperature GC%: percentage of GC bases.

Step	Temperature	Time	Cycle
Initial denaturation	94° C	2’	-
Denaturation	95° C	45’’	
Annealing	58° C	45”	40X
Extension	72° C	45’’	
Final extension	72° C	7’	-
Cooling	10° C	∞	-

Table 3.9 Thermal cycling conditions for amplification of exons 59, 60 and 61 of CDH23 gene.

4. RESULTS

4.1 Cohort of recruited subjects

A total of 130 blood samples from subjects with apparent non-syndromic hearing loss clinically ascertained and clinically characterized by many Medical Genetics Centers and departments of Audiology and Otorhinolaryngology mainly located in the North-East of Italy were collected.

All the selected individuals had previously been tested for *GJB2* mutations/*GJB6* deletions, and resulted negative.

A subset of 78 selected individuals have so far been tested with our customized NGS targeted panel approach. The analyzed cohort includes 54 females and 24 males, with ages ranging from 1 to 68-years: average age 20-years (SD ± 18.5); mode 7-years and median age 13-years.

In particular the cohort was stratified as follow: 14 children with ages from 0-5 years old; 18 subjects with ages from 6-10 years old; 16 individuals with ages from 11-15 years old; 4 subjects with ages from 16-20 years old; 6 individuals with ages from 21-30 years old; 5 subjects with ages from 31-40 years old; 8 subjects with ages from 41-50 years old; 8 subjects with ages from 51-60 years old and 4 individuals with more than 60 years.

In the selected cohort, the age of onset of the hearing loss was as follows: 38 cases with documented congenital hearing loss, 7 with peri-lingual HL, 23 with post-lingual and 10 with unspecified onset of the hearing deficit. Among the 78 subjects tested, 4 individuals with bilateral and asymmetric HL were included, all the other had bilateral symmetric HL.

The course of the HL was prevalently stable (42/78 subjects), while it was progressive in 22/78; in only 1 individual HL was fluctuating and the course of the defect was unknown in 13 individuals.

In 28 of the tested subjects the degree of HL was profound, it was severe in 23, moderate in 20 and mild in only 4 individuals.

4.2 Targeted gene panel

4.2.1 Results of gene-selection

The targeted gene-panel, developed to analyze mainly genes responsible for non-syndromic deafness or hearing loss with non-syndromic onset in Caucasian, was obtained after an accurate gene-selection process, explained in detail in section 3.4. The final targeted gene-panel comprised 59 genes, listed in Table 4.1. The targeted gene-panel included 26 autosomal recessive (AR) genes, 17 autosomal dominant (AD) genes, 4 X-linked genes, 11 Usher syndrome genes (inherited as autosomal recessive) and the auditory neuropathy gene (inherited as autosomal dominant).

Autosomal recessive genes (n=26)	*CLDN14; COL11A2; DFNB59; ESPN; GIPC3; GJB2; GJB3; GJB6; GRXCR1; HGF; MYO15A; MYO6; OTOA; OTOF; OTOG; OTOGL; PTPRQ; RDX; SLC26A4; SLC26A5; STRC; TBC1D24; TECTA; TMC1; TMPRSS3; TPRN*
Usher syndrome genes (n=11)	*CDH23; CIB2; CLRN1; GPR98; MYO7A; PCDH15; PDZD7; USH1C; USH1G; USH2A; WHRN*
Autosomal dominant genes (n=17)	*ACTG1; CCDC50; CEACAM16; COCH; DFNA5; EYA4; GRHL2; HOMER2; KCNQ4; LEDGF; MIR96; MYH14; MYH9; OSBPL2; P2RX2; POU4F3; SLC17A8*
Auditory neuropathy gene (n=1)	*DIAPH3*
X-linkedgenes (n=4)	*COL4A6; POU3F4; PRPS1; SMPX*

Table 4.1 Genes included in the targeted-gene panel.

4.1.1 Ampliseq designer results

The Ampliseq designer v.4.4.1 designed for the selected-genes, a total of 1,646 amplicons divided in two primers pools. The total covered bases were 226.8 Kb that ensured overall targeted region

coverage of 98.6%. The missed bases (3,153 bp) were located in particular in exon 1 or in GC-rich regions.

In particular there were 36 genes covered at 100%, 13 genes had a coverage ranging between 97 and 99%, 8 genes were covered ranging between 90 and 95% and 2 genes (*ESPN, P2RX2*) were covered less than 85%, as reported in Table 4.2.

gene name	# exons	# amplicons	total bp	covered bp	missed bp	overall coverage %
ACTG1	5	7	1,228	1,169	59	95.2
CCDC50	12	16	1,689	1,689	0	100
CDH23	73	91	11,849	11,702	147	98.8
CEACAM16	6	10	1,398	1,398	0	100
CIB2	7	7	684	650	34	95
CLDN14	1	4	740	740	0	100
CLRN1	9	8	1,051	1,051	0	100
COCH	11	14	1,873	1,873	0	100
COL11A2	67	63	6,606	6,572	34	99.5
COL4A6	47	49	6,078	6,078	0	100
DFNA5	10	11	1,671	1,671	0	100
DFNB59	6	8	1,179	1,179	0	100
DIAPH3	31	34	4,290	4,290	0	100
ESPN	13	18	2,825	2,339	486	82.8
EYA4	20	20	2,421	2,421	0	100
GIPC3	6	7	1,059	1,055	4	99.6
GJB2	1	4	701	701	0	100
GJB3	1	4	833	833	0	100
GJB6	1	3	806	806	0	100
GPR98	90	128	20,721	20,707	14	99.9
GRHL2	16	19	2,198	2,198	0	100
GRXCR1	4	6	953	953	0	100
HGF	21	20	2,583	2,562	21	99.2
HOMER2	10	10	1,245	1,220	25	98
KCNQ4	14	18	2,368	2,326	42	98.2
LEDGF	16	17	1938	1920	18	99,1
MIR96	1	2	97	97	0	100
MYH14	42	50	6,951	6,951	0	100
MYH9	40	50	6,683	6,683	0	100
MYO15A	64	84	11,873	11,146	727	93.9
MYO6	34	39	4,538	4,538	0	100
MYO7A	51	65	7,642	7,642	0	100
OSBPL2	16	13	1,773	1,773	0	100
OTOA	30	31	4,058	4,056	2	99
OTOF	50	60	7,259	6,964	295	95.9

OTOG	55	73	9,878	9,624	254	97.4
OTOGL	58	65	8,195	8,195	0	100
P2RX2	19	12	1,718	1,448	270	84.3
PCDH15	43	52	8,284	8,284	0	100
PDZD7	17	22	3,474	3,422	52	98.5
POU3F4	1	6	1,106	1,106	0	100
POU4F3	2	6	1,057	1,057	0	100
PRPS1	8	8	1,097	1,097	0	100
PTPRQ	42	54	7,224	7,224	0	100
RDX	19	18	2,099	2,099	0	100
SLC17A8	12	14	2,010	2,010	0	100
SLC26A4	20	21	2,743	2,743	0	100
SLC26A5	20	19	2,652	2,652	0	100
SMPX	3	3	327	327	0	100
STRC	29	37	5,908	5,897	11	99.8
TBC1D24	7	14	1,820	1,820	0	100
TECTA	23	35	6,928	6,928	0	100
TMC1	20	21	2,683	2,654	29	98.9
TMPRSS3	15	12	1,688	1,688	0	100
TPRN	4	12	2,216	2,088	128	94.2
USH1C	29	33	3,334	3,100	234	93
USH1G	4	8	1,446	1,302	144	90
USH2A	72	92	17,043	17,043	0	100
WHRN	14	18	2,964	2,841	123	95.9

Table 4.2 Results obtained from Ampliseq Designer v.4.4.1. Number (#) of exons, number (#) of amplicon; total base pair (bp), covered bp, missed bp and overall coverage expressed as percentage for the selected genes.

4.2 Coverage analysis plug-in results

The summary of the average values output of Ion Torrent Coverage analysis plug-in is reported in Table 4.3. In particular, for each individual an average of 522,405.12 reads were mapped, of which 94.37% were on-target. For each amplicon, an average of 295.55 reads were aligned. The average base coverage obtained was 249.03X and 91.11% of target bases were covered at least 20X, while 68.45% bases were covered 100X.

The mean amplicon length was 236.46 bp, as expected from the AmpliSeq design, which calculate a mean amplicon length of 228 bp.

	number of mappedreads	% reads on target	Averagereads per amplicon	% amplicons with at least 20 reads	% amplicons with at least 100 reads	%amplicons with no strand bias	% amplicons reading end-to-end	average base coveragedepth	% uniformity of base coverage	% target base coverage at 20x	% target base coverage at 100X	% end-to-end reads	Meanampliconlenght
mean allsamples	522405.12	94.37	295.55	93.62	76.07	94.55	46.27	249.03	85.36	91.11	68.45	70.03	236.46
SD-all	197253.83	1.99	114.19	8.63	16.27	2.39	15.86	105.92	11.05	10.84	17.45	8.94	15.98

Table 4.3 Mean value and standard deviation (SD) of the data obtained from Ion Torrent coverage analysis plug-in.

4.3 Variant caller plug-in results

The variant Caller plug-in produced an average of 499 (SD ± 47) variants (SNVs and INDELs) per individual, of which 131 (SD ± 47) were exonic and 368 (SD ± 41) were intronic.

According to the criteria explained in detail in section 3.5.6, the final number of candidate variants was reduced to a mean of 9 variants per subjects (SD ± 4.3).

Some of the filtered variants resulted to be common alignment errors and were further excluded.

A total of 1,311 unique variants, present in only one individual of the analyzed cohort, were identified; among which, 871 were covered higher than 20X. 448 of the filtered variants had an rs accession number reported in dbSNP and only 8 were reported in ClinVar database as pathogenic or likely pathogenic.

Over the 871 variants there were 297 exonic (162 non-synonymous SNV, 92 synonymous SNV, 13 frameshift deletions, 10 frameshift insertions, 7 non-frameshift deletions, 2 non-frameshift insertions, 3 non-frameshift substitutions, 4 stop-gain, 4 unknown) and 574 intronic (5 splice-site mutation, 5 localized at 5'UTR and 18 localized at 3'UTR). The "unknown" called variants were further manually annotated considering the main isoform of the gene.

After an accurate evaluation and filtration process, 43 variants were reported as pathogenic, from which 18 were novel (41.8%), never described in literature, nor present in the public databases of healthy population or disease-specific. All the likely pathogenic variants identified are reported in Table 4.4, while the list of novel mutations with in silico prediction of pathogenicity is reported in Table 4.5.

The most frequently mutated gene in the analysed cohort was *CDH23*, with 13 likely causative variants, of which 4 were never previously described, as reported in Figure 4.1. The three *GJB2* mutations identified in two samples were previously missed with the Sanger sequencing analysis.

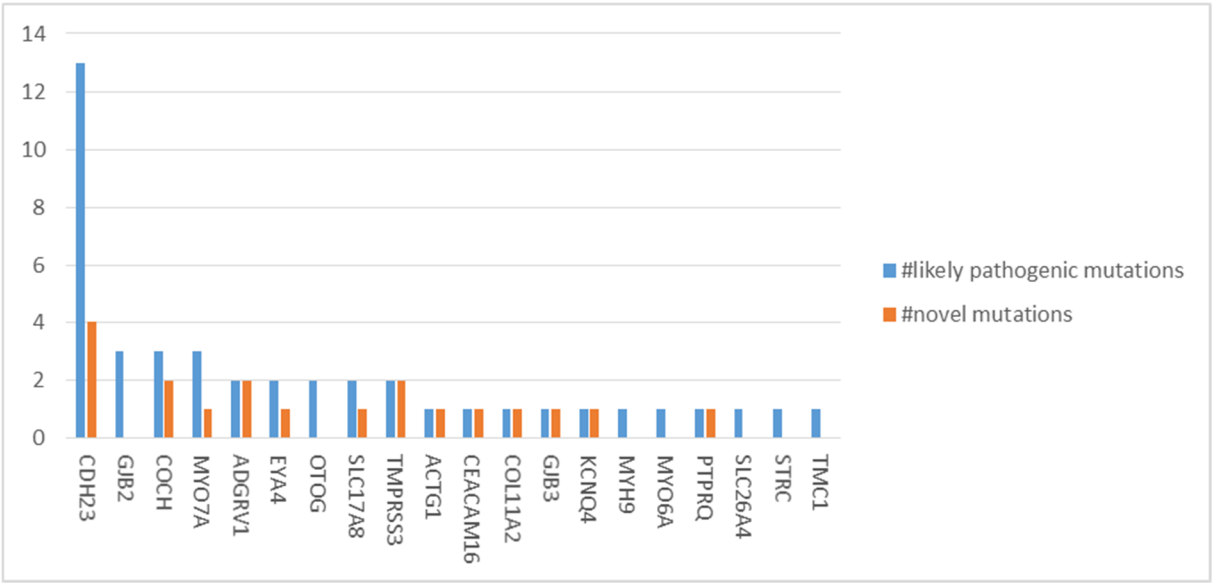
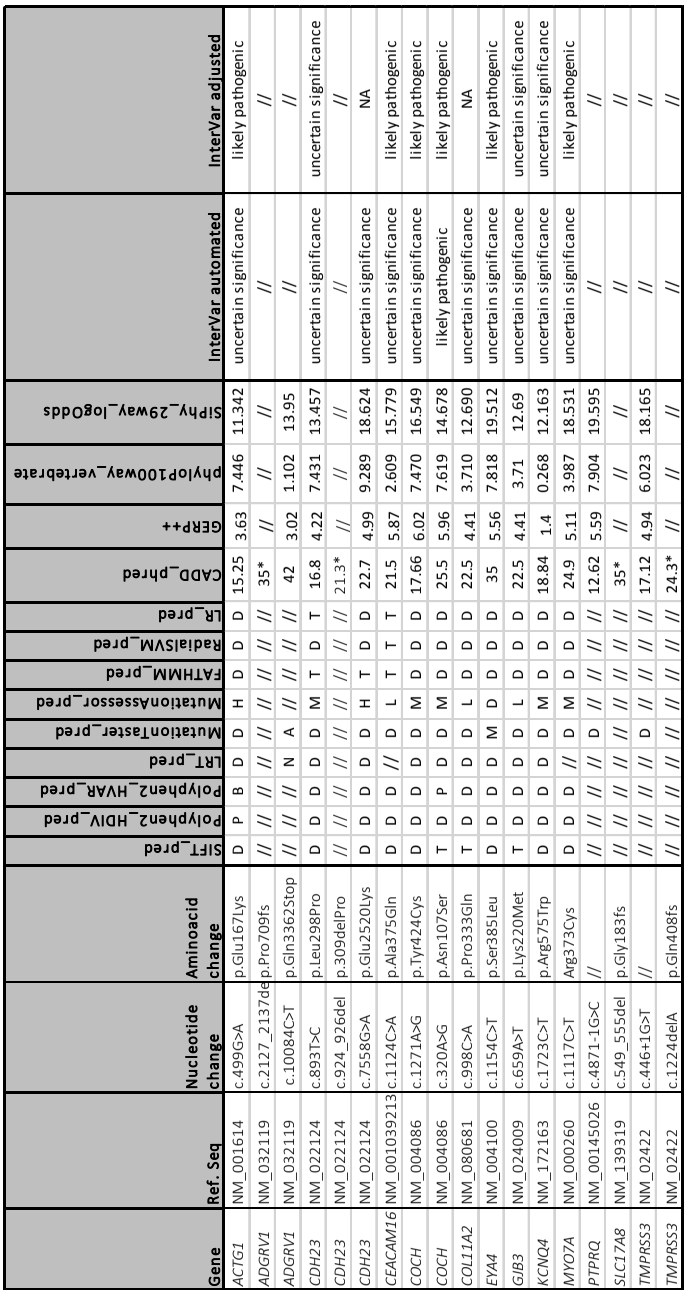


Figure 4.1 Histogram of the frequently mutated genes. X-axis reports the genes with likely causative variants. Y-axis represents the number of mutations. Blue bar indicates the total identified mutations, orange bars represents the novel one.

					Allele #1			Allele #2			
ID #	Sex	Age	Gene	Ref. Seq	Nucleotide	Aminoacid	rs dbSNP	Nucleotide	Aminoacid	rs dbSNP	Inheritance
91	m	4	*ACTG1*	NM_001614	c.499G>A	p.Glu167Lys	//	//	//	//	AD
125	f	4	*ADGRV1*	NM_032119	c.2127_2137del	p.Pro709fs	//	c.10084C>T	p.Gln3362Stop	//	AR *
10	m	11	*CDH23*	NM_022124	c.893T>C	p.Leu298Pro	//	c.924_926del	p.309delPro	//	AR
118	f	1	*CDH23*	NM_022124	c.2263C>T	p.His755Tyr	rs181255269	c.2263C>T	p.His755Tyr	rs181255269	AR *
151	f	13	*CDH23*	NM_022124	c.3293A>G	p.Asn1098Ser	rs41281310	c.1096G>A	p.Ala366Thr	rs143282422	AR *
160	f	53	*CDH23*	NM_022124	c.7558G>A	p.Glu2520Lys	//	c.7558G>A	p.Glu2520Lys	//	AR *
8	f	16	*CDH23*	NM_022124	c.6530C>A	p.Pro2177His	rs748946014	c.8966-1G>C	//	//	AR
34	m	3	*CDH23*	NM_022124	c.2304+1G>A	//	rs41281316	c.6050-9G>A	//	rs367928692	AR Usher syndrome
77	f	8	*CDH23*	NM_022124	c.7558G>A	p.Glu2520Lys	//	c.1515-12G>A	//	rs3693967013	AR
81	f	18	*CDH23*	NM_022124	c.4625G>A	p.Gly1542Asp	rs781339262	c.7361C>T	p.Thr2454Met	rs772949926	AR
141	m	30	*CEACAM16*	NM_001039213	c.1124C>A	p.Ala375Gln	//	//	//	//	AD
9	f	13	*COCH*	NM_004086	c.1348A>G	p.Ile450Val	rs139503327	//	//		AD
23	f	41	*COCH*	NM_004086	c.1271A>G	p.Tyr424Cys	//	//	//	//	AD
122	m	15	*COCH*	NM_004086	c.320A>G	p.Asn107Ser	//	//	//	//	AD
147	f	47	*COL11A2*	NM_080681	c.998C>A	p.Pro333Gln	//	//	//	//	AD
14	f	9	*EYA4*	NM_004100	c.1154C>T	p.Ser385Leu	//	//	//	//	AD
38	f	14	*EYA4*	NM_004100	c.925A>G	p.Thr309Ala	rs556335059	//	//	//	AD
76	f	7	*GJB2*	NM_004004	c.71G>A	p.Trp24Stop	rs104894306	c.71G>A	p.Trp24Stop	rs104894306	AR
56	m	9	*GJB2*	NM_004004	c.101T>C	p.Met34Thr	rs35887622	c.35delG	p.Gly12fs	rs80338939	AR
11	f	30	*GJB3*	NM_024009	c.659A>T	p.Lys220Met	//	//	//	//	AD
138	m	10	*KCNQ4*	NM_172163	c.1723C>T	p.Arg575Trp	//	//	//	//	AD
153	f	7	*MYH9*	NM_002473	c.3485+6C>T	//	rs867754177	//	//	//	AD
114	m	6	*MYO6*	NM_004999	c.3554C>G	p.Pro2285Arg	rs762361071	//	//	//	AD
99	f	1	*MYO7A*	NM_000260	c.4921G>A	p.Glu1641Lys	rs767975012	//	//	//	AD
5	f	12	*MYO7A*	NM_000260	c.730C>T	p.Arg244Cys	rs373942326	c.1117C>T	Arg373Cys	//	AR
68	f	15	*OTOG*	NM_001292063	c.6713G>A	p.Gly2238asp	rs528937385	c.7693+1G>A	//	rs548496846	AR
51	f	47	*PTPRQ*	NM_00145026	c.4871-1G>C	//	//	c.4871-1G>C	//	//	AR
132	m	33	*SLC17A8*	NM_139319	c.861A>G	p.Ile287Met	rs770205149	//	//	//	AD
86	f	1	*SLC17A8*	NM_139319	c.549_555del	p.Gly183fs	//	//	//	//	AD
15	f	32	*SLC26A4*	NM_000441	c.1001+1G>A	//	rs80338849	//	//	//	AR
109	f	65	*SLC26A4*	NM_000441	c.1001+1G>A	//	rs80338849	//	//	//	AR
120	m	10	*STRC*	NM_153700	c.4917_4918delACinsCT	//	rs727503441	c.4917_4918delACinsCT	//	rs727503441	AR *
3	m	26	*TMC1*	NM_138691	c.2230C>T	p.Arg744Stop	rs150738413	//	//	rs150738413	AD
83	f	7	*TMPRSS3*	NM_02422	c.1224delA	p.Gln408fs	//	c.446+1G>T	//	//	AR *

*Table 4.4 Mutations identified in positive cases. ID # indicates the identification number for each subject. M: male; F: female. The mutations are nominated according to HGVS nomenclature. // indicates that no rs number or no mutations are present. AD: autosomal dominant; AR: autosomal recessive. * Family segregation not yet performed. The mutations written in blue are novel.*



Gene	Ref. Seq	Nucleotide change	Aminoacid change	SIFT_pred	Polyphen2_HDIV_pred	Polyphen2_HVAR_pred	LRT_pred	MutationTaster_pred	MutationAssessor_pred	FATHMM_pred	RadialSVM_pred	LR_pred	CADD_phred	GERP++	phyloP100way_vertebrate	SiPhy_29way_logOdds	InterVar automated	InterVar adjusted
ACTG1	NM_001614	c.499G>A	p.Glu167Lys	D	P	B	D	D	H	D	D	D	15.25	3.63	7.446	11.342	uncertain significance	likely pathogenic
ADGRV1	NM_032119	c.2127_2137del	p.Pro709fs	//	//	//	//	//	//	//	//	//	35*	//	//	//	//	//
ADGRV1	NM_032119	c.10084C>T	p.Gln3362Stop	//	//	//	N	A	//	//	//	//	42	3.02	1.102	13.95	//	//
CDH23	NM_022124	c.893T>C	p.Leu298Pro	D	D	D	D	D	M	T	D	T	16.8	4.22	7.431	13.457	uncertain significance	uncertain significance
CDH23	NM_022124	c.924_926del	p.309delPro	//	//	//	//	//	//	//	//	//	21.3*	//	//	//	//	//
CDH23	NM_022124	c.7558G>A	p.Glu2520Lys	D	D	D	D	D	H	T	D	D	22.7	4.99	9.289	18.624	uncertain significance	NA
CEACAM16	NM_001039213	c.1124C>A	p.Ala375Gln	D	D	D	//	D	L	T	T	T	21.5	5.87	2.609	15.779	uncertain significance	likely pathogenic
COCH	NM_004086	c.1271A>G	p.Tyr424Cys	D	D	D	D	D	M	D	D	D	17.66	6.02	7.470	16.549	uncertain significance	likely pathogenic
COCH	NM_004086	c.320A>G	p.Asn107Ser	T	D	P	D	D	M	D	D	D	25.5	5.96	7.619	14.678	likely pathogenic	likely pathogenic
COL11A2	NM_080681	c.998C>A	p.Pro333Gln	T	D	D	D	D	L	D	D	D	22.5	4.41	3.710	12.690	uncertain significance	NA
EYA4	NM_004100	c.1154C>T	p.Ser385Leu	D	D	D	D	M	D	D	D	D	35	5.56	7.818	19.512	uncertain significance	likely pathogenic
GJB3	NM_024009	c.659A>T	p.Lys220Met	T	D	D	D	D	L	D	D	D	22.5	4.41	3.71	12.69	uncertain significance	uncertain significance
KCNQ4	NM_172163	c.1723C>T	p.Arg575Trp	D	D	D	D	D	M	D	D	D	18.84	1.4	0.268	12.163	uncertain significance	uncertain significance
MYO7A	NM_000260	c.1117C>T	Arg373Cys	D	D	D	//	D	M	D	D	D	24.9	5.11	3.987	18.531	uncertain significance	likely pathogenic
PTPRQ	NM_00145026	c.4871-1G>C	//	//	//	//	//	D	//	//	//	//	12.62	5.59	7.904	19.595	//	//
SLC17A8	NM_139319	c.549_555del	p.Gly183fs	//	//	//	//	//	//	//	//	//	35*	//	//	//	//	//
TMPRSS3	NM_02422	c.446+1G>T	//	//	//	//	//	D	//	//	//	//	17.12	4.94	6.023	18.165	//	//
TMPRSS3	NM_02422	c.1224delA	p.Gln408fs	//	//	//	//	//	//	//	//	//	24.3*	//	//	//	//	//

*Table 4.5 List of novel mutations identified with prediction of pathogenicity. * manually annotated. // no prediction. NA: not available because no information on segregation. SIFT: D: deleterious; T: tolerated. PolyPhen2 HDIV/PolyPhen2 HVAR: D: Probably damaging; P: possibly damaging; B: benign. LRT: D: Deleterious; N: Neutral; U: Unknown. MutationTaster: A: disease_causing_automatic; D: disease_causing; N: polymorphism; P: polymorphism_automatic. MutationAssesor: H: high; M: medium; L: low; N: neutral. H/M means functional and L/N means non-functional. FATHMM/RadialSVM/LR: D: Deleterious; T: Tolerated.*

4.4 Genotype - phenotype correlation

The NGS analysis resulted negative in 13/78 subjects (16.7%), in which no pathogenic or likely pathogenic variants were found. We defined as "non-conclusive" 21/78 subjects (26.9%). Among the latter group, 14/21 subjects (66.7%) carried a monoallelic variant in a gene reportedly responsible of a recessive form of hearing loss; in 2/21 of the "non-conclusive" subjects (9.5%) two variants of the same gene were found and determined, by segregation analysis, to be *in cis,* while in 5/21 (23.8%) despite a pathogenic prediction, the identified variants did not seem to segregate with the deafness phenotype in the family.

In 10/78 (12.8%) cases in which the possible prediction of pathogenicity for the identified mutations was not so strong, the highly recommended family segregation analysis is still pending.

34 out of 78 studied subjects carried either a single pathogenic variant in a gene associated with a dominantly transmitted hearing loss or two pathogenic variants in a gene with autosomal recessive inheritance and so were classified as positive cases. Our analysis therefore resulted to have a diagnostic yield of 43.6%.

The family history and audiological information on all the 34 positive-cases are reported in Table 4.6, while a graphical representation is reported in Figure 4.2.

In particular, 20 of them had a positive family history of hearing loss, and in the vast majority (29/34) no consanguinity was reported in the family. The hearing loss was bilateral and symmetric in 31 subjects, while it was bilateral and asymmetric in 3 subjects.

The HL was predominantly congenital (14/34) or post-verbal (12/34). Stable HL was reported in 17 of these subjects, while progressive in 10 cases. In only one subject a fluctuating course was documented. The HL was predominantly profound (12/34), then moderate (9/34), and severe (8/34); in only 2 cases it was reported as mild.

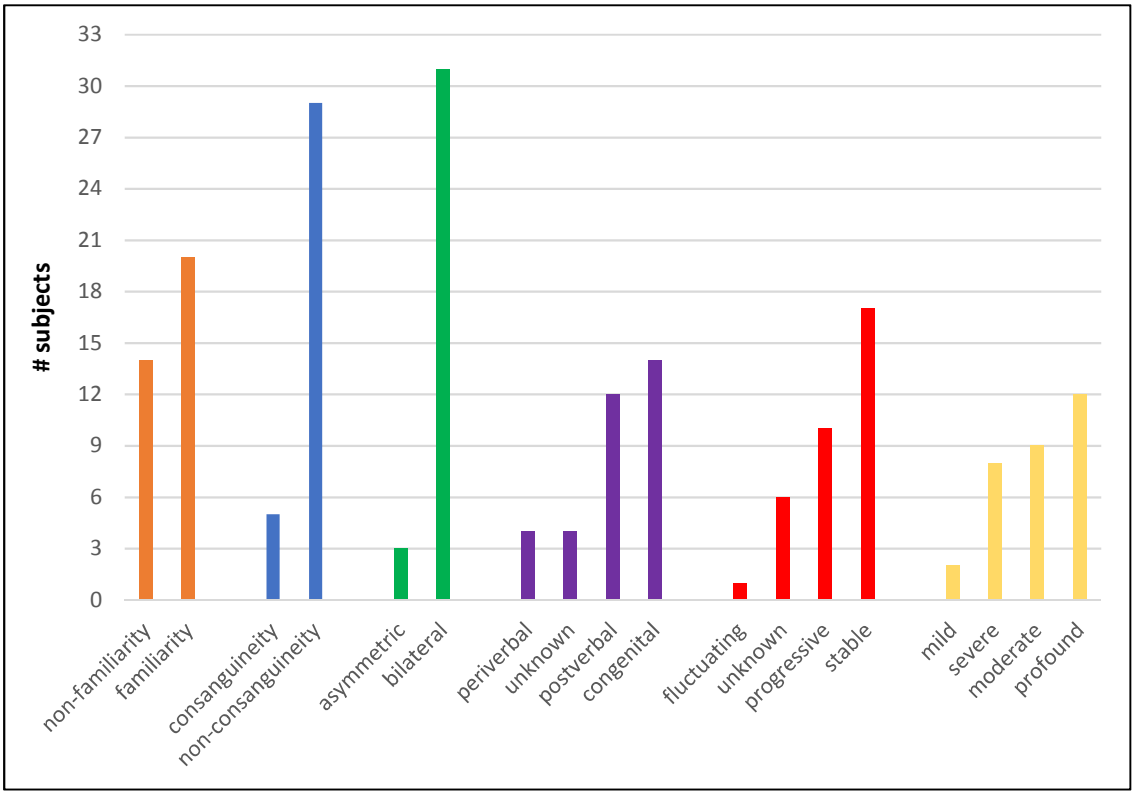


Figure 4.2 Familiar and audiological information of the NGS positive cases. X-axis reports the condition; Y-axis indicates the number (#) of subjects.

ID #	Sex	Age	Gene	familiarity	consanguineity	onset	laterality	progression	severity	audiogram shape
91	m	4	*ACTG1*	no	no	congenital	asymmetric	stable	moderate dx severe sx	unknown
125	f	4	*ADGRV1*	no	no	congenital	bilateral	stable	moderate	flat
10	m	11	*CDH23*	no	no	post-lingual	bilateral	progressive	moderate	U-shaped
118	f	1	*CDH23*	yes	no	congenital	bilateral	stable	severe	down-sloping
151	f	13	*CDH23*	no	no	post-lingual	bilateral	unknown	mild	flat
160	f	53	*CDH23*	yes	yes	post-lingual	bilateral	fluctuating	profound	unknown
8	f	16	*CDH23*	yes	no	peri-lingual	bilateral	stable	profound	unknown
34	m	3	*CDH23*	no	no	congenital	bilateral	stable	profound	unknown
77	f	8	*CDH23*	no	no	unknown	bilateral	progressive	profound	flat
81	f	18	*CDH23*	yes	no	post-lingual	asymmetric	progressive	mild dx anacusia sx	unknown
141	m	30	*CEACAM16*	yes	no	congenital	bilateral	unknown	profound	unknown
9	f	13	*COCH*	yes	no	congenital	bilateral	stable	profound	unknown
23	f	41	*COCH*	yes	no	post-lingual	bilateral	unknown	severe	down-sloping
122	m	15	*COCH*	yes	no	peri-lingual	bilateral	stable	moderate	flat
147	f	47	*COL11A2*	yes	no	post-lingual	bilateral	progressive	moderate	flat
14	f	9	*EYA4*	yes	no	post-lingual	bilateral	progressive	profound	unknown
38	f	14	*EYA4*	no	no	unknown	bilateral	unknown	moderate	unknown
76	f	7	*GJB2*	yes	yes	congenital	bilateral	stable	severe	down-sloping
56	m	9	*GJB2*	yes	yes	peri-lingual	bilateral	stable	moderate	U-shaped
11	f	30	*GJB3*	yes	no	peri-lingual	bilateral	progressive	severe	down-sloping
138	m	10	*KCNQ4*	yes	no	post-lingual	bilateral	stable	moderate	U-shaped
153	f	7	*MYH9*	no	no	unknown	bilateral	progressive	severe	down-sloping
114	m	6	*MYO6*	no	no	congenital	bilateral	stable	severe	flat
99	f	1	*MYO7A*	no	yes	congenital	bilateral	unknown	severe	down-sloping
5	f	12	*MYO7A*	no	no	post-lingual	bilateral	stable	moderate	flat
68	f	15	*OTOG*	yes	no	post-lingual	bilateral	progressive	moderate	down-sloping
51	f	47	*PTPRQ*	yes	yes	post-lingual	bilateral	progressive	profound	unknown
132	m	33	*SLC17A8*	yes	no	congenital	bilateral	stable	profound	flat
86	f	1	*SLC17A8*	no	no	congenital	bilateral	stable	severe	unknown
15	f	32	*SLC26A4*	yes	no	congenital	bilateral	unknown	profound	unknown
109	f	65	*SLC26A4*	yes	no	congenital	bilateral	progressive	profound	flat
120	m	10	*STRC*	no	no	unknown	bilateral	stable	mild	down-sloping
3	m	26	*TMC1*	yes	no	congenital	bilateral	stable	profound	unknown
83	f	7	*TMPRSS3*	no	no	post-lingual	asymmetric	stable	profound dx severe sx	ski-slope

Table 4.6 Genotype-phenotype correlation of the positive cases. ID # indicates the identification number for each subject. M: male; F: female. dx: right sx: left.

4.5 Case reports

4.5.1 Case 1: a case of consistent genotype-phenotype correlation

The DNA sample of a Caucasian 47-year-old woman (II-1), heterozygote for a maternally transmitted *GJB2* c.35delG mutation, was analyzed with the targeted gene panel. She presented bilateral sensorineural profound post-lingual hearing loss, with vestibular dysfunctions and referred onset at 10 years of age. Her parents were consanguineous as illustrated in Figure 4.3 and both had congenital sensorineural hearing loss; the father also had vestibular dysfunctions. The two sisters of the proband were normal hearing.

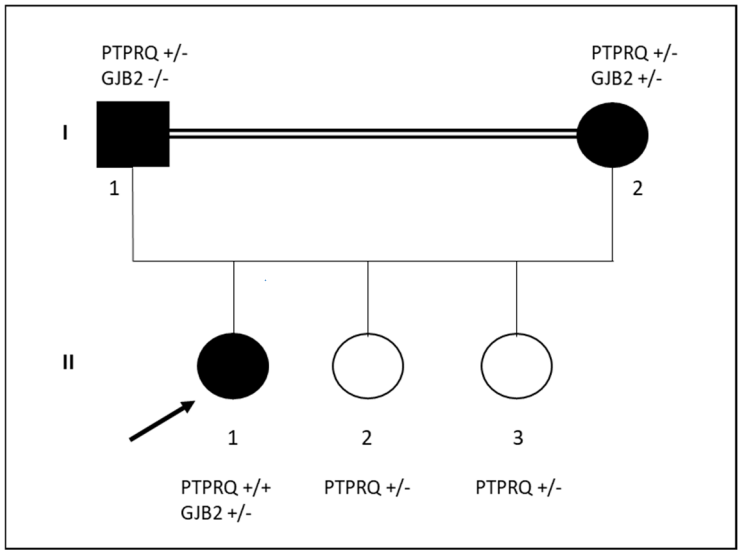


Figure 4.3 Family pedigree. The arrow indicates the proband (II-1). Square indicates males and circle female individuals. Solid symbol represents hearing-impaired individuals. +/- Heterozygous; +/+ homozygous; -/- wild-type

A total of 524 variants were obtained for the proband, of which 389 intronic and 135 exonic. After the filtration process, a novel homozygous splice-site mutation in the *PTPRQ* gene (NM_001145026) was identified. The mutation, c.5390-1G>C, had never been described in literature or reported in

public databases. According to the Human Splicing Finder tool, this is an alteration of the exonic splicing silencer (ESS) site. This mutation, which had a total coverage of 88X, was visualized with IGV; and Sanger validated, as illustrated in Figure 4.4.

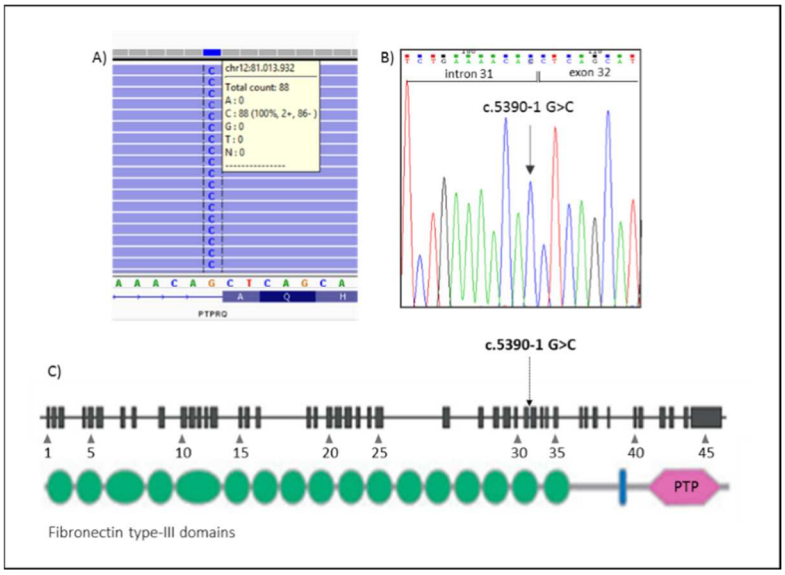


Figure 4.4 A) IGV visualization of the reads. B) Electropherogram of the intron 31/exon 32 of the PTPRQ gene. The arrow indicates the c.5390-1G>C splice site mutation. C) Schematic representation of the splice-site variant location. Green circle represents the fibronectin type III domains. PTP: protein-tyrosin phosphatase

The protein-tyrosine phosphatase, receptor-type, Q (*PTPRQ*) had both protein-tyrosine phosphatase activity and phosphatidylinositol phosphatase activity; it is associated with autosomal recessive non-syndromic sensorineural hearing loss with vestibular dysfunction (DFNB84A; OMIM # 613391) [54]. The parents of the proband were both found to be heterozygote for of the splice site mutation.

4.5.2 Case 2: Early diagnosis of Usher Syndrome

A 3-year-old Caucasian boy (II-1), born from non-consanguineous normal-hearing parents, as illustrated in Figure 4.5, was referred to our center because of congenital sensorineural bilateral profound hearing loss and peripheral retinal alteration.

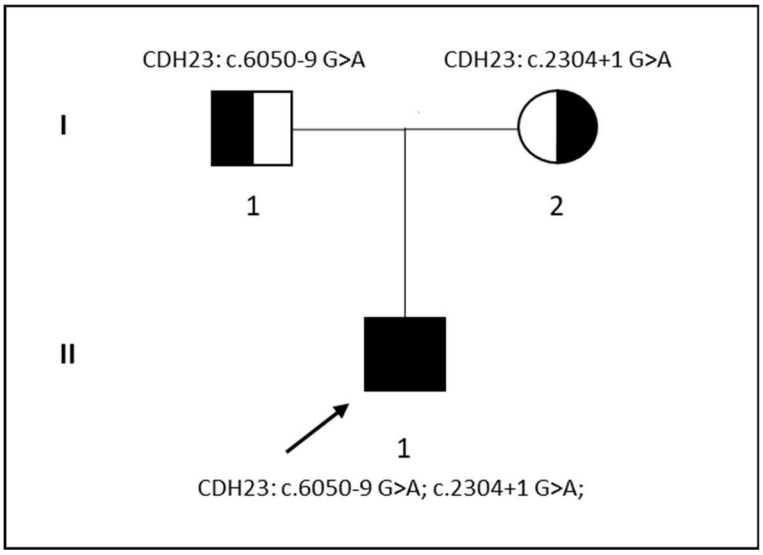


Figure 4.5 Family pedigree. The arrow indicates the proband (II-1). Square indicates males and circle female individuals. Solid symbol represents hearing-impaired individual. Half-coloured symbols indicate normal hearing subjects, carrier of the mutation.

The targeted gene panel approach allowed the identification of 476 variants, of which 127 exonic and 349 intronic; the filtration process pointed out two heterozygous likely pathogenic alterations in the *CDH23* gene (NM_022124) c.2304+1G>A and c.6050-9G>A.

The c.2304+1 G>A splice-site mutation, was localized in intron 22 and reported in dbSNP as rs769433759; it was not listed in the 1000 Genomes database but was reported in the disease-specific database (Deafness Variation Database and LOVD) as pathogenic and associated with Usher syndrome type 1D, based on published literature. This variant, previously known as IVS20+1G>A had originally been identified in a Dutch family with atypical Usher syndrome type I and in an European and a Spanish family with Usher syndrome type 1 [55].

The pathogenic role of this variant was tested *in vitro by* minigene assay which demonstrated this alteration to generate two different transcripts, one of which did not recognize the main splice donor site and created a new donor site including the first149 nucleotides of intron 21; the second transcript produced the skipping of exon 21 [56].

The second *CDH23* variant identified, c.6050-9G>A, was localized in intron 46 and reported in dbSNP as rs367928692; it was reported in public databases with a very low frequency ($4{,}969 \times 10^{-5}$ gnomAD browser, last accessed October 2017). This mutation, previously indicated as IVS45-9G>A, was reported in disease-specific databases as pathogenic and associated with Usher syndrome type 1D. The two identified variants with a coverage of 74X and 82X respectively, were Sanger validated and evaluated with IGV, as illustrated in Figure 4.6

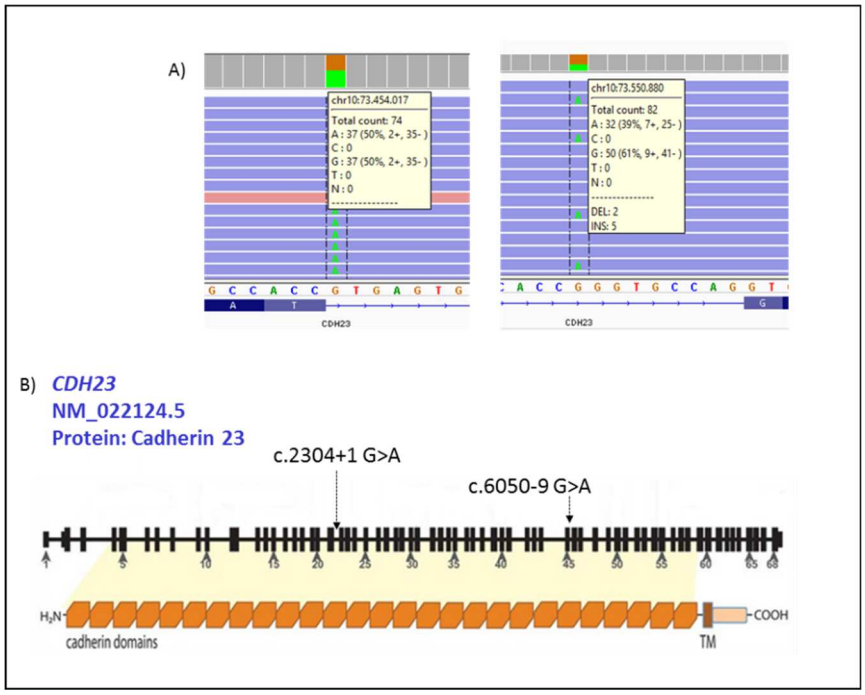


Figure 4.6 A) IGV visualization of the reads with the variants. B) Schematic representation of splice-site variants position within the CDH23 gene. TM: transmembrane domain.

The *CDH23* gene encodes a member of the cadherin superfamily, which comprises calcium-dependent cell-cell adhesion glycoproteins. The gene is associated both with autosomal recessive

non-syndromic hearing loss (DFNB12; OMIM # 601386) and with Usher syndrome type 1D (OMIM # 601067). It is reported in literature that *CDH23* missense mutations tend to produce a less severe phenotype, both in Usher syndrome type I and in NSHL, whereas nonsense, splice-site, and frameshift mutations produce a severe Usher syndrome type I phenotype [57].

Two splice-site mutations were identified in the child tested; as expected these were inherited from his unaffected parents, who were both carriers, as illustrated in Figure 4.5. According to the data reported in literature, this genotype could be associated with a severe Usher syndrome.

4.5.3 Case 3: detection of a novel splice-site mutation and in vitro demonstration of its pathogenicity

The proband (II-2), an 11-year-old Caucasian girl, presented sensorineural seemingly non-syndromic peri-lingual bilateral HL; the same phenotype was present in her 16-year-old sister, while the parents were normal hearing, as illustrated in Figure 4.7.

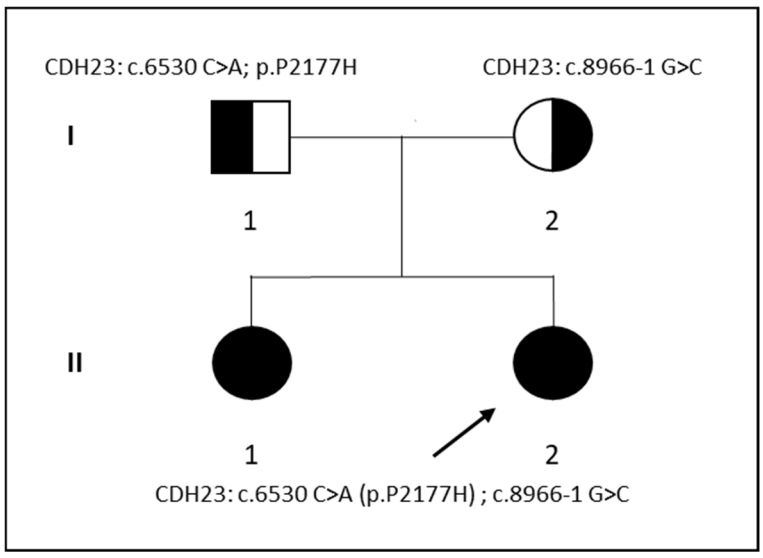


Figure 4.7 Family pedigree. The arrow indicates the proband (II-2). Square indicates males and circle females individuals. Solid symbol represents hearing-impaired individuals. Half-coloured symbols indicate normal hearing subjects, carrier of the mutation.

The targeted gene panel approach allowed the identification of 457 variants in total, divided in 107 exonic and 350 intronic variants; among which, the likely causative variants: c.6530C>A (p.Pro2177His) and c.8966-1G>C were located in the *CDH23* gene (NM_022124). In Figure 4.8 is reported the IGV visualization of the reads and the positions of the gene variants.

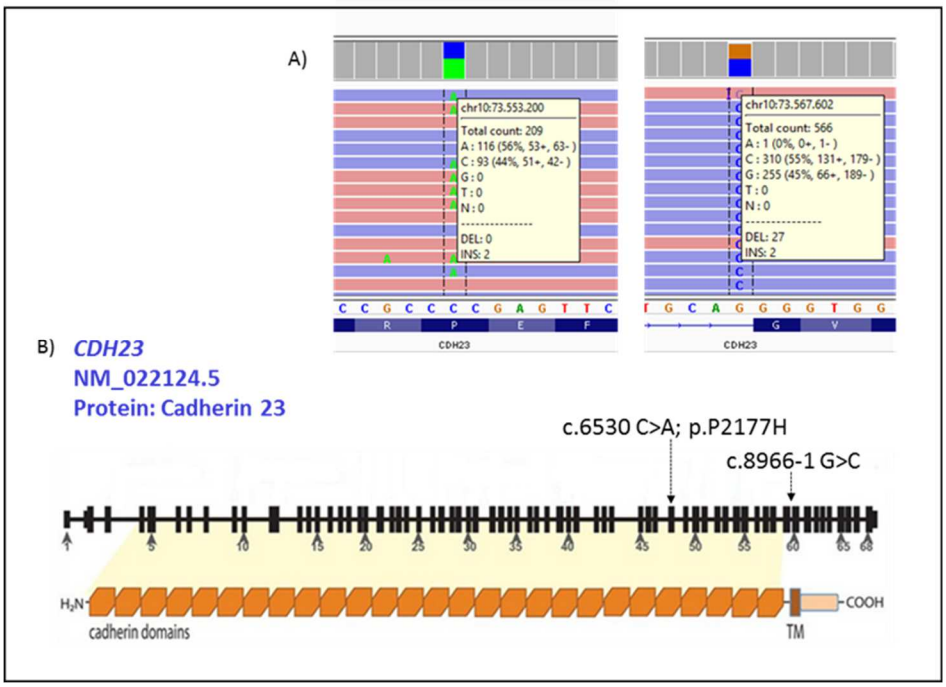


Figure 4.8 A) IGV visualization of the reads with the variants. B) Schematic representation of non-synonymous and splice-site variants within the CDH23 gene. TM: transmembrane domain.

The non-synonymous SNV c.6530C>A (p.Pro2177His), located in exon 48, was covered 209 X and was reported in dbSNP with rs748946014. It was identified in the general population with a low frequency (4.067×10^{-6} gnomAD browser, last accessed October 2017) and the aminoacid change was predicted deleterious by 8/9 *in silico* predictor tools, described in detail in section 1.10 and reported in Table 4.7. The Pro2177 residue was located in a highly phylogenetic conserved region (GERP++: 5.28; CADD_Phred: 22) in a cadherin-domain of the protein, as shown in Figure 4.9. InterVar tool automatically predicted this mutation as of "unknown significance".

SIFT_pred	Polyphen2_HDIV_pred	Polyphen2_HVAR_pred	LRT_pred	MutationTaster_pred	MutationAssessor_pred	FATHMM_pred	RadialSVM_pred	LR_pred	CADD_phred	GERP++	phyloP100way_vertebrate	SiPhy_29way_logOdds
D	D	D	D	D	H	//	D	D	22	5,28	7,786	18,903

Table 4.7 Results of in silico predictor tools. D: deleterious. H: high. // no prediction

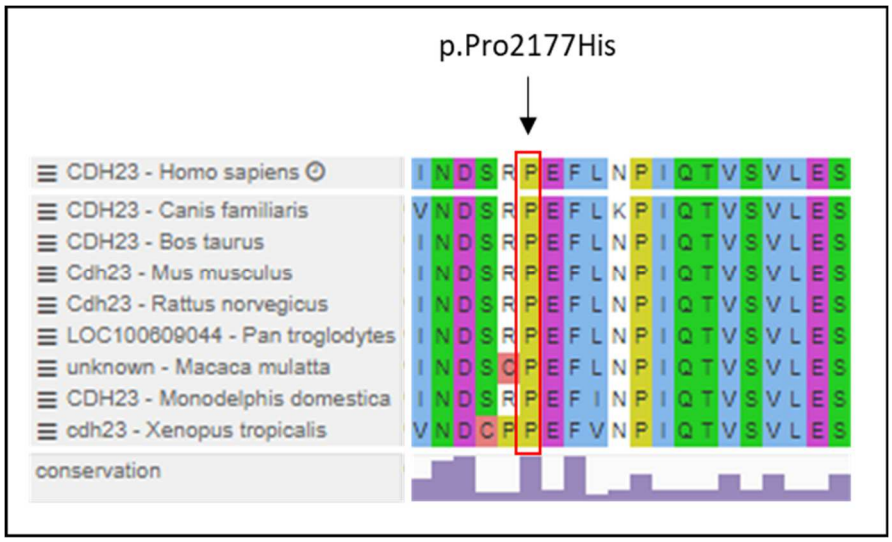


Figure 4.9 Multiple alignments of the CHD23 protein aminoacid sequence among species. Pro2177 is squared in red.

The other identified variant, c.8966-1G>C, covered 566X was a novel splice-site mutation, not present in public databases, nor in disease specific databases. It was located in intron 59 and was predicted by human splicing finder tool to alter an intronic ESS site. The alterations seemed to determine a deletion of 90 bp in exon 60.

Family segregation analysis was performed and demonstrated that the father carried the non-synonymous SNV, the mother carried the splice-site mutation, while the sisters had both mutations. Considering this information on family segregation, InterVar parameters for the non-synonymous variant were manually adjusted and the missense variant could be classified as “likely pathogenic”.

To investigate on the specific effect of the novel splice site mutation, mRNA from the two affected sisters and the parents was retrotranscribed as described in section 3.7. The cDNA was amplified with specific primers, and the products were Sanger sequenced. As illustrated in Figure 4.10 the cDNA from the father, defined as wild-type, presented the entire sequences of the exons 59, 60, 61; while the cDNA of the sisters and of the mother showed skipping of the entire exon 60.

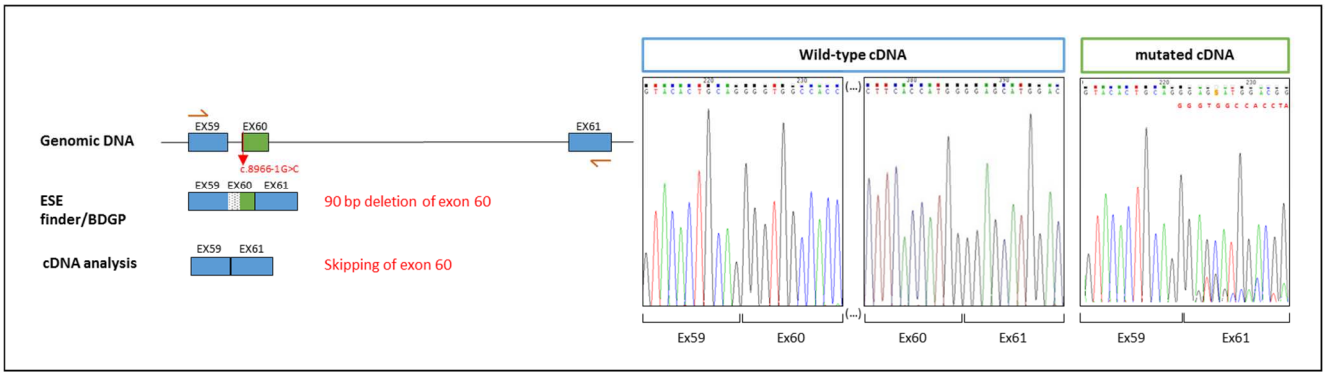


Figure 4.10 Schematic representation of the splice-site position. Red arrows represent the primers used. Prediction of ESE finder tool and results of cDNA analysis. On the right, the electropherogram of wild-type and mutant cDNA.

In this case, the targeted gene panel lead to the identification in the affected relatives of two variants that could be assigned to the pathogenic category. The exact pathogenic effect of the novel splice-site mutation was demonstrated *in vitro*.

4.5.4 *Case 4: Dual molecular diagnosis in a family in which segregate non-syndromic HL and Waardenburg syndrome type 1*

The proband, a 9-year-old Caucasian girl (III-1) was referred to our department for an etiological evaluation of her sensorineural hearing loss. Her family history was positive for deafness as reported in Figure 4.11, with both congenital and post-lingual cases reported. No cardiac problems were present in the family.

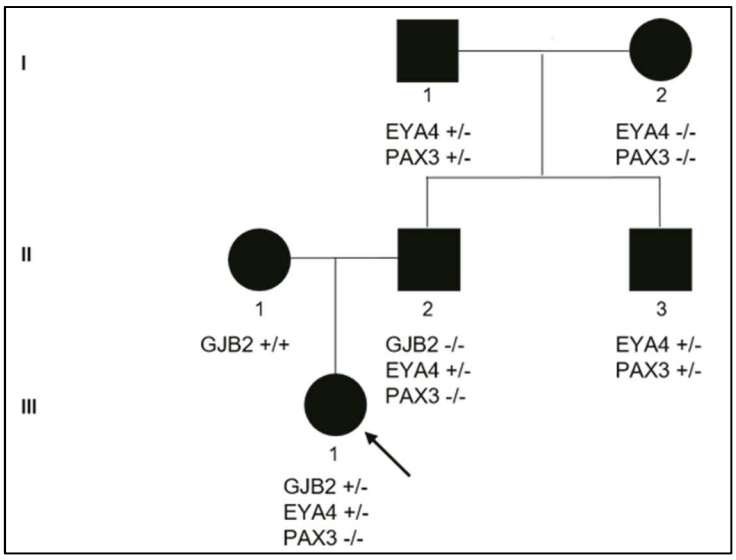


Figure 4.11 Family pedigree. The arrow indicates the proband (III-1). Square indicates males and circle female individuals. Solid symbol represents hearing-impaired individuals. +/- Heterozygous; +/+ homozygous; -/- wild-type.

Pure tone audiometry revealed mild hearing loss, which was reported and documented as being post-lingual. *GJB2* analysis on the proband's gDNA revealed the presence of a heterozygous c.35delG mutation, transmitted from her congenitally deaf homozygote mother (II-1), while the father tested negative.

NGS analysis with the 59-targeted gene-panel was performed on the gDNA of the proband. The bioinformatics analysis finally highlighted a novel heterozygous mutation in exon 13 of the *EYA4*

gene: NM_004100.4: c.1154C>T (NP_004091.3: p.Ser385Leu). This variant wasn't present in ExAC database (http://exac.broadinstitute.org/; last accessed September 2017) or in gnomAD browser (http://gnomad.broadinstitute.org/; last accessed September 2017). It was never described as associated with hearing phenotype nor it was present in deafness variation database (http://deafnessvariationdatabase.org; last accessed September 2017).

All the utilized *in silico* tools, predicted this alteration as damaging since substituting a polar aminoacid with a non-polar one in a highly phylogenetically conserved residue (CADD phred: 35) within the functionally relevant eya-homologous region (eyaHR) as reported in Figure 4.12.

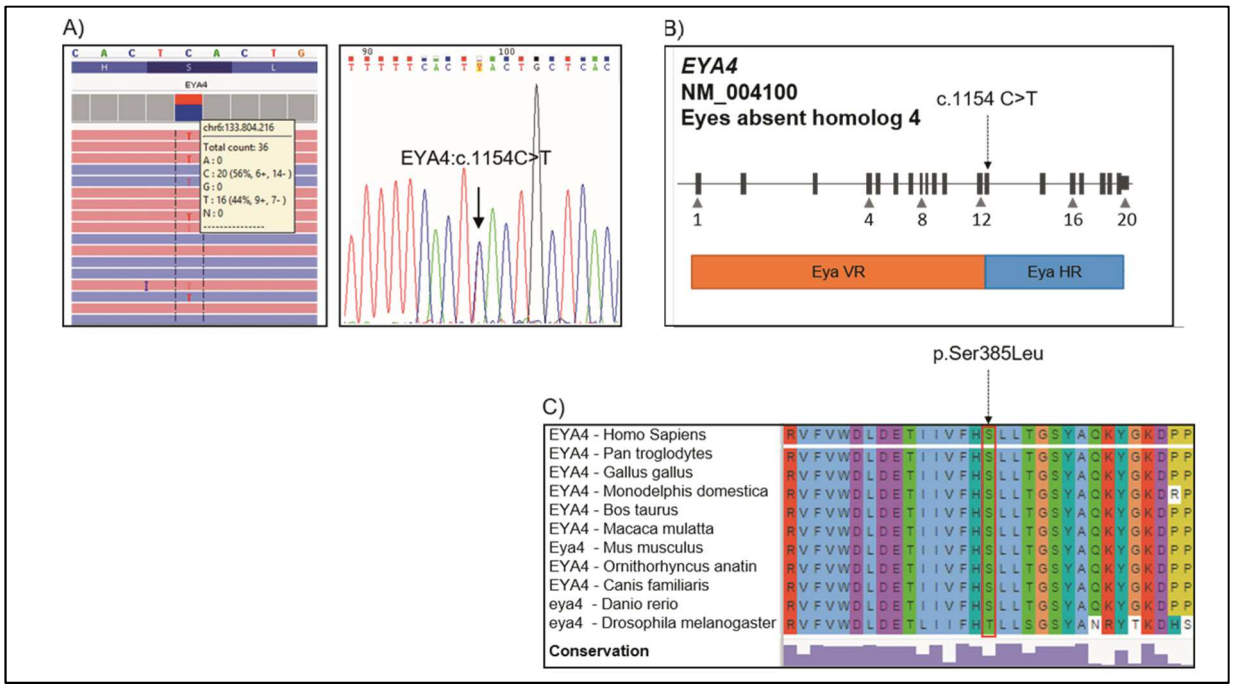


Figure 4.12 A) On the left IGV visualization; on the right Sanger sequencing electropherogram. The arrow indicates the c.1154C>T mutation. B), On the top EYA4 schematic exon representation and on the bottom protein domains: EyaVR (Eya-variable region) and Eya HR (Eya-homologous region). The EYA4 mutation is localized in exon 13 on the EYA HR domain. C) Multiple alignments of the EYA4 protein aminoacid sequence among species. Ser385 is squared in red.

Segregation analysis was performed on gDNA samples of the proband and other hearing-impaired family members, who also underwent a complete audiological and clinical phenotype evaluation.

The proband's grandmother (I-2) was a 67-year-old woman affected by profound hearing loss with post-lingual onset and progressive course. She also reported unspecified familial HL, although no other affected members of her family were available for the study. The 74-year-old grandfather (I-1) had congenital profound hearing loss, brilliant blue eyes, no facial dysmorphic features and was referred as having precocious white hair with a white forelock. The 39-year-old uncle (II-3) had congenital profound hearing loss, brilliant blue eyes, facial dysmorphisms (*dystopia canthorum*, downslanting palpebral fissures and malar hypoplasia) and pigmentary alterations, among which partially white eyebrows and early white hair with referred white forelock. These individuals (I-1 and II-3) who clearly met major diagnostic criteria for Waardenburg syndrome type 1 according to the WS Consortium criteria [58], had not been previously diagnosed or offered appropriate molecular testing.

Based on the clinical evaluations we proceeded with *PAX3* gene testing in the two subjects with WS1 phenotype (I-1 and II-3) together with segregation analysis of the *EYA4* variant.

The proband (III-1), her father (II-2), her uncle (II-1) and grandfather (I-1) were all found to be heterozygote for the *EYA4* mutation, while the grandmother (I-2), was negative.

Comprehensive molecular and clinical data for all the family-member are summarized in Table 4.8.

			Phenotype			Genotype	
Pedigree position	Gender	Age	HL Onset	Severity	WS signs	*EYA4*	*PAX3*
I-1	M	74	congenital	profound	blu eyes	c.1154C>T; p.Ser385Leu	c.321+1G>A
I-2	F	67	post-lingual	profound	//	//	//
II-2	M	42	post-lingual	profound	//	c.1154C>T; p.Ser385Leu	//
II-3	M	39	congenital	profound	blu eyes, dystopia cantorum	c.1154C>T; p.Ser385Leu	c.321+1G>A
III-1	F	9	post-lingual	mild	//	c.1154C>T; p.Ser385Leu	//

Table 4.8 Summary of phenotypic features and genotypes. M: male; F: female. Age of the subjects at evaluation. WS indicates Waardenburg syndrome. The symbol // indicates absence of WS clinical signs or mutation.

The entire coding sequence of the *PAX3* gene was analyzed in individuals I-1 and II-3 with Sanger sequencing, as described previously. A novel splice-site mutation NM_181459.3: c.321+1G>A was identified in these subjects. The mutation, never previously described in literature, nor reported in the public databases as ExAC database (http://exac.broadinstitute.org/; last accessed September 2017) or gnomAD browser (http://gnomad.broadinstitute.org/; last accessed September 2017). This splicing mutation was predicted by Human splicing finder to alter the splice-donor site of PAX3 intron 2 as reported in Figure 4.13.

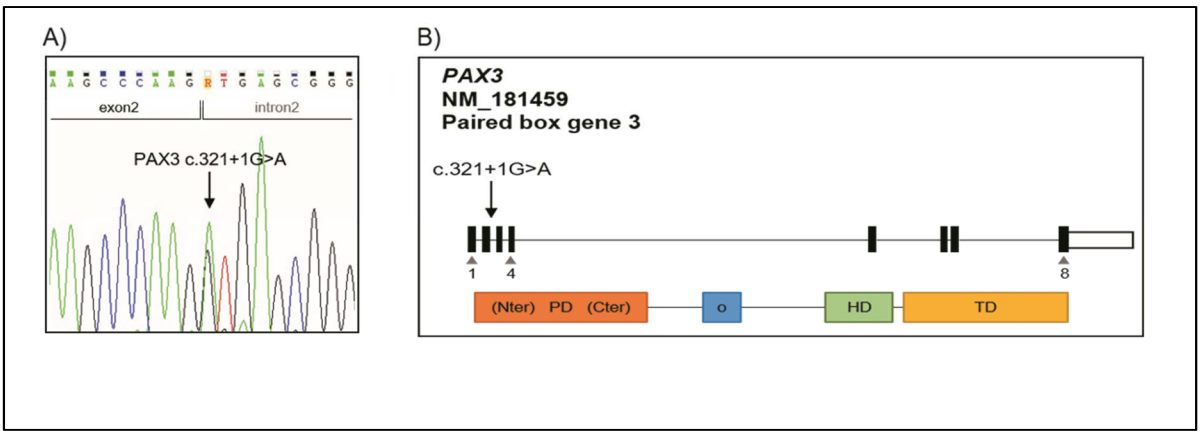


Figure 4.13 A) Electropherogram of exon/intron 2 of PAX3 gene. The arrow indicates the c.321+1G>A splice-site mutation. B) On the top PAX3 schematic exon representation and on the bottom protein domains: PD (paired domain); o (octapeptide motif); HD (homeodomain); TD (transactivation domain). The PAX3 mutation is located in intron 2.

The novel *EYA4* mutation detected by targeted NGS approach in the proband (III-1) could be responsible for the type of hearing loss she presented and, with obvious age differences, seemed to be typical of the audiological history reported for her father.

At the clinical evaluation, the 42-year-old proband's father, whose hearing loss was post-lingual, with referred onset before the age of 10, was found to be profoundly deaf. The proband, 9-year-old at the time of evaluation, also had post-lingual hearing loss that was measured as mild and still prevalently involving mid frequencies.

A novel *PAX3* splice-site mutation was identified in the affected uncle (II-3) and grandfather (I-1) of the proband, both of whom had WS1 phenotype never previously investigated.

These two subjects also carried the familial *EYA4* mutation, but the milder clinical manifestations and the progressive course of the DFNA10 phenotype had obviously been masked by the effect of the novel *PAX3* gene splice-site mutation, which by itself is the likely cause of their congenital profound hearing impairment [59] (Appendix 3).

5. DISCUSSION

Hearing loss is a highly heterogeneous condition, challenging both from a clinical and genetic standpoint. The cause of non-syndromic sensorineural deafness can be extremely difficult to investigate based on solely clinical data, especially when the hearing loss is pre-lingual and profound. On the other hand, a deafness syndrome can be caused by alterations in different genes and different mutations in the same gene may result both in SHL and NSHL [60]. Finally, a specific genetic cause may present with extremely variable phenotypic expression, even within members of the same family [61], and often the same gene can sustain both dominant and recessive conditions.

In this study, a highly selected targeted panel approach was decided with the aim of developing a translational tool with high diagnostic relevance for NSHL and for some forms of SHL with onset as isolated HL, such as Pendred and Usher syndrome. In these conditions, the earlier age of onset of the hearing loss versus other later clinical manifestations may be the cause of diagnostic pitfalls or late diagnoses. In young individuals, these disorders may be therefore clinically indistinguishable from isolated hearing loss but, in view of possible later complications, especially the loss of vision in Usher syndrome, an early etiological diagnosis is absolutely essential for a proper clinical follow-up and the establishment of adequate rehabilitation strategies.

The rationale for the choice of a targeted strategy to comprehensively investigate the genetic bases of hereditary hearing loss lies in several levels of considerations: first of all, the genetic heterogeneity of the condition as a whole, even though within an apparently manageable number of disease-causing or highly candidate disease genes.

The use of other NGS approaches, namely whole genome or whole exome sequencing, would give the possibility of including in the analysis a much higher number of genes, and even noncoding sequences, versus the limited number of genes in the targeted approach. Obviously, a targeted panel

would have the disadvantage of not accounting for new genes possibly found to be responsible of HL, which should eventually be accordingly implemented.

The advantages of the number of genes and sequences though need to be weighed against the difficulties of data analysis and variants interpretation, the risk of unsolicited or incidental findings and, ultimately but not less importantly, the lower depth of coverage that these latter approaches offer.

A comprehensive targeted panel puts together all, or a large and highly prioritized set of genes, with recognized pathogenic role within the investigated clinical area and guarantees the highest coverage of any sequence of interest that may have been introduced.

A targeted panel may combine advantageous features of the different NGS strategies and may therefore result more efficient and actionable.

The Ion Torrent PGM was considered the NGS ideal platform both for its innovative sequencing chemistry and for technical features that respond to the needs of a single laboratory.

Developing a comprehensive NGS targeted panel is a rather demanding process which essential pre-requirement is a solid clinical knowledge about the disorders to be studied. A series of different steps should take place, going from literature interrogation and intense datamining to gene prioritization, in which expression, function, networking and “omic” aspects of each gene need to be examined in depth.

The process of gene-selection we undertook was performed also according to evidences of pathogenicity and possible impact on hearing loss in Caucasian individuals, with particular attention to early onset conditions. This type of choice was dictated by the characteristics of the population sample that would mainly be evaluated during the project, largely Caucasian and of pediatric age.

An extensive review of the literature and of databases was therefore performed, with a particular attention in integrating clinical information with genomic and transcriptomic data. Different

bioinformatics tools were exploited for this gene selection phase in which the respective information on gene and protein function were compared.

The 59 genes finally selected accounted for more than 90% of all the deafness-associated mutations reported in Caucasian by the literature at the time of gene selection phase. (Deafness Variation Database, accessed on May 2015).

The use of targeted gene panel in combination with massive parallel sequencing and a tailored bioinformatics pipeline of data analysis, allowed us to identify the genetic cause of hearing loss in 34 subjects out of the 78 tested in total. The diagnostic yield of the 59-gene panel that we have developed reached a value of 43.6%.

The obtained diagnostic rate is slightly higher than the mean value of 41% reported in studies evaluating massive parallel sequencing for hearing loss [62]; in these studies targeted panels comprising different number of genes were used to evaluate different population samples. Among these studies, those performed in Caucasians or in European populations, report diagnostic rates ranging from 83% [31] to 38% [63]. The higher value is nevertheless referred to a population of only six samples, while the 38% value refers to 24 samples tested for 34 genes. Currently no studies have reported a profile of diagnostic yield on a large Italian cohort.

In our work 78 Caucasian subjects, mainly from Italy, were tested for 59 highly selected genes, so we believe our data may have a greater impact in our population, despite the possibility of having missed pathogenic mutations possibly located in genes not included in the panel, since unknown or never reported mutated in Caucasians at the time of genes-selection.

Furthermore, the diagnostic rate we report in our study is higher than the mean value of 33% reported in a very large cohort of 1119 samples from deaf individuals subjected to whole exome sequencing [33].

Taken together these data indicate how our targeted approach allowed to develop a successful diagnostic tool and how this tool actually responds to the original purpose of the project.

Despite the estimated frequency for autosomal recessive (75-80%) and autosomal dominant (20%) HL, 52.9 % of positive subjects of our cohort carried biallelic mutations in autosomal recessive genes, while 47.1% were heterozygotes for mutations in autosomal dominant genes. Our interpretation of the relatively high number of dominant cases is a possible selection bias, which might depend on the easier recruitment of cases with obvious familial transmission versus apparently sporadic cases in prevalently small families.

We found in the tested population a high number of subjects carrying only one likely causative variant in genes associated with recessive forms of HL (41.2%). This higher carrier frequency in subjects with hearing loss with respect to control populations is an interesting observation common to other studies [64].

Among the possible causes of this higher carrier rate could be a non-optimal overall coverage of some genes (i.e. *ESPN*, *P2RX2*) due to the difficulty of the AmpliSeq Designer to design amplicons; despite we filled the gaps with Sanger Sequencing in cases in which a recessive form of hearing loss was suspected and only one mutation was identified in a gene, and so we actually covered the whole sequence, we didn't found any likely causative mutations these heterozygotes.

These data open new perspectives on the research of other molecular mechanism that could be responsible of HL. Considering that one of the limits of targeted gene approach is the difficult to investigate gene deletions/duplications or chromosome rearrangements, it could be useful to integrate different technologies to obtain a more comprehensive information. Other mechanisms that could underlie the "missed" mutation data and that represent a future challenge for genomics studies are

non-coding sequence variants that could possibly alter intronic regions or regulatory elements or the presence of mutations in related genes sustaining a digenic inheritance.

All the variants identified in the course of the study were collected in an *in-house* database that became an invaluable tool for the identification of recurrent or novel variants, with their relative information (e.g. allele frequency, frequency in healthy population as reported in 1000 Genomes database, dbSNP accession number and *in silico* pathogenic prediction), for the identification of possible alignment errors, and for further stratification and correlation between genotype and phenotype. The local database also collects fundamental information about Sanger validation of detected variants and about family segregation analysis.

The database, developed by our laboratory, permits to evaluate the mutation's frequency among the cohort of selected subjects, prevalently Caucasians and mainly of developmental age, and it can be used to assess epidemiological data in our population.

In the overall, among the unique variants identified, 871 had a coverage higher than 20X; which was the lower cut-off established for the study; of these 448 had an "rs accession number" reported in dbSNP.

After the accurate filtration process and data analysis, 43 variants were actually classified as likely causative, 41.8% of which were novel (18/43), i.e. never described in literature, nor present in the public databases of healthy population or disease-specific cohorts. An integrated combination of nine *in silico* predictor tools that evaluate the type of aminoacid change and the aminoacid conservation among species was used to characterize each detected variant. The InterVar tool was also used to classify variants according to the criteria recommended by the ACMG guidelines and establish their patogenicity and it was also possible to adjust the parameters according to information on segregation analysis.

It is interesting to note that only 8 of the variants which we have actually been able to classify as disease-associated were already reported in the ClinVar database as pathogenic or likely pathogenic. All the novel pathogenic variants and the associated phenotypes will be submitted to the public databases in order to contribute to a better knowledge about the genetic bases of hearing loss.

The likely causative variants were found to alter the following genes included in the panel: *CDH23, GJB2, COCH, MYO7A, ADGRV1, EYA4, OTOG, SLC17A8, TMPRSS3, ACTG1, CEACAM16, COL11A2, GJB3, KCNQ4, MYH9, MYO6, PTPRQ, SLC26A4, STRC, TMC1*. The fact that only 33% of the total number of tested genes was found to harbor mutations could represent an epidemiological information given we have mainly tested a population from the North-East part of Italy, although a larger study would be necessary to confirm our results.

Among the patogenic variants identified, we found four *GJB2* mutations, that had been previously missed by Sanger sequencing analysis, either due to PCR technical limits or to the low sensitivity of Sanger sequencing in comparison with the NGS technology.

The most frequently mutated gene in our selected cohort was *CDH23*, which was found altered in 8 patients with 13 different pathogenic variants, of which 3 novel SNVs.

Even if the high frequency of *CDH23* gene mutations was already reported in two large Japanese cohorts [65] [66], our data for the first time report such a high number of mutations in *CDH23,* in a Caucasian cohort, resulting in a frequency of 23.5% (8/34).

In this work, seven individuals, compound heterozygote for *CDH23* mutations, presented NSHL prevalently characterized by post-lingual, moderate to profound hearing loss, consistently with what reported in the OMIM database and in literature for the DFNB12 phenotype (OMIM # 601386).

One subject presented Usher syndrome type 1D (OMIM # 601067).

For all the subjects with *CDH23* variants, it was possible to perform genotype-phenotype correlation that was compared and found to be consistent with the pertinent literature.

Two of the pathogenic *CDH23* mutations were detected in two sisters with sensorineural non-syndromic peri-lingual bilateral HL, without other symptoms. We identified, in these subjects, a known non-synonymous SNV and a novel splice-site variant of which we characterized the precise pathogenic mechanism *in vitro* and demonstrated to cause a complete exon-skipping.

We were able to substantiate with molecular data a suspect of Usher syndrome in a young child, who presented with bilateral congenital profound HL and suggestive early retinal alterations. Given the young age of the boy, the early identification of two splice-site pathogenic mutations of *CDH23* gene represented an extremely important prognostic information for the planning of a proper clinical follow-up and future audiological and clinical management as well as for the fundamentally important genetic counseling in the family.

We have also been able to establish a straightforward genotype-phenotype correlation in the case of a woman, born from consanguineous parents, who presented with bilateral sensorineural profound post-lingual hearing loss, with referred onset at 10 years of age, and vestibular dysfunctions. A novel *PTPRQ* homozygous splice-site mutation was detected in her DNA. This finding contributes to expanding the so far limited list of known pathogenic variants of this gene. Furthermore, we could confirm, based on the phenotype of the proband and the data reported in the OMIM database and in literature (DFNB84A, OMIM # 613391), a consistent genotype-phenotype correlation.

Our work also gave us the possibility of disentangling a complicate pattern of transmission in a deafness family with many affected members, as reported in section 4.5.4.

In this family, we observed the occurrence of assortative or non-random mating by deafness state, a well-documented condition in which deaf individuals partner together. Due to the known wide genetic

heterogeneity in hearing loss, assortative mating increases the likelihood of co-segregation of two or more distinct forms of deafness in the same family [18].

In the family we describe, two novel mutations in different deafness genes, *EYA4* (DFNB10) and *PAX3*, causing respectively non-syndromic and syndromic autosomal dominant HL were identified. A dual molecular diagnosis was obtained since these mutations, one of which explained the non-syndromic hearing loss of a father and his daughter, were found to co-occur in two other relatives showing a previously misdiagnosed Waardenburg syndrome type 1 phenotype that clearly masked the expected DFNA10 audiological profile.

Until now only 7 DFNA10 families have been reported in literature [67] and variable audiogram configurations have been described. Nevertheless, most of the individuals carrying heterozygous missense mutations in the *EYA4* gene show bilateral, progressive sensorineural hearing loss with variable age of onset ranging between the pre-lingual period to slightly over 50 years of age. DFNA10 typically involves mid frequencies at onset and progresses thereafter to affect all the frequencies [68]. Since one of the main features of the DFNA10 phenotype is progression, precocious identification of the *EYA4* mutation had a very relevant impact both for the prognosis and for the program of clinical follow-up of the young girl we describe.

Heterozygous mutations in the *PAX3* gene associate with a very variable phenotype, as in the case of the family described in this work. The individuals we report had in fact both congenital profound hearing loss and brilliant blue eyes, but facial dysmorphic features were present only in one subject. The milder clinical manifestations and the progressive course of the DFNA10 phenotype caused by the *EYA4* mutation they were also carrying, had obviously been masked by the effect of the novel *PAX3* gene splice-site mutation, which by itself is the likely cause of their congenital profound hearing impairment.

The targeted gene panel approach that we developed proved to be a very useful tool in a diagnostic setting for hearing loss due to its many advantages, such as high coverage of targeted regions, high throughput and straightforward bioinformatics data analysis and lower costs with respect to single gene sequencing,

Our work demonstrates how the combination of a highly efficient NGS technology, accurate gene-selection and consistent bioinformatics analysis, coupled with a thorough clinical evaluation allows to efficiently investigate the molecular causes of deafness. This integrated approach provides, together with a better genetic counseling and risk of recurrence for each affected individual, extremely important prognostic and follow-up information for a better and more precise audiological management of individuals with hearing loss. The integration of advanced technology and bioinformatics knowledge within the modern process of clinical reasoning is at the base of the new era of personalized medicine.

6. REFERENCES

[1] J.O. Pickles, Auditory pathways: anatomy and physiology, Handb. Clin. Neurol. 129 (2015) 3–25. doi:10.1016/B978-0-444-62630-1.00001-9.

[2] R. Fettiplace, C.M. Hackney, The sensory and motor roles of auditory hair cells, 7 (2006) 19–29. doi:10.1038/nrn1828.

[3] P. Kazmierczak, U. Mu, Sensing sound : molecules that orchestrate mechanotransduction by hair cells, 35 (2012). doi:10.1016/j.tins.2011.10.007.

[4] A.A. Dror, K.B. Avraham, Review Hearing Impairment : A Panoply of Genes and Functions, Neuron. 68 (2010) 293–308. doi:10.1016/j.neuron.2010.10.011.

[5] H. Zeng, Y. Zhao, Sensing movement: microsensors for body motion measurement., Sensors (Basel). 11 (2011) 638–660. doi:10.3390/s110100638.

[6] M.D. Eisen, D.K. Ryugo, Hearing molecules: Contributions from genetic deafness, Cell. Mol. Life Sci. 64 (2007) 566–580. doi:10.1007/s00018-007-6417-3.

[7] C.C. Morton, W.E. Nance, Newborn hearing screening--a silent revolution., N. Engl. J. Med. 354 (2006) 2151–2164. doi:10.1056/NEJMra050700.

[8] A.Y. Yamamoto, M.M. Mussi-Pinhata, M. de L. Isaac, F.R. Amaral, C.G. Carvalheiro, D.C. Aragon, A.K. da S. Manfredi, S.B. Boppana, W.J. Britt, Congenital cytomegalovirus infection as a cause of sensorineural hearing loss in a highly immune population., Pediatr. Infect. Dis. J. 30 (2011) 1043–1046. doi:10.1097/INF.0b013e31822d9640.

[9] G. Conti, R. Gallus, A.R. Fetoni, B.M. Martina, E. Muzzi, E. Orzan, G. Bastanza, Early definition of type, degree and audiogram shape in childhood hearing impairment., Acta Otorhinolaryngol. Ital. 36 (2016) 21–28. doi:10.14639/0392-100X-1074.

[10] M. Mazzoli, G. Van Camp, V. Newton, G. N, D. F, P. A, Recommendations for the Description of Genetic and Audiological Data for Families with Nonsyndromic Hereditary Hearing Impairment, Audiol. Med. 1 (2003) 148–150. doi:10.1080/16513860301713.

[11] 1997_Connexin 26 mutation in NSHL_Kelsell.pdf, (n.d.).

[12] A.A. Dror, K.B. Avraham, Review Hearing Impairment : A Panoply of Genes and Functions, Neuron. 68 (2010) 293–308. doi:10.1016/j.neuron.2010.10.011.

[13] T. Koffler, K. Ushakov, K.B. Avraham, Genetics of Hearing Loss: Syndromic., Otolaryngol. Clin. North Am. 48 (2015) 1041–1061. doi:10.1016/j.otc.2015.07.007.

[14] M. Parker, M. Bitner-Glindzicz, Genetic investigations in childhood deafness., Arch. Dis. Child.

0 (2014) 1–8. doi:10.1136/archdischild-2014-306099.

[15] W.J. Kimberling, M.S. Hildebrand, A.E. Shearer, L. Maren, J.A. Halder, E.S. Cohn, R.G. Weleber, M. Edwin, R.J.H. Smith, Frequency of Usher Syndrome in Two Pediatric Populations: Implications for genetic screening of Deaf and Hard of Hearing Children, 12 (2011) 512–516. doi:10.1097/GIM.0b013e3181e5afb8.Frequency.

[16] V. Pingault, D. Ente, F. Dastot-Le Moal, M. Goossens, S. Marlin, N. Bondurand, Review and update of mutations causing Waardenburg syndrome., Hum. Mutat. 31 (2010) 391–406. doi:10.1002/humu.21211.

[17] N.E. Morton, Genetic epidemiology of hearing impairment., Ann. N. Y. Acad. Sci. 630 (1991) 16–31.

[18] S. Angeli, X.I. Lin, X.U.E.Z. Liu, Genetics of Hearing and Deafness, Anat Rec. 295 (2015) 1812–1829. doi:10.1002/ar.22579.Genetics.

[19] A. Murgia, E. Orzan, R. Polli, M. Martella, C. Vinanzi, E. Leonardi, E. Arslan, F. Zacchello, Cx26 deafness: mutation analysis and clinical variability., J. Med. Genet. 36 (1999) 829–832.

[20] M. Amorini, P. Romeo, R. Bruno, F. Galletti, C. Di Bella, P. Longo, S. Briuglia, C. Salpietro, L. Rigoli, Prevalence of Deafness-Associated Connexin-26 (GJB2) and Connexin-30 (GJB6) Pathogenic Alleles in a Large Patient Cohort from Eastern Sicily, Ann. Hum. Genet. 26 (2015) 341–349. doi:10.1111/ahg.12120.

[21] P. Gasparini, R. Rabionet, G. Barbujani, S. Melçhionda, M. Petersen, K. Brøndum-Nielsen, A. Metspalu, E. Oitmaa, M. Pisano, P. Fortina, L. Zelante, X. Estivill, High carrier frequency of the 35delG deafness mutation in European populations. Genetic Analysis Consortium of GJB2 35delG., Eur. J. Hum. Genet. 8 (2000) 19–23. doi:10.1038/sj.ejhg.5200406.

[22] K. Cryns, E. Orzan, A. Murgia, P.L.M. Huygen, F. Moreno, I. del Castillo, G. Parker Chamberlin, H. Azaiez, S. Prasad, R.A. Cucci, E. Leonardi, R.L. Snoeckx, P.J. Govaerts, P.H. Van de Heyning, C.M. Van de Heyning, R.J.H. Smith, G. Van Camp, A genotype-phenotype correlation for GJB2 (connexin 26) deafness, J. Med. Genet. 41 (2004) 147 LP – 154. http://jmg.bmj.com/content/41/3/147.abstract.

[23] I. del Castillo, M. Villamar, M.A. Moreno-Pelayo, F.J. del Castillo, A. Alvarez, D. Telleria, I. Menendez, F. Moreno, A deletion involving the connexin 30 gene in nonsyndromic hearing impairment., N. Engl. J. Med. 346 (2002) 243–249. doi:10.1056/NEJMoa012052.

[24] F.J. del Castillo, M. Rodriguez-Ballesteros, A. Alvarez, T. Hutchin, E. Leonardi, C.A. de Oliveira,

H. Azaiez, Z. Brownstein, M.R. Avenarius, S. Marlin, A. Pandya, H. Shahin, K.R. Siemering, D. Weil, W. Wuyts, L.A. Aguirre, Y. Martin, M.A. Moreno-Pelayo, M. Villamar, K.B. Avraham, H.-H.M. Dahl, M. Kanaan, W.E. Nance, C. Petit, R.J.H. Smith, G. Van Camp, E.L. Sartorato, A. Murgia, F. Moreno, I. del Castillo, A novel deletion involving the connexin-30 gene, del(GJB6-d13s1854), found in trans with mutations in the GJB2 gene (connexin-26) in subjects with DFNB1 non-syndromic hearing impairment., J. Med. Genet. 42 (2005) 588–594. doi:10.1136/jmg.2004.028324.

[25] J. Rodriguez-Paris, M.L. Tamayo, N. Gelvez, I. Schrijver, Allele-Specific impairment of GJB2 expression by GJB6 deletion del(GJB6-d13s1854), PLoS One. 6 (2011). doi:10.1371/journal.pone.0021665.

[26] R.L. Alford, K.S. Arnos, M. Fox, J.W. Lin, C.G. Palmer, A. Pandya, H.L. Rehm, N.H. Robin, D. a Scott, C. Yoshinaga-Itano, American College of Medical Genetics and Genomics guideline for the clinical evaluation and etiologic diagnosis of hearing loss., Genet. Med. 16 (2014) 347–55. doi:10.1038/gim.2014.2.

[27] A.E. Shearer, M.S. Hildebrand, C.M. Sloan, R.J.H. Smith, Deafness in the genomics era., Hear. Res. 282 (2011) 1–9. doi:10.1016/j.heares.2011.10.001.

[28] B. Vona, I. Nanda, M.A.H. Hofrichter, W. Shehata-dieler, T. Haaf, Non-syndromic hearing loss gene identi fi cation : A brief history and glimpse into the future *, Mol. Cell. Probes. (2015). doi:10.1016/j.mcp.2015.03.008.

[29] S. De Keulenaer, J. Hellemans, S. Lefever, J.-P. Renard, J. De Schrijver, H. Van de Voorde, M.A. Tabatabaiefar, F. Van Nieuwerburgh, D. Flamez, F. Pattyn, B. Scharlaken, D. Deforce, S. Bekaert, W. Van Criekinge, J. Vandesompele, G. Van Camp, P. Coucke, Molecular diagnostics for congenital hearing loss including 15 deafness genes using a next generation sequencing platform., BMC Med. Genomics. 5 (2012) 17. doi:10.1186/1755-8794-5-17.

[30] M. Sommen, I. Schrauwen, G. Vandeweyer, N. Boeckx, J.J. Corneveaux, J. van den Ende, A. Boudewyns, E. De Leenheer, S. Janssens, K. Claes, M. Verstreken, N. Strenzke, F. Predohl, W. Wuyts, G. Mortier, M. Bitner-Glindzicz, T. Moser, P. Coucke, M.J. Huentelman, G. Van Camp, DNA Diagnostics of Hereditary Hearing Loss: A Targeted Resequencing Approach Combined with a Mutation Classification System., Hum. Mutat. 37 (2016) 812–819. doi:10.1002/humu.22999.

[31] A.E. Shearer, A.P. DeLuca, M.S. Hildebrand, K.R. Taylor, J. 2nd Gurrola, S. Scherer, T.E. Scheetz,

R.J.H. Smith, Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing., Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 21104–21109. doi:10.1073/pnas.1012989107.

[32] H.L. Rehm, S.J. Bale, P. Bayrak-toydemir, J.S. Berg, K.K. Brown, J.L. Deignan, M.J. Friez, B.H. Funke, ACMG Practice Guidelines ACMG clinical laboratory standards for next-generation sequencing, 15 (2013). doi:10.1038/gim.2013.92.

[33] C. Zazo Seco, M. Wesdorp, I. Feenstra, R. Pfundt, J.Y. Hehir-Kwa, S.H. Lelieveld, S. Castelein, C. Gilissen, I.J. de Wijs, R.J. Admiraal, R.J. Pennings, H.P. Kunst, J.M. van de Kamp, S. Tamminga, A.C. Houweling, A.S. Plomp, S.M. Maas, P.A. de Koning Gans, S.G. Kant, C.M. de Geus, S.G. Frints, E.K. Vanhoutte, M.F. van Dooren, M.-J.H. van den Boogaard, H. Scheffer, M. Nelen, H. Kremer, L. Hoefsloot, M. Schraders, H.G. Yntema, The diagnostic yield of whole-exome sequencing targeting a gene panel for hearing impairment in The Netherlands, Eur. J. Hum. Genet. 25 (2017) 308–314. doi:10.1038/ejhg.2016.182.

[34] B. Merriman, I. Torrent, J.M. Rothberg, Progress in Ion Torrent semiconductor chip based sequencing, Electrophoresis. 33 (2012) 3397–3417. doi:10.1002/elps.201200424.

[35] M.L. Metzker, Sequencing technologies the next generation, Nat. Rev. Genet. 11 (2010) 31–46. doi:10.1038/nrg2626.

[36] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data., Nucleic Acids Res. 38 (2010) e164. doi:10.1093/nar/gkq603.

[37] X. Chang, K. Wang, wANNOVAR: annotating genetic variants for personal genomes via the web, J. Med. Genet. 49 (2012) 433 LP – 436. http://jmg.bmj.com/content/49/7/433.abstract.

[38] H. Yang, K. Wang, Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR, Nat. Protoc. 10 (2015) 1556–1566. doi:10.1038/nprot.2015.105.

[39] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm., Nat. Protoc. 4 (2009) 1073–1081. doi:10.1038/nprot.2009.86.

[40] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations., Nat. Methods. 7 (2010) 248–249. doi:10.1038/nmeth0410-248.

[41] J.M. Schwarz, C. Rodelsperger, M. Schuelke, D. Seelow, MutationTaster evaluates disease-

causing potential of sequence alterations., Nat. Methods. 7 (2010) 575–576. doi:10.1038/nmeth0810-575.

[42] B. Reva, Y. Antipin, C. Sander, Predicting the functional impact of protein mutations: application to cancer genomics., Nucleic Acids Res. 39 (2011) e118. doi:10.1093/nar/gkr407.

[43] S. Chun, J.C. Fay, Identification of deleterious mutations within three human genomes., Genome Res. 19 (2009) 1553–1561. doi:10.1101/gr.092619.109.

[44] H.A. Shihab, J. Gough, D.N. Cooper, I.N.M. Day, T.R. Gaunt, Predicting the functional consequences of cancer-associated amino acid substitutions., Bioinformatics. 29 (2013) 1504–1510. doi:10.1093/bioinformatics/btt182.

[45] M. Kircher, D.M. Witten, P. Jain, B.J. O'Roak, G.M. Cooper, J. Shendure, A general framework for estimating the relative pathogenicity of human genetic variants., Nat. Genet. 46 (2014) 310–315. doi:10.1038/ng.2892.

[46] M. Pertea, G.M. Pertea, S.L. Salzberg, Detection of lineage-specific evolutionary changes among primate species, BMC Bioinformatics. 12 (2011) 274. doi:10.1186/1471-2105-12-274.

[47] E. V Davydov, D.L. Goode, M. Sirota, G.M. Cooper, A. Sidow, S. Batzoglou, Identifying a high fraction of the human genome to be under selective constraint using GERP++., PLoS Comput. Biol. 6 (2010) e1001025. doi:10.1371/journal.pcbi.1001025.

[48] M. Garber, M. Guttman, M. Clamp, M.C. Zody, N. Friedman, X. Xie, Identifying novel constrained elements by exploiting biased substitution patterns., Bioinformatics. 25 (2009) i54–62. doi:10.1093/bioinformatics/btp190.

[49] P. Jehl, J. Manguy, D.C. Shields, D.G. Higgins, N.E. Davey, ProViz-a web-based visualization tool to investigate the functional and evolutionary features of protein sequences., Nucleic Acids Res. 44 (2016) W11–5. doi:10.1093/nar/gkw265.

[50] F.-O. Desmet, D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres, C. Beroud, Human Splicing Finder: an online bioinformatics tool to predict splicing signals., Nucleic Acids Res. 37 (2009) e67. doi:10.1093/nar/gkp215.

[51] S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W.W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H.L. Rehm, Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology., Genet. Med. 17 (2015) 405–424. doi:10.1038/gim.2015.30.

[52] Q. Li, K. Wang, InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines, Am. J. Hum. Genet. 100 (2017) 267–280. doi:10.1016/j.ajhg.2017.01.004.

[53] J.T. den Dunnen, R. Dalgleish, D.R. Maglott, R.K. Hart, M.S. Greenblatt, J. McGowan-Jordan, A.-F. Roux, T. Smith, S.E. Antonarakis, P.E.M. Taschner, HGVS Recommendations for the Description of Sequence Variants: 2016 Update., Hum. Mutat. 37 (2016) 564–569. doi:10.1002/humu.22981.

[54] M. Schraders, J. Oostrik, P.L.M. Huygen, T.M. Strom, E. van Wijk, H.P.M. Kunst, L.H. Hoefsloot, C.W.R.J. Cremers, R.J.C. Admiraal, H. Kremer, Mutations in PTPRQ are a cause of autosomal-recessive nonsyndromic hearing impairment DFNB84 and associated with vestibular dysfunction., Am. J. Hum. Genet. 86 (2010) 604–610. doi:10.1016/j.ajhg.2010.02.015.

[55] L.M. Astuto, J.M. Bork, M.D. Weston, J.W. Askew, R.R. Fields, D.J. Orten, S.J. Ohliger, S. Riazuddin, R.J. Morell, S. Khan, S. Riazuddin, H. Kremer, P. van Hauwe, C.G. Moller, C.W.R.J. Cremers, C. Ayuso, J.R. Heckenlively, K. Rohrschneider, U. Spandau, J. Greenberg, R. Ramesar, W. Reardon, P. Bitoun, J. Millan, R. Legge, T.B. Friedman, W.J. Kimberling, CDH23 Mutation and Phenotype Heterogeneity: A Profile of 107 Diverse Families with Usher Syndrome and Nonsyndromic Deafness, Am. J. Hum. Genet. 71 (2002) 262–275. doi:10.1086/341558.

[56] M.J. Aparisi, G. Garcia-Garcia, E. Aller, M.D. Sequedo, C. Martinez-Fernandez de la Camara, R. Rodrigo, M. Armengot, J. Cortijo, J. Milara, M. Diaz-LLopis, T. Jaijo, J.M. Millan, Study of USH1 splicing variants through minigenes and transcript analysis from nasal epithelial cells., PLoS One. 8 (2013) e57506. doi:10.1371/journal.pone.0057506.

[57] L.M. Astuto, J.M. Bork, M.D. Weston, J.W. Askew, R.R. Fields, D.J. Orten, S.J. Ohliger, S. Riazuddin, R.J. Morell, S. Khan, S. Riazuddin, H. Kremer, P. van Hauwe, C.G. Moller, C.W.R.J. Cremers, C. Ayuso, J.R. Heckenlively, K. Rohrschneider, U. Spandau, J. Greenberg, R. Ramesar, W. Reardon, P. Bitoun, J. Millan, R. Legge, T.B. Friedman, W.J. Kimberling, CDH23 Mutation and Phenotype Heterogeneity: A Profile of 107 Diverse Families with Usher Syndrome and Nonsyndromic Deafness, Am. J. Hum. Genet. 71 (2002) 262–275. doi:10.1086/341558.

[58] L.A. Farrer, K.M. Grundfastj, J. Amos, K.S. Arnos, J.H. Asher, P. Beighton, S.R. Diehl, T.T.J. Fex, C. Foy, T.B. Friedman, J. Greenberg, C. Hotht, M. Marazita, A. Milunskyj, R. Morell, W. Nance, V. Newton, R. Ramesar, T.B. San, A. James, S. Cathy, R.G. Wagner, E.R. Wilcox, I. Winship, A.P. Readh, Waardenburg Syndrome (WS) Type I Is Caused by Defects at Multiple Loci , One of Which Is Near ALPP on Chromosome 2 : First Report of the WS Consortium, Am. J. Hum.

Genet. 50 (1992) 902–913.

[59] F. Cesca, E. Bettella, R. Polli, E. Cama, P. Scimemi, R. Santarelli, A. Murgia, A novel mutation of the EYA4 gene associated with post-lingual hearing loss in a proband is co-segregating with a novel PAX3 mutation in two congenitally deaf family members, Int. J. Pediatr. Otorhinolaryngol. 104 (2018) 88–93. doi:10.1016/j.ijporl.2017.10.042.

[60] D. Yan, X.-Z. Liu, Cochlear molecules and hereditary deafness., Front. Biosci. 13 (2008) 4972–4983.

[61] T. Hutchin, N.N. Coy, H. Conlon, E. Telford, K. Bromelow, D. Blaydon, G. Taylor, E. Coghill, S. Brown, R. Trembath, X.Z. Liu, M. Bitner-Glindzicz, R. Mueller, Assessment of the genetic causes of recessive childhood non-syndromic deafness in the UK - implications for genetic testing., Clin. Genet. 68 (2005) 506–512. doi:10.1111/j.1399-0004.2005.00539.x.

[62] A.E. Shearer, R.J.H. Smith, Massively Parallel Sequencing for Genetic Diagnosis of Hearing Loss, Otolaryngol. Neck Surg. 153 (2015) 175–182. doi:10.1177/0194599815591156.

[63] I. Schrauwen, M. Sommen, J.J. Corneveaux, R. a. Reiman, N.J. Hackett, C. Claes, K. Claes, M. Bitner-Glindzicz, P. Coucke, G. Van Camp, M.J. Huentelman, A sensitive and specific diagnostic test for hearing loss using a microdroplet PCR-based approach and next generation sequencing, Am. J. Med. Genet. Part A. 161 (2013) 145–152. doi:10.1002/ajmg.a.35737.

[64] C.M. Sloan-Heggen, A.O. Bierer, A.E. Shearer, D.L. Kolbe, C.J. Nishimura, K.L. Frees, S.S. Ephraim, S.B. Shibata, K.T. Booth, C.A. Campbell, P.T. Ranum, A.E. Weaver, E.A. Black-Ziegelbein, D. Wang, H. Azaiez, R.J.H. Smith, Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss, Hum. Genet. 135 (2016) 441–450. doi:10.1007/s00439-016-1648-8.

[65] K. Mizutari, H. Mutai, K. Namba, Y. Miyanaga, A. Nakano, Y. Arimoto, S. Masuda, N. Morimoto, H. Sakamoto, K. Kaga, T. Matsunaga, High prevalence of CDH23 mutations in patients with congenital high-frequency sporadic or recessively inherited hearing loss., Orphanet J. Rare Dis. 10 (2015) 60. doi:10.1186/s13023-015-0276-z.

[66] M. Miyagawa, S. Nishio, S. Usami, Prevalence and clinical features of hearing loss patients with CDH23 mutations: a large cohort study., PLoS One. 7 (2012) e40366. doi:10.1371/journal.pone.0040366.

[67] E. VanBeelen, A.M.M. Oonk, J.M. Leijendeckers, E.H. Hoefsloot, R.J.E. Pennings, I. Feenstra, H. Dieker, P.L.M. Huygen, A.F.M. Snik, H. Kremer, H.P.M. Kunst, Audiometric Characteristics

of a Dutch DFNA10 Family With Mid-Frequency Hearing Impairment, Ear Hear. 37 (2016) 103–111. doi:10.1097/AUD.0000000000000217.

[68] M. Verstreken, F. Declau, I. Schatteman, D. Van Velzen, K. Verhoeven, G. Van Camp, P.J. Willems, E.W. Kuhweide, E. Verhaert, P. D'Haese, F.L. Wuyts, P.H. Van de Heyning, Audiometric analysis of a Belgian family linked to the DFNA10 locus., Am. J. Otol. 21 (2000) 675–681.

Appendix 1

Table 1. List of syndromic forms of hearing loss

Syndrome	Gene	Protein	OMIM
Alport Syndrome	*COL4A5*	Collagen, type IV, alpha 5	303630
	COL4A3	Collagen, type IV, alpha 3	120070
	COL4A4	Collagen, type IV, alpha 4	120131
Branchio-oto-renal Syndrome	*EYA1*	Eyes absent homolog 1	601653
	SIX5	Homeobox protein SIX5	600963
	SIX1	Homeobox protein SIX1	601205
CHARGE Syndrome	*SEMA3E*	Semaphorin 3E	608166
	CHD7	Chromodomain-helicase-DNA-binding protein 7	608892
Jervell&Lange-Nielsen Syndrome	*KCNQ1*	Potassium channel, voltage gated KQT-like subfamily Q, member 1	607542
	KCNE1	Potassium channel, voltage gated subfamily E regulatory beta subunit 1	176261
Norrie Disease	*NDP*	Norrie disease protein	300658
Pendred Syndrome	*SLC26A4*	Pendrin	605646
	FOXI1	Forkhead box protein I1 ATP-sensitive inward rectifier	601093
	KCNJ10	Potassium channel 10	602208
Stickler Syndrome	*COL2A1*	Collagen, type II, alpha-1	120140
	COL11A1	Collagen, type XI, alpha-1	120280
	COL11A2	Collagen, type XI, alpha-2	120290
	COL9A1	Collagen, type IX, alpha-1	120210
	COL9A2	Collagen, type IX, alpha-2	120260
Treacher Collins Syndrome	*TCOF1*	Treacher Collins-Franceschetti syndrome 1	606847
	POLR1D	Polymerase I Polypeptide D	613715
	POLR1C	Polymerase I Polypeptide C	610060
Usher Syndrome	*MYO7A*	Myosin VIIA	276903
	USH1C	Harmonin	605242
	CDH23	Cadherin 23	605516
	PCDH15	Protocadherin 15	605514
	SANS	Scaffold protein containing ankyrin repeats and sam domain	607696
	CIB2	Calcium and integrin binding protein 2	605564
	USH2A	Usherin	608400
	VLGR1	Very large G-coupled protein receptor isoform b	602851
	WHRN	Whirlin	607928
	CLRN1	Clarin 1	606397
	HARS	Histidyl tRNA synthetase	142810
	PZDZ7	PDZ domain containing 7	NA
	CEP250	250	NA
Waardenburg Syndrome	*PAX3*	Paired box 3	606567
	SNAI2	Snailhomolog 2	602150

	EDN3	Endothelin 3	131242
	EDNRB	Endothelin receptor type B	131244
	MITF	Microphthalmia-associated transcription factor	156845
	SOX10	SRY box10	602229
Perrault Syndrome	*HSD17B4*	Hydroxysteroid (17-beta) dehydrogenase 4	601860
	HARS2	Histidyl-TRNA Synthetase 2	600783
	CLPP	Caseinolytic mitochondrial matrix peptidase proteolytic subunit	601119
	LARS2	Leucyl-tRNA synthetase2	604544

Genes and proteins involved in SHL with specific OMIM number. NA: not available.

Appendix 2

Table 2. List of genes related to NSHL

Mode of inheritance	Locus	Gene	Protein
AD	NA	CRYM	Crystallin, Mu
AD	DFNA1	DIAPH1	Diaphanous 1
AD	DFNA2A	KCNQ4	KCNQ4
AD	DFNA2B	GJB3 (Cx31)	Connexin 31
AD	DFNA3A	GJB2 (Cx26)	Connexin 26
AD	DFNA3B	GJB6 (Cx30)	Connexin 30
AD	DFNA4	MYH14	Nonmuscle myosin heavy chain XIV
AD	DFNA4	CAECAM16	Carcinogenic antigen-related cell adhesion molecule 16
AD	DFNA5	DFNA5	Gasdermin E
AD	DFNA6/14/38	WFS1	Wolframin
AD	DFNA8/12	TECTA	A-tectorin
AD	DFNA9	COCH	Cochlin
AD	DFNA10	EYA4	Eyes absent 4
AD	DFNA11	MYO7A	Myosin VIIa
AD	DFNA13	COLL11A2	Type XI collagen α2
AD	DFNA15	POU4F3	Class 3 POU
AD	DFNA17	MYH9	Nonmuscle myosin heavy chain IX
AD	DFNA20/26	ACTG1	γ-actin
AD	DFNA22	MYO6	Myosin VI
AD	DFNA23	SIX1	Sine Oculis Homeobox, Drosophila, homolog of,1
AD	DFNA25	SLC17A8	VGLUT-3
AD	DFNA28	GRHL2	Grainyhead-like 2
AD	DFNA36	TMC1	Transmembrane channel-like protein 1
AD	DFNA41	P2RX2	Purinergic receptor P2X ligand-gated ion channel 2
AD	DFNA44	CCDC50	Coiled-coil domain-containing protein 50
AD	DFNA50	MIR96	MicroRNA96
AD	DFNA51	TJP2	Tight junction protein 2
AD	DFNA56	TNC	Tenascin-C
AD	DFNA64	SMAC/DIABLO	Second Mitochondria-Derived Activator of Caspase/Direct Inhibitor of Apoptosis protein Binding protein with a low pI
AD	DFNA65	TBC1D24	Tbc1 domain family, member 24
AD	DFNA66	CD164	Cd164 Antigen
AD	DFNA67	OSBPL2	Oxysterol-binding Protein-like Protein 2
AD	NA	HOMER2	Homer, drosophila, homolog of 2
AD	NA	MCM2	Minichromosome maintenance complex component
AD	NA	KITLG	Kit ligand
AD	NA	DMXL2	Dmx-like 2
AR	DFNB1A	GJB2 (Cx26)	Connexin 26
AR	DFNB1B	GJB6 (Cx30)	Connexin 30
AR	DFNB2	MYO7A	Myosin VIIa
AR	DFNB3	MYO15A	Myosin Xva
AR	DFNB4	SLC26A4	Pendrin
AR	DFNB6	TMIE	Transmembrane inner ear-expressed gene
AR	DFNB7/11	TMC1	Transmembrane channel-like protein 1
AR	DFNB8/10	TMPRSS3	Transmembrane protease, serine 3

AR	DFNB9	OTOF	Otoferlin
AR	DFNB12	CDH23	Cadherin 23
AR	DFNB15/72/95	GIPC3	GIPC PDZ domain-containing family member 3
AR	DFNB16	STRC	Stereocilin
AR	DFNB18	USH1C	Harmonin
AR	DFNB18B	OTOG	Otogelin
AR	DFNB21	TECTA	α-tectorin
AR	DFNB22	OTOA	Otoancorin
AR	DFNB23	PCDH15	Protocadherin 15
AR	DFNB24	RDX	Radixin
AR	DFNB25	GRXCR1	Glutaredoxin, cysteine-rich 1
AR	DFNB28	TRIOBP	Trio-binding protein
AR	DFNB29	CLDN14	Claudin 14
AR	DFNB30	MYO3A	Myosin IIIa
AR	DFNB31	WHRN	Whirlin
AR	DFNB35	ESRRB	Oestrogen-related receptor β
AR	DFNB36	ESPN	Espin
AR	DFNB37	MYO6	Myosin VI
AR	DFNB39	HGF	Hepatocyte growth factor
AR	DFNB42	ILDR1	Immunoglobulin-like domain containing receptor 1
AR	DFNB44	ADCY1	Adenylate cyclase 1
AR	DFNB48	CIB2	Calcium and integrin binding protein 2
AR	DFNB49	MARVELD2	Tricellulin
AR	DFNB49	BDP1	B-double prime 1, subunit of RNA polymerase III transcription initiation factor
AR	DFNB53	COL11A2	Type XI collagen α2
AR	DFNB59	PJVK	DFNB59 gene
AR	DFNB60	SLC22A4	Solute carrier family 22 (organic cation transporter)
AR	DFNB61	SLC26A5	Prestin
AR	DFNB63	LRTOMT/COMT2	Leucine rich transmembrane o-methlytransferase
AR	DFNB67	LHFPL5	Tetraspan membrane protein
AR	DFNB68	S1PR2	Sphingosine -1 phosphatase receptor 2
AR	DFNB70	PNPT1	Polyribonucleotide nucleotidytransferase 1
AR	DFNB73	BSND	Barttin
AR	DFNB74	MSRB3	Methionine sulfoxide reductase B3
AR	DFNB76	SYNE4	Spectrin repeat-containing nuclear envelope protein 4
AR	DFNB77	LOXHD1	Lypoxygenase homology domain containing 1
AR	DFNB79	TPRN	Taperin
AR	DFNB82	GPSM2	G-protein signaling modulator 2
AR	DFNB84	PTPRQ	Protein tyrosine phosphate receptor Q
AR	DFNB84	OTOGL	Otogelin like protein
AR	DFNB86	TBC1D24	Tbc1 domain family, member 24
AR	DFNB88	ELMOD3	Elmo/CED12 domain-containing protein 3
AR	DFNB89	KARS	Lysyl-tRNA synthetase
AR	DFNB91	SERPINB6	Serpin Peptidase inhibitor clade B, member 6
AR	DFNB93	CABP2	Calcium-binding protein 2
AR	DFNB94	NARS2	Asparaginyl-tRNA synthetase 2
AR	DFNB97	MET	MET protooncogene
AR	DFNB98	TSPEAR	Thrombospondin-type laminin g domain and ear repeats
AR	DFNB99	TMEM132E	Transmembrane protein 132e
AR	DFNB101	GRXCR2	Glutaredoxin, cysteine-rich 2

AR	DFNB102	EPS8	Epidermal growth factor receptor pathway substrate 8
AR	DFNB103	CLIC5	Chloride intracellular channel 5
AR	DFNB105	CDC14A	Cell division cycle 14, S. Cerevisiae homolog A
AR	NA	FAM65B	Family with sequence similarity 65, member b
AR	NA	EPS8L2	EPS8-like protein 2
AR	NA	WBP2	WW Domain binding protein 2
AR	NA	ROR1	Receptor tyrosine kinase-like orphan receptor 1
X-linked	DFNX1	PRPS1	Phosphoribosylpyrophosphate synthetase 1
X-linked	DFNX3	POU3F4	Class 3 POU
X-linked	DFNX4	SMPX	Small muscle protein X-linked
X-linked	DFNX5	AIFM1	Apoptosis-inducing factor, mitochondria-associated, 1
X-linked	DFNX6	COL4A6	Collagen, type IV, alpha-6

List of NSHL genes from Hereditary Hearing Loss Homepage. (http://hereditaryhearingloss.org/; last accessed October 2017). AD: autosomal dominant; AR: autosomal recessive; NA: not available.

Appendix 3

International Journal of Pediatric Otorhinolaryngology 104 (2018) 88–93

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

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

A novel mutation of the *EYA4* gene associated with post-lingual hearing loss in a proband is co-segregating with a novel *PAX3* mutation in two congenitally deaf family members

Federica Cesca[a], Elisa Bettella[a], Roberta Polli[a], Elona Cama[b,c], Pietro Scimemi[b,c], Rosamaria Santarelli[b,c], Alessandra Murgia[a,c,*]

[a] Laboratory of Molecular Genetics of Neurodevelopment, Department of Women's and Children's Health, University of Padova, Italy
[b] Audiology and Phoniatrics Service, Treviso Regional Hospital, Italy
[c] Neuroscience Department, University of Padova, Italy

ARTICLE INFO

Keywords:
EYA4
Non-syndromic hearing loss
PAX3
Waardenburg syndrome type 1
Next Generation Sequencing

ABSTRACT

Objectives: This work was aimed at establishing the molecular etiology of hearing loss in a 9-year old girl with post-lingual non-syndromic mild sensorineural hearing loss with a complex family history of clinically heterogeneous deafness.

Methods: The proband's DNA was subjected to NGS analysis of a 59-targeted gene panel, with the use of the Ion Torrent PGM platform. Conventional Sanger sequencing was used for segregation analysis in all the affected relatives. The proband and all the other hearing impaired members of the family underwent a thorough clinical and audiological evaluation.

Results: A new likely pathogenic mutation in the *EYA4* gene (c.1154C > T; p.Ser385Leu) was identified in the proband and in her 42-year-old father with post-lingual non-syndromic profound sensorineural hearing loss. The *EYA4* mutation was also found in the proband's grandfather and uncle, both showing clinical features of Waardenburg syndrome type 1. A novel pathogenic splice-site mutation (c.321+1G > A) of the *PAX3* gene was found to co-segregate with the *EYA4* mutation in these two subjects.

Conclusion: The identified novel *EYA4* mutation can be considered responsible of the hearing loss observed in the proband and her father, while a dual molecular diagnosis was reached in the relatives co-segregating the *EYA4* and the *PAX3* mutations. In these two subjects the DFNA10 phenotype was masked by Waardenburg syndrome.

The use of NGS targeted gene-panel, in combination with an extensive clinical and audiological examination led us to identify the genetic cause of the hearing loss in members of a family in which different forms of autosomal dominant deafness segregate. These results provide precise and especially important prognostic and follow-up information for the future audiologic management in the youngest affected member.

1. Introduction

Hearing loss (HL) is the most common birth defect and sensory disorder in the industrialized countries, with a prevalence of two to four in 1000 newborns [1]. It can be determined by environmental causes (trauma, medications, medical problems, environmental exposure) or by genetic causes (50%).

Hereditary hearing loss (HHL) can be syndromic hearing loss (SHL) in about 25% of cases, in which deafness is associated with other signs and/or symptoms, and non-syndromic hearing loss, or isolated (NSHL) in 75% of cases. Non-syndromic hearing loss generally follows simple Mendelian inheritance and is more frequently autosomal recessive (75–80%) than autosomal dominant (20%) or X-linked (2–5%), and in few cases (1%) NSHL is due to mitochondrial mutations [2].

Hereditary hearing loss is characterized by a vast genetic heterogeneity with 160 loci described in humans but only 100 genes so far identified (http://hereditaryhearingloss.org/).

The *GJB2* gene (Cx26) located at the DFNB1 locus plays a major pathogenic role worldwide, with a frequency of mutant alleles that varies between countries and even between regions of the same country. In Italy about 50% of preverbal non-syndromic autosomal recessive HL cases are due to *GJB2* mutations [3]. The remaining cases,

* Corresponding author. Department of Women's and Children's Health, University of Padova, Via Giustiniani 3, 35128, Padova, Italy.
E-mail address: alessandra.murgia@unipd.it (A. Murgia).

https://doi.org/10.1016/j.ijporl.2017.10.042
Received 1 August 2017; Received in revised form 25 October 2017; Accepted 27 October 2017
Available online 31 October 2017

without an etiological explanation, still need to be characterized at the molecular level.

Testing all NSHL associated genes, and the most common forms of SHL, Usher and Pendred/EVA syndrome genes [4] by Sanger sequencing has so far been hampered by technical limits, but the advent of Next Generation Sequencing (NGS) technologies has made possible to simultaneously sequence multiple genes in a single reaction, which couples an unprecedented diagnostic capacity with significant reductions in time and costs for genetic testing.

EYA4, at the DFNA10 locus, was first identified in a large American family as a causal gene for autosomal dominant NSHL [5]. Hearing impairment associated with *EYA4* mutations is characterized by post-lingual onset and progressive course (DFNA10 OMIM #601316) [6]; in only one case reported in literature the hearing defect has been described in association with dilatative cardiomyopathy [7].

The *EYA4* gene is composed of 20 exons encoding a 639 amino-acid protein which includes a highly conserved 271 amino acid carboxy-terminus called the eya-homologous region (eyaHR) and a phylogenetically more divergent proline-serine-threonine (PST)-rich transactivation domain at the amino-terminus, called eya-variable region (eyaVR) [8].

Waardenburg syndrome (WS) is a clinically and genetically heterogeneous condition primarily characterized by congenital HL and pigmentation anomalies in eyes, skin, and hair; its estimated frequency is 1/40,000 in the general population [9]. Four clinical types of WS have been described (WS1–WS4). The most frequent phenotypes, WS type 1 (WS1, OMIM #193500) and type 2 (WS2, OMIM #193510) are distinguished by dystopia canthorum present only in WS1. Six disease associated genes (*PAX3, MITF, SNAI2, EDN3, EDNRB, SOX10*) have so far been involved in this syndrome [10].

WS1 is exclusively caused by mutation in the *PAX3* gene (OMIM #606597), localized on chromosome 2q35 and composed of 10 exons. The 479 amino acid PAX3 protein contains two DNA binding domains (the paired domain and a homeo-domain), a conserved octapeptide located in between (mediating protein–protein interactions), and a Ser/Thr/Pro-rich carboxy-terminal transactivation domain [11].

In this study, we report the identification of a new likely pathogenic mutation of the *EYA4* gene in a 9-year old girl, with mild sensorineural post-lingual hearing loss, and member of a large family of hearing impaired individuals with reportedly wide phenotypic differences. The *EYA4* mutation was transmitted by her profoundly deaf father with a history of non-syndromic post-lingual progressive hearing loss. A dual molecular diagnosis was instead obtained in the paternal grandfather and uncle in whom the *EYA4* mutation was found to co-segregate with a novel *PAX3* splicing mutation, causing Waardenburg syndrome type 1, that masked DFNA10 phenotype due to *EYA4* mutation.

2. Material and methods

2.1. Subjects and phenotype

A three-generation family with six members affected by hearing loss with different clinical features were subjected to thorough audiological evaluations.

A complete clinical evaluation for identification of other associated clinical features was performed.

All the studied individuals provided written informed consent for molecular genetic testing. The work was conducted according to the ethical standards as defined by the Helsinki Declaration and according to indications from the local institutional ethics committee.

2.2. Audiometry

Hearing thresholds were tested at frequencies from 250 to 8000 Hz (Grason-Stadler GSI 61 audiometer) in a sound-attenuating room. The degree of hearing impairment was defined by the pure tone average (PTA) threshold levels at 0.5, 1, 2 and 4 kHz, and was classified as mild (PTA 21–40 dB HL), moderate (PTA 41–70 dB HL), severe (PTA 71–95 dB HL) and profound (PTA > 95 dB HL) [12].

2.3. Molecular characterization

Genomic DNA (gDNA) was extracted from peripheral blood leukocytes with Maxwell automatic extractor (PROMEGA Madison, WI USA), according to the manufacturer's instructions. DNA quantification was performed with Qubit Fluorometer using Qubit dsDNA HS Assay Kit (ThermoFisher SCIENTIFIC Waltham, MA USA).

The *GJB2* gene (NCBI reference sequences RefSeq NG_008358.1) was pre-screened by direct Sanger sequencing and *GJB6* known deletions were tested as previously reported [13].

2.4. Hearing loss targeted NGS panel

The gene-panel, which was developed to mainly target genes responsible of non-syndromic deafness or hearing loss with non-syndromic onset, comprised a total of 59 genes of which 37 autosomal recessive (AR), 18 autosomal dominant (AD) and 4 X-linked genes. Among AR genes, the panel included one Pendred (*SLC26A4*) and 11 Usher syndrome genes; the autosomal dominant auditory neuropathy gene (*DIAPH3*) was also included. The entire gene coding region and 10 exon-flanking intronic bases were analyzed.

Ion AmpliSeq Designer v.4.4.1 software (ThermoFisher SCIENTIFIC Waltham, MA USA) was used to select primers for preparation of DNA libraries. A total of 1646 amplicons divided in two pools were obtained. The amplicon size-range is 125–375 bp (228 bp mean). The designed gene-panel ensured overall targeted region coverage of 98.6% spanning through 456.8 Kb of sequence.

2.5. Library preparation and Next Generation Sequencing protocol

The libraries were constructed using 10 ng of gDNA, amplified with Ion AmpliSeq DNA and RNA Library Kit 2.0 (ThermoFisher SCIENTIFIC Waltham, MA USA) and indexed with Ion Xpress Barcode Adapters Kit (ThermoFisher SCIENTIFIC Waltham, MA USA) according to the manufacturer's instructions. The template and the enrichment preparation were performed using the Ion PGM Hi-Q One Touch 2 System and Ion PGM Hi-Q OT2 chemistry kit. Single-end 400 base-read sequencing runs were carried out with the use of Ion PGM Hi-Q sequencing kit with Ion 316 Chip v2 on the Ion PGM system (ThermoFisher SCIENTIFIC Waltham, MA USA).

2.6. NGS data analysis

Sequencing data were processed using the Ion Torrent Suite v4.4; variant calling was performed by the Variant Caller plug-in v4.4. Single Nucleotide Variants (SNV) and small insertion/deletions (INDELs) were then annotated using the free web tool wANNOVAR [14]. Reads alignment was visualized using Integrative Genome Viewer (IGV) software.

The variants were filtered based on: i) minor allele frequency (MAF) in healthy individuals (< 1%) as reported in public database (dbSNP, 1000genomes, ESP6500, ExAC, gnomAD) to exclude possible polymorphisms. ii) depth coverage higher than 20X. iii) type of molecular alteration (frameshift variants, non-synonymous SNV, stop-gain mutation, splice-site mutations).

The pathogenicity of missense mutation was evaluated with a cluster of *in silico* predictor tools: SIFT [15], Polyphen-2 [16], MutationTaster [17], MutationAssessor [18], LRT [19], FATHMM [20], RadialSVM, LR, PhyloP [21] GERP++ [22], SiPhy [23], CADD phred [24].

ACMG guidelines [25] and InterVar database [26] were used to determine the pathogenicity of the mutations, which were named

according to HGVS recommended nomenclature. NCBI reference sequences RefSeq NM_004100 and NP_004091.3 were used for the numbering of the *EYA4* mutation and NCBI reference sequences RefSeq NM_181459.3 for the numbering of the *PAX3* mutation.

The prediction of splice-site mutations was performed with Human splicing finder [27].

The possible pathogenic variants were Sanger validated and family segregation analysis was performed.

2.7. *EYA4* and *PAX3* Sanger sequencing analysis

Exon 13 of the *EYA4* gene (OMIM #601316; NCBI reference sequences RefSeq NG_011596.1), containing the likely pathogenic variant, was amplified with AmpliTaq Gold (ThermoFisher SCIENTIFIC Waltham, MA USA) according to manufacturing instruction with designed ad hoc primers.

The entire coding sequence of the *PAX3* gene (OMIM #606597; NCBI reference sequences RefSeq NG_011632.1) and exon-intron junctions were PCR-amplified from gDNA using designed ad hoc primers.

PCR products were purified using ExoSAP-IT TM (ThermoFisher SCIENTIFIC Waltham, MA USA) and sequencing reactions were performed with BigDye® Terminator v3.1 kit according to manufacturer's instruction. The sequencing products were analyzed with 3130 Genetic Analyzer capillary sequencer (ThermoFisher SCIENTIFIC Waltham, MA USA).

3. Results

The proband, a 9-year-old Caucasian girl (III-1) was referred to our department for an etiological evaluation of her sensorineural hearing loss. Her family history was positive for deafness as reported in Fig. 1, with both congenital and post-lingual cases reported. No cardiac problems were present in the family.

Pure tone audiometry revealed mild hearing loss, which was reported and documented as being post-lingual.

GJB2 analysis on the proband's gDNA revealed the presence of a heterozygous c.35delG mutation, transmitted from her congenitally deaf homozygote mother (II-1), while the father tested negative.

NGS analysis with the 59-targeted gene-panel described in section 2.4 was performed on the gDNA of the proband. The bioinformatic analysis finally highlighted a novel heterozygous mutation in exon 13 of the *EYA4* gene: NM_004100.4: c.1154C > T (NP_004091.3: p.Ser385Leu). This variant wasn't present in ExAC database (http://exac.broadinstitute.org/; last accessed September 2017) or in gnomAD browser (http://gnomad.broadinstitute.org/; last accessed September 2017). It was never described as associated with hearing phenotype nor it was present in deafness variation database (http://deafnessvariationdatabase.org; last accessed September 2017).

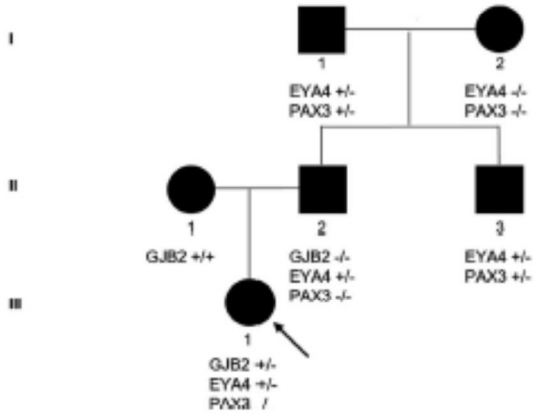


Fig. 1. Family pedigree. The arrow indicates the proband (III-1). Square indicates males and circle female individuals. Solid symbol represents hearing-impaired individuals. +/− heterozygous; +/+ homozygous; −/− wild-type.

All the utilized *in silico* tools, as reported in section 2.6, predicted this alteration as damaging since substituting a polar aminoacid with a non-polar one in a highly phylogenetically conserved residue (CADD phred: 35) within the functionally relevant eya-homologous region (eyaHR) as reported in Fig. 2.

Segregation analysis was performed on gDNA samples of the proband and other hearing-impaired family members, who also underwent a complete audiological and clinical phenotype evaluation.

The proband's grandmother (I-2) was a 67-year-old woman affected by profound hearing loss with post-lingual onset and progressive course. She also reported unspecified familial HL, although no other affected members of her family were available for the study. The 74-year-old grandfather (I-1) had congenital profound hearing loss, brilliant blue eyes, no facial dysmorphic features and was referred as having precocious white hair with a white forelock. The 39-year-old uncle (II-3) had congenital profound hearing loss, brilliant blue eyes, facial dysmorphisms (dystopia canthorum, downslanting palpebral fissures and malar hypoplasia) and pigmentary alterations, among which partially white eyebrows and early white hair with referred white forelock. These individuals (I-1 and II-3) who clearly met major diagnostic criteria for Waardenburg syndrome type 1 according to the WS Consortium criteria [28], had not been previously diagnosed or offered appropriate molecular testing.

Based on the clinical evaluations we proceeded with *PAX3* gene testing in the two subjects with WS1 phenotype (I-1 and II-3) together with segregation analysis of the *EYA4* variant. Comprehensive molecular and clinical data are summarized in Table 1.

The proband (III-1), her father (II-2), her uncle (II-1) and grandfather (I-1) were all found to be heterozygote for the *EYA4* mutation, while the grandmother (I-2), was negative.

The entire coding sequence of the *PAX3* gene was analyzed in individuals I-1 and II-3 with Sanger sequencing, as described previously. A novel splice-site mutation NM_ 181459: c.321 + 1G > A was identified in these subjects. The mutation, never previously described in literature, nor reported in the public databases as ExAC database (http://exac.broadinstitute.org/; last accessed September 2017) or gnomAD browser (http://gnomad.broadinstitute.org/; last accessed September 2017). This splicing mutation was predicted by Human splicing finder to alter the splice-donor site of *PAX3* intron 2 as reported in Fig. 3.

4. Discussion

Hearing loss is a highly heterogeneous condition, challenging both from a clinical and genetic stand point. The cause of a non-syndromic sensorineural deafness can be extremely difficult to investigate based on solely clinical data, especially when the hearing loss is pre-lingual and profound. A deafness syndrome can be caused by alterations in different genes and different mutations in the same gene may result both in SHL and NSHL [29]. Finally, a specific genetic cause may present with extremely variable phenotypic expression, even within members of the same family [30], and often the same gene can sustain both dominant and recessive conditions.

Furthermore, the occurrence of assortative or non-random mating by deafness state (i.e., a condition in which deaf individuals partner together) is well-documented and may cause the segregation of two or more distinct forms of deafness in the same nuclear family [31]. Assortative mating represents a further challenge for the effort of disentangling complicate patterns of transmission in deafness families with many affected members.

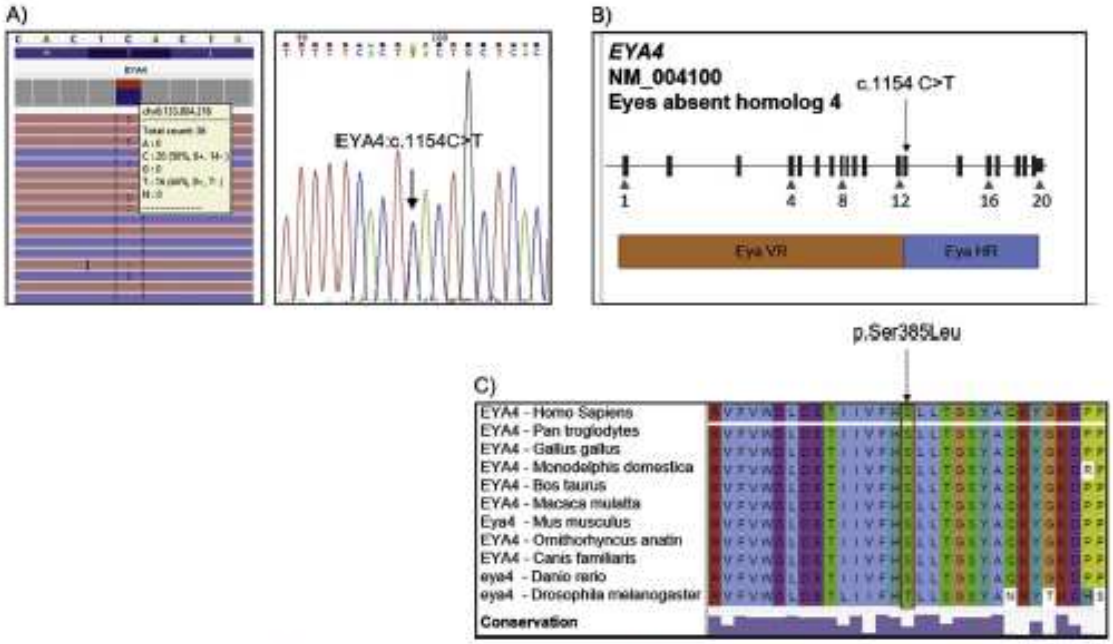


Fig. 2. A, on the left IGV visualization; on the right Sanger sequencing electropherogram. The arrow indicates the c.1154C > T mutation. B, On the top *EYA4* schematic exon representation and on the bottom protein domains: EyaVR (Eya-variable region) and EyaHR (Eya-homologous region). The *EYA4* mutation is localized in exon 13 on the EYA HR domain. C, Multiple alignments of the *EYA4* protein aminoacid sequence among species. Ser385 is squared in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In this study, we describe a family representing a classical example of assortative mating, in which we identified novel mutations of different deafness genes, causing both syndromic and non-syndromic autosomal dominant HL.

We were able to make a dual diagnosis since we found that these mutations co-occurred in two affected relatives who showed a phenotype of Waardenburg syndrome type 1, clearly masking the expected DFNA10 audiological profile.

The novel *EYA4* mutation detected by targeted NGS approach in the proband (III-1) could be responsible for the type of hearing loss she presented and, with obvious age differences, seemed to be typical of the audiological history reported for her father. According to ACMG guidelines and to InterVar database this mutation could be defined as "likely pathogenic".

Until now only 7 DFNA10 families have been reviewed in literature [32] and variable audiogram configurations have been described. Most of the subjects of these families carrying heterozygous missense mutations in the *EYA4* gene showed bilateral, progressive sensorineural hearing loss with variable age of onset ranging between the prelingual period to slightly over 50 years. DFNA10 typically involves mid frequencies at onset and progresses thereafter to affect all the frequencies [33].

At our clinical evaluation, the 42-year-old proband's father, whose hearing loss was post-lingual, with referred onset before the age of 10, was found to be profoundly deaf. The proband, 9-year-old at the time of evaluation, also had post-lingual hearing loss, that was measured as mild and still prevalently involving mid frequencies. Since one of the main features of the DFNA10 phenotype is progression, identification of the *EYA4* mutation had a very relevant impact both for the prognosis and for the program of clinical follow-up in this girl.

We identified a novel *PAX3* splice-site mutation in the affected uncle (II-3) and grandfather (I-1) of the proband, both of whom had WS1 phenotype never previously investigated.

Heterozygous mutations in the *PAX3* gene are found either as de novo or as dominantly inherited alterations, showing a very variable phenotypic expression as in the case of the family described in this work. WS1 is thought to be fully penetrant when considering at least one sign, although penetrance of each feature may be not complete [10]. The individuals we describe had in fact both congenital profound hearing loss and brilliant blue eyes, but facial dysmorphisms were only

Table 1
Summary of phenotypic features and genotypes.

Pedigree position	Gender	Age	Phenotype			Genotype	
			HL Onset	Severity	WS signs	*EYA4*	*PAX3*
I-1	M	74	congenital	profound	blu eyes	c.1154C > T; p.Ser385Leu	c.321 + 1G > A
I-2	F	67	post-lingual	profound	//	//	//
II-2	M	42	post-lingual	profound	//	c.1154C > T; p.Ser385Leu	//
II-3	M	39	congenital	profound	blu eyes, dystopia canthorum	c.1154C > T; p.Ser385Leu	c.321 + 1G > A
III-1	F	9	post-lingual	mild	//	c.1154C > T; p.Ser385Leu	//

M: male; F: female. Age of the subjects at evaluation. WS indicates Waardenburg syndrome. The symbol // indicates absence of WS clinical signs or mutation.

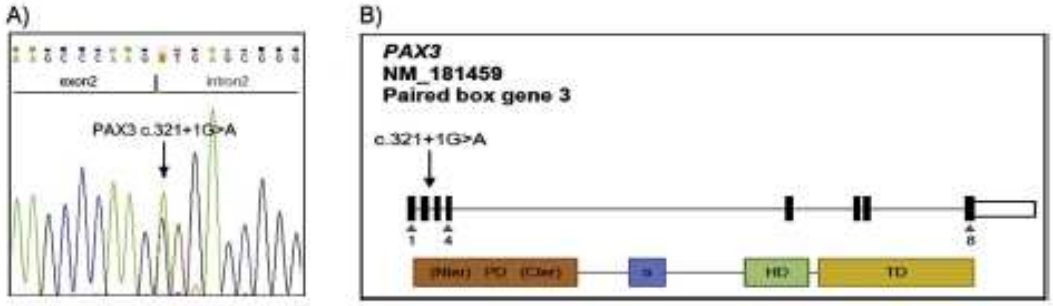


Fig. 3. A, Electropherogram of exon/intron 2 of *PAX3* gene. The arrow indicates the c.321 + 1G > A splice-site mutation. B, On the top *PAX3* schematic exon representation and on the bottom protein domains: PD (paired domain); o (octapeptide motif); HD (homeodomain); TD (transactivation domain). The *PAX3* mutation is localized in intron 2.

present in subject II-3.

These two subjects also carried the familial *EYA4* mutation, but the milder clinical manifestations and the progressive course of the DFNA10 phenotype had obviously been masked by the effect of the novel *PAX3* gene splice-site mutation, which by itself is the likely cause of their congenital profound hearing impairment.

The multiple affected subjects and their different phenotypes had so far been obstacles to reaching correct clinical diagnoses and search for possible etiological bases of the hearing loss in this family, where finally two different molecular causes were found to segregate.

The combination of a highly efficient NGS technology and thorough clinical evaluation has allowed to understand the causes of deafness and provide, together with a better genetic counseling and risk of recurrence for each affected individual, extremely important prognostic and follow-up information for the future audiologic management of the proband.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

The Authors wish to thank the family members for their active participation in the study.

References

[1] C.C. Morton, W.E. Nance, Newborn hearing screening–a silent revolution, N. Engl. J. Med. 354 (2006) 2151–2164, http://dx.doi.org/10.1056/NEJMra050700.
[2] R.J.H. Smith, J.P.J. Bale, K.R. White, Sensorineural hearing loss in children, Lancet (London, Engl. 365 (2005) 879–890, http://dx.doi.org/10.1016/S0140-6736(05)71047-3.
[3] M. Amorini, P. Romeo, R. Bruno, F. Galletti, C. Di Bella, P. Longo, S. Briuglia, C. Salpietro, L. Rigoli, Prevalence of deafness-associated Connexin-26 (GJB2) and Connexin-30 (GJB6) pathogenic alleles in a large patient cohort from eastern sicily, Ann. Hum. Genet. 79 (2015) 341–349, http://dx.doi.org/10.1111/ahg.12120.
[4] M. Parker, M. Bitner-Glindzicz, Genetic investigations in childhood deafness, Arch. Dis. Child. 0 (2014) 1–8, http://dx.doi.org/10.1136/archdischild-2014-306009.
[5] M.E. ONeill, J. Maesetta, D. Nishimura, S. Wayne, G. Van Camp, L. Van Laer, C. Negrini, R.R. Wilcox, A. Chen, K. Fukushima, L. Ni, V.C. Sheffield, R.J. Smith, A gene for autosomal dominant late-onset progressive non-syndromic hearing loss, DFNA10, maps to chromosome 6., Hum. Mol. Genet. 5 (1996) 851–856.
[6] C. Frykholm, J. Klar, H. Arnason, A.-C. Rehnman, M. Lodahl, U. Wedin, N. Dahl, L. Tranebjærg, N.D. Rasmussen, Phenotypic variability in a seven-generation Swedish family segregating autosomal dominant hearing impairment due to a novel EYA4 frameshift mutation, Gene 563 (2015) 10–16, http://dx.doi.org/10.1016/j.gene.2015.02.022.
[7] J. Schonberger, L. Wang, J.T. Shin, S. Do Kim, F.S.S. Depreux, H. Zhu, L. Zon, A. Pizard, J.B. Kim, C.A. Macrae, A.J. Mungall, J.G. Seidman, C.E. Seidman, Mutation in the transcriptional coactivator EYA4 causes dilated cardiomyopathy and sensorineural hearing loss, Nat. Genet. 37 (2005) 418–422, http://dx.doi.org/10.1038/ng1527.
[8] F. Liu, J. Hu, W. Xia, L. Hao, J. Ma, D. Ma, Z. Ma, Exome sequencing identifies a mutation in EYA4 as a novel cause of autosomal dominant non-syndromic hearing loss, PLoS One 10 (2015), http://dx.doi.org/10.1371/journal.pone.0126602.
[9] V. Pingault, E. Bodereau, V. Baral, S. Gherbi, N. Lourdon, V. Coubignac, F. Denoyelle, N. Noel-Petroff, H. Duncan Le Pointe, M. Elmaleh-Bergès, N. Sznajderal, S. Marlin, SOX10 mutations mimic isolated hearing loss, Clin. Genet. 88 (2015) 352–359, http://dx.doi.org/10.1111/cge.12506.
[10] V. Pingault, D. Ente, F. Dastot-Le Moal, M. Goossens, S. Marlin, N. Bondurand, Review and update of mutations causing Waardenburg syndrome, Hum. Mutat. 31 (2010) 391–406, http://dx.doi.org/10.1002/humu.21211.
[11] A.E. Lalwani, J.R. Brister, J. Fex, K.M. Grundfast, S. Popkin, T.B. San Agustin, R.R. Wilcox, Further elucidation of the genomic structure of PAX3, and identification of two different point mutations within the PAX3 homeobox that cause Waardenburg syndrome type 1 in two families, Am. J. Hum. Genet. 56 (1995) 75–83.
[12] M. Mazzoli, G. Van Camp, V. Newton, N. Giarbini, F. Declau, A. Parving, Recommendations for the description of genetic and audiological data for families with nonsyndromic hereditary hearing impairment, Audiol. Med. 1 (2003) 148–150, http://dx.doi.org/10.1080/16513860310713.
[13] F.J. del Castillo, M. Rodríguez-Ballesteros, A. Alvarez, T. Hutchin, E. Leonardi, C.A. de Oliveira, H. Azaiez, Z. Brownstein, M.R. Avenarius, S. Marlin, A. Pandya, H. Shahin, R.R. Siemering, D. Weil, W. Wuyts, L.A. Aguirre, Y. Martin, M.A. Moreno-Pelayo, M. Villamar, K.B. Avraham, H.-H.M. Dahl, M. Kanaan, W.E. Nance, C. Petit, R.J.H. Smith, G. Van Camp, E.L. Sartorato, A. Morgia, F. Moreno, I. del Castillo, A novel deletion involving the connexin-30 gene, del (GJB6-d13s1854), found in trans with mutations in the GJB2 gene (connexin-26) in subjects with DFNB1 non-syndromic hearing impairment, J. Med. Genet. 42 (2005) 588–594, http://dx.doi.org/10.1136/jmg.2004.028324.
[14] X. Wang, K. Wang, wANNOVAR: annotating genetic variants for personal genomes via the web, J. Med. Genet. 49 (2012) 433 LP – 436 http://jmg.bmj.com/content/49/7/433.abstract.
[15] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4 (2009) 1073–1081, http://dx.doi.org/10.1038/nprot.2009.86.
[16] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249, http://dx.doi.org/10.1038/nmeth0410-248.
[17] J.M. Schwarz, C. Rodelsperger, M. Schuelke, D. Seelow, MutationTaster evaluates disease-causing potential of sequence alterations, Nat. Methods 7 (2010) 575–576, http://dx.doi.org/10.1038/nmeth0810-575.
[18] B. Reva, Y. Antipin, C. Sander, Predicting the functional impact of protein mutations: application to cancer genomics, Nucleic Acids Res. 39 (2011), http://dx.doi.org/10.1093/nar/gkr407 e118.
[19] S. Ghus, J.C. Fay, Identification of deleterious mutations within three human genomes, Genome Res. 19 (2009) 1553–1561, http://dx.doi.org/10.1101/gr.092619.109.
[20] H.A. Shihab, J. Gough, D.N. Cooper, I.N.M. Day, T.R. Gaunt, Predicting the functional consequences of cancer-associated amino acid substitutions, Bioinformatics 29 (2013) 1504–1510, http://dx.doi.org/10.1093/bioinformatics/btt182.
[21] M. Pertea, G.M. Pertea, S.L. Salzberg, Detection of lineage-specific evolutionary changes among primate species, BMC Bioinforma. 12 (2011) 274, http://dx.doi.org/10.1186/1471-2105-12-274.
[22] E.V Davydov, D.L. Goode, M. Sirota, G.M. Cooper, A. Sidow, S. Batzoglou, Identifying a high fraction of the human genome to be under selective constraint using GERP + +, PLoS Comput. Biol. 6 (2010), http://dx.doi.org/10.1371/journal.pcbi.1001025 e1001025.
[23] M. Garber, M. Guttman, M. Clamp, M.C. Zody, N. Friedman, X. Xie, Identifying novel constrained elements by exploiting biased substitution patterns,

Bioinformatics 25 (2009) i54–62, http://dx.doi.org/10.1093/bioinformatics/btp190.
[24] M. Kircher, D.M. Witten, P. Jain, B.J. O'Roak, G.M. Cooper, J. Shendure, A general framework for estimating the relative pathogenicity of human genetic variants, Nat. Genet. 46 (2014) 310–315, http://dx.doi.org/10.1038/ng.2892.
[25] S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W.W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H.L. Rehm, Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the american college of medical genetics and genomics and the association for molecular pathology, Genet. Med. 17 (2015) 405–424, http://dx.doi.org/10.1038/gim.2015.30.
[26] Q. Li, K. Wang, InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines, Am. J. Hum. Genet. 100 (2017) 267–280, http://dx.doi.org/10.1016/j.ajhg.2017.01.004.
[27] F.-O. Desmet, D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres, C. Beroud, Human Splicing Finder: an online bioinformatics tool to predict splicing signals, Nucleic Acids Res. 37 (2009), http://dx.doi.org/10.1093/nar/gkp215 e67.
[28] L.A. Farrer, R.M. Grundfast, J. Amos, R.S. Arnos, J.H. Asher, P. Brighton, G.R. Diehl, T.T.J. Fox, C. Foy, T.B. Friedman, J. Greenberg, C. Hoehn, M. Maestrini, A. Mikosukey, R. Morell, W. Nance, V. Newton, R. Ramesar, T.B. Sasa, A. James, S. Cadby, R.G. Wagner, E.R. Wilcox, I. Winship, A.P. Read, Waardenburg syndrome (WS) type 1 is caused by defects at multiple loci, one of which is near ALPP on chromosome 2: first report of the WS consortium, Am. J. Hum. Genet. 50 (1992) 902–913.
[29] D. Yan, X.-Z. Liu, Cochlear molecules and hereditary deafness, Front. Biosci. 13 (2008) 4972–4981.
[30] T. Hutchin, N.N. Coy, H. Conlon, E. Telford, K. Bromelow, D. Blaydon, G. Taylor, S. Coghill, S. Brown, R. Trembath, X.Z. Liu, M. Bitner-Glindzicz, R. Mueller, Assessment of the genetic causes of recessive childhood non-syndromic deafness in the UK - implications for genetic testing, Clin. Genet. 68 (2005) 506–512, http://dx.doi.org/10.1111/j.1399-0004.2005.00539.x.
[31] S. Angeli, X.L. Lin, X.U.Z.Z. Liu, Genetics of hearing and deafness, Anat. Rec. 295 (2012) 1801–1829, http://dx.doi.org/10.1002/ar.22579 Genetics.
[32] E. VanBeelen, A.M.M. Oonk, J.M. Leijendekkers, S.H. Hoefsloot, R.J.E. Pennings, I. Feenstra, H. Dieker, F.L.M. Huygen, A.F.M. Snik, H. Kremer, H.P.M. Kunst, Audiometric characteristics of a Dutch DFNA10 family with mid-frequency hearing impairment, Ear Hear 37 (2016) 103–111, http://dx.doi.org/10.1097/AUD.0000000000000217.
[33] M. Verstreken, F. Declau, I. Schatteman, D. Van Velzen, K. Verhaeven, G. Van Camp, P.J. Willems, E.W. Kuiwenhe, E. Verbaert, P. D'Haese, F.L. Wuyts, P.H. Van de Heyning, Audiometric analysis of a Belgian family linked to the DFNA10 locus, Am. J. Otol. 21 (2000) 675–681.

Ringraziamenti

- ✓ *Desidero ringraziare la professoressa Murgia, per avermi dato l'opportunità di svolgere questo progetto di tesi, arricchendolo con importanti esperienze congressuali. Desidero inoltre ringraziarla per aver creduto costantemente nelle mie potenzialità e stimolata scientificamente con discussioni molto interessanti.*

- ✓ *Un sincero ringraziamento va alle mie colleghe, Roberta, Elisa, Emanuela e Mariacristina, per aver condiviso con me questo percorso di crescita sia scientifica che umana; per essermi state accanto nei momenti di felicità e per avermi sostenuta nei momenti di difficoltà. Un grazie speciale va ad Elisa che con pazienza e grande professionalità mi ha fatto da tutor, supportandomi notevolmente in tutto il percorso ed in particolare alla fine.*

- ✓ *Ringrazio mia mamma che con il suo grande amore mi ha sempre incoraggiata a "non mollare", anche quando sarebbe stato facile farlo. Ringrazio mio fratello che ora è molto presente nella mia vita e mi consiglia sempre la cosa giusta da fare.*

- ✓ *Ringrazio mio papà, a cui dedico questo lavoro, so che sarebbe molto orgoglioso di me per questo traguardo e con il suo amore incondizionato so che mi sta aiutando e proteggendo.*

- ✓ *Ringrazio di cuore Paolo che con il suo amore, la sua pazienza e la sua dolcezza rappresenta la mia roccia ed il mio punto fisso. Lo ringrazio tanto perché senza la sua forza non sarei riuscita a raggiungere questo traguardo, per questo dedico anche a lui questa tesi.*