Unravelling the genetic bases of hearing loss: functional characterisation of pathogenic variants and novel candidate genes identified by whole-exome sequencing

Thesis

UNRAVELLING THE GENETIC BASES OF HEARING LOSS: FUNCTIONAL CHARACTERISATION OF PATHOGENIC VARIANTS AND NOVEL CANDIDATE GENES IDENTIFIED BY WHOLE-EXOME SEQUENCING

Thesis submitted by

Chiara Chiereghin

IRCCS Istituto Clinico Humanitas, Rozzano (MI), Italy

Affiliated Research Centre to "The Open University" Milton Keynes, UK

For the degree of Doctor of Philosophy

School of Life, Health and Chemical Sciences

Director of studies: Prof. Rosanna Asselta

Supervisor: Dr. Giulia Soldà

External supervisor: Prof. Karen Steel

Thesis submission date: 11 January 2019

Abstract

Hearing Loss (HL) is the most common sensory disorder in humans, and more than half of cases are due to genetic factors. In 30% of cases, hereditary HL is associated with additional clinical features, and it is defined syndromic (SHL), whereas in 70% of cases HL is the only symptom, and it is considered nonsyndromic (NSHL). HL is characterised by an extreme genetic heterogeneity, with more than 150 loci currently associated and at least 100 genes identified, making extremely challenging to obtain a molecular diagnosis with traditional screening methods. For these reasons, whole-exome sequencing (WES) has been introduced to search for mutations and novel genes underlying the disease.

In this frame, this thesis describes the study of the genetic bases of deafness in 11 HL (3 SHL and 8 NSHL) families applying WES and an accurate functional validation of the candidate mutations.

This approach allowed the identification of the causative mutations in 7 out of the 11 families analysed, while for 3 NSHL families no strong candidate pathogenic variants emerged from our WES data analyses.

In addition, in one NSHL family we identified a missense variant in the candidate deafness-causing gene *DIAPH2*, coding for a protein involved in actin filament elongation. My *in-vitro* studies indicate a possible functional impairment of the mutant DIAPH2 protein, which might be very relevant in hearing function. In fact, immunohistochemical studies indicate that the mouse ortholog protein Diap2 is expressed in the cochlea in the actin-rich stereocilia of the sensory outer hair cells during development. Auditory brainstem response measurements to evaluate the hearing phenotype of *Diaph2* knock-out and knock-in mice were also undertaken. However, at least at the time points analysed, no hearing impairment was detected in the mouse models. Consequently, the role of *DIAPH2* in HL in humans still needs to be further clarified.

Contents

List of figures

List of tables

Abbreviations

6-FAM	6-Carboxyfluorescein
6TM	Six transmembrane domains
ABR	Auditory brainstem responses
AD	Autosomal dominant
ANOVA	Analysis of variance
AR	Autosomal recessive
AS	Alport syndrome
ASPA	Animals (Scientific Procedures) Act
AUNA1	Auditory neuropathy, autosomal dominant, 1
BM	Basilar membrane
bp	base pairs
BWA	Burrows-Wheeler Aligner
CADD	Combined Annotation Dependent Depletion
cAMP	Cyclic adenosine monophosphate
CC	Coiled coil
CNBD	Cyclic nucleotide binding domain
CNG	Cyclic nucleotide-gated
CNV	Copy number variant
CRISPR	Clustered regularly interspaced short palindromic repeats
Cas9	CRISPR-associated 9
DAB	3,3'-Diaminobenzidine

DAD	*Diaphanous* autoregulatory domain
DAPI	4',6-Diamidino-2-Phenylindole
dB	Decibel
DD	Dimerization domain
DID	*Diaphanous* inhibitory domain
DNA	Deoxyribonucleic acid
dpc	Days post coitum
dpf	Days post fertilisation
DRF	*Diaphanous*-related formins
E	Embryonic day
EDTA	Ethylenediaminetetraacetic acid
EEG	Electroencephalography
EGFP	Enhanced Green Fluorescent Protein
EIEE24	Epileptic encephalopathy, early infantile, 24
EP	Endocochlear potential
FBS	Fetal bovine serum
FH	Formin homology
GATK	Genome Analysis Toolkit
GBD	GTPase binding domain
GDB	Gelatin dilution buffer
GDPR	General Data Protection Regulation
GER	Greater epithelium ridge
GME	Greater Middle East Variome study cohort

gnomAD	Genome Aggregation Database
gRNA	Guide RNA
HB	Hindbrain
HC	Hair cell
HCN	Hyperpolarization-activated cyclic nucleotide-gated
HL	Hearing loss
HRM	High-Resolution Melting
HS	High salt
IFN	Interferon
IHC	Inner hair cell
IMPC	International Mouse Phenotyping Consortium
Indel	Insertion-deletion
IQ	Intelligence quotient
ISO	International Standard Organization
ivd	Intervening domains
KO	Kölliker's organ
LCCL	Limulus factor C, cochlin, and late gestation lung protein Lgl1
LER	Lesser epithelium ridge
LOD	Logarithm of odds
LS	Low salt
MAF	Minor allele frequency
MEM	Minimum Essential Medium
MET	Mechanoelectrical transduction

MGP	Mouse genetics project
miRNA	microRNA
MLPA	Multiplex Ligation Probe Amplification
mRNA	Messenger RNA
N.A.	Numerical Aperture
NADf	North African Deafness
NB	Neuroblasts
NGS	Next-generation sequencing
NHLBI-ESP	National Heart, Lung, and Blood Institute Exome Sequencing Project
NSHL	Nonsyndromic hearing loss
OC	Organ of Corti
OCTN	Organic cation transporter
OHC	Outer hair cell
P	Postnatal day
PAM	Protospacer adjacent motif
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDS	Pendred syndrom
POF	Premature ovarian failure
PRLTS	Perrault syndrome
RM	Reissner membrane
RNA	Ribonucleic acid
ROH	Runs of homozygosity

rRNA	Ribosomal RNA
RT-PCR	Reverse transcriptase-polymerase chain reaction
SAG	Statoacustic ganglion
SA-HRP	Streptavidin-horseradish peroxidase
SC	Supporting cell
SCBMS	Seizures, cortical blindness, microcephaly syndrome
SD	Standard deviation
SEM	Standard error of the mean
SGN	Spiral ganglion neuron
SHL	Syndromic hearing loss
SM	Scala media
SNV	Single nucleotide variants
SP	Signal peptide
SPL	Sound pressure level
ST	Scala tympani
SV	Stria vascularis
SVe	Scala vestibuli
TF	Transcription factor
TGN	Trans-Golgi Network
TM	Tectorial membrane
VUS	Variant of unknown significance
vWFA	von Willebrand factor A
WES	Whole-exome sequencing

1. Introduction

1.1. The ear anatomy and the hearing process

The ear is responsible for the sense of hearing and equilibrium and it can be divided in three anatomical components: the outer, the middle and the inner ear (Fig. 1.1A).

The outer ear includes the auricle (or pinna), which collects the sound waves, and the ear canal, which channels the sound to the tympanic membrane (or eardrum). The middle ear is an air-filled compartment, whose principal role is to amplify the air sound waves coming from the outer ear and transfer them into the motion of the fluid in the inner ear. This task is accomplished through the three ossicles of the middle ear (malleus, incus and stapes) that form an ossicular chain which connects the eardrum to the oval window of the cochlea.

The inner ear comprises the vestibular system, responsible for balance, and the cochlea, which contains the organ of Corti (OC), the receptor organ for hearing [Dror and Avraham, 2010].

The cochlea is a snail-shaped duct filled with fluid which consist of three parallel membranous compartments called scalae: the scala media, containing the OC, lies between an upper chamber, the scala vestibuli and a lower chamber, the scala tympani (Fig. 1.1B).

The scalae vestibuli and tympani are connected to the middle ear by the oval and round window respectively. Both chambers are filled with a high Na^+ and low K^+ and Ca^{2+} fluid called perilymph and communicate at the apex of the cochlea through an opening called helicotrema. The scala media is isolated from the scalae vestibuli and tympani respectively by Reissner and basilar membranes (RM and BM) and contains endolymph, a fluid rich in K^+ and low in Na^+. The stria vascularis (SV), located in the lateral wall of the scala media, is responsible for the maintenance of the composition of the endolymph and for the generation of the endocochlear potential (EP), which is essential for the sensory transduction in the cochlea [Fuchs, 2010].

The OC, which resides on the basilar membrane, includes the sensory hair cells (HC) and the supporting cells (SC). The apical surfaces of these cells are joined together by

intercellular junction complexes and form a reticular lamina which restricts the ion flow between the endolymph and the perilymph-like fluid beneath these cells. The hair cells are not in direct contact with the BM and are surrounded by several diverse SCs, including inner border and inner phalangeal cells, pillar cells, Deiters' cells, Hensen's cells and Claudius' cells (Fig. 1.1C). The SCs not only provide a structural support to HCs, but also play an important role during development, in cochlear homeostasis and in the response to trauma [Gale and Jagger, 2010].

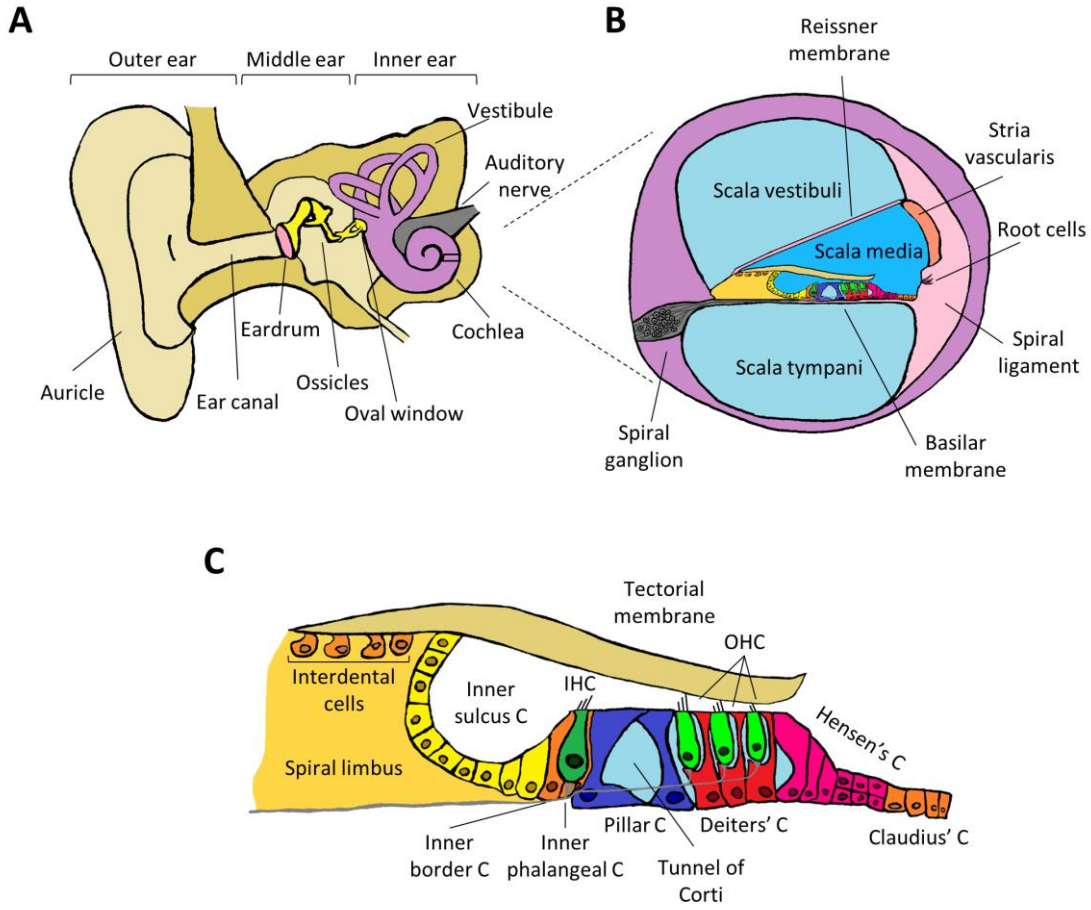


Figure 1.1. Diagram of the human ear. A. Schematic illustration of the three anatomical components of the ear. **B.** Cross section of the cochlear duct showing the three scalae and the organ of Corti, which is further detailed in C. **C.** The organ of Corti with the overlying tectorial membrane is illustrated. Inner hair cells (IHC), outer hair cells (OHC) and the main supporting cells (C) subtypes are shown. Inner border and inner phalangeal cells surround IHCs, while OHC are supported by Deiters' cells. Pillar cells form the tunnel of Corti. The interdental cells, connecting the tectorial membrane to the spiral limbus, are also illustrated. Adapted from Dror and Avraham, 2010.

The hair cells are arranged in a highly organised conformation of one inner row (inner hair cells, IHC) and three outer rows of hair cells (outer hair cells, OHC) along the duct (Fig. 1.1C). The apical side of the hair cells contains actin-rich projections organised in rows called stereocilia (or hair bundle) that bath in the endolymph. The apical ends of stereocilia are linked to the side of the neighboring taller stereocilia through tip links, composed of the transmembrane proteins protocadherin 15 and cadherin 23 [McGrath *et al.*, 2017] (Fig. 1.2). The tallest row of OHC stereocilia is in physical contact with the overlying tectorial membrane (TM). The TM, an acellular gelatinous matrix composed of collagen and glycoproteins, is attached to the spiral limbus through the interdental cells. It is secreted by the cells of the Kölliker's organ, a transient structure which is present during the development and progressively regresses to form the inner sulcus in the mature organ of Corti (Fig. 1.1C) [Kelley, 2007]. When the sound wave is transmitted to the oval window by the stapes, it travels along the scalae vestibuli and tympani causing the perilymph movement which, in turn, induces the vibration of the BM. The displacement of the BM results in the movement of the HC towards the TM and in the deflection of the hair bundle in the direction of the taller stereocilia. The motion of the hair bundle is coupled to the opening of mechanoelectrical transduction (MET) channels - which are localised in the tips of stereocilia - by the tip links (Fig. 1.2). The EP, generated by the SV, provides the main driving force for sensory transduction. In fact, it helps to drive K^+ ions through the MET channel rapidly. This results in the depolarisation of the HCs that triggers the opening of voltage-sensitive calcium channels and the release of neurotransmitter at the base of the HCs, leading to the mechanoelectrical transduction of the sound. Deflection of the hair bundle in the opposite direction closes the MET channels, stopping the neurotransmitter release [Wangemann, 2002; Richardson *et al.*, 2011]. While the IHCs are the true sensory cells, making synaptic connections with afferent spiral ganglion neuron (SGN) terminals, the OHCs amplify the signal enhancing the sensitivity and selectivity of

the cochlea. In fact, the depolarisation of the OHC results in the change of the cell length that amplifies the motion of the OC in that specific region. The length change is made possible by the concerted voltage-induced conformational change of prestin protein, localised at the basolateral membrane of the OHCs [Zheng *et al.*, 2000].

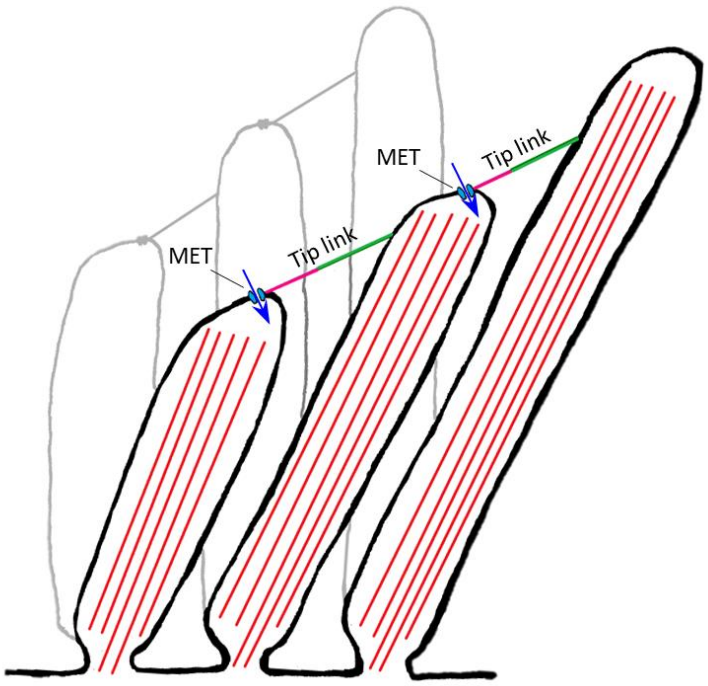


Figure 1.2. Schematic representation of hair cell stereocilia during deflection. The tip-link proteins protocadherin 15 (in pink) and cadherin 23 (in green) and the MET channels are indicated. The blue arrows indicate the entrance of cations through the MET channels. The red lines represent actin filaments. Adapted from Richardson *et al.*, 2011.

The sensory information transduced by the IHCs is conveyed by the SGNs, whose central axons form the vestibulocochlear nerve (cranial nerve VIII) and project to the cochlear nuclei in the brainstem. Subsequently, the auditory information is transmitted to the brainstem superior olivary complex and then travels through the inferior colliculus in the midbrain and to the medial geniculate nucleus of the thalamus, and finally to the auditory cortex [Butler and Lomber, 2013].

The cochlea is characterised by a tonotopic organisation, transducing high frequencies at the base and low frequencies at the apex. This organization results from several morphological gradients existing along the cochlea. For example, the OHCs and their

stereocilia become longer from the base to the apex and both the BM and the TM membrane become wider and thicker towards the apex, so that the HCs are activated by different frequencies depending on their localisation along the sensory epithelium [Son *et al.*, 2012]. This tonotopic organisation is also maintained in all relays of the ascending auditory pathway [Butler and Lomber, 2013].

1.2. The development of the inner ear in mammals

Although the mouse is most often the preferred animal model, the principal cochlear developmental events are similar in all mammals [Bianchi and Fuchs, 2010]. The mammalian inner ear forms from a thickening of the ectoderm on either side of the rhombencephalon (hindbrain), called the otic placode (Fig. 1.3A). The otic placode, evident in the mouse embryo at 8-8.5 days *post coitum* (dpc) [Kaufman, 1999] and in the human embryo at ~20 days *post fertilization* (dpf) [O'Rahilly, 1983], starts to invaginate to form the otic pit (Fig. 1.3B) and, by 10 dpc in the mouse [Kaufman, 1999] and ~28 dpf in the human [O'Rahilly, 1983], it completely closes, forming the otic vescicle or otocyst (Fig. 1.3C). During the otocyst formation, the cochlear and vestibular neurons originate by delamination of the anteroventral region of the otic vescicle and form a single structure, called statoacustic ganglion (SAG) (Fig. 1.3B-D), from which the cochlear and vestibular ganglia will arise later in the development [Torres and Giráldez, 1998].

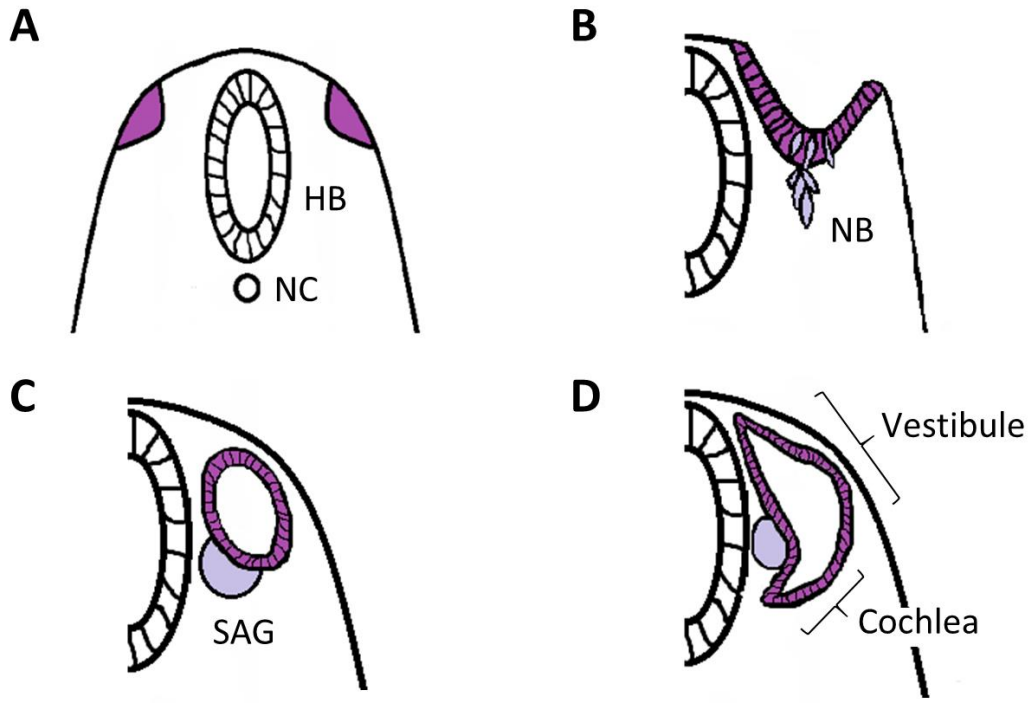


Figure 1.3. Schematic illustration of mammalian inner ear early development. A. Transverse section of the head region of a developing mammalian embryo at the level of the hindbrain (HB, dorsal to the top). The otic placodes are shown in purple. NC: notochord. **B.** Invagination of the otic placode into the otic pit. The neuroblasts (NB) delaminating from the otic epithelium are shown in lilac. **C.** The otic vescicle or otocyst and the statoacustic ganglion (SAG) are shown. **D.** Otocyst differentiation: the dorsal region will give rise to the vestibular structures, whereas the ventral region will elongate and coil to form the cochlea. Adapted from Torres and Giráldez, 1998.

Numerous signal molecules originating from the hindbrain and the periotic mesenchyme contribute to the complex morphogenic changes that establish the dorsoventral and anteroposterior axes. As a result, at 10.75 dpc (~32 dpf in human) the cochlear portion starts to protrude from the ventral region of the otocyst (Fig. 1.3D) and as it elongates it starts to coil, achieving the 1.75 turns of the mature mouse cochlea by 17 dpc and the final 2.5 turns at 57 dpf in the human [Cantos *et al.*, 2000; O'Rahilly, 1983]. At this stage, the scala vestibuli and scala tympani begin to open starting from the base towards the apex. As a result, at the apex only one chamber – the developing scala media – can be seen in a cross section of a 16.5 dpc mouse cochlea (Fig. 1.4).

While these gross morphological changes are occurring, the epithelial cells of the emerging cochlear duct, which have proliferated from the apex to the base of the cochlea (12-16 dpc in the mouse), begin to differentiate in a base-to-apex direction. In addition, there is a modiolar to strial gradient of differentiation, so that IHCs differentiate before OHCs. For this reason, at 15 dpc IHCs can be identified in the mid-basal turn of the mouse cochlea, while cell differentiation only starts at 18 dpc in the apical turn [Bianchi and Fuchs, 2010]. Differences in the dorsal and ventral wall of the cochlear duct can already be observed at 12.5 dpc, as the dorsal wall epithelium appears thicker than the ventral wall (6 *versus* 2-3 cell layers). By 16 dpc the dorsal epithelium differentiates in two cellular ridges, called the greater epithelium ridge (GER), which comprises the Kölliker's organ and will develop into the inner sulcus, spiral limbus and IHCs, and the lesser epithelium ridge (LER), from which OHCs, Deiters' cells and the cells of the outer sulcus will emerge. The pillar cells develop at the boundary of the two epithelial ridges and will form the "walls" of the tunnel of Corti only after birth in the mouse. In fact, besides being distinguishable at 16 dpc based on their position, supporting cells start their morphological differentiation only postnatally [Bianchi and Fuchs, 2010], so that the cochlea reaches its mature configuration approximately at 2 postnatal weeks in the mouse. Therefore, mice start to hear around postnatal day 12 (P12) [Mikaelian and Ruben, 1965]. In humans, instead, by 25 weeks *post* conception the structures of the cochlea have a near-adult maturation, so that neonates are able to respond to sound stimuli at birth [Bianchi and Fuchs, 2010].

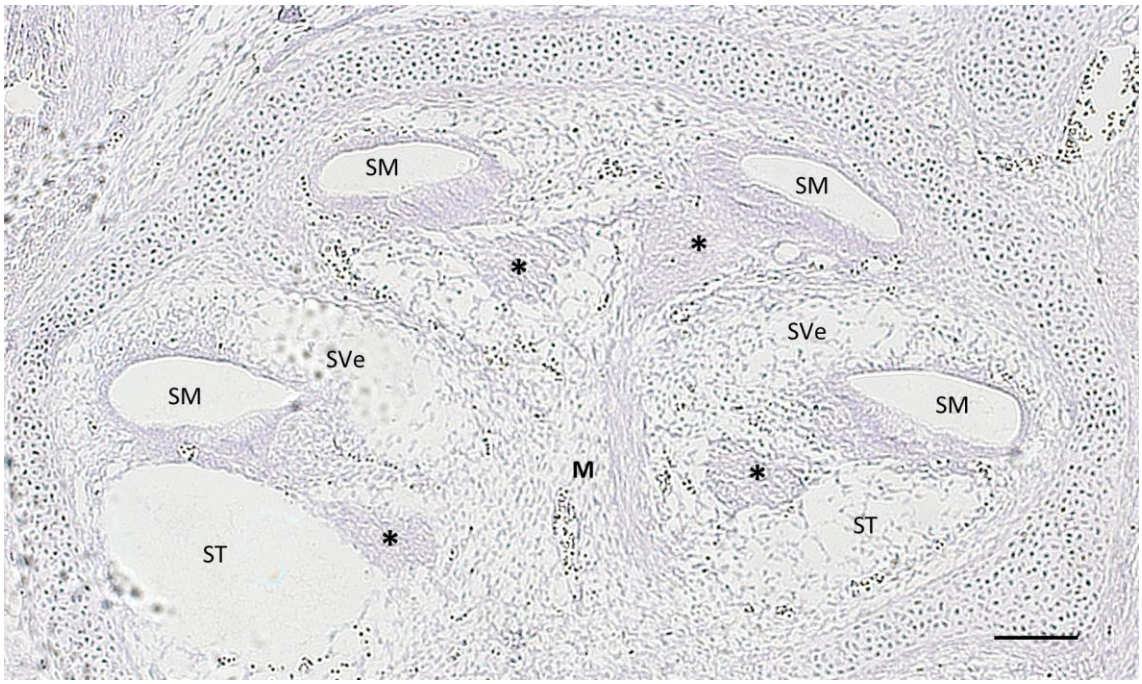


Figure 1.4. Mid-modiolar section of a 16.5 dpc mouse foetus. The scalae tympani (ST) and vestibuli (SVe) start to open in a base-to-apex direction. Only the scala media (SM) can be observed in the apical turn. Asterisks indicate the spiral ganglia. M: modiolus. Dorsal to the bottom. The image was acquired with a Zeiss Axioskop microscope. Scale bar: 100 μm.

1.3. Hearing Loss

Hearing loss (HL), affecting at least 1 in 500 newborns, is the most common sensory disorder in humans. Over the 5% of the world's population has disabling hearing loss and it is estimated that by 2050 1 in 10 people will be affected by disabling hearing loss (World Health Organization, 2018, http://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss). Hearing loss may have a dramatic impact on the quality of life at any age. If not promptly diagnosed in newborns, it could have major effects in speech acquisition and social development. When occurring later in adult life, HL could impact the social and working life. Also, in elderly people, deafness could lead to isolation and depression [Bitner-Glindzicz, 2002; Morton and Giersch, 2010].

Hearing loss is a heterogeneous disease that can have environmental or genetic causes. Noise exposure, ototoxicity (e.g. aminoglycosides, platinum derivatives), congenital (cytomegalovirus, rubella, toxoplasmosis) or acquired (meningitis, measles) infections,

prematurity, head trauma are all environmental causes of deafness. However, it is estimated that at least half of HL cases has a genetic basis [Schrijver, 2004]. To complicate matters, some genetic variants can predispose to environmental and age-related hearing loss [Tu and Friedman, 2018].

Given its clinical heterogeneity, HL can be further classified considering several criteria, such as the severity, being divided in mild (when the hearing loss is between 20 and 39 dB), moderate (40-69 dB), severe (70-89 dB), profound (over 90 dB) HL (Fig. 1.5) [Schrijver, 2004]. HL is defined "disabling" when the hearing loss is greater than 40 dB in adults (from 15 years old) or greater than 30 dB in children (0-14 years old) in the better hearing ear (World Health Organization, 2018, http://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss). HL can also be classified based on the age of onset as prelingual or post-lingual, with reference to the commencement of speech. Based on the site of damage, HL can be defined conductive (when the HL is caused by outer and/or middle ear abnormalities, such as in otosclerosis) or sensorineural (caused by malfunction of the inner ear, the VIII nerve, brainstem, or auditory cortex). When HL is characterised both by conductive and sensorineural defects, it is defined mixed. Finally, genetic cases of HL can be divided into nonsyndromic (NSHL), when HL is the only symptom, or syndromic (SHL), when additional clinical features are present [Korver *et al.*, 2017].

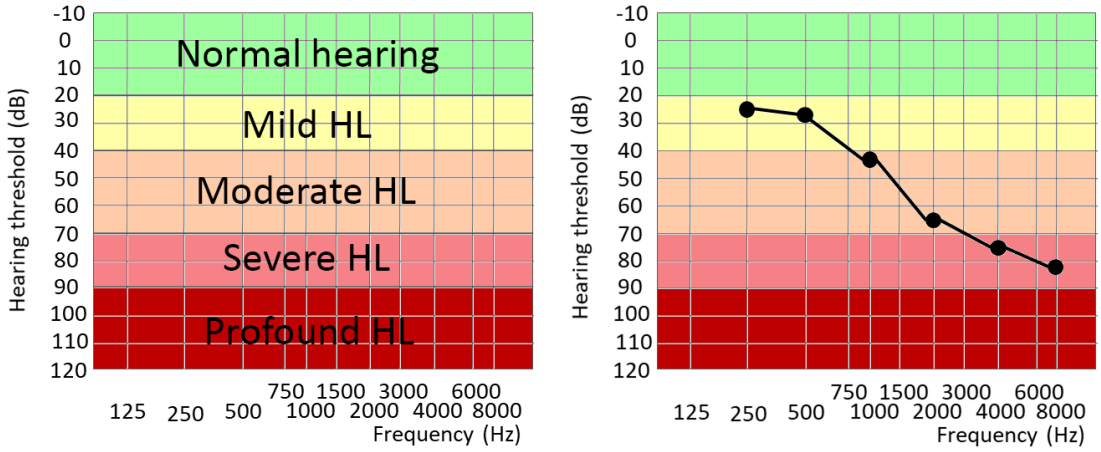


Figure 1.5. Severity of HL. The audiogram is a graph showing the audible thresholds for each frequency tested. On the left, audiogram showing the different degrees of hearing loss. On the right, example of audiogram from a hearing-impaired individual, showing that the severity of HL can vary at different frequencies. The threshold average for the right and left ear is represented.

1.4. Genetics of hearing loss

The clinical heterogeneity of HL is paralleled by its extreme genetic and allelic heterogeneity, reflecting the anatomical and functional complexity of the ear. Not surprisingly, it is estimated that up to 1% of all human genes are involved in hearing function [Nance 2003]. Approximately 70% of genetic cases are nonsyndromic, while in the remaining 30% of cases the presence of additional clinical features associated with HL can contribute to the diagnosis of one of the over 400 syndromes that have hearing loss as a component [Schrijver, 2004; Korver *et al.*, 2017].

1.4.1. Nonsyndromic hearing loss

To date, more than 150 loci have been associated with NSHL and over 100 responsible genes have been identified (Hereditary Hearing Loss Homepage, https://hereditaryhearingloss.org/, last accessed: October 2018) (Fig. 1.6). NSHL is most often sensorineural and the inheritance pattern can be autosomal recessive (ARNSHL, 80% of cases), autosomal dominant (ADNSHL, 15-20%), X-linked (1.5%) or mitochondrial (1%) [Schrijver, 2004]. In most cases, recessive NSHL is prelingual and severe, whereas

dominant NSHL is post-lingual and progressive [Petit, 1996]. NSHL loci are named based on the mode of inheritance as DFNA (AD), DFNB (AR), DFNX (X-linked). One locus (DFNY1) for Y-linked NSHL was also mapped [Wang *et al.*, 2004], but no responsible gene has been identified yet. In addition, DFNM denotes loci associated with modifier genes, whose mutation can affect the phenotype resulting from a mutation in another gene [Riazuddin *et al.*, 2000; Bykhovskaya *et al.*, 2000; Yousaf *et al.*, 2018a]. Finally, specific symbols are used for mitochondrial NSHL (MTRNR, MTTS), auditory neuropathy (AUNA) and otosclerosis (OTSC). A number is assigned to each locus, reflecting the chronological order in which they were mapped [Dror and Avraham, 2010]. However, the same gene can be associated with multiple loci, as different mutations in the same gene might be responsible for both ARNSHL and ADNSHL or even cause SHL (Fig. 1.6, Table 1.1, Appendix 6.1). The extreme genetic and clinical heterogeneity of the disease has hindered a full comprehension of genotype-phenotype correlations in NSHL patients. In fact, the onset, severity and progression of NSHL associated with different underlying genetic causes are often indiscernible [Vona *et al.*, 2015] (Table 1.1). For instance, mutations in many ADNSHL genes with diverse functions cause post-lingual progressive NSHL beginning with loss of high frequencies and are characterised by a downsloping audiogram (Table 1.1). On the other hand, other ADNSHL genes have more distinctive audiogram profiles and thus the phenotype observed might be predictive of the underlying gene. Indeed, the AudioGene tool has been developed in order to prioritise candidate genes for mutation screening on the basis of the audioprofile [Hildebrand *et al.*, 2009]. In addition, a relatively small subset of ARNSHL genes is frequently associated with vestibular symptoms or cochlear abnormalities. Therefore, the observation of such clinical features might be suggestive of mutations in specific genes (Table 1.1, Appendix 6.1) and thus provide a guidance for genetic testing.

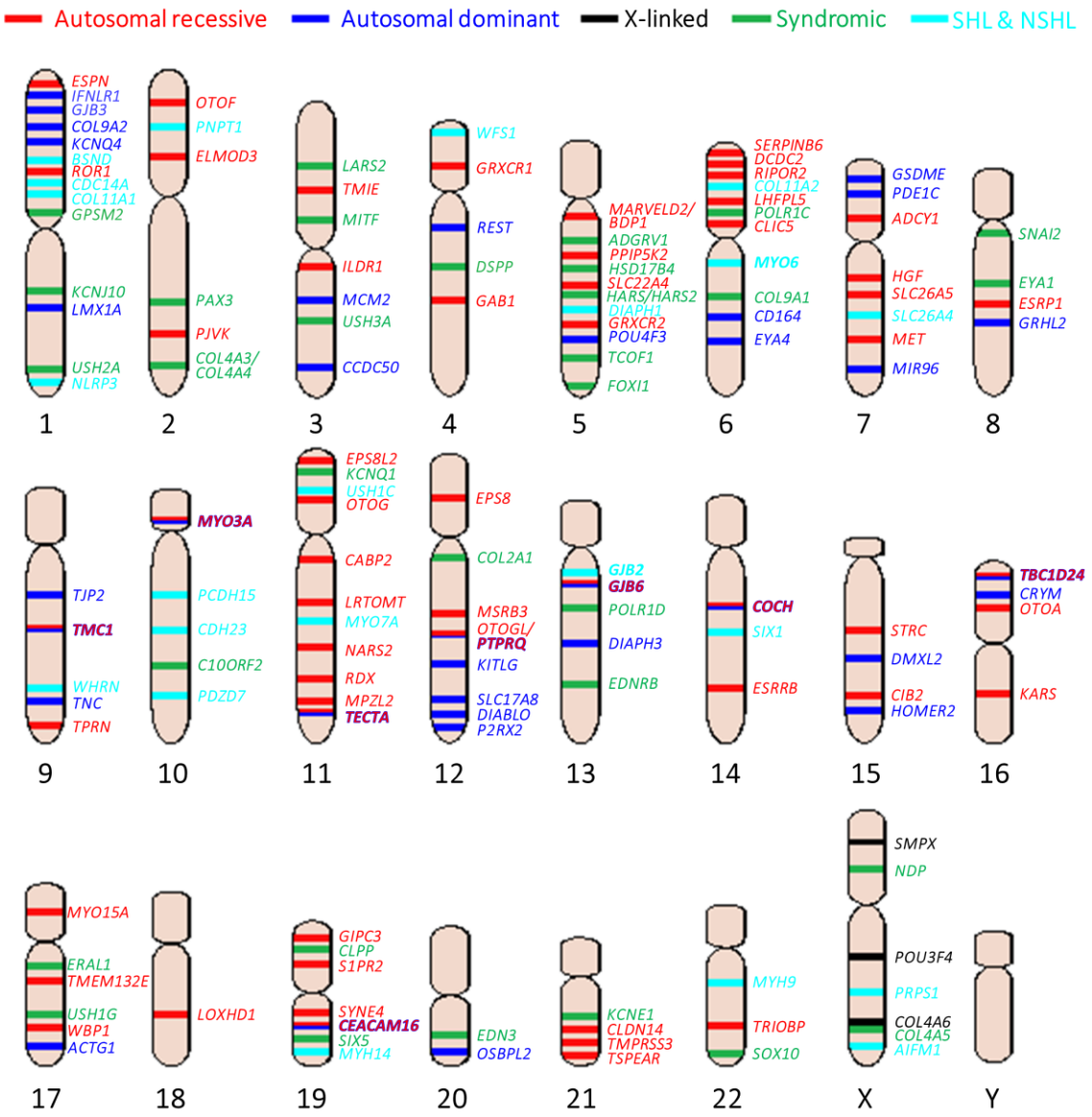


Figure 1.6. Schematic representation of the localisation of the genes associated with HL. The genes associated with autosomal recessive (in red), autosomal dominant (in blue), X-linked (in black) NSHL and with some syndromic forms of HL (in green) are indicated. The genes both associated with NSHL and SHL are shown in light blue. Updated (based on Hereditary Hearing Loss Homepage, https://hereditaryhearingloss.org/, last accessed: October 2018) from Raviv *et al.*, 2010.

In the last 25 years, the identification of NSHL genes through linkage studies in large families, characterisation of mouse models (see Paragraph 1.6) and, more recently, next-generation sequencing (see Paragraph 1.5) allowed the discovery of numerous molecular processes fundamental for hearing. Therefore, the genes responsible for NSHL can be grouped in broad functional categories, including: gene regulation, mechanotransduction,

ion homeostasis.

Gene regulation – The exact otocyst patterning and cochlear morphogenesis involve the spatiotemporally controlled interaction of many transcription factors (TFs) [Cantos *et al.*, 2000], therefore several genes encoding TFs are associated with NSHL in humans.

Examples include two TFs of the POU domain superfamily, *POU4F3* and *POU3F4*, which are associated with progressive HL. *POU4F3* was shown to be fundamental for OHCs survival and maturation as the corresponding *Pou4f3* knock-out mouse show OHC degeneration [Hertzano *et al.*, 2004]. Mutations in this gene cause ADNSHL DFNA15 (MIM#602459) [Vahava *et al.*, 1998] in humans. The mouse orthologue of *POU3F4*, instead, is expressed in the periotic mesenchyme and mutant mice show impaired endocochlear potential due to defects in spiral ligament fibrocytes [Braunstein *et al.*, 2008]. *POU3F4* is associated with X-linked HL DFNX2 (MIM#304400), which can be sensorineural or mixed, and might be also associated with bony labyrinth defects [de Kok *et al.*, 1995]. Besides TFs, other important regulatory molecules are microRNAs (miRNAs), which act at the post-transcriptional level by controlling the expression of target genes. In particular, the miR-183 family of miRNAs (comprising miR-183, miR-182 and miR-96) plays an important role in vertebrate hair cell development: indeed, mutations affecting *MIR96* gene lead to deafness in both mice (*Diminuendo* mutant) and humans (autosomal dominant deafness DFNA50) [Lewis *et al.*, 2009; Mencía *et al.*, 2009].

Mechanotransduction – The process by which hair cells convert the mechanical stimulus of sound waves into an electrical signal rely on the integrity of all the cochlear structures contributing to this process. Therefore, it should come as no surprise that mutations in genes encoding hair bundle proteins, channels and extracellular matrix proteins cause HL. The core of stereocilia is composed of packed bundles of crosslinked actin filaments, which are regulated by several proteins, such as diaphanous 1 (involved in actin polymerization), espin (required for stereocilia integrity and length regulation), unconventional myosins

(essential for stereocilia organisation and for the transport of other structural proteins along the stereocilia). Mutations in *ACTG1* (encoding γ-actin, the predominant actin isoform in hair cells), *DIAPH1*, *ESPN* genes as well as mutations in the genes encoding motor proteins (*MYO3A*, *MYO6*, *MYO7A*, *MYO15A*, *MYH9*) have all been linked to different types of inherited NSHL [respectively Zhu *et al.*, 2003, van Wijk *et al.* 2003; Lynch *et al.*, 1997; Naz *et al.*, 2004; Walsh *et al.*, 2002, Grati *et al.*, 2016; Melchionda *et al.*, 2001, Ahmed *et al.*, 2003a; Liu *et al.*, 1997a, Liu *et al.*, 1997b; Wang *et al.*, 1998; Lalwani *et al.*, 2000] and, in the case of *MYO7A*, also to SHL [Weil *et al.*, 1995]. The tip of each stereocilium is connected to the side of the adjacent taller stereocilium with a tip link, which is composed by cadherin 23 and protocadherin 15 proteins and is fundamental to couple the displacement of the hair bundle to the opening of MET channels (Fig. 1.2). Although the exact composition of the MET channels is still unclear, TMC1 has been recently demonstrated to form their pore, likely with the contribution of other proteins, such as TMIE and LHFPL5 [Pan *et al.*, 2018]. The genes encoding these proteins (*CDH23*, *PCDH15*, *TMC1*, *TMIE*, *LHFPL5*) have also been associated with NSHL [respectively Bork *et al.*, 2001; Ahmed *et al.*, 2003b; Kurima *et al.*, 2002; Naz *et al.*, 2002; Kalay *et al.*, 2006].

Finally, another essential structural component for mechanotransduction is the tectorial membrane: genes encoding for some of the extracellular matrix protein composing it, namely *TECTA* [Mustapha *et al.*, 1999], *COL11A2* [Chen *et al.*, 2005], *OTOA* [Zwaenepoel *et al.*, 2002], *OTOG* [Schraders *et al.*, 2012] and *OTOGL* [Yaritz *et al.*, 2012] have also been linked to NSHL.

Ion homeostasis – The mechanotransduction would not take place without the endocochlear potential and the precise ion composition existing in the endolymphatic fluid. Not surprisingly, several genes coding for proteins involved in the maintenance of ion gradients (*SLC26A4*), in the recycling of potassium (*KCNQ4*, *GJB2*, *GJB3*, *GJB6*) or contributing to

the physical separation of endolympatic and perylimphatic spaces (*CLDN14*) are associated with HL. For instance, the Cl^-/HCO^{3-} exchanger pendrin, encoded by *SLC26A4* gene and expressed in the spiral prominence (the projecting structure at the lower margin of the stria vascularis) and in the outer sulcus (between the spiral prominence and the supporting cells of Claudius), was found to be important for the maintenance of EP, as demonstrated by loss of EP in *Slc26a4*$^{-/-}$ mice. In fact, lack of pendrin leads to the acidification of endolymph which, in turn, inhibits the acid-sensitive calcium channels of the stria vascularis, causing an increase of endolymphatic calcium concentration [Wangemann *et al.*, 2007]. Concerning K^+ recycling, the potassium entered in the HCs upon the opening of MET channels, is then released into the perilymph from basolateral K^+ channels, including KCNQ4. The potassium is then recycled through the outer sulcus root cells and the spiral ligament fibrocytes, and it is transported from cell to cell via gap junctions to the stria vascularis [Wangemann, 2002; Jagger *et al.*, 2010]. Each gap junction between two cells is composed of two hexameric hemichannels or connexons, made up of the transmembrane protein connexins. Connexin 26 (encoded by *GJB2* gene), 30 (*GJB6*) and 31 (*GJB3*) form the gap junctions in the inner ear. Importantly, recessive mutations in *GJB2* account for half of the cases of prelingual ARNSHL (DFNB1A, MIM#220290) and *GJB2/GJB3* or *GJB2/GJB6* digenic inheritance has also been described [Del Castillo *et al.*, 2003; Liu *et al.*, 2009]. Finally, ion homeostasis is preserved thanks to the separation of compartments containing fluids with different ion composition. This is made possible by the tight junctions that form a semipermeable seal at the apical side of sensorial and supporting cells. Tight junctions are composed of several molecules including claudins, occludins and tricellulin. Both *CLDN14* (encoding claudin 14) and *MARVELD2* (coding for tricellulin) mutations have been associated with ARNSHL [Wilcox *et al.*, 2001; Riazuddin *et al.*, 2006].

Table 1.1. Example of NSHL genes and their clinical manifestations.

Gene	Locus	Onset	NSHL phenotype	Additional features (SHL)
ACTG1	DFNA20/26	Post-lingual (1st/2nd decade)	Downsloping audiogram; moderate-to-profound; progressive	
CDH23	DFNB12	Prelingual	Downsloping; moderate-to-profound; progressive	Usher syndrome (profound congenital HL, retinitis pigmentosa, constant vestibular dysfunction)
CLDN14	DFNB29	Prelingual	Severe-to-profound; stable	
COL11A2	DFNA13	Post-lingual (2nd/4th)	Moderate-to-severe; progressive	Otospondylomegaepiphyseal dysplasia (sensorineural HL and short extremities with large epiphyses)
	DFNB53	Prelingual	Severe-to-profound; stable	
DIAPH1	DFNA1	Post-lingual (1st)	Low-frequency; progresses to profound HL involving all frequencies	High-frequency progressive HL with thrombocytopenia
ESPN	DFNB36	Prelingual	Profound with or without vestibular areflexia	
GJB2	DFNA3A	Prelingual/post-lingual (1st/2nd)	Downsloping; mild-to-profound (variable); progressive	Different syndromic forms causing HL and skin problems
	DFNB1A	Prelingual	Mild-to-profound (variable); progressive (~50%). Vestibular symptoms (~50%).	
GJB3	DFNA2B	Post-lingual (4th)	Downsloping; progressive	
GJB6	DFNA3B	Prelingual	Downsloping; mild-to-profound; progressive	
	DFNB1B	Prelingual	Severe; stable. Possible vestibular symptoms.	
KCNQ4	DFNA2A	Post-lingual (1st/2nd)	Downsloping; progresses, involving all frequencies	
LHFPL5	DFNB66/67	Prelingual	Severe-to-profound; stable	
MARVELD2	DFNB49	Prelingual	Moderate-to-profound; stable	
MIR96	DFNA50	Post-lingual (1st)	Flat; mild-to-profound; progressive	
MYH9	DFNA17	Post-lingual (1st/3rd)	Downsloping; progressive	Macrothrombocytopenia and granulocyte inclusions with or without nephritis or sensorineural HL
MYO15A	DFNB3	Prelingual	Severe-to-profound; stable	
MYO3A	-	Post-lingual (1st)	Moderate-to-profound; progressive	
	DFNB30	Prelingual	Downsloping; moderate-to-profound; progressive	
MYO6	DFNA22	Post-lingual (1st)	Mild-to-profound; progressive	HL and hypertrophic cardiomyopathy
	DFNB37	Prelingual	Severe-to-profound; progressive	
MYO7A	DFNA11	Post-lingual (1st/5th)	Flat/gently downsloping; moderate; progressive. Mild vestibular symptoms in some patients.	
	DFNB2	Prelingual/post-lingual (1st/2nd)	Downsloping; severe-to-profound. Vestibular symptoms in some patients.	Usher syndrome
OTOA	DFNB22	Prelingual	Moderate-to-profound; stable	
OTOG	DFNB18B	Prelingual	Downsloping; mild-to-moderate; stable	
OTOGL	DFNB84B	Prelingual	Downsloping; moderate; stable	
PCDH15	DFNB23	Prelingual	Severe-to-profound; stable	Usher syndrome
POU3F4	DFNX2	Prelingual	Sensorineural or mixed; progresses to profound. Female carriers have a later onset. Possible bony labirinth defects.	
POU4F3	DFNA15	Post-lingual (2nd/6th)	Flat/downsloping; mild-to-severe; progressive	
SLC26A4	DFNB4	Prelingual	Moderate-to-profound. EVA or Mondini dysplasia; vestibular symptoms.	Pendred syndrome (HL and goiter)
TECTA	DFNA8/12	Prelingual/post-lingual (1st/2nd)	U-shaped; moderate-to-severe; stable/progressive	
	DFNB21	Prelingual	Flat/U-shaped; severe-to-profound; stable	
TMC1	DFNA36	Post-lingual (1st/3rd)	Downsloping; severe-to-profound; progresses, involving all frequencies	
	DFNB7/11	Prelingual	Profound; stable	
TMIE	DFNB6	Prelingual	Severe-to-profound; stable	

1.4.2. Syndromic hearing loss

More than 400 syndromes that include hearing loss have been described. However, in some syndromes, HL is a minor component, whereas in others it is an important symptom for clinical diagnosis [Morton and Giersch, 2010]. Examples of syndromes in which HL is an important component include Pendred syndrome (PDS), Alport syndrome (AS) and Perrault syndrome (PRLTS) [Parker and Bitner-Glindzicz, 2015].

PDS, characterised by deafness, thyroid dysfunction and possible cochlear and vestibular malformations, is one of the most common form of autosomal recessive SHL [Koffler *et al.*, 2015]. It is most frequently caused by biallelic pathogenic variants in *SLC26A4*, but cases of double heterozygosity for one mutation in *SLC26A4* and one mutation in either *KCNJ10* or *FOXI1* genes have also been described [Yang *et al.*, 2007; Yang *et al.*, 2009]. While KCNJ10 potassium channels and pendrin (Paragraph 1.4.1) are essential for the generation and maintenance of the EP [Yang *et al.*, 2009], *FOXI1* is a transcription factor that regulates *SLC26A4* expression [Yang *et al.*, 2007]. In addition, pendrin is expressed in thyroid follicular cells, where it participates to the apical iodide secretion [Bizhanova and Kopp, 2009]. Mutations in these genes are also associated with NSHL (DFNB4, MIM#600791). However, since the association with syndromic features such as goitre could be delayed, a first diagnosis of NSHL could change in PDS over time [Parker and Bitner-Glindzicz, 2015].

AS is a progressive inherited renal disease characterised by sensorineural hearing loss and ocular anomalies. It is caused by mutations in the *COL4A3*, *COL4A4*, and *COL4A5* genes, encoding the α3, α4, and α5 chains of collagen type IV. Indeed, the α3α4α5(IV) collagen triple helix is a major structural component of the glomerular basement membrane and of the basement membranes in the cochlea and eye. Mutations in these genes cause the formation of abnormal collagen chains which fail to be incorporated in the basement membranes and result in the persistence of the foetal collagen α1α1α2(IV) heterotrimer,

which is more susceptible to endoproteolysis. Therefore, the arrest of the developmental isoform switch undermines the integrity and mechanical stability of the basement membranes, consequently leading to the clinical phenotype observed in AS patients [Hudson *et al.*, 2003; Kashtan, 2019]. While the most common form of AS is X-linked (MIM#301050, ~65% of cases) and it is caused by mutations in the *COL4A5* gene, mutations in *COL4A3* and *COL4A4* genes are associated with both autosomal dominant (MIM#104200, ~20%) and autosomal recessive (MIM#203780, ~15%) AS. Renal disease usually progresses from microscopic hematuria to end-stage renal disease in all males with X-linked AS and in all males and females with autosomal recessive AS [Kashtan, 2019]. Some heterozygous mutations in *COL4A3* and *COL4A4* genes can cause benign familial hematuria (MIM#141200), a milder phenotype which is characterised by persistent microscopic hematuria, rarely combined with progressive proteinuria and end-stage renal disease [Stokman *et al.*, 2016]. AS is characterised by a high allelic heterogeneity, which is paralleled by a high inter- and intra-familial phenotypic variability [Lemmink *et al.,* 1997].

PRLTS (MIM #233400), is an autosomal recessive disorder that, to date, has been associated with mutation in 6 genes (*CLPP*, *C10ORF2*, *ERAL1, HARS2*, *HSD17B4* and *LARS2*) [Newman *et al.*, 2018]. It is characterised by sensorineural HL in both sexes and ovarian dysgenesis in females, so that it could be confused with NSHL in males from clinical diagnosis [Parker and Bitner-Glindzicz, 2015]. As the clinical diagnosis is mostly based on suggestive findings and family history, the genetic diagnosis can be very useful to confirm a suspect of PRLTS and to exclude other causes of HL and ovarian dysfunction.

1.5. Next-generation sequencing and its application in studying the genetic bases of human hearing loss

The high level of genetic and allelic heterogeneity of hearing loss and the lack of a precise phenotype-genotype correlation make its genetic diagnosis with traditional screening

methods (i.e. Sanger sequencing) extremely challenging. In fact, while mutations in the *GJB2* gene are the most common cause of autosomal recessive NSHL in many populations, the vast majority of deafness-causing variants are private [Duman and Tekin, 2012]. Moreover, Sanger sequencing is particularly expensive and time-consuming, as many HL genes have numerous and/or long exons. The development of next-generation sequencing (NGS) technologies, allowing the massively parallel analysis of millions of sequences in a single run, has dramatically reduced the cost and the time of DNA sequencing [Bamshad *et al.*, 2011]. In addition, as 85% of disease-associated variants affects exonic or splice-site regions, exome represents a subset of the genome where to efficiently search for pathogenic variants. Indeed, whole-exome sequencing (WES) has provided a fast and cost-effective alternative screening method to search for alleles underlying Mendelian disorders [Bamshad *et al.*, 2011].

Despite being characterised by different chemistry and protocols, the past and currently available NGS platforms share a similar workflow for WES. First, the genomic DNA is randomly fragmented, and the ends of the obtained fragments are ligated to platform-specific adaptors. Then, the genomic regions of interest (e.g. the exome) are enriched, either by PCR amplification or by capturing by hybridization, and the resulting DNA molecules are clonally amplified. Each DNA molecule is amplified in an isolated space, generally by emulsion PCR (Roche 454 system, ThermoFisher Ion Torrent) or bridge PCR (Illumina platforms), and sequenced in a parallel fashion. The obtained sequences, called "reads", are usually generated through sequencing-by-synthesis reactions, in which the nucleotides are detected as they are incorporated into a nascent DNA strand. For Ion Torrent, this is made possible by the detection of the changes in pH resulting from the incorporation of the different nucleotides. On the contrary, for Roche 454 and Illumina platforms, the sequence analysis is based on the use of optical detection, whether of light via pyrosequencing (Roche) or of the fluorescence emitted by specific reversible

terminators (Illumina) [Shendure and Ji, 2008; Levy and Myers, 2016]. Massive parallel sequencing allows to sequence at the same time many randomly overlapping DNA fragments, so that each nucleotide can be included in different fragments, which are analysed several times. This strategy allows to increase the depth of coverage and, consequently, the sequencing accuracy [Lin *et al.*, 2012]. In addition, Illumina technology enables the "paired-end" sequencing, in which each DNA fragment is sequenced from both ends, improving the mapping of reads to the reference sequence [Levy and Myers, 2016]. Even if Illumina is currently the predominantly used NGS technology, the so-called third-generation technologies are emerging, with the advantage of overcoming the requirement of DNA amplification and allowing the sequencing of longer reads. Indeed, platforms such as the Oxford Nanopore are valuable for the resolution of structural variations in the genome, which are hardly detected using short-read sequencing [Levy and Myers, 2016; Heather and Chain, 2016].

Concerning HL, NGS has allowed not only the creation of clinical panels to test known deafness genes [Brownstein *et al.*, 2012], but also the discovery of novel HL genes. For instance, NGS has enabled the identification of the causative gene in "orphan" NSHL loci, which were mapped through linkage analyses in large families before the introduction of NGS technologies [Rehman *et al.*, 2010; Santos-Cortez *et al.*, 2014]. In addition, WES alone is particularly helpful for small pedigrees, which are not suitable for linkage analysis [Vona *et al.*, 2015], and has been successfully applied to the discovery of novel NSHL and SHL genes and mutations [Chatzispyrou *et al.*, 2017; Yousaf *et al.*, 2018b]. Importantly, since the first publications describing the discovery of novel deafness genes using NGS in 2010 [Rehman et al., 2010; Walsh et al., 2010] a total of 46 novel NSHL genes have been identified by targeted sequencing or WES – alone or coupled with linkage analyses [Vona *et al.*, 2015; Hereditary Hearing Loss Homepage, https://hereditaryhearingloss.org/]. However, the pace of the discovery of novel deafness genes is not always followed by a

clear understanding of their role in the hearing function. For this purpose, animal models – primarily mice - can be particularly useful.

1.6. The mouse as a model to study the genetics of hearing loss

The study of the physiology of hearing and of the pathological mechanisms of hearing loss in humans is limited by the inaccessibility of the inner ear in living humans and demands the use of animal models. The laboratory mouse provides several advantages that are generally convenient for the study of the genetics bases of human disease, such as the relatively short gestation time (3 weeks), the short life span (enabling to follow the evolution of a disease in a contracted time period), the availability of inbred strains (isogenic mice, that ensures the study of a genetic feature in a homogenous genetic background), the high protein sequence conservation between mouse and human (~99% of genes in mice have a human orthologue), the mouse-human conserved synteny relationship (conservation of gene content and order on chromosomes) and the ease of genetic manipulation [Mouse Genome Sequencing Consortium *et al.*, 2002; Kikkawa *et al.*, 2012]. In addition, the morphological and functional similarity of the hearing structures [Kikkawa *et al.*, 2012] and the conservation of the principal inner ear development stages in mice and humans (Paragraph 1.2) make the mouse a model particularly suitable for hearing research.

Mouse mutants have proven a highly effective research tool to investigate auditory function, as well as to characterise the molecular processes that are impaired once a mutation occurs in one of the critical genes for hearing.

In many cases, the discovery of the ortholog mouse deafness gene has led to the subsequent identification of the human deafness gene. For example, a mutation in *Myo6* gene was first identified as the cause of deafness in the spontaneous mouse model Snell's waltzer [Avraham *et al.*, 1995]. Only some years later *MYO6* was associated with autosomal

dominant (DFNA22) [Melchionda *et al.*, 2001] and autosomal recessive (DFNB37) [Ahmed *et al.*, 2003a] hearing loss in humans.

In addition to the study of spontaneously occurring mutations, the investigation of chemically-induced mouse mutants has allowed the discovery of many genes involved in hearing and balance defect, thus helping to elucidate the physiological process of hearing and the pathological mechanism underlying deafness [Hardisty-Hughes *et al.*, 2010].

Finally, knock-out and knock-in models, allowing the targeted mutation of specific genes, have enabled the understanding of the function of specific components of the hearing apparatus and the study of their interaction. For example, the targeted knock-out of *Slc26a4* gene provided insights into the function of pendrin, which had already been linked to DFNB4 and PDS [Everett *et al.*, 2001]. Importantly, in 2006 the Sanger Institute's Mouse Genetics Project (MGP) was established with the aim of producing mouse knock-outs of every gene and screening them for a large set of traits, including hearing defects [White *et al.*, 2013]. This approach has already led to the identification of several novel genes associated with deafness and the data collected are made publicly available on the database of the International Mouse Phenotyping Consortium (IMPC, http://www.mousephenotype.org), of which the MGP is a founder member.

Besides these international phenotyping efforts, the generation of mouse knock-outs and knock-ins is particularly useful to study the effect of human mutations identified in novel candidate deafness genes and, ultimately, to demonstrate the involvement of these genes in hearing loss. The process of generating specific mouse mutants harboring the corresponding mutations identified in human genes has been facilitated by the development of the CRISPR/Cas9 genome-editing technology, which enables the fast and efficient generation of these mouse models [Wang *et al.*, 2013].

Despite the undeniable role of mouse models in the advancement of the knowledge in the auditory field, it is necessary to point out that there are also examples of mouse models that

don't recapitulate the human phenotype. The lack of a suitable mouse model hampers the elucidation of the function of these genes, as well as of the pathophysiological mechanisms leading to deafness [Hosoya *et al.*, 2016].

1.7. Thesis aims

The overall goal of this project is to provide new insights into the genetic bases of deafness focussing on 11 families affected by hereditary hearing loss and applying whole-exome sequencing as a comprehensive screening for candidate causative variants.

In the cases of identification of variants of unknown significance in known deafness genes, this work aims at functionally characterise the candidate mutations by specific *in-silico* and *in-vitro* studies.

In the case of variants identified in novel candidate deafness genes, my project aims at demonstrating their involvement in the pathogenesis of hearing loss. In particular, *DIAPH2* gene was identified as a novel candidate NSHL gene in one Italian family with a recessive inheritance pattern. *DIAPH2* codes for a member of the *Diaphanous*-related formins, which are involved in the nucleation of linear unbranched actin filaments. Interestingly, other members of the gene family, namely *DIAPH1* and *DIAPH3*, are associated with different types of hearing loss. This work aims at exploring the role of this gene in hearing function by studying the effect of the mutation by *in-vitro* studies and by using mouse models. The specific aims of this part of the work are:

1. Evaluate the effect of the *DIAPH2* candidate mutation
2. Investigate whether *DIAPH2* is mutated in other deaf individuals
3. Analyse the expression of the mouse ortholog protein Diap2 in the cochlea
4. Evaluate the hearing phenotype of *Diaph2* knock-out and knock-in mouse models.

2. Materials and Methods

2.1. Patients

A total of 11 HL families, including 8 families affected by dominant or recessive NSHL and 3 families diagnosed with Alport syndrome, were enrolled in the study described in the present thesis (Table 2.1).

The study was approved by local Ethical Committees of the Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico of Milan and of ASST Grande Ospedale Metropolitano Niguarda and was performed according to the Declaration of Helsinki. Signed informed consent was obtained from all participants and from parents of subjects younger than 18 years. Patients' samples were anonymised and sensitive data were treated according to the Italian legislation and the European General Data Protection Regulation (GDPR).

All available patients underwent ear, nose and throat examinations, and pure-tone audiometry, in accordance with International Standard Organization (ISO 8253-1-3) protocols. Nongenetic causes of deafness were excluded based on clinical history. Alport syndrome was primarily diagnosed by renal biopsy.

All recruited probands were negative for mutations in the gap-junction proteins connexin 26 and 30 (*GJB2*, *GJB6*), and the mitochondrial 12S rRNA (*MTRNR1*) genes. The enrolment of patients and the ascertainment of their family history and phenotype, the collection of biological samples and the screening of *GJB2*, *GJB6* and *MTRNR1* genes were performed by the Audiology Unit of the Policlinico Hospital of Milan and by the Laboratory of Medical Genetics of the Niguarda Hospital of Milan.

For screening purposes, a cohort of 125 Italian audiologically-tested normal-hearing controls (mean age at withdrawal 32±9) and a cohort of ~650 deaf individuals were included in the study.

Table 2.1. HL families selected in the study.

Family	Inheritance	Country	HL phenotype	Additional information
NSHL3	AR/XR	Italy	Prelingual, downsloping audiogram	
NSHL7	AR	Morocco	Post-lingual, severe-to-profound, progressive	Consanguineous family
NSHL16	AD	Italy	Post-lingual, high intrafamiliar variability	
NSHL17	AD	Italy	Post-lingual, late-onset, asymmetric	
NSHL18	AR	Italy	Early-onset, moderate-to-severe	Consanguineous family
NSHL19	AD	Italy	Post-lingual, ski-slope audiogram	
NSHL20	AD	Italy	Post-lingual, late-onset, ski-slope audiogram	
NSHL22	AR/XR	Italy	Early-onset, severe; vertigo reported	
AS1	AD/XD	Italy	Post-lingual	Alport syndrome
AS2	AD/XD	Italy	Post-lingual	Alport syndrome
AS3	AR	Italy	Post-lingual	Alport syndrome

AR/AD: autosomal recessive/dominant, XR/D: X-linked recessive/dominant.

2.2. Whole-exome sequencing (WES)

Genomic DNA was extracted from peripheral blood according to standard protocols.

For family NSHL7, WES was performed starting from 1 μg of genomic DNA and using the SeqCap EZ Human Exome Library v.2.0 (Roche NimbleGen) kit, following the manufacturer's instructions. Paired-end 75-bp (base pairs) libraries were sequenced on an HiSeq 2000 (Illumina) at the service facility of the Yale Center for Genome Analysis. For all the other families I used the Nextera Rapid Capture Exome Enrichment kit (Illumina) to prepare 150-bp paired-end libraries, starting from 50 ng of genomic DNA and following the manufacturer's instructions. Briefly, genomic DNA was fragmented and linked to adapter sequences and subsequently PCR-amplified. Then, the amplified DNA library was hybridised with capture probes to select the targeted region of interest (the exome). Up to

12 different libraries were pooled, subjected to a second round of hybridisation, to ensure high specificity of the captured regions, and PCR amplified. Finally, the obtained exome libraries were sequenced on a NextSeq500 platform (Illumina) at the Humanitas Genomic facility. Our WES data analysis pipeline included: 1) run quality control (with FastQC), 2) alignment to hg19 reference genome (using the BWA program), 3) PCR duplicates removal and realignment (Picard and GATK software), 4) variant calling and annotation (GATK and Annovar programs), and 5) variant prioritisation (in-house pipeline). In particular, common single nucleotide variants (SNVs) and insertion-deletions were excluded by filtering against 1000 Genome Project, National Heart, Lung, and Blood Institute Exome Sequencing Project (NHLBI-ESP; 6500 exomes, http://evs.gs.washington.edu/EVS/), and dbSNP135 (except variants flagged as clinically relevant) databases (see Paragraph 3.1).

2.3. Linkage analysis

Thirteen individuals (III7, III8, III10, III11, IV1 to IV8, IV10, Fig. 3.2A) from the NSHL7 family were genotyped using the Infinium OmniExpressExome-8 v1.4 BeadChip array (Illumina), following the manufacturer's instruction. This array comprises more than 950,000 polymorphic markers, providing a comprehensive coverage of the genome for linkage analysis and association studies. In brief, genomic DNA was amplified by whole-genome amplification, fragmented and subsequently hybridised onto an array containing millions of beads coated with copies of DNA oligonucleotide probes specific to a locus of interest. The hybridised DNA was used as template for an enzymatic single-base extension of each probe, in correspondence with a specific SNP. Then, the specific staining of the added nucleotide with fluorescence dyes ensured the discrimination between the different alleles for each bead type. Finally, the detection of fluorescence signal intensities on the array was performed with the iScan System (Illumina) and intensity files were converted into genotype calls by Genome Studio (Illumina). Data were analysed using the

EasyLinkage program, and the GeneHunter algorhitm [Lindner and Hoffmann, 2005]. Linkage analyses were performed by Dr. Michela Robusto.

2.4. Mutational screening

All candidate variants identified by WES were confirmed by Sanger sequencing.

Briefly, the genomic regions carrying the variants were PCR amplified using sets of primers designed based on the known GenBank sequence (Table 2.2). PCRs were performed on 10–20 ng of genomic DNA, following standard procedures. Direct sequencing of PCR amplicons was performed on both strands using the BigDye Terminator Cycle Sequencing Ready Reaction Kit v.1.1 and an automated ABI-3500DX DNA sequencer (Applied Biosystem). The Variant Reporter software (Applied Biosystems) was used for variant detection.

Additionally, a High-Resolution Melting (HRM) assay to screen *SLC22A4* exon 1 was performed on our in-house database of 125 audiologically-tested normal-hearing controls. A new specific primer pair (Table 2.2) was designed to PCR amplify a 265-bp genomic DNA fragment encompassing the *SLC22A4* c.338G>A mutation, and PCR amplicons were screened for mutations by HRM on a LightCycler 480 using the HRM Master kit (Roche) and a touch-down protocol. Amplicons were analysed with the Gene Scanning Software (Roche). All *SLC22A4* exons and adjacent splice junctions were also screened - at the Laboratory of Medical Genetics of the Niguarda Hospital of Milan - in 7 additional unrelated NSHL patients of North African origin (2 from Morocco, 4 from Egypt and 1 from Tunisia).

For family NSHL18, serial PCRs were performed to better define the extent of the deleted region encompassing the entire *OTOA* gene (Table 2.2).

Table 2.2. Primers used in mutational screening.

Family	RefSeq Accession No	PCR assay	Primer sequence (5'-3')	Size (bp)
NSHL3	NM_006729.4	DIAPH2	AACCTCATTGTGGGGGTATG TCTGCTCAATCACTTTGAATCC	495
NSHL7	NM_003059.2	SLC22A4	CCAATGGCTTCAATGGTATGT CCGCTGCTGGAAGTATGAAC	440
		SLC22A4_HRM	CTGGCGCAACAACAGTGTC ACGCAGAGGGAGGGTCAG	265
NSHL16	NM_005422.2	TECTA	CTTCGACGGCCACTACTACA CGGGCTTCTTCTTGTTGATG	320
	NM_172163.2	KCNQ4	GAGCCATGCGTCTCTGAGC CAGCACGTTGTAGACCCAGT	356
	NM_001172411.1	VANGL1	GTTTCCCACCCTCGTGGTAT GAGGGTACTGGCATCAGGAG	216
	NM_000817.2	GAD1	GGATGGCAGTTTTTCCAAGA CCACATCAGCCAGAACTTGA	325
	NM_001369.2	DNAH5	CACCCAACTGACCACAACAG GAGGCAAGTGTGACCCAAAT	381
NSHL17	NM_138691.2	TMC1	TATGCAGAACAATTTGCCTTTCAG AGACACCGATTGTATTCTCCTC	244
NSHL18	NM_170664.2	5'_OTOA1	CCTTGAGTAGCGTGATGTTCC GGCTCCTACTGTTGCTTCCA	2732
		5'_OTOA2	CACGCCTATCACAGAAAGCA GAGTGCAACTGGATGGGACT	5184
		5'_OTOA3	ACGGTGAATTTGGTGGAGAG GTGGGCTCCAGTTGATGAAT	4991
		OTOA	ATGAGGTGATGCCAGGACTC TCTCCCCATTTCTCTGGCTA	467
		3'_OTOA1	GGGGTAAAACTGTCCCTGGT TGTTTCCACCCTTTCTTTGG	3360
		3'_OTOA2	CGCCCAGTCAACAGATAGGT AGCAGGTGGGTGAAGCTCT	5158
	NM_021072.3	HCN1	TGATTTCTGATACACCCTAACAATG TGCATACTGAGGACAATAAATCTT	231
NSHL20	NM_001135058.1	COCH	TCTTACCACATGCCCAACTG ATGAGGCGGAAATTGCTATC	170
AS1	NM_000495.4	COL4A5	CCTGTTTCCAATCCTTCCAT TGCAGTGACAGCCTCCATAC	359
AS2	NM_000495.4	COL4A5	AGTCGGCCTCTATCGTTGTC GTAAGCCAGGTTCCCCCTTC	369
AS3	NM_000091.4	COL4A3	GACCGAGCCCTACAAAACC GTGGAGGAGGGATGGAAGTG	282

2.5. Expression vector preparation

For the functional characterisation of candidate splicing mutations, the genomic DNA regions containing the identified variants were PCR amplified from patients' genomic DNA (with primers including *NdeI* restriction sites; Table 2.3) and cloned in the hybrid alpha-globin-fibronectin minigene plasmid (modified pBS-KS) [Baralle *et al.*, 2003].

For the functional characterisation of the COL4A3 signal-peptide deletion, the genomic DNA region coding for the predicted signal peptide was PCR amplified from the patient's genomic DNA (with primers including *XhoI* and *KpnI* restriction sites; Table 2.3) and cloned in-frame upstream of the EGFP (Enhanced Green Fluorescent Protein) coding region in the pEGFP-N1 plasmid (Clontech).

Finally, expression vectors containing human *DIAPH2*, *SLC22A4* and *HCN1* coding sequences were kindly provided, respectively, by the laboratory of Professor Marino Zerial (pCMV_HA-hDIA2B) [Gasman *et al.*, 2003], of Professor Giovambattista Pani (pCDNA3.1_OCTN1) [Martini *et al.*, 2012] and of Professor Christel Depienne (pCDNA3.1_hHCN1) [Nava *et al.*, 2014]. Point variations were introduced by site-direct mutagenesis, using specific sets of primers (Table 2.3) and the QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies).

The plasmid encoding a constitutively active form of RhoA, used in DIAPH2 immunolocalisation studies, was provided by the laboratory of Professor Taroh Iiri (pCDNA3.1_Myc-RhoA_G14V) [Oishi *et al.*, 2012].

Recombinant plasmids were extracted with the PureYield Plasmid Midiprep System (Promega), and verified by sequencing.

Table 2.3. Plasmids used in transfection experiments.

Family	Experiment	Plasmid	Cloning/mutagenesis primer sequence (5'-3')*
NSHL3	splicing	pBS-KS-*DIAPH2*_ex8_wt pBS-KS-*DIAPH2*_ex8_mut	ggaattccatatgCAAGGAGCAACAACTCCAAGA ggaattccatatgTCTGCTCAATCACTTTGAATCC
	immunolocalisation	pCMV_HA-hDIA2B_wt	Gasman *et al.*, 2003
		pCMV_HA-hDIA2B_mut	GTTGGAGAAGAGAACgTTCTAGATAAACTTT AAAGTTTATCTAGAAcGTTCTCTTCTCCAAC
	immunolocalisation	pcDNA3.1_Myc-RhoA_G14V	Oishi *et al.*, 2012
NSHL7	immunolocalisation	pCDNA3.1_Myc_OCTN1_wt	Martini *et al.*, 2012
		pCDNA3.1_Myc_OCTN1_mut	TGGAGCAGGAGAGCTaCCTGGATGGCTGGGA TCCCAGCCATCCAGGtAGCTCTCCTGCTCCA
NSHL17	splicing	pBS-KS-*TMC1*_ex19_wt pBS-KS-*TMC1*_ex19_mut	ggaattcCATATGCAGAACAATTTGCCTTTCAG ggaattccatatgAGACACCGATTGTATTCTCCTC
NSHL18	immunolocalisation, patch-clamp	pCDNA3.1_HCN1_wt	Nava *et al.*, 2014
		pCDNA3.1_HCN1_Y417C	AGCAAGTGGAACAATgCATGTCATTCCATAAG TTATGGAATGACATGcATTGTTCCACTTGCTT
	patch-clamp	pCDNA3.1_HCN1_Y417F	AGCAAGTGGAACAATtCATGTCATTCCATAAG CTTATGGAATGACATGaATTGTTCCACTTGCT
AS1	splicing	pBS-KS-*COL4A5*_ex29_wt pBS-KS-*COL4A5*_ex29_mut	ggaattccatatgACCCTGTTTCCAATCCTTCCA ggaattccatatgGCCGGGCCATGATTTTATT
AS3	immunolocalisation	COL4A3-SP-wt-hybEGFP COL4A3-SP-mut-hybEGFP	ggcctcgaGGTGGCCTGAGAGCCTGA agaggtaccTGGAGGAGGGATGGAAGTG

* Lowercase letters indicate nucleotides added to the primers to introduce the restriction sites (note that, for pBS-KS-*TMC1*_ex19 plasmids, the *NdeI* restriction site located in intron 18 was used). The red lowercase letter in mutagenesis primers indicate the variant nucleotide.

2.6. Cell cultures and transfections

HeLa and HEK293 cells were cultured in Dulbecco's modified Eagle medium supplemented with 2 mM L-glutamine, 10% fetal bovine serum (FBS) and antibiotics (100 U/ml penicillin and 100 µg/ml streptomycin; Euroclone) and grown at 37°C in a humidified atmosphere of 5% CO_2 and 95% air, according to the standard procedures.

UB/OC-2 cells, deriving from the organ of Corti of E13 temperature-sensitive Immortomouse [Rivolta and Holley, 2002], were cultured in Minimum Essential Medium (MEM; ThermoFisher Scientific) supplemented with GlutaMAX, 10% FBS and 50 U/ml γ-IFN (gamma-interferon, PeproTech) and grown at 33°C in a humidified atmosphere of 5%

CO_2. These conditions maintain cells in the proliferative state. To differentiate the cells, γ-IFN was rinsed off and cells were cultured in MEM with only GlutaMAX and 10% FBS at 33°C for 2 days and then moved to 39°C in a humidified atmosphere of 5% CO_2.

For splicing studies, an equal number of cells ($2.5*10^5$ HeLa or UB/OC-2 in proliferative state; $3*10^5$ HEK293 or UB/OC-2 in differentiating conditions) were transiently transfected with 1 μg of either wild-type or mutant recombinant pBS-KS vector (Table 2.3).

For immunofluorescence assays, $2*10^5$ HeLa or $2.5*10^5$ HEK293 cells were seeded on 22x22 mm glass coverslip and transfected with 500 ng of plasmid encoding either the wild-type or the mutant protein under study. For immunolocalisation of DIAPH2 protein, an equimolar quantity of the pcDNA3.1_Myc_RhoA-G14V vector was also cotransfected.

The Jet-PRIME reagent (Euroclone) was used for transfections in HeLa and HEK293 cells, whereas Lipofectamine 2000 (Thermo Fisher scientific) was used as transfection reagent in the UB/OC-2 cell line.

For electrophysiology analyses, $2.5*10^5$ HEK293 cells were seeded on 20x20 mm glass coverslip and cotransfected with 1 μg of plasmid encoding either the wild-type or one of the mutant HCN1 channels (Table 2.3) and with 0.5 μg of pEGFP using the Lipofectamine 2000 reagent (Thermo Fisher scientific) and following the manufacturer's instructions.

2.7. Splicing analyses

Total RNA was isolated from cells 24 hours after transfection, using the EuroGold TriFast reagent (Euroclone). Random hexamers and the ImProm-II Reverse Transcriptase (Promega) were used to perform first-strand cDNA synthesis, starting from 500 ng of total RNA, according to the manufacturer's instructions. Of a total of 20 μL of the RT reaction, 1 μL was used as template for amplifications, using primers annealing to the flanking α-globin/FN1 exonic sequences (α2-3: 5'-CAACTTCAAGCTCCTAAGCCACTGC-3' and

Bra2: 5'-CAATGGATGGGGGTGGAG-3'). RT-PCRs were performed under standard conditions using the GoTaq DNA Polymerase (Promega) on a Mastercycler EPgradient (Eppendorf).

Competitive-fluorescent RT-PCRs were performed on RNA from transfected cells, using the same oligonucleotide pair adopted for splicing assays, with the reverse Bra2 primer labelled with 6-FAM (6-Carboxyfluorescein). Amplified fragments were separated on an ABI-3500DX sequencer and quantitated by the GeneMapper v4.0 software (Applied Biosystems). The sum of all fluorescence peak areas in a single run was set equal to 100%, and the relative quantity of each transcript expressed as a fraction of the total.

2.8. Cell immunofluorescence experiments

Cells seeded on coverslips were fixed 24 or 48 hours after transfection using 4% paraformaldehyde for 10 minutes at room temperature. The cells were washed twice for 15 minutes in 100 mM glycine and permeabilised/blocked for 30 minutes in GDB (gelatin dilution buffer: 0.1% gelatin, 0.3% Triton X-100, 450 mM NaCl, 20 mM phosphate buffer pH 7.4). Coverslips were washed in HS (high salt: 500 mM NaCl, 20 mM phosphate buffer pH 7.4) buffer and incubated for 2 hours with primary antibodies diluted in GDB (Table 2.4). The cells were then washed 3 times in HS buffer and incubated 1 hour with fluorophore-conjugated secondary antibodies (Table 2.4). After 3 washes in HS buffer, cell nuclei were counterstained with DAPI (4',6-Diamidino-2-Phenylindole) for 5 minutes and cells were washed once in HS and LS (low salt: 150 mM NaCl, 10 mM phosphate buffer pH 7.4) buffers and bidistilled water. Coverslips were mounted with FluorPreserve reagent (Merck Millipore). Alternatively, ProLong Diamond Antifade Mountant with DAPI (Molecular Probes) was used for mounting and nuclei counterstaining. For actin staining, cells were labelled with ActinGreen 488 ReadyProbes Reagent (Molecular Probes) before mounting, following the manufacturer's instructions. Confocal images were acquired using

a 60x UPLSAPO oil-immersion objective (N.A. 1.35, Olympus) with an Olympus FluoView FV1000 confocal microscope or a 63x HC PL APO 63x oil-immersion objective (N.A. 1.40, Leica) with a Leica TCS SP8 confocal microscope at a resolution of 1 airy unit. Identical gain, offset, exposure, and laser-power settings were applied to wild-type and mutant conditions in every experiment.

Images were analysed with Fiji (Fiji is just ImageJ) software [Schindelin *et al.*, 2012].

Table 2.4. Antibodies used in cell immunofluorescence experiments.

Host species	Antibody (id, manufacturer)	Working dilution	Type
mouse	Anti-HA probe (sc-7392, Santa-Cruz Biotechnology)	1:75	primary
rabbit	Anti-Myc tag (ab9106, Abcam)	1:100	primary
rabbit	Anti-HCN1 (APC-056, Alomone Labs)	1:100	primary
mouse	Anti-calnexin (sc-23954, Santa-Cruz Biotechnology)	1:75	primary
mouse	Anti-TGN38 (sc-166594, Santa-Cruz Biotechnology)	1:75	primary
goat	Dylight 549 anti-mouse (115-505-174, Jackson ImmunoResearch)	1:200	secondary
goat	Alexa Fluor 488 anti-rabbit (ab150077, Abcam)	1:200	secondary
goat	Alexa Fluor 633 anti-rabbit (A21072, Molecular Probes)	1:200	secondary

2.9. Patch-clamp analyses

Whole-cell patch clamp recordings were performed on GFP-expressing cells 24 hours after transfection using an Axopatch 200B amplifier and a Digidata 1440 analog/digital interface (Molecular Devices) as described previously [Nava *et al.*, 2014]. The extracellular solution contained 130 mM NaCl, 15 mM KCl, 0.5 mM $MgCl_2$, 1.8 mM $CaCl_2$, 10 mM glucose and 5 mM HEPES, adjusted to pH 7.4 with NaOH. The pipette solution contained 120 mM potassium aspartate, 10 mM NaCl, 10 mM KCl, 1 mM $CaCl_2$, 10 mM EGTA, 2 mM MgATP and 10 mM HEPES, adjusted to pH 7.2 with KOH. Microelectrodes were pulled

from borosilicate glass (uncoated, PG165T-10; Harvard Apparatus) and had a resistance of 2–4 MΩ. Liquid junction potential was not corrected. Electrophysiological recordings were carried out at room temperature (21–23 °C). Whole-cell voltage-clamp mode was used to investigate the ion currents through the membrane of transfected cells. A series of test pulses, ranging from −140 to +40 mV in 10-mV increments from a holding potential of −20 mV, was applied to the cells for 1 second. Currents were measured in steady state at the end of the test pulses. Data were acquired and analysed with Clampfit 10.7 (Molecular Devices) software. Patch-clamp experiments were performed by Mr. Filippo Mirabella.

2.10. Mouse studies

Mouse studies were performed at the laboratory of Professor Karen Steel (King's College London, London, UK) in accordance with the UK Home Office regulations and the UK Animals (Scientific Procedures) Act 1986 (ASPA) under UK Home Office licences, and the study was approved by the King's College London Ethical Review Committee.
Mice were culled using methods approved under these licences to minimise any possibility of suffering.

2.11. Mutant mice

The *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice (PEND colony) were produced at the Wellcome Trust Sanger Institute and have a C57BL/6N genetic background. *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice carry a CreERT2_EGFP knock-in allele that both interrupts *Slc26a4* transcription and expresses EGFP and CreERT2 fusion protein under the control of *Slc26a4* promoter. CreERT2 comprises the Cre recombinase from bacteriophage P1 fused with a variant of the human estrogen receptor harboring three mutations. These mutations abolish the binding of the physiologic ligand (17β-estradiol), but not of tamoxifen, which allows the translocation of Cre in the nucleus. When crossed with mice containing a loxP-flanked

(floxed) sequence, tamoxifen-induced Cre recombination results in the deletion of the floxed sequence only in the cells expressing *Slc26a4* of the offspring. Therefore, this mouse model is useful for the creation of cell-type-specific tamoxifen-inducible conditional knock-outs.

To assess the function of the new candidate human deafness gene *DIAPH2*, found mutated in family NSHL3, the *Diaph2*^{em2Kcl} knock-out mice (DIAH colony) and the *Diaph2*^{em3Kcl} knock-in mice (DIA colony) were generated by the Genome Editing & Embryology Core (GEEC, King's College London, London, UK) using the CRISPR-Cas9 technology. Both mice have a C57BL/6J genetic background. *Diaph2*^{em2Kcl} mice display a 19-bp deletion (NM_172493.2:c.855_873del) resulting in a frameshift and hence creating a premature stop codon in exon 9. *Diaph2*^{em3Kcl} mice display the missense point mutation NM_172493.2:c.877A>G; NP_766081.1:p.I293V in exon 8, corresponding to the NM_006729.4:c.868A>G; NP_006720.1:p.I290V variant identified in family NSHL3. Of note, three additional synonymous variants were introduced (NM_172493.2:c.852T>C, NM_172493.2:c.855T>A, NM_172493.2:c.867G>A) in order to destroy the gRNA binding and PAM site - thus preventing the re-cutting of Cas9 after successful integration of the oligo donor.

2.12. Mouse genotyping

Ear clip tissue was lysed overnight in lysis solution (200 mM NaCl, 100 mM Tris HCl pH 8.0, 5 mM EDTA pH 8.0, 0.2% SDS) with 1 mg/mL proteinase K at 55ºC. Lysed samples were diluted 3 times in MilliQ water and, after centrifugation, the supernatant was diluted 20 times. PCRs were performed on 2 μL of diluted supernatant using the DreamTaq Hot Start Green PCR Master Mix (Thermo Fisher Scientific) and specific sets of primers (Table 2.5, Fig. 2.1). PCR products were separated by a 2% agarose gel electrophoresis.

For mice of PEND colony, three PCR assays (Table 2.5) were used to identify the wild-type (PEND wt) or mutant (PEND mut) allele and the presence of the cassette (by

detecting the Cre sequence). In the *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ homozygous mice, the wild-type-specific PCR assay fails due to the insertion of the cassette (Fig. 2.1A).

For DIAH and DIA colonies, PCR products were digested overnight at 37°C, with *MboII* and *NdeI* restriction enzymes (New England BioLabs), respectively. In fact, the 19-bp deletion in the *Diaph2*^{em2Kcl} allele abolishes a *MboII* restriction site present in the wild-type allele, whereas one of the synonymous variants introduced in the *Diaph2*^{em3Kcl} allele creates a new *NdeI* restriction site (Fig. 2.1B). As a result, the restriction digestion reaction of the 330 bp PCR amplicon yields two products (89 and 241 bp) for the wild-type allele in DIAH colony and two products (62 and 268 bp) for the *Diaph2*^{em3Kcl} allele in DIA colony, thus allowing to recognise each genotype after gel electrophoresis.

Table 2.5. Primers used in mouse genotyping.

Mouse colony	PCR assay	Primer sequence (5'-3')	Size (bp)	
			WT	MUT
PEND	PEND wt	ACCAATGGCAGATAAGGAGGAC TCTGACCCAGCCAGAAAATG	597	/
	PEND mut	ACCAATGGCAGATAAGGAGGAC TCGTGGTATCGTTATGCGCC	/	326
	Cre	CATTTGGGCCAGCTAAACAT TAAGCAATCCCCAGAAATGC	/	233
DIAH/DIA	DIA2	AAGAGCAATTGACCCGAAAC TGATGGCTTTAGTGGGACAA	330/330	311*/330

* *Diaph2*^{em2Kcl} allele displays a 19-bp deletion.

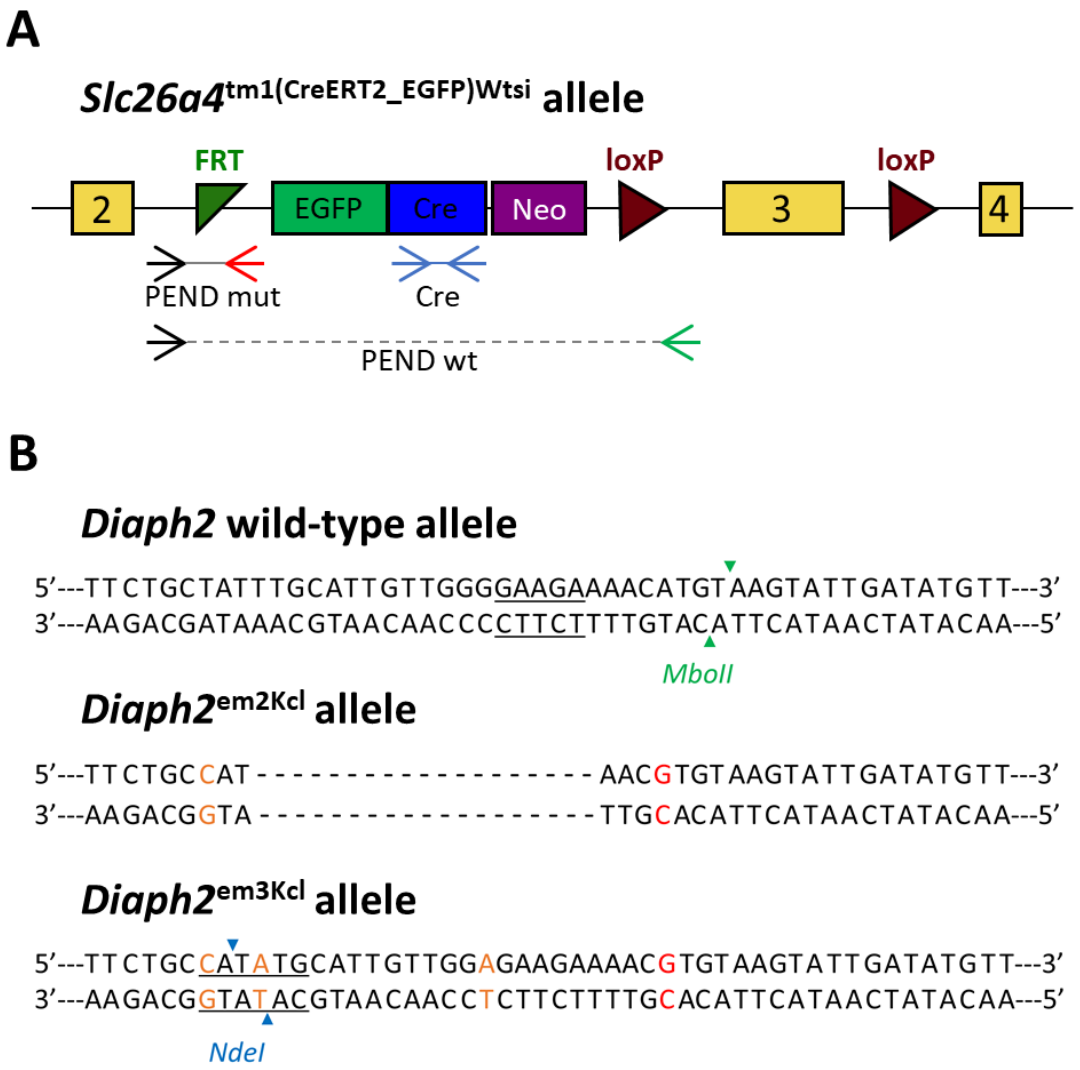


Figure 2.1. Genotyping strategies for the PEND, DIAH and DIA mouse colonies. **A.** For *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice, the three PCR assays used to verify the correct insertion of the cassette are shown (arrows indicate the primer pairs). PEND wt PCR fails in the mutant allele due to the presence of the cassette. FRT: flippase recognition target; loxP: locus of X(cross)-over in P1 (recognized by Cre recombinase). **B.** For DIAH and DIA colonies the same set of primers was used, and each genotype was assessed after a specific restriction digestion reaction. The sequence recognised by each restriction enzyme is underlined and the cutting sites are indicated by arrowheads. The dashes in *Diaph2*^{em2Kcl} allele indicate the 19-bp deletion. The missense point mutation NM_172493.2:c.877A>G; NP_766081.1:p.I293V is indicated in red, while the synonymous variants are shown in orange. Note that *Diaph2*^{em2Kcl} allele also carries the missense mutation and one of the synonymous variants.

2.13. Auditory brainstem response recordings

Auditory brainstem responses (ABR) are synchronous neural responses to auditory stimuli. ABR measurements were performed to estimate the hearing sensitivity of wild-type and mutant mice to broadband click and frequency-specific stimuli. The evoked electrical potentials of the brainstem auditory neural activity are recorded from the scalp of anesthetised animals using electrodes. The protocol used for ABR, developed by Dr. Neil Ingham, is extensively described in his 2011 publication [Ingham *et al.*, 2011].

Briefly, the mouse was weighed and tested for the presence of a Preyer reflex (a backwards movement of the pinnae elicited by the sound) using the "click-box", which emits a ~20 kHz sound at over 90 dB SPL (sound pressure level). Then, the mouse was anaesthetised by an intraperitoneal injection of ketamine/xylazine (respectively, 1 mg and 0.01 mg/g of mouse body weight) and placed in an empty cage on a heating pad until fully anaesthetised, as assessed by testing righting and pedal withdrawal reflexes. The animal was then placed in a sound-attenuating chamber on a heating blanked, in order to maintain the body temperature. The electrodes were placed subdermally at the vertex of the animal (between the ears, active electrode), in the patch of bare skin overlying the bulla behind the left ear (reference electrode), and right ear (ground electrode) (Fig. 2.2A). The mouse was laid in a natural position facing the loudspeaker, at a distance of 20 cm.

To check the correct functioning of the system, at the beginning of ABR measurements a test ABR trace was recorded presenting a 70 dB SPL click to the mouse. The heart rate was also recorded to monitor the status of the mouse. First, a series of click stimuli was presented to the mouse at increasing SPLs from 0 to 95 dB SPL in 5-dB steps, to determine the ABR threshold to the click stimuli. Then, the responses to tone pips of several frequencies (3, 6, 12, 18, 24, 30, 36, 42 kHz) and SPL (0-95 dB SPL in 5-db steps) were recorded. The tone pips had 5 msec duration with 1 msec rise/fall time and were presented at 42.6/sec rate. Finally, the 70 dB SPL test ABR and heart rate recordings were

repeated, to ensure that no deterioration in the response and in the mouse condition had taken place during the measurements.

The recovery was promoted with an intraperitoneal injection of atipamezole (0.001 mg/g body weight) and the mouse was placed back to its cage and monitored until fully conscious.

The averaged ABR waveforms, resulting from the average of 256 sweep record for each stimulus, were visualized using TraceView custom software. For each mouse, the ABR threshold to the click stimuli and to tone stimuli of each frequency presented was visually determined as the lowest stimulus level where any characteristic peak or depression of the waveform could still be recognised (Fig. 2.2B).

As the ABR thresholds for each stimulus were not normally distributed, they were analysed by the Mann-Whitney test (comparison between 2 genotype groups) or by the Kruskal-Wallis test (comparison among >2 genotype groups).

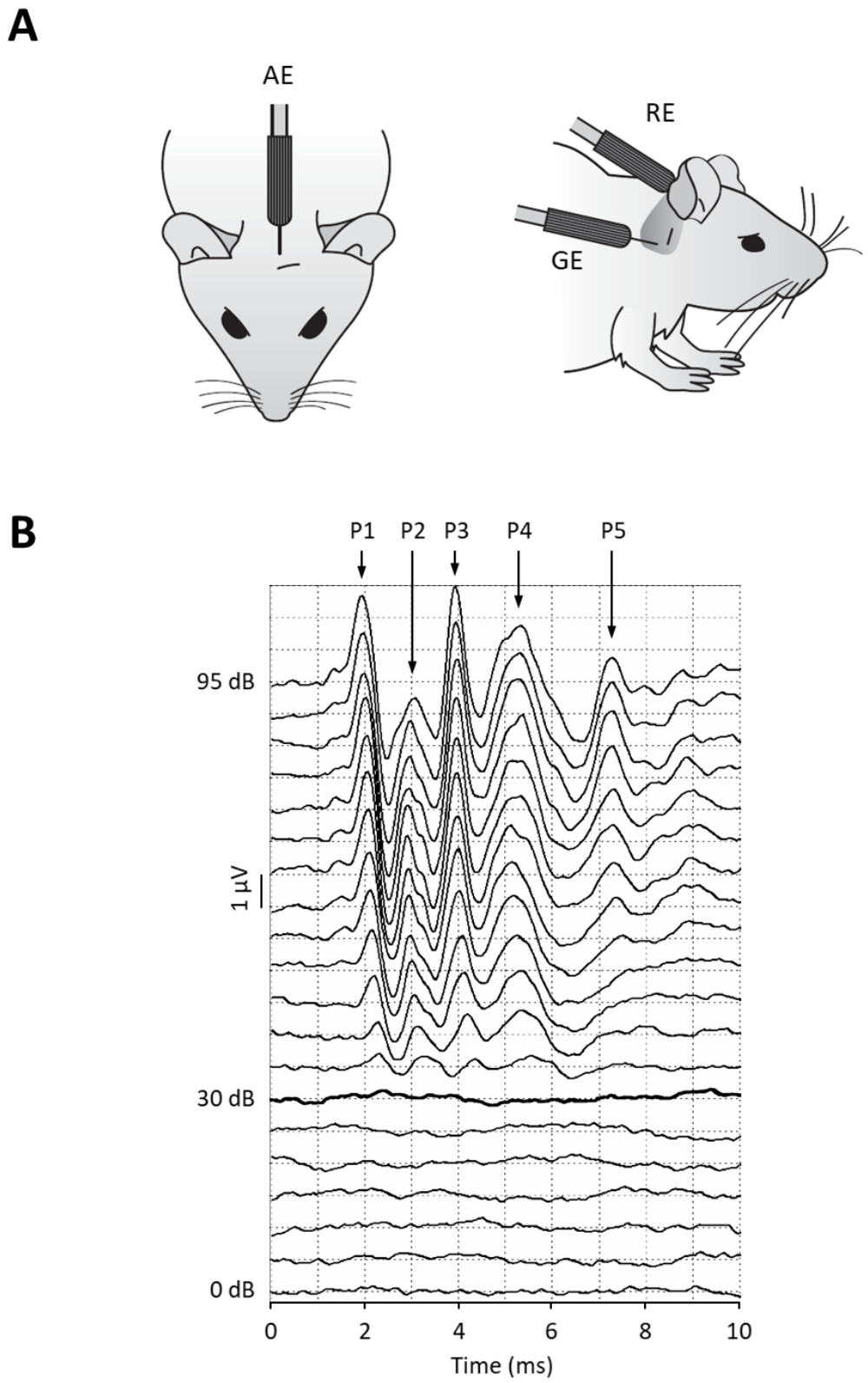


Figure 2.2. ABR measurements. A. Positioning of the active electrode (AE) and of the reference (RE) and ground electrodes (GE). Adapted from Ingham *et al.*, 2011. **B.** Example of visual determination of ABR threshold. The 5 characteristic peaks of the ABR waveform (P1-P5) are indicated by arrows. P1 is considered to represent the activities from the auditory nerve, while the other peaks are thought to be generated at higher level, within the central auditory system, even if the precise origins are not yet clear [Zhou *et al.*, 2006]. The estimated threshold ABR (30 dB) is plotted as a heavy trace.

2.14. Immunohistochemistry on wax sections

Immunohistochemical studies were performed to evaluate the expression of Diap2 protein in the mouse inner ear. For all experiments E14.5, E16.5 and P5 mice of the C57BL/6 strain and E18.5 and P0 mice with a C3HeB/FeJ genetic background were used. For each age, a minimum of three mice was tested.

Formalin-fixed paraffin-embedded wild-type mouse heads were sectioned sagittally at 8 μm thick using the Leica RM2255 microtome, mounted on Superfrost Plus slides (Thermo Fisher Scientific) and incubated at 42°C overnight. The Ventana Discovery machine and reagents (all from Roche) were used for immunohistochemistry, following the manufacturer's instructions. Briefly, the machine performed a deparaffinisation step using the EZ prep solution and antigen retrieval using the CC1 (cell conditioning 1) reagent in a heated step. The slides were incubated for 1 hour with the goat anti-Dia2 antibody (1:50, sc-10892, Santa Cruz Biotechnology) and subsequently incubated for 16 minutes with the Biotin-SP AffiniPure Donkey Anti-Goat IgG (H+L) secondary antibody (1:100, 705-065-147, Jackson ImmunoResearch). Primary and secondary antibodies were diluted in a phosphate-buffered saline (PBS) solution containing 10% FBS, 2% bovine serum albumin, 0.1% Triton-X100, and 10 mM sodium azide. After incubation with SA-HRP (streptavidin-horseradish peroxidase) the immunohistochemistry reaction was revealed using DAB (3,3'-Diaminobenzidine). Finally, slides were counterstained with haematoxylin. After staining, the slides were removed from Ventana machine and manually dehydrated using ascending ethanol concentrations (70%, 85%, 100%) and subsequently washed in acetone and xylene before mounting with Eukitt mounting medium (Sigma-Aldrich) and 24x50 mm coverslips. Brightfield images of the sections were acquired using a Zeiss Axioskop microscope connected with the Axiovision 3.0 software.

3. Results and discussion

3.1. Whole-exome sequencing data analysis of HL families

A total of 26 individuals from 11 families with HL and no molecular diagnosis were selected for WES, in collaboration with the Audiology Unit of the Policlinico Hospital of Milan and with the Laboratory of Medical Genetics of the Niguarda Hospital of Milan.

Among them, 8 families were affected by dominant or recessive NSHL and were negative for mutations in the *GJB2* and *GJB6* genes. The other 3 families were diagnosed with Alport syndrome (AS) and had been subjected to extensive and repeated screenings of the three AS-causing genes (*COL4A3*, *COL4A4*, *COL4A5*) with no conclusive results.

At least 2 affected individuals in each NSHL family were subjected to WES. In addition to the proband, we decided to sequence at least one sibling in recessive families. In dominant families, when available, we selected for WES the affected individuals most distantly related to the proband, in order to reduce the number of shared variants. Finally, in AS families, we sequenced only the proband.

For NSHL7 family, WES was performed on two individuals through an external sequence provider (the Yale Center for Genome Analysis, YCGA, Yale University, USA) using a SeqCap EZ exome kit (44 Mb target region) and a HiSeq2000 platform (Illumina). For the other 10 families, I prepared WES libraries with the Nextera Rapid Capture kit (37 Mb target region) and sequenced them on a NextSeq500 platform (Illumina) at the Humanitas Genomic facility. Overall, we obtained an average of 8.4 Gb of high-quality sequence data per exome with a mean coverage of ~100X and ~99.6% coverage of the target.

Reads were aligned to hg19 reference genome using the BWA (Burrows-Wheeler Aligner) program and variant calling and annotation were performed, respectively, with GATK (Genome Analysis Toolkit) and Annovar software. Detected variants were then prioritised using our in-house pipeline. In particular, to remove common variations, nonsynonymous/splice-site single nucleotide variants (SNVs) and insertion-deletions (indels) were filtered against 1000 Genome Project, National Heart, Lung, and Blood

Institute Exome Sequencing Project (NHLBI-ESP; 6500 exomes), and dbSNP135 (except variants flagged as clinically relevant) databases. The threshold for the maximum minor allele frequency (MAF) was set to 1% in case of recessive inheritance and to 0.1% for dominant NSHL [Bamshad *et al.*, 2011]. Passing-filter variants were then prioritised depending on the inheritance pattern of the family. The availability of exome data from multiple members of the same family allowed to select only the variants shared between affected individuals and to filter out the variants also present in normal-hearing relatives. Furthermore, a homozygosity mapping analysis was performed on WES data from consanguineous families to identify runs of homozygosity (ROH) shared by affected relatives [Pippucci *et al.*, 2011]. The frequency of the prioritised/candidate variants was further checked in the Genome Aggregation Database (gnomAD, http://gnomad-beta.broadinstitute.org/) and in an in-house cohort of 3541 Italian control exomes, to identify population-specific variants. In addition, the pathogenicity of the variations was predicted using several *in-silico* programs (CADD, Condel, FATHMM, Mutation Assessor, MutationTaster, Polyphen 2 HumDiv, Polyphen 2 HumVar and SIFT for nonsynonymous changes; NNSPLICE, Human Splicing Finder and NetGene for splice-site variants). In parallel, the EXCAVATOR2 tool [D'Aurizio *et al.*, 2016] was used to detect possible copy number variants (CNV) that escaped our standard SNV/indel analysis pipeline (Fig. 3.1). We confirmed all prioritised variants by Sanger sequencing and evaluated their segregation in the available family members.

Finally, we selected the most promising pathogenic variants focussing first on already-known deafness-causing genes and then considering novel candidates for further functional analyses. For the latter, we evaluated the expression of the encoded proteins in the cochlea and the connection with genes and pathways involved in deafness pathogenesis.

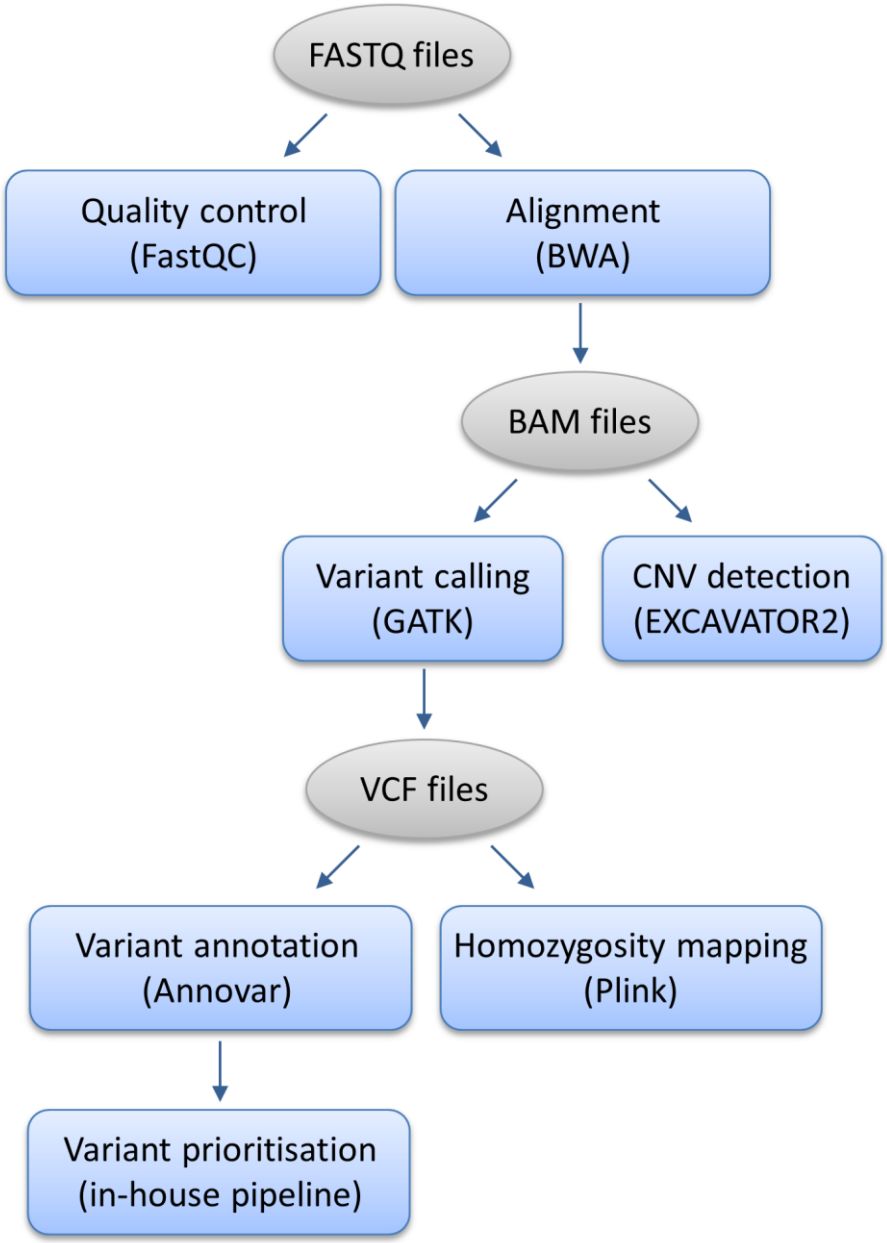


Figure 3.1. WES data analysis pipeline. The main steps of WES data analysis are shown in square shapes, with the program used in parentheses. The grey round shapes indicate the file produced and used in the subsequent steps of the analysis. The homozygosity mapping analysis was performed only in consanguineous families.

The above described approach was fundamental for the identification of pathogenic variants in known deafness genes (*SLC22A4*, *OTOA*, *COCH*, *SLC26A4*) in four NSHL families (NSHL7, NSHL18, NSHL20, and NSHL22, respectively) and highlighted the candidate NSHL-causing gene *DIAPH2* in family NSHL3 (Table 3.1).

Table 3.1. Top variants identified by WES data analyses in selected NSHL families.

			Variants		
Family	**Inheritance**	**Gene**	**cDNA localisation**	**Protein localisation**	**Status**
NSHL3	AR/XR	*DIAPH2*	NM_006729.4: c.868A>G	NP_006720.1: p.I290V	Novel candidate gene, novel variant
NSHL7	AR	*SLC22A4*	NM_003059.2: c.338G>A	NP_003050.2: p.C113Y	Known pathogenic variant in NSHL gene*
NSHL18	AR	*OTOA*	NM_170664.2: ~250-Kb deletion	/	Deletion spanning NSHL gene
NSHL20	AD	*COCH*	NM_004086.2: c.1096G>A	NP_001128530.1: p.V366M	Novel variant in NSHL gene
NSHL22	AR/XR	*SLC26A4*	NM_000441.1: c.1001G>T; c.1229C>T	NP_000432.1: p.G334V; p.T410M	Known pathogenic variants in NSHL gene

AR: autosomal recessive, XR: X-linked recessive, AD: autosomal dominant.

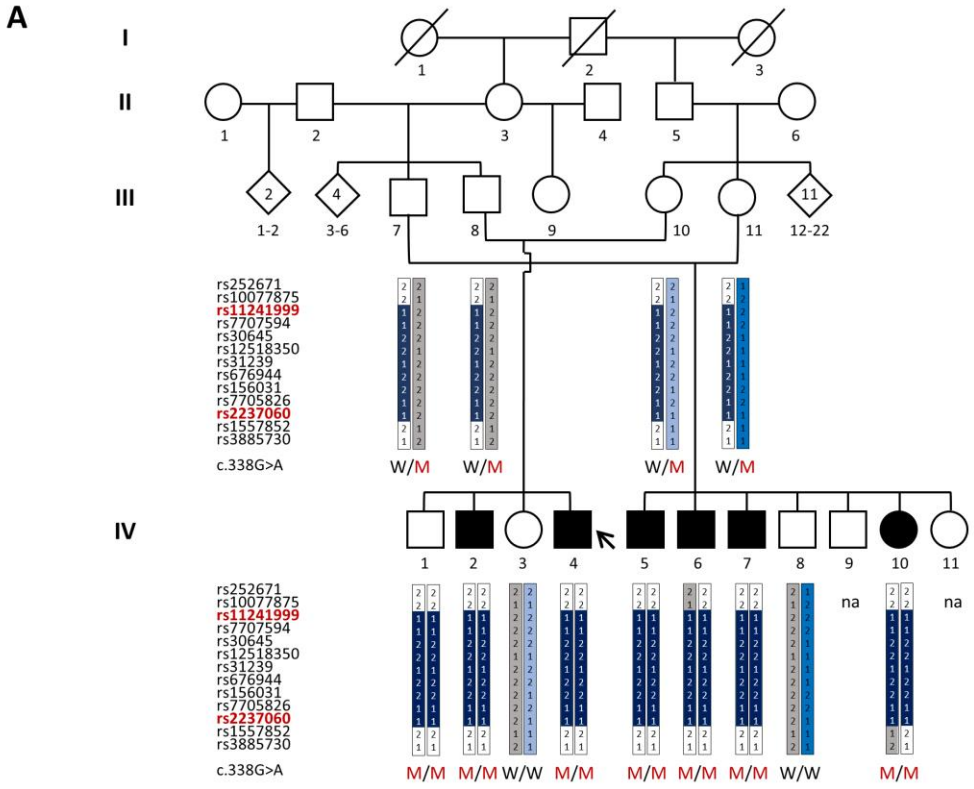
* At the time we started to study *SLC22A4*, the gene had not yet been associated with NSHL.

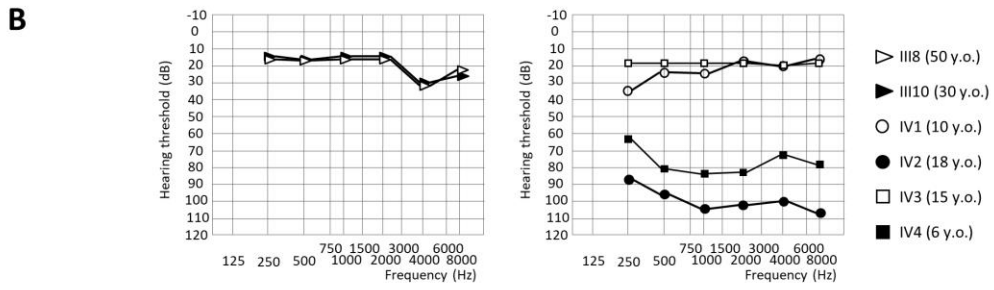
3.2. NSHL7 family: the *SLC22A4* gene

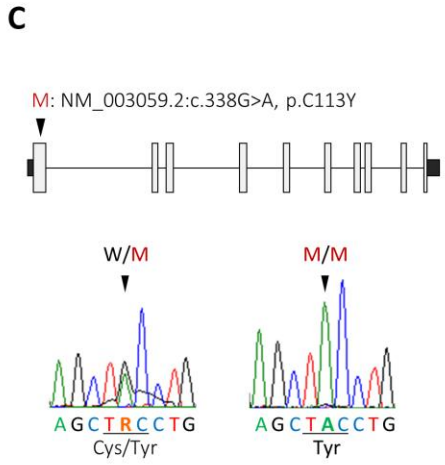
Family NSHL7 is a large Moroccan consanguineous family affected by post-lingual bilateral severe-to-profound progressive NSHL, with a recessive pattern of inheritance. The proband (IV4) is a 12-year-old male with a 24-year-old deaf brother (IV2) and a normal-hearing brother (IV1, 26 years old) and sister (IV3, 21 years old). The four deaf (three males, IV5-7, and one female, IV10) and three unaffected (two males, IV8-9, and one female, IV11) cousins of the proband, currently living in Morocco, were also available to the study (Fig. 3.2A-B). The exomes of IV2 and IV4 affected siblings were sequenced in 2014 at the service facility of the Yale Center for Genome Analysis. We ruled out the presence of pathogenic variants in known deafness genes, with particular focus on the 58

recurrent mutations included in the North African Deafness (NADf) chip [Chakchouk *et al.*, 2015]. Single nucleotide analyses of WES data prioritised 6 homozygous variants shared by IV2 and IV4 and affecting genes not associated with deafness. Considering the predicted pathogenicity of the variants and the function of the genes affected, we selected *SLC22A4* as the most interesting candidate NSHL gene to further investigate. The NM_003059:c.338G>A (p.C113Y) variation within *SLC22A4* exon 1 (Fig. 3.2C) is present in the homozygous state in all affected family members, but also in one normal hearing individual (IV1). Importantly, the variant is located in the only run of homozygosity shared by the two affected brothers, as assessed by the homozygosity mapping performed on WES data. To corroborate *SLC22A4* candidacy as NSHL gene in the analysed family, a linkage study was performed - by Dr. Michela Robusto - in the 13 available NSHL7 family members using the Illumina Infinium OmniExpressExome-8 v1.4 BeadChip kit. Parametric multi-point SNPs genotyping data analysis with the GeneHunter v2.1 r5 software evidenced a unique strong linkage signal peak (LOD>3.5) located in an interval of about 3.2 Mb on chromosome 5q23.3-q31.1 (chr5:128789554-131970885, hg19 Human reference genome coordinates), delimited by polymorphic markers rs11241999 and rs2237060 and encompassing *SLC22A4* gene (Fig. 3.2A). Besides, the c.338G>A variant is absent in our in-house database, as well as in a cohort of 125 Italian audiologically-tested normal-hearing controls (mean age at withdrawal 32±9), tested with a specific high-resolution melting assay. The substitution is also absent in the Greater Middle East Variome study cohort (GME, http://igm.ucsd.edu/gme/data-browser.php), which includes subjects from the North African population (85 from the Northwest Africa region and 423 from Northeast Africa) [Scott *et al.*, 2016]. The exons and splice junctions of *SLC22A4* gene were also screened in 7 additional unrelated NSHL patients of North African origin (2 from Morocco, 4 from Egypt and 1 from Tunisia; screening performed at the Laboratory of Medical Genetics of the Niguarda Hospital of Milan), but no putative

pathogenic variants were detected. Conversely, the variation is reported, in the heterozygous state, in 5 out of 114412 individuals in the gnomAD database, with an estimated allele frequency of 2.185e-5.







	SLC22A4 variant
Predictor	**p.C113Y**
SIFT	Damaging
Polyphen 2 HumDiv	Probably Damaging
Polyphen 2 HumVar	Probably Damaging
MutationTaster	Disease causing
Mutation Assessor	FI H
FATHMM	Damaging
Condel	D
CADD	32

Figure 3.2. Identification of the c.338G>A variant in *SLC22A4* gene within a linkage region in family NSHL7. A. Pedigree showing the co-segregation of NSHL with a haplotype spanning from polymorphic marker rs11241999 to rs2237060 (in red) within the chromosomal region 5q23.3-q31.1. The blue vertical bar indicates the disease-associated haplotype shared by all affected individuals. The genotype of available individuals is indicated below the corresponding symbols and the arrow points to the proband. W: wild-type allele, M: mutant allele. na: not analysed by genetic linkage study. Linkage analyses were performed by Dr. Michela Robusto. **B.** Audiograms of the normal-hearing parents (III8, III10, on the left) and of the 4 siblings (IV1-4) (average hearing loss for the right and left ears is shown). IV3 is not compliant for further audiometric evaluations. Age at audiometric evaluation is shown. y.o.: years old. **C.** Schematic representation of *SLC22A4* gene, where exons are indicated by rectangles and introns by black lines. Electropherograms showing the sequence surrounding the mutated nucleotide in a normal-hearing carrier (left) and in the proband (right). The position of the variant identified in family NSHL7 is indicated by an arrowhead. R: A or G. **D.** *In-silico* analyses of the p.C113Y missense variant identified in *SLC22A4* gene. The amino acid sequence alignments of SLC22A4 orthologs in the region surrounding the mutant residue (p.C113) is shown on the top. Protein sequences were retrieved from UniProt, and alignments were generated with Clustal Omega. The amino acid residue affected by the mutation is shaded in red and indicated by a red arrow. Identical amino acids are marked by an asterisk, while partially conserved residues are indicated by a colon. At the bottom, pathogenicity prediction of the p.C113Y missense variant with 8 commonly-used software. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$). A scaled C-score of greater or equal to 10 indicates that these are predicted to be the 10% most deleterious substitutions that you can do to the human genome, a score of greater or equal to 20 indicates the 1% most deleterious and so on.

SLC22A4 is located on chromosome 5q31.1, is composed of 10 exons and codes for OCTN1 (organic cation transporter 1). The organic cation/carnitine transporter OCTN1 is a 551-amino acid protein with 12 transmembrane domains [Burckhardt and Wolff, 2000] which is ubiquitously expressed in the body [Yabuuchi *et al.*, 1999]. The *SLC22A4* gene is also expressed in the cochlea, as reported in the SHIELD (Shared Harvard Inner-Ear Laboratory Database, https://shield.hms.harvard.edu) [Shen *et al.*, 2015], and further supported by the result of a specific RT-PCR assay on a RNA sample of organ of Corti derived from a postnatal day 4 wild-type mouse (our data, not shown). More specifically, it was demonstrated to have a diffuse expression across the inner ear epithelia, including the

hair cells and the apical surface plasma membrane of the endothelial cells of the stria vascularis [Ben Said *et al.*, 2016]. At the subcellular level, the encoded OCTN1 protein was described to localise in mitochondria, where it would participate in the maintenance of intracellular carnitine homeostasis [Lamhonwah and Tein, 2006]. In addition, *in-vitro* and *in-vivo* studies identified ergothioneine, a potent antioxidant, as the key substrate of OCTN1. Indeed, a metabolome analysis of blood and several organs of *Octn1*$^{-/-}$ mice indicated complete deficiency of ergothioneine and greater susceptibility to intestinal inflammation [Gründemann *et al.*, 2005; Kato *et al.*, 2010]. These findings seem to provide support to genetic analyses in humans, where polymorphisms in the *SLC22A4* gene have been previously linked to Crohn's disease [Peltekova *et al.*, 2004], colorectal cancer [Martini *et al.*, 2012; Li *et al.*, 2017], and susceptibility to rheumatoid arthritis [Tokuhiro *et al.*, 2003]. In the case of family NSHL7, however, no clinical symptoms or signs were reported from patients that could suggest the presence of inflammatory and/or intestinal diseases. Interestingly, the hearing phenotype of a different *Slc22a4* knock-out mouse was recently made available by the International Mouse Phenotyping Consortium (*Slc22a4*$^{tm1.1(KOMP)Vlcg}$ allele; http://www.mousephenotype.org/data/genes/MGI:1353479): the evoked auditory brainstem responses are not significantly different between wild-type mice and homozygous mutants, indicating no apparent signs of hearing loss.

The cysteine 113, affected by the mutation, is highly conserved across species (Fig. 3.2D) and it is one of the four highly-conserved cysteines that are present in all human organic cation transporter (OCTN1-2, OCT1-3), suggesting the importance of this residue for the protein structure [Burckhardt and Wolff, 2010]. Not surprisingly, the p.C113Y variation is predicted to be deleterious by all the software used (Fig. 3.2D).

To evaluate the impact of the amino acid substitution on OCTN1 protein localisation, I performed immunolocalisation studies on HeLa and HEK293 cells overexpressing a Myc-tagged isoform of the wild-type and mutant OCTN1 (pCDNA3.1_OCTN1-myc-wt,

pCDNA3.1_OCTN1-myc-mut) 48 hours after transfection. However, in contrast with the findings of Lamhonwah and Tein [Lamhonwah and Tein, 2006], who suggest a specific mitochondrial expression of the protein, my immunolocalisation studies do not show an exact co-localisation with the mitochondrial marker MitoTracker Orange, being the transporter seemingly confined in other intracellular compartments. In fact, the protein seems to localise along the secretory pathway, in perinuclear compartments as well as in an interconnected network that extend thorough the cytoplasm. It is thus plausible that, at later time points, the protein could be observed at the membrane, as postulated by other authors [Frigeni *et al.*, 2016]. Nonetheless, our studies suggest that the p.C113Y variant does not alter OCTN1 localisation at the time point analysed (Fig. 3.3).

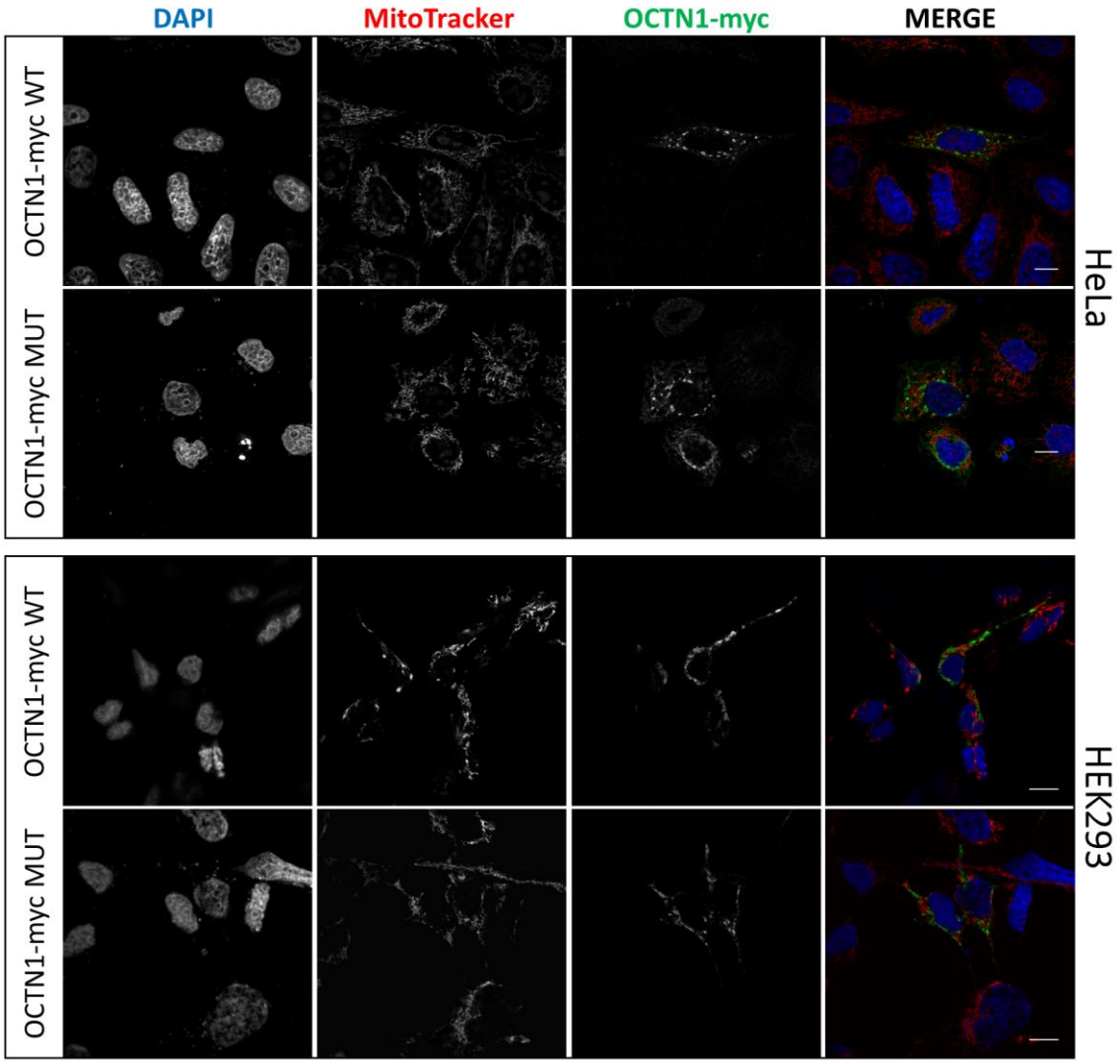


Figure 3.3. OCTN1 immunolocalisation studies. Co-localisation studies were performed on HeLa (upper panel) and HEK293 (lower panel) cells 48 hours after transfection with plasmids coding for Myc-tagged isoforms of OCTN1 wild-type or mutant proteins (OCTN1-myc WT, OCTN1-myc MUT). Images were acquired with Olympus FV1000 Inverted Confocal IX81 Microscope and are representative of 3 independent experiments. DAPI: 4',6-diamidino-2-phenylindole. Single confocal sections are shown. Scale bars: 10 μm.

In 2016, *SLC22A4* was identified as the DFNB60-associated gene responsible for NSHL in two unrelated consanguineous families of Tunisian ancestry [Ben Said *et al.*, 2016]. Intriguingly, the same NM_003059.2:c.338G>A (p.C113Y) variation was detected in both Tunisian families, as in family NSHL7. The authors showed that the p.C113Y mutation alters the ergothioneine uptake activity, and consequently the fatty-acid-based metabolic energy production pathway. Starting from this finding and considering the protein's

peculiar pattern of expression, the authors speculate that OCTN1 could be crucial for carnitine and ergothioneine transport into the inner ear. In addition, being the mutation located in the same haplotype in the two unrelated Tunisian families and found in an additional carrier among 122 Tunisian normal-hearing individuals, a founder effect of the p.C113Y mutation was hypothesised [Ben Said *et al.*, 2016].

Importantly, the linkage region associated with the hearing impairment in the Moroccan family we analysed partially overlaps with the linkage regions of the previously-reported Tunisian families [Ben Said *et al.*, 2016] (Fig. 3.4), thus further strengthening the hypothesis of a founder effect of the mutation.

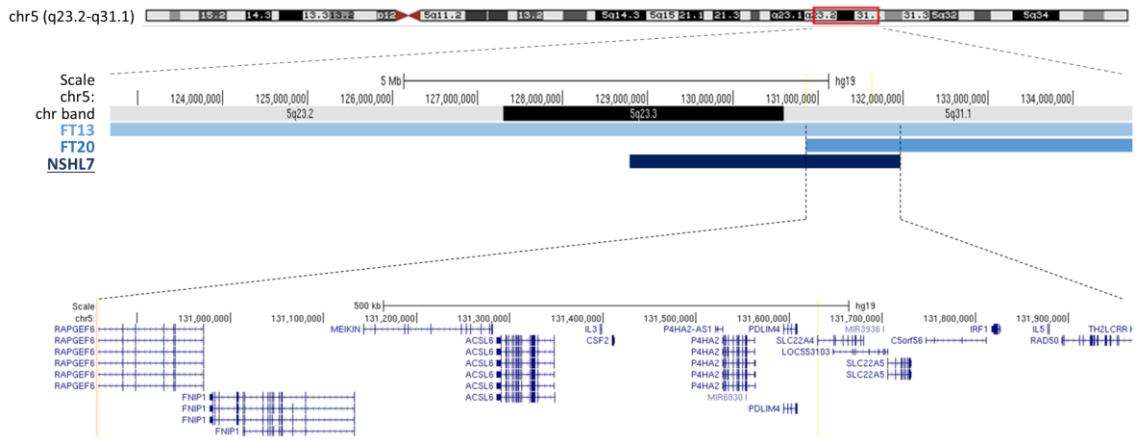


Figure 3.4. **Schematic representation of the deafness-associated haplotype in the three unrelated North African families reported to carry the NM_003059.2:c.338G>A (p.C113Y) missense variation.** A comparison between the linkage regions associated with the hearing impairment in the two previously-reported Tunisian families (FT13 and FT20) [Ben Said *et al.*, 2016], and the Moroccan NSHL7 family (in dark blue) on chromosome 5q23.2–q31.1 is reported, and the shared region containing *SLC22A4* is shown (image adapted from the UCSC Genome Browser).

3.2.1. Discussion and conclusions on family NSHL7

The frequency of the mutations associated with deafness can vary markedly in different ethnic groups. In the case of North Africans, the mutational landscape of NSHL is only partially overlapped with the one found in Caucasians. Indeed, NSHL-causing mutations identified in North African countries currently involve a relatively small set of genes, i.e. *GJB2*, *MYO7A*, *MYO15A*, *SLC26A4*, *TMC1*, *TMPRSS3*, *DFNB31*, *ESRRB*, *ESPN*, *DFNB59*, *LRTOMT*, *LHFPL5*, *PNPT1*, *TPRN*, and *MT-RNR1* [Chakchouk *et al.*, 2015]. Besides the bias due to the low number of genetically-analysed patients, this could also be related to the frequency of consanguineous marriages among North African populations, which inevitably leads to a clustering of rare recessive genetic defects [Rudman *et al.*, 2017]. In addition, Northwest African samples from the GME study (comprising individuals from Morocco, Algeria, and Tunisia) were among the least admixed ones, and it is thus suggested that these three countries represent a unique founder population [Scott *et al.*, 2016].

The association of the p.C113Y variant in *SLC22A4* gene with recessive NSHL was published in 2016 [Ben Said *et al.*, 2016], while I was also in the process of elucidating its role in the deafness of family NSHL7. The identification of the same mutation in family NSHL7 not only corroborates the involvement of the *SLC22A4* gene in deafness, but also suggests that the p.C113Y substitution is possibly a founder NSHL mutation with a broader distribution than previously thought. This could have important implications for the genetic diagnosis of deafness at least in patients/families with known Northwest African origin.

However, it should be noted that the p.C113Y variant was also found, in the homozygous state, in a normal-hearing NSHL7 family member, suggesting incomplete penetrance. Indeed, incomplete/reduced penetrance of the hearing phenotype in humans has been reported [Vona *et al.*, 2015], and attempts were made to identify modifier genes that could explain this phenomenon. For example, the modifier gene *METTL13* (DFNM1 locus)

have been identified as a dominant suppressor of recessive DFNB26/*GAB1* deafness [Riazudin *et al.*, 2000; Yousaf *et al.*, 2018a]. By analogy with *METTL13*, it is tempting to hypothesise the involvement of a dominant modifier of *SLC22A4*, which could suppress the onset of deafness in the non-penetrant individual IV1 from family NSHL7. Unfortunately, the genetic mapping of such a modifier is currently unfeasible due to the limited number of subjects with non-penetrant phenotype associated with the *SLC22A4* p.C113Y variant.

Another point that deserve further consideration is the discrepancy of the knock-out mouse phenotype. In fact, the ABR screening performed as part of the IMPC phenotyping pipeline did not show any sign of hearing impairment in the *Slc22a4* knock-out mouse model. Although the number of analysed animals was very limited (8 homozygous mutants only), and the panel of experiments to test the hearing function could be certainly broadened, these results do not help in shedding light on the role of *SLC22A4* in the normal hearing function. However, given 1) the absence of mutations in other NSHL-associated genes (including the 58 pathogenic variants of the NADf chip [Chakchouk *et al.*, 2015]); 2) the geographical origins of the NSHL7 family; 3) the identification of a unique significant linkage signal peak in family NSHL7, encompassing *SLC22A4* gene, and 4) the absence of more plausible pathogenic variants within the linkage region, it can be assumed that the p.C113Y substitution in *SLC22A4* is the likely cause of deafness in family NSHL7 and that it is recurrently associated with NSHL in Northwest African population. Further analysis in other NSHL families from the same geographical area would help in deciphering if this variation could be considered a solid marker of the disease, whose systematic screening could favorably influence deafness diagnosis. Also, it would be interesting to evaluate the impact of different genetic backgrounds on the expressivity of *SLC22A4*-dependent hearing phenotype studying wider patient populations with different ethnic backgrounds.

3.3. NSHL18 family: the *HCN1* and *OTOA* genes

Family NSHL18 is a consanguineous family, segregating early-onset moderate-to-severe HL with a likely autosomal recessive inheritance pattern. The proband (IV3) is a 23-year-old male with a 28-year-old deaf sister (IV1) and a 24-year-old normal-hearing brother (IV2). The proband's parents (III1, III2) are first cousins and have normal hearing. Individuals III5 and III12, cousins of III1 and III2, also show HL. Unfortunately, the II6 branch of the family was not available to the study. We subjected to WES the two affected siblings (IV2 and IV4) (Fig. 3.5A-B). I first performed a homozygosity mapping analysis to identify ROH shared by the affected siblings (Table 3.2). However, no likely pathogenic variants were identified in the two shared ROH. After ruling out the presence of mutations in known deafness genes, single nucleotide variant analyses of exome data prioritised 11 possibly pathogenic variants present in the homozygous state in the two affected siblings. Then, pathogenicity predictions together with segregation studies selected the NM_021072:c.1250A>G (p.Y417C) variant within exon 5 of the *HCN1* gene as the most promising candidate (Fig. 3.5C). Of note, *HCN1* gene is the last gene on the short arm of chromosome 5 before the centromere. Therefore, we hypothesised that the gap of flanking SNV at the 3' end of the gene could have prevented the detection of ROH in this genomic region. The c.1250A>G variation is present in the homozygous state in both affected siblings and in the heterozygous state in normal hearing relatives (III1 to 3, IV2) (Fig. 3.5A). In addition, it is absent in the gnomAD database, as well as in our in-house database.

Table 3.2. Runs of homozygosity shared between affected siblings.

ID	chr	start	end	Mb	NSNP	PHOM
IV1	16	55862901	58757683	2.89	85	0.988
IV2	16	55862901	58757683	2.89	86	0.988
IV1	20	60293919	62737318	2.44	114	0.991
IV2	20	60293919	62714783	2.42	84	0.988

The localisation of ROHs, their length (Mb), the number of SNPs (NSNP) and the proportion of homozygous calls within the run (PHOM) are indicated.

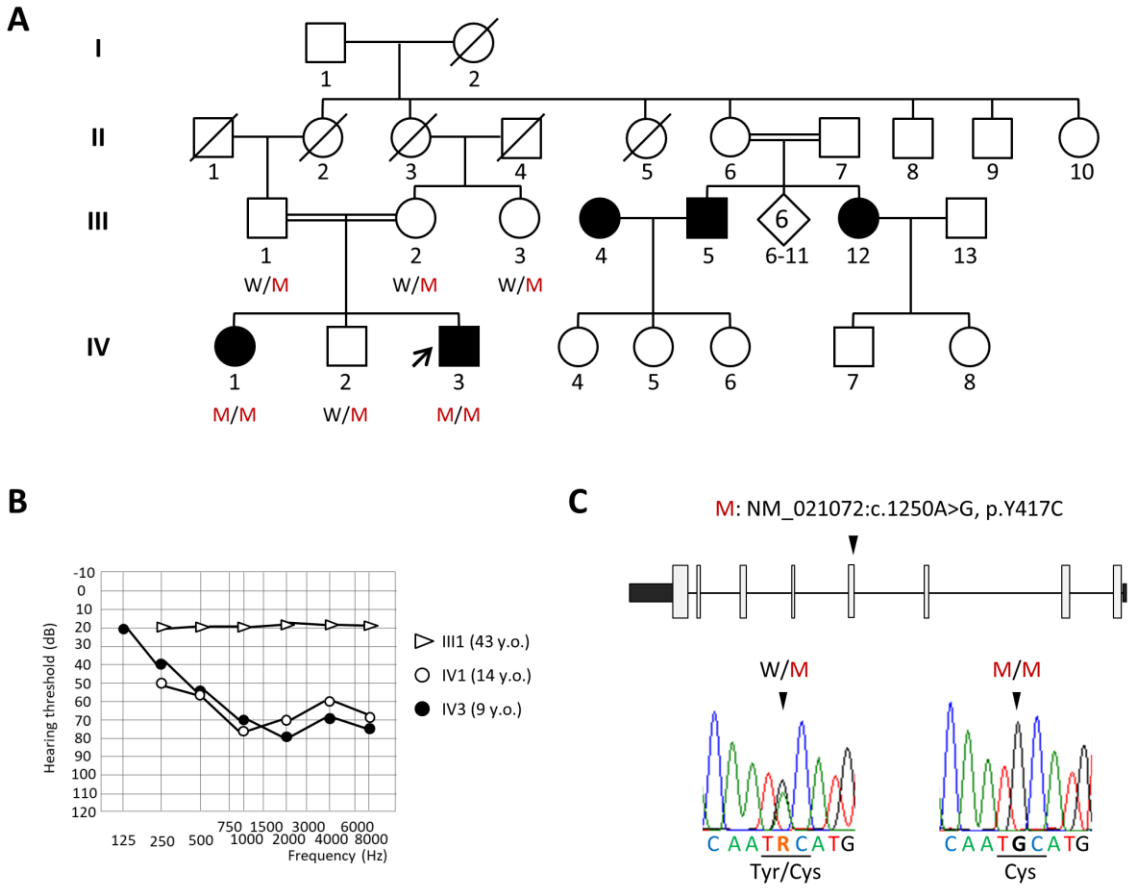
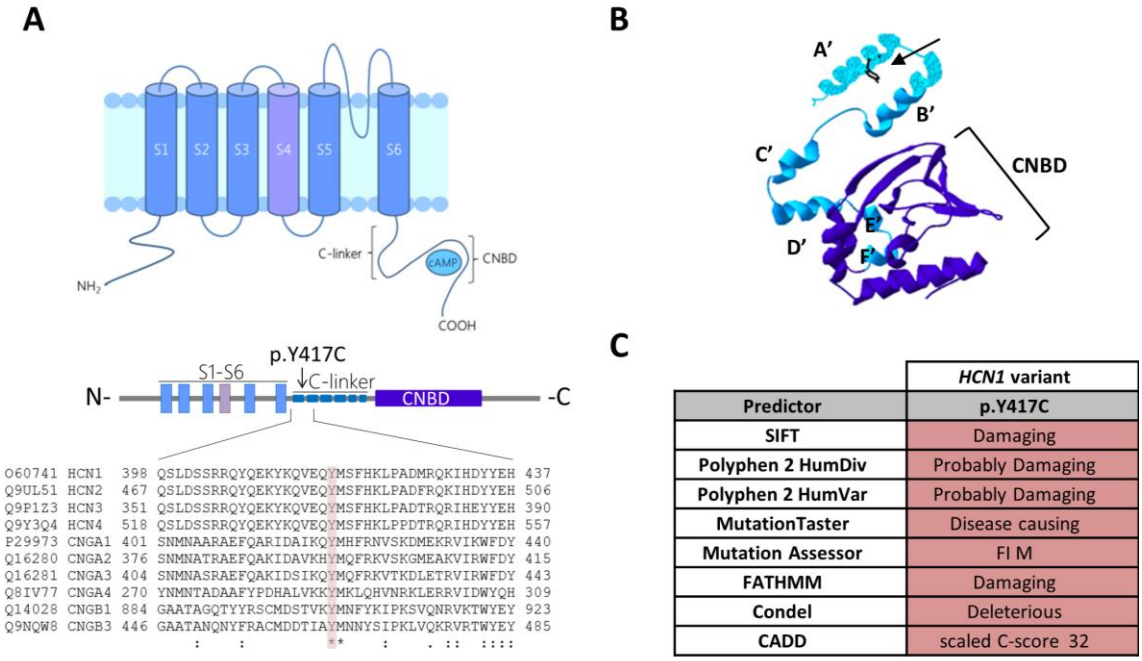


Figure 3.5. WES identifies one novel missense variant in the *HCN1* gene. A. The pedigree of NSHL18 family members is shown. The genotypes of all available individuals are indicated below the corresponding symbols and the arrow points to the proband. W: wild-type allele, M: mutant allele. **B.** Audiograms showing the threshold average for the right and left ear and the age at audiometric evaluation for each individual. **C.** Schematic representation of *HCN1* gene, where exons are indicated by rectangles and introns by black lines. Electropherograms showing the sequence surrounding the mutated nucleotide in one carrier (left) and one homozygous (right) family member. The position of the variant identified in family NSHL18 is indicated by an arrowhead. R: A or G.

HCN1 is located on chromosome 5p12, comprises 8 exons and codes for HCN1, one of the four members of the potassium/sodium hyperpolarization-activated cyclic nucleotide-gated (HCN) channels, which have a pivotal role in cellular excitability. HCNs belong to the 6TM (six transmembrane domains) ion channel superfamily and assemble in homo/hetero-tetrameric complexes. Each HCN subunit is thus composed of six transmembrane domains (S1-S6) with intracellular amino and carboxyl termini. The S4 domain, containing positively charged amino acids, is the channel voltage sensor, while the

region spanning from S5 to S6 (including the reentrant pore loop) forms the pore region. The S6 domain is connected to the C-terminal cyclic nucleotide binding domain (CNBD) by the C-linker, which is composed of six α-helices (A'-F') separated by short loops [Craven and Zagotta, 2006; Lee and MacKinnon, 2017] (Fig. 3.6A-B). HCNs are principally activated by hyperpolarization and modulated by binding of cyclic AMP (cAMP). In particular, cAMP enhances the channel activity binding to the CNBD and promoting its conformational change. This conformational change is transmitted to the C-linker and results in the release of the tonic inhibitions exerted by the CNBD on the channel pore. As a consequence of cAMP binding, the channel activation voltage shifts to more positive potentials [Wainger *et al.*, 2001; Lolicato *et al.*, 2011]. The opening probability is also enhanced by a Src-mediated tyrosine phosphorylation [Zong *et al.*, 2005].

Tyrosine 417 is located in the A'-helix of the C-linker trait and it is highly conserved not only among HCNs, but also among the CNG (cyclic nucleotide-gated) channel subfamily (Fig. 3.6A-B), which are also regulated by cyclic nucleotide binding and share the same structural organization of HCNs [Craven and Zagotta, 2006]. In addition, the substitution with a cysteine is likely to disrupt the local α-helix structure, as calculated by the GOR IV secondary-structure prediction software (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html, Fig. 3.8B), and is classified as deleterious by all of the 8 prediction software used (Fig. 3.6C).



	HCN1 variant
Predictor	**p.Y417C**
SIFT	Damaging
Polyphen 2 HumDiv	Probably Damaging
Polyphen 2 HumVar	Probably Damaging
MutationTaster	Disease causing
Mutation Assessor	FI M
FATHMM	Damaging
Condel	Deleterious
CADD	scaled C-score 32

Figure 3.6. *In-silico* analyses of the novel *HCN1* p.Y417C missense variant identified in family NSHL18. A. Schematic representation of HCN1 protein showing the six transmembrane helices (S1-S6), the C-linker and the cyclic nucleotide-binding domain (CNBD). The tyrosine affected by the mutation is located in the A' helix of C-linker domain and it is highly conserved among HCN channels as well as in the CNG channel family (lower panel). The amino acid residue affected by the mutation is shaded in red. Identical amino acids are marked by an asterisk, while partially conserved residues are indicated by a colon. Clustal Omega alignment of amino acid sequences retrieved from UniProt is shown. **B.** Ribbon representation of the mouse C-linker/CNBD Hcn1 subunit (PDB entry: 3UOZ). The mutated tyrosine is indicated by a black arrow. A'-F': C-linker α-helices. **C.** Pathogenicity prediction of the p.Y417C missense variant with 8 commonly-used software. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$).

The *HCN1* gene has been previously associated with epileptic encephalopathy, early infantile, 24 (EIEE24, MIM #615871), where five heterozygous de-novo mutations with a gain-of-function or a dominant-negative effect have been described [Nava *et al.*, 2014]. However, I deemed *HCN1* a good candidate for NSHL, since it is expressed in the mammalian cochlear hair cells and specifically interacts with stereociliary tip-link protocadherin 15 [Ramakrishnan *et al.*, 2012]. Furthermore, HCN1 is the most expressed subunit in spiral ganglion neurons and functionally contribute to their signal transmission, so that Hcn1 deficient mice show longer latencies in auditory brainstem responses and have deficits in spatial acuity [Kim and Holt, 2013; Ison *et al.*, 2017]. In addition, Ison and colleagues found that *Hcn1*$^{-/-}$ mice had higher ABR thresholds as compared to wild-types only at high frequencies [Ison *et al.*, 2017]. Interestingly, the hearing impairment in the two affected siblings of family NSHL18 is less severe in the low frequencies (Fig. 3.5B).

With this background, the p.Y417C variant was considered worth to be further investigated. First, to identify another patient carrying a mutation in the same region of HCN1 C-linker, I screened *HCN1* exon 5 in 656 individuals from our cohort of deaf patients, but no variants were found. Second, *in-vitro* functional studies were performed - in collaboration with the laboratory of Professor Michela Matteoli (Humanitas Clinical and Research Center) - to evaluate the impact of the identified missense variant on HCN1 channel function. In particular, over-expression experiments, using pCDNA3.1_hHCN1-WT or pCDNA3.1_hHCN1-p.Y417C constructs, were conducted in HEK293 cells first to assess the localisation of the wild-type and the p.Y417C mutant channels, and then to carry out whole-cell patch-clamp analyses. Immunofluorescence experiment showed that the missense variant does not impair HCN1 membrane localisation (data from Mr. Filippo Mirabella, not shown). Conversely, whole-cell patch-clamp recordings showed a strikingly decreased current of the p.Y417C mutant channel compared to the wild-type (Fig. 3.7A). Indeed, the current density for cells overexpressing p.Y417C protein was significantly

reduced as compared to cells overexpressing the wild-type channel (one-way ANOVA, p-value < 0.01), while it did not differ significantly from endogenous current density (control cells, transfected with GFP) (Fig. 3.7B). Taken together, these data suggested a loss of function effect of the amino acid substitution.

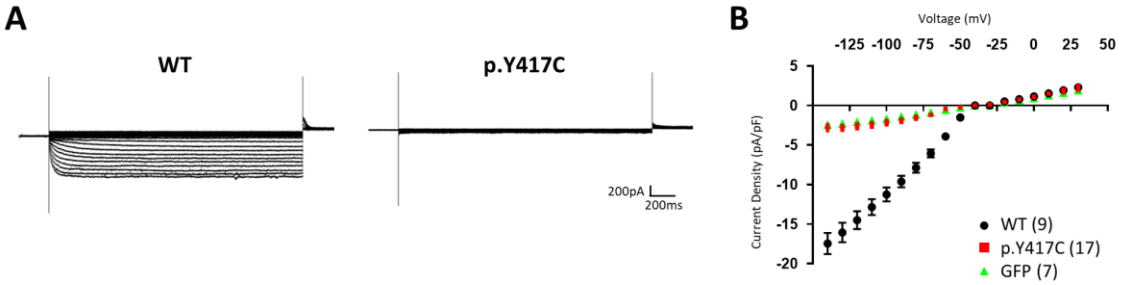


Figure 3.7. Electrophysiological analyses on HEK293 cells overexpressing wild-type or mutant HCN1 protein. A. Representative traces of whole-cell currents recorded in HEK293 cells transfected with vector expressing the wild-type (WT) or p.Y417C HCN1 protein. **B.** Plot of mean current density as a function of test voltage for cells overexpressing WT or p.Y417C HCN1; cells transfected with GFP were used as control. Data are presented as mean ± SEM (standard error of the mean) with the numbers of experiments indicated in parentheses. Electrophysiological analyses were performed by Mr. Filippo Mirabella.

To further characterise the mechanism by which the p.Y417C mutation causes the loss of channel function, I introduced a conservative mutation in HCN1 generating the mutant construct p.Y417F, in which the tyrosine 417 is substituted with a phenylalanine. Indeed, while the substitution of a tyrosine with a cysteine is predicted to disrupt HCN1 A' α-helix structure by the GOR IV software, the amino acid phenylalanine differs from tyrosine just for the absence of the hydroxyl group in the para position of the benzene ring and it is not expected to alter HCN1 secondary structure (Fig. 3.8A-B). As opposed to cells overexpressing the p.Y417C HCN1 mutant, currents were recorded in HEK293 cells overexpressing the p.Y417F mutant protein (Fig. 3.8C); this observation suggested that the substitution with a cysteine could indeed disrupt the C-linker structure, possibly altering the normal regulation of the channel gating.

Figure 3.8. Evaluation of HCN1 p.Y417F mutant channel. A. Amino acid structures showing that phenylalanine differs from tyrosine for the absence of the hydroxyl group in the para position of the benzene ring. **B.** GOR IV software secondary structure prediction for WT, p.Y417C and p.Y417F proteins. The region corresponding to A' α-helix is squared and the position of the substitution is indicated by a black arrow. Only the p.Y417C mutation is predicted to disrupt A' α-helix and has also an effect on B' α-helix. Random coil (c), α-helix (h), extended strand (e). **C.** Representative trace of whole-cell currents recorded in HEK293 cells overexpressing the p.Y417F HCN1 protein. Electrophysiological analyses were performed by Mr. Filippo Mirabella.

Despite these encouraging data, the implementation of WES data analysis with the detection of structural variants (using the EXCAVATOR2 tool) led to the identification of a putative homozygous deletion encompassing the *OTOA* gene in the proband and his affected sister.

OTOA gene is composed by 28 exons and is located on chromosome 16p12.2 within a 68 Kb segmental duplication with a high sequence identity [Shahin *et al.*, 2010]. It encodes the otoancorin, a noncollogeneous glycoprotein specifically expressed in the inner ear and necessary for the attachment of the tectorial membrane to the spiral limbus and, consequently, for the stimulation of inner hair cells [Zwaenepoel *et al.*, 2002; Lukashkin *et al.*, 2012]. ABR recordings of *Otoa* knock-out mice reported by the IMPC (*Otoa*$^{tm1.1(KOMP)Vlcg}$ allele; http://www.mousephenotype.org/data/genes/MGI:2149209) show a severe HL at all frequencies, which is more accentuated at high frequencies. Furthermore, loss of function mutations in *OTOA* have been associated with autosomal recessive deafness DFNB22 (MIM #607039). Interestingly, two different large homozygous deletions including the *OTOA* gene were previously identified. The first one, spanning 320-550 Kb, was reported in a consanguineous Palestinian family and its carrier frequency in that

population was 1% [Shahin *et al.*, 2010]. A smaller deletion, shorter than 190 Kb, was identified in a consanguineous Turkish family with mild-to-moderate sensorineural HL [Bademci *et al.*, 2014]. In addition, a heterozygous 228.5 Kb-deletion in *OTOA*, together with a missense mutation on the other allele, was recently identified in an Italian family with prelingual severe HL [Fontana *et al.*, 2017].

Serial PCR analyses, performed on genomic DNA from affected siblings and their parents, confirmed the presence of a 230-275 Kb homozygous deletion encompassing *OTOA* and the upstream and downstream nearest neighboring genes *METTL9* and *RRN3P1* in individuals IV1 and IV3 (Fig. 3.9). The presence of *OTOA* deletion was also independently confirmed by MLPA (Multiplex Ligation Probe Amplification) analysis at the Laboratory of Medical Genetics of the Niguarda Hospital of Milan (data not shown). Therefore, the loss of *OTOA* function was considered the most likely cause of deafness in the family.

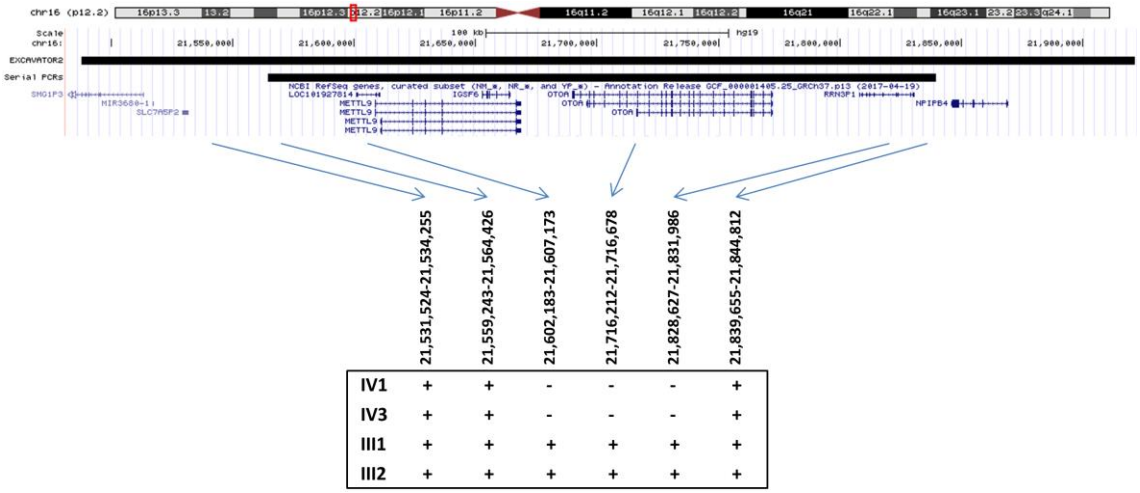


Figure 3.9. WES data analysis with EXCAVATOR2 identifies a homozygous deletion in *OTOA* gene. EXCAVATOR2 tool detected a ~433 Kb deletion encompassing the *OTOA* gene. Serial PCR analyses narrowed the deleted region to ~230-275 Kb. The position of the PCR amplicons is shown. Plus and minus symbols indicate whether the PCR amplification yielded a product of the expected size or not (image adapted from the UCSC Genome Browser).

3.3.1. Discussion and conclusions on NSHL18 family

The alignment of short reads on the reference genome is extremely challenging for highly similar sequences, since these reads cannot be mapped to a single region. Low mapping quality for exon 20-28 of *OTOA* gene is a well-known problem of the NGS approach [Mandelker *et al.*, 2014], and is due to the peculiar location of these exons within a 68 Kb segmental duplication with more than 99% sequence identity [Shahin *et al.*, 2010]. For the same reason, this region represents a hot spot for CNVs, occurring as a result of non-allelic homologous recombination [Shearer *et al.*, 2014]. Several CNVs involving *OTOA* have been identified [Shahin *et al.*, 2010; Shearer *et al.*, 2014; Bademci *et al.*, 2014], making this gene the second HL gene in which CNVs are most commonly detected [Shearer *et al.*, 2014]. Even if exome data are most suited to detect SNVs and short indels, specific tools have been designed to extract information on larger structural variants. Those programs are mostly based on a read count approach and are affected by the nonuniform efficiency of exome capture [D'aurizio *et al.*, 2016]. For sequences with low mapping quality it is even more difficult to rely on the output of these tools, which perform poorly in case of repetitive regions [Pirooznia *et al.*, 2015]. Nonetheless, in this case the use of the EXCAVATOR2 algorithm was fundamental to detect the large deletion including the *OTOA* gene, which escaped our first analyses. Given the important contribution of CNVs to NSHL [Shearer *et al.*, 2014], it is thus essential to include a bioinformatic pipeline for their detection exploiting whole-exome data and, ideally, a specific genetic testing of CNV using other available technologies (e.g. arrayCGH) to integrate WES data [Moteki *et al.*, 2016].

Concerning the p.Y417C HCN1 variant, even though our *in-vitro* studies indicate a functional impact on the channel activity, it is difficult to establish its contribution to the phenotype of the NSHL18 family, whose deafness - at the current state of knowledge - can be entirely attributed to the homozygous deletion of the *OTOA* gene.

However, we speculated that the loss of HCN1 function in the two deaf siblings could cause an unreported phenotype, distinct to deafness. Based on the association of *HCN1* with neurological anomalies, further clinical investigations were proposed to family NSHL18. The proband had an electroencephalography (EEG) and an intelligence quotient (IQ) test. While no waveform abnormalities were found in the EEG, the IQ test diagnosed a mild mental retardation in the proband. Interestingly, learning and memory disabilities were reported in the corresponding *Hcn1* knock-out mouse model [Nolan *et al.*, 2003]. Based on all these evidences, it is indeed plausible that the p.Y417C variant in *HCN1* might be the cause of the mild intellectual disability diagnosed in the proband. However, it has also been suggested that hearing impairment in children - especially if not promptly diagnosed – could contribute to cognitive deficits [Quittner *et al.*, 2004]. Therefore, it would be important to evaluate the segregation of the p.Y417C variant with mild cognitive impairment in all NSHL18 family member, as well as to identify other unrelated individuals with a similar phenotype, carrying loss of function mutations in HCN1.

3.4. NSHL20 family: identification of a novel pathogenic variant in the *COCH* gene

Family NSHL20 segregates late-onset hearing loss with an autosomal dominant inheritance pattern. The proband (IV2) is a 38-year-old man affected by bilateral deafness. The hearing in the family is characterised by a ski-slope loss, with a mild impairment in the low frequencies and a moderate-to-severe deafness in the middle and high frequencies. In individual II2, audiometrically tested at 82 years, the degree of hearing loss is more pronounced with a severe deafness in the low frequencies and a profound deafness at higher frequencies (Fig. 3.10A-B). The analysis of WES data from individuals III3, IV1 and IV2 pointed out 108 heterozygous variants shared only between deaf individuals. Among these, the NM_004086:c.1096 G>A transition (p.V366M) in the *COCH* gene was the only variant identified in a known NSHL gene and was thus considered the most likely cause of deafness in the family (Fig. 3.10C).

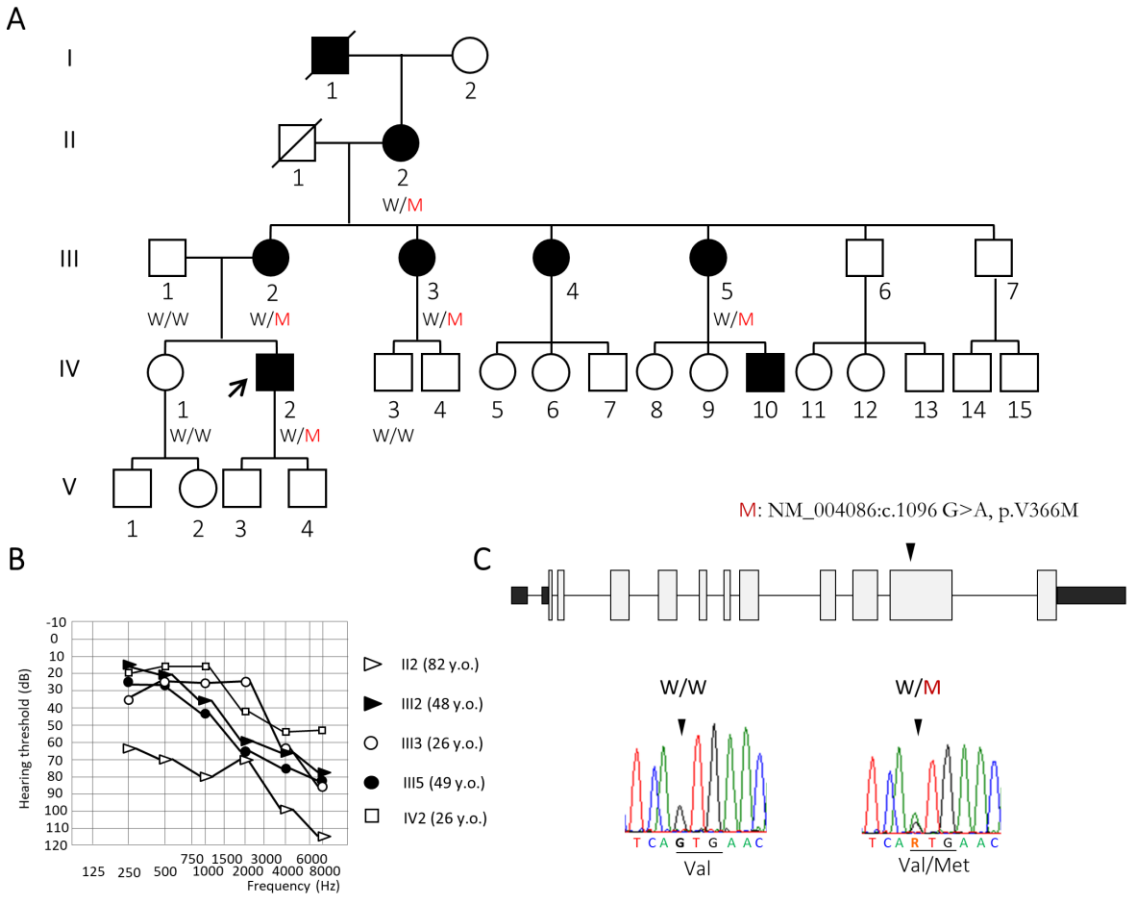


Figure 3.10. WES identifies a novel mutation in the *COCH* gene. A. The pedigree of NSHL20 family is shown. The genotypes of all available individuals are indicated below the

corresponding symbols. The arrow points to the proband. W: wild-type allele, M: mutant allele**. B.** Audiograms showing the threshold average for the right and left ear and the age at audiometric evaluation for each individual. **C.** Schematic representation of the *COCH* gene, where exons are indicated by rectangles and introns by black lines. Electropherograms showing the sequence surrounding the mutated nucleotide in one wild-type (left) and one heterozygous (right) family member. The position of the missense variant identified in family NSHL20 is indicated by an arrowhead. R: A or G.

The *COCH* gene is located on chromosome 14q12 and is responsible for autosomal dominant deafness DFNA9 (MIM #601369), which is characterised by an adult onset (in the 2nd to 3rd decade, starting from high frequencies) and a progressive decline of hearing to anacusis (by the 6th or 7th decade) and could be accompanied by vestibular impairment [Robertson *et al.*, 1998; JanssensdeVarebeke *et al.*, 2018]. Although the real worldwide incidence of mutation in *COCH* is not known [Bae *et al.*, 2014], DFNA9 has been reported to be the third most common type of autosomal dominant NSHL [Hildebrand *et al.*, 2010]. *COCH* comprises 12 exons and encodes cochlin, an extracellular matrix protein highly expressed in the inner ear. Cochlin is composed of a signal peptide (SP), a LCCL (*Limulus* factor C, cochlin, and late gestation lung protein Lgl1) domain, two intervening domains (ivd) and two von Willebrand factor A (vWFA)-like domains [Robertson *et al.*, 2001; Bae *et al.*, 2014] (Fig. 3.11).

To date, 25 mutations in the *COCH* gene have been associated with ADNSHL [Bae *et al.*, 2014; Tsukada *et al.*, 2015; Jung *et al.*, 2015; Gu *et al.*, 2016], mainly affecting LCCL or vWFA domains (Fig. 3.11). Interestingly, it has been reported that mutations in the LCCL domain cause HL with the presence of vestibular symptoms, whereas pathogenic variants in the vWFA domains are mostly associated with hearing defects only [Bae *et al.*, 2014]. All the variants are thought to have a dominant negative or gain-of-function effect, presumably causing impaired protein secretion and the concomitant aggregation and accumulation of cochlin intracellularly. Recently, a homozygous nonsense *COCH* variant (c.292C>T; p.R98X) was reported in a consanguineous family of Moroccan origin segregating

ARNSHL. In contrast to the 25 heterozygous variants causing DFNA9, the homozygous variant is expected to have a loss of function effect and seems to be associated with an earlier age of onset (Fig. 3.11) [JanssensdeVarebeke *et al.*, 2018].

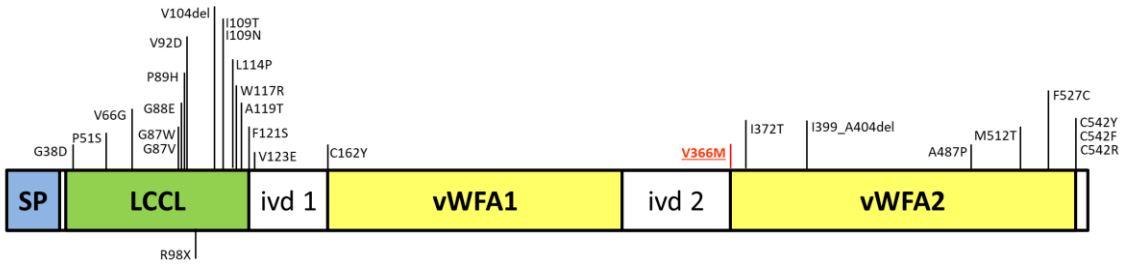


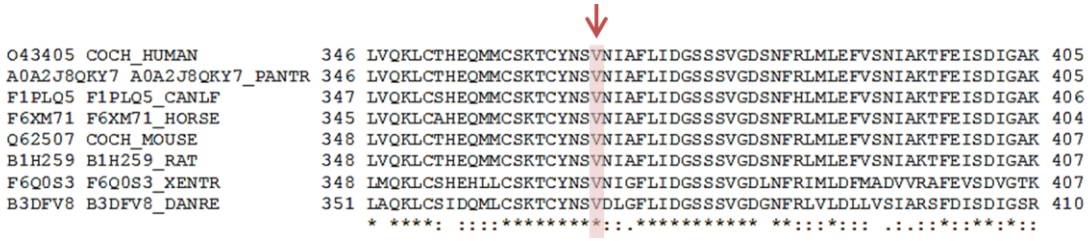
Figure 3.11. Schematic representation of cochlin domains. SP: signal peptide, LCC: *Limulus* factor C, cochlin, and Lgl1, vWFA1/2: von Willebrand factor A-like 1/2, ivd 1/2: intervening domain 1/2. The positions of the previously reported mutations are indicated in black (autosomal dominant up, autosomal recessive down), whereas the missense mutation identified in NSHL20 family is shown in red. Modified from Bae *et al.*, 2014.

The p.V366M variant is located in the vWFA2 domain of cochlin protein and, in agreement with the aforementioned correlation, the affected individuals of the NSHL20 family are reported solely with hearing impairment. In addition, although a follow-up of patients' hearing thresholds is not available, a correlation between the severity of HL and age among the affected individuals in the family can be observed, with younger subjects showing a milder impairment (Fig. 3.10B).

The missense mutation affects a highly conserved residue, which is invariant across vertebrates, and it is predicted to significantly impact on protein function by 7 out of 8 commonly used programs to predict the deleteriousness of a given amino-acid substitution (Fig. 3.12). Furthermore, this variation is extremely rare both worldwide and in Italy, being absent in the gnomAD database and in our in-house collection of more than 3500 Italian exomes. Interestingly, a different substitution in the same nucleotide (c.1096 G>T, p.V366L) is reported, at the heterozygous state, in just 1 out of 15490 individuals from the gnomAD (allelic frequency 3.228e-5).

In conclusion, based on the genotype-phenotype correlation and on the *in-silico* predictions, we propose that the p.V366M mutation is responsible for NSHL in family 20.

A



B

	COCH variant
Predictor	**p.V366M**
SIFT	Damaging
Polyphen 2 HumDiv	Probably Damaging
Polyphen 2 HumVar	Possibly Damaging
MutationTaster	Disease causing
Mutation Assessor	FI M
FATHMM	Tolerated
Condel	Deleterious
CADD	scaled C-score 25.4

Figure 3.12. *In-silico* analyses of the novel *COCH* p.V366M missense variant identified in family NSHL20. A. Amino acid sequence alignments of COCH orthologs in the region surrounding the mutant residue (p.V366). Protein sequences were retrieved from UniProt, and alignments were generated with Clustal Omega. The amino acid residue affected by the mutation is shaded in red and indicated by a red arrow. Identical amino acids are marked by an asterisk, while partially conserved residues are indicated by a colon. **B.** Pathogenicity prediction of the p.V366M missense variant generated with 8 commonly-used software. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$).

3.5. NSHL22 family: identification of two known pathogenic variants in *SLC26A4* gene

The NSHL22 pedigree is compatible with a recessive (autosomal or X-linked) inheritance (Fig. 3.13A). The two siblings (II2 and II3) were diagnosed with early-onset severe bilateral HL. II2 is a 39-year-old man who reports vertigo, which worsened in 2008 after cochlear implantation. II3 is 37 years old and was diagnosed with NSHL at 3 years of age (Fig. 3.13A-B). We subjected both siblings to WES, and analysed the data following the pipeline described in paragraph 3.1. We identified possibly pathogenic variants in the homozygous or compound heterozygous state affecting 9 genes. Among these putative mutations, we found two known pathogenic variants in the *SLC26A4* gene: NM_000441:c.1001G>T (p.G334V; rs146281367; ClinVar: 189039) [Walsh *et al.*, 2006; Kahrizi *et al.*, 2009] and NM_00441:c.1129C>T (p.T410M; rs111033220; ClinVar: 43498) [López-Bigas *et al.*, 2001; Arellano *et al.*, 2005] (Fig. 3.13C).

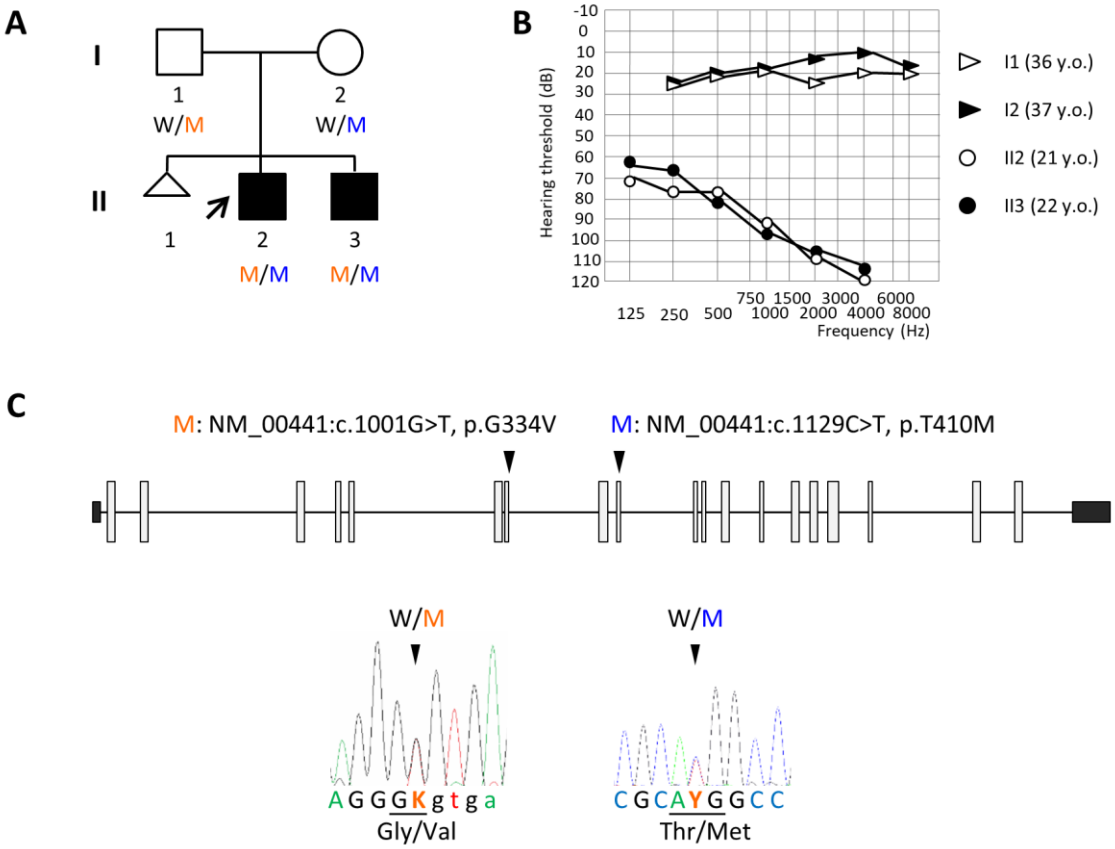


Figure 3.13. Identification of two known pathogenic variants in the *SLC26A4* gene. **A.** Family NSHL22 pedigree. The genotypes are indicated below the corresponding symbol and the arrow points to the proband. W: wild-type allele, M: mutant alleles. The two

mutations are indicated with different colors. The segregation analysis was performed at the Laboratory of Medical Genetics of the Niguarda Hospital of Milan. **B.** Pure-tone air-conduction thresholds of NSHL22 family members. For each individual the threshold average for the right and left ear and the age at audiometric evaluation is shown. **C.** Schematic representation of *SLC26A4* gene, where exons are indicated by rectangles and introns by black lines. Electropherograms showing the sequence surrounding the mutated nucleotides in the proband. The position of the variants identified in family NSHL22 is indicated by arrowheads. K: G or T; Y: C or T. Exonic nucleotides are indicated with uppercase, while intronic nucleotides are indicated with lowercase letters.

The *SLC26A4* gene, comprising 21 exons, is located on chromosome 7q22.3. It codes for the 780-amino-acids Cl^-/HCO_3^- exchanger pendrin, which is expressed in different tissues including the inner ear and the thyroid [Dror *et al.*, 2011]. In the cochlea, pendrin is expressed in non-sensory epithelial cells of the outer sulcus and spiral prominence [Royaux *et al.*, 2003], and it mediates the secretion of HCO_3^- in the endolymph, thus contributing to the maintenance of the fluid pH [Wangemann *et al.*, 2007]. Pendrin also participates to the apical iodide secretion in thyroid follicular cells [Bizhanova and Kopp, 2009]. Not surprisingly, mutations in the gene have been associated with both autosomal recessive NSHL with EVA - enlarged vestibular aqueduct - (DFNB4, MIM#600791) and Pendred syndrome (PDS, MIM #274600), which is characterised by congenital deafness and impaired iodide organification, possibly leading to goitre. However, PDS is often under-recognised, as goitre usually manifest in adulthood [Hilgert *et al.*, 2009]. Moreover, EVA can occur alone or together with Mondini dysplasia, a malformation of the cochlear turns. Despite the variety of clinical manifestation associated with *SLC26A4* mutations, a detailed genotype-phenotype correlation has not been defined, since some of the mutations have been associated with different phenotypes [Dror *et al.*, 2011].

Both variants are predicted to be deleterious by 8 out of 8 software used and affect amino acids highly conserved across vertebrates (Fig. 3.14).

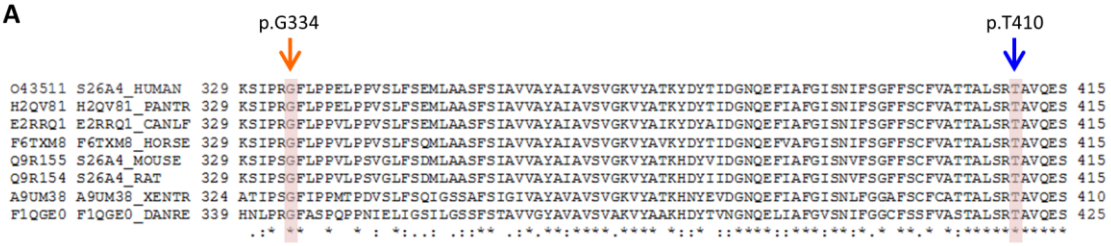
The c.1001G>T variant is reported, in the heterozygous state, in 2 out of 122538 individuals in the gnomAD database and it was firstly described in the homozygous state in

a consanguineous Palestinian family segregating severe-to-profound HL with EVA. In particular, the mutation affects the last nucleotide of exon 8, and causes both an amino acid substitution (NP_000432.1:p.G334V) and a splicing defect. Indeed, it was shown to cause the retention of the first 41 nucleotides of intron 8, leading to a premature stop codon immediately after glycine 334 [Walsh *et al.*, 2006]. The homozygous c.1001G>T mutation was also associated with a PDS phenotype in an Iranian consanguineous family [Kahrizi *et al.*, 2009] and it was shown to dramatically impair pendrin transport activity [Dossena *et al.*, 2011].

Similarly, the p.T410M variant (present in the heterozygous state in 51/138276 from the gnomAD database and in 3/3541 Italian control exomes) was described in the homozygous state both in individuals with deafness and normal thyroid function [López-Bigas *et al.*, 2001] and in PDS patients; this phenotypic variability has also been reported within the same family [Arellano *et al.*, 2005]. Taylor and colleagues demonstrated that the mutant protein fails to reach the plasma membrane, being retained in the endoplasmic reticulum [Taylor *et al.*, 2002].

The analysis of the segregation of the two variants within family NSHL22 confirmed that each variant was inherited from one of the normal-hearing parents (performed at the Laboratory of Medical Genetics of the Niguarda Hospital of Milan). Following molecular diagnosis, one of the affected sibling self-reported "thyroid problems". Considering that the variants were also described in PDS patients, II2 and II3 were referred to a thyroid endocrinologist and subjected to further clinical evaluation and management. In this family, WES not only was fundamental to provide a genetic diagnosis, but also allowed to address the patients to the more suitable monitoring and treatment.

A



B

	***SLC26A4* variants**	
Predictor	**p.G334V**	**p.T410M**
SIFT	Damaging	Damaging
Polyphen 2 HumDiv	Probably Damaging	Probably Damaging
Polyphen 2 HumVar	Probably Damaging	Probably Damaging
MutationTaster	Disease causing	Disease causing
Mutation Assessor	FI M	FI M
FATHMM	Damaging	Damaging
Condel	Deleterious	Deleterious
CADD	scaled C-score 34	scaled C-score 27.2
Database		
gnomAD	2/250276 (7.991e-6)	48/282236 (1.701e-4)
In-house Italian exome collection	0/3541	3/3541 (8.47e-4)

Figure 3.14. *In-silico* analyses of the two *SLC26A4* missense variants (p.G334V and p.T410M) identified in family NSHL22 in the compound heterozygous state. A. Amino acid sequence alignments of SLC26A4 orthologs in the region surrounding the two mutant residues (p.G334 and p.T410). Protein sequences were retrieved from UniProt, and alignments were generated with Clustal Omega. The amino acid residues affected by the mutations are shaded in red and indicated by an orange and a blue arrow respectively. Identical amino acids are marked by an asterisk, while partially conserved residues are indicated by a colon. **B.** Pathogenicity prediction of the p.G334V and p.T410M missense variants with 8 commonly-used software and presence in databases. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$).

3.5.1. Mouse models of DFNB4

The study of *Slc26a4* knock-out and knock-in mouse models have facilitated the understanding of pendrin function and its role in DFNB4 and PDS hearing loss. The lack of pendrin in mice prevent the development of hearing function and causes the enlargement of the vestibular aqueduct, mimicking the early-onset deafness and EVA

observed in humans [Wangemann, 2013]. In contrast to PDS patients, however, *Slc26a4*$^{-/-}$ mice show no signs of thyroid abnormalities [Everett *et al.*, 2001], thus not recapitulating the complete PDS phenotype.

The *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mouse analysed in this work carries a CreERT2_EGFP cassette in *Slc26a4* intron 2 which interrupts the gene transcription and ensures the expression of EGFP and CreERT2 under the control of *Slc26a4* promoter (see Paragraph 2.11 and Fig. 2.1A). Therefore, *Slc26a4* expression is abolished in *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ homozygous mice and CreERT2 is expressed - only in cells expressing *Slc26a4* – both in homozygous and heterozygous mice. Therefore, the *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice can be used to create conditional knock-outs (i.e. heterozygous for *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ allele and homozygous for a floxed sequence in a gene of interest) that don't express a specific gene of interest only in the cells where pendrin is expressed, and in a tamoxifen-inducible manner.

3.5.1.1. Auditory phenotype of *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice

The auditory function of *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ heterozygous and homozygous mice and of wild-type littermates was evaluated using ABR recordings at different ages. In particular, mice of each genotype were tested at 3, 4 and 8 weeks of age.

As expected, *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ homozygous mice are profoundly deaf already at 3 weeks of age, showing no responses at the maximum sound level presented (95 dB SPL) at all frequencies tested (3-42 kHz) (Fig. 3.15). Mutant mice did not show any evident sign of vestibular impairment (e.g. circling, head tilting) at the ages tested for ABRs. The early-onset of deafness in the homozygous mutant mice is consistent with the prelingual HL reported in the NSHL22 family. However, the severity of hearing impairment in the deaf siblings varies at different frequencies, being moderate in the low frequencies and severe-to-profound in the middle and high frequencies. This difference in the severity of the phenotype in mice and humans have already been reported [Choi *et al.*, 2011] and might be

in part explained by the different pathogenic mechanisms associated with the different mutations in the gene.

Of note, no statistically significant difference was observed in ABR thresholds from *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ heterozygous and wild-type groups at all frequencies and at any of the age tested. As heterozygous mice have normal hearing, this mouse model can be "safely" used to generate cell-type specific tamoxifen-inducible conditional knock-outs, to study the function of a gene of interest in hearing, in a region-specific and temporally-restricted fashion.

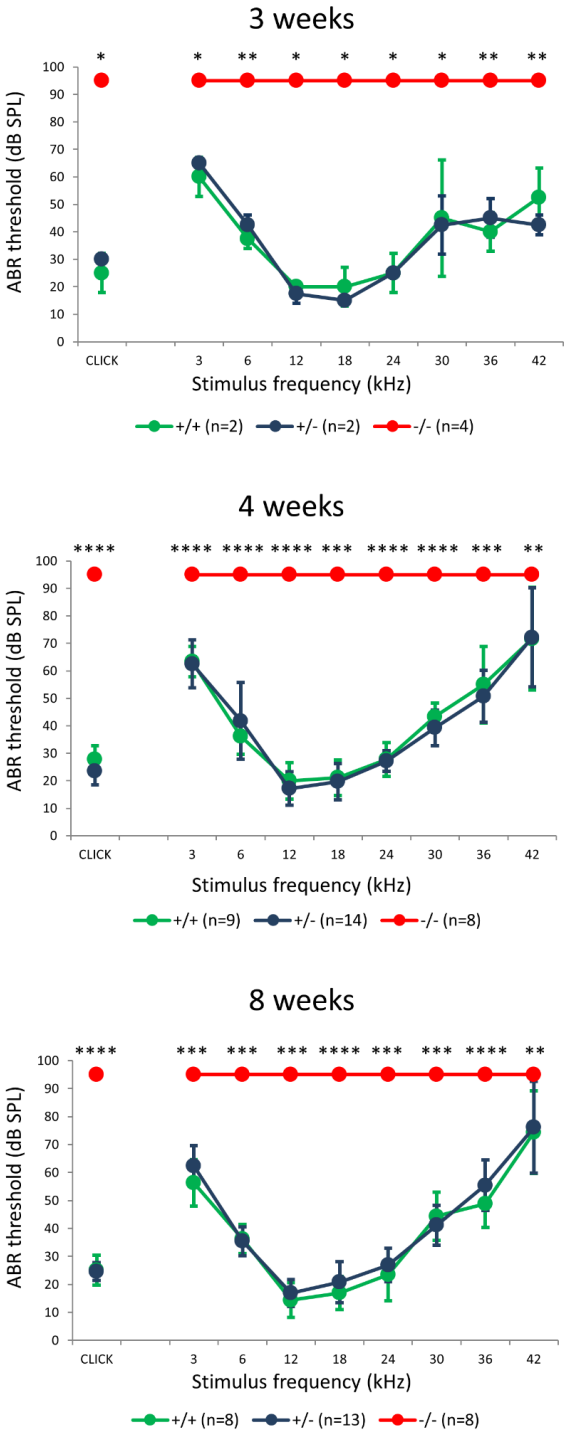


Figure 3.15. ABR recordings of *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mutant mice at 3, 4 and 8 weeks. Mean ABR thresholds (± SD, standard deviation) for clicks and tone pips are plotted for wild-type (+/+, in green), heterozygous (+/-, in blue) and homozygous (-/-, in red) mice aged 3, 4 and 8 weeks. For each stimulus, data were analysed by the Kruskal-Wallis test followed by the Dunn's *post-hoc* analysis. Asterisks indicate p-values (*: $p<0.05$; **: $p<0.01$; ***: $p<0.001$; ****: $p<0.0001$). Statistical analyses were performed with GraphPad Prism 8.0. I performed all the ABR measurements at 3 weeks and in 9 and 12 mice at 4 and 8 weeks, respectively. The remaining ABR recordings and threshold calls (at 4 and 8 weeks) had already been performed by Dr. Jing Chen and Ms. Elysia James. ABR measurements for each individual mouse I personally tested are shown in Appendix 6.2.

3.6. NSHL3 family: the *DIAPH2* gene

Family NSHL3 pedigree is compatible with recessive (autosomal or X-linked) inheritance (Fig. 3.16A). The proband (III5) is a 15-year-old male with a 23-year-old deaf brother (III3) and a 20-year-old normal-hearing brother (III4). Both affected siblings have pre-lingual bilateral hearing loss which is more severe in the middle and high frequencies. HL in III5 is moderate in the low frequencies and severe-to-profound in middle-high frequencies, whereas in III3 is severe in low frequencies and profound in middle-high frequencies (Fig. 3.16B). The three siblings (III3, III4 and III5) were subjected to WES, and data analysis (Table 3.3), following the pipeline described in paragraph 3.1, lead to the identification of the NM_006729.4:c.868A>G (p.I290V) variant within exon 8 of the *DIAPH2* gene on chromosome X (Fig. 3.16C). The putative mutation co-segregates with the phenotype in the family, being present in the heterozygous state in female carriers and in the hemizygous state in the affected brothers. In addition, it is absent in the normal hearing brother and in one normal hearing male cousin, who is the son of a carrier individual (Fig. 3.16A). *DIAPH2* has not yet been associated with NSHL, but the identified variation is absent both in our in-house database of about 3500 Italian exomes, and in a cohort of 125 Italian audiologically-tested normal-hearing controls (data from Dr. Michela Robusto). In addition, it is reported in the gnomAD database in the heterozygous state only in 3 European non-Finnish females (allele frequency: 1.816e-5).

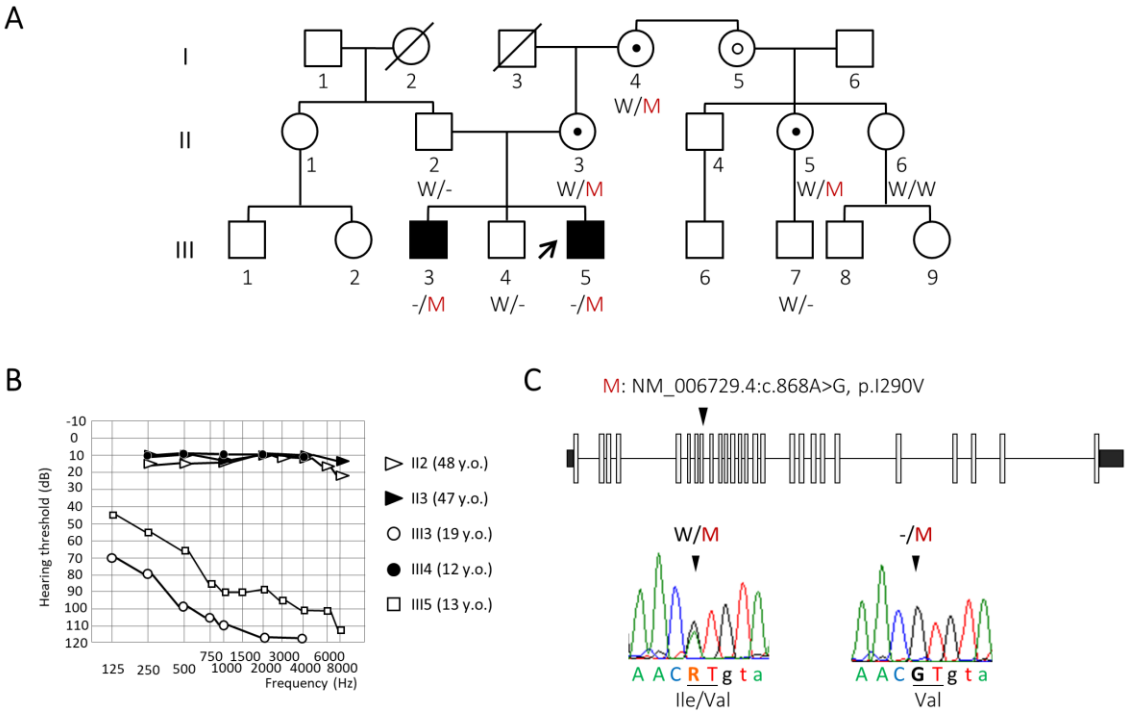


Figure 3.16. Identification of a variant in the novel candidate NSHL gene *DIAPH2*. **A.** Pedigree of family NSHL3, showing the segregation of the variant within the family. The genotypes of available individuals are indicated below the corresponding symbol and the proband is indicated by a black arrow. W: wild-type allele, M: mutant allele. ○: obligate carrier, ●: carrier of the mutation confirmed by Sanger sequencing. **B.** Pure-tone air-conduction thresholds of NSHL3 family members. For each subject the threshold average for the right and left ear and the age at audiometric evaluation is shown. **C.** Schematic representation of *DIAPH2* gene, where exons are indicated by rectangles and introns by black lines. Electropherograms showing the sequence surrounding the mutated nucleotide in the proband (right) and his mother (heterozygous carrier, left). The position of the variant identified in family NSHL3 is indicated by an arrowhead. R: A or G. Exonic nucleotides are indicated with uppercase, while intronic nucleotides are indicated with lowercase letters.

Table 3.3. Prioritisation of variants identified with WES in family NSHL3.

Filter type	**Nr of variants**		
	III3	III4	III5
Tot	42647	42532	43528
NS/SS/I	13963	13801	14239
Not in 1000Genomes	3412	3412	3539
Not in NHLBI-ESP	3256	3267	3384
Not in dbSNP135	1137	1154	1164
Hom/Hemi/compound Het	142	143	140
Shared by the affected siblings	4*		
Within known ARNSHL genes	/		

NS: non-synonymous variant, SS: splice-site variant, I: indel. * The remaining 3 variants were excluded for the high frequency in later releases of dbSNP, for the presence of hemizygous individuals in the gnomAD database, or for lack of segregation with phenotype.

DIAPH2, comprising 27 exons, is located on chromosome Xq21.33 and codes for a member of the *Diaphanous*-related formins (DRFs), which are involved in the nucleation of linear unbranched actin filaments [Campellone and Welch, 2010] and in the alignment and stabilisation of microtubules [Yasuda *et al.*, 2004]. As the two other DRFs, DIAPH1 and DIAPH3, the DIAPH2 protein can be functionally divided into two halves. The C-terminal half is required for actin assembly and includes three structural and functional domains: i) the profiling-binding FH1 (formin homology 1) domain; ii) the actin-binding FH2 (formin homology 2) domain; and iii) the DAD (*diaphanous* autoregulatory domain), which can interact with the N-terminal portion. The N-terminal half is the regulatory region of DRFs and it is divided in the GTPase binding domain (GBD) and the *diaphanous* inhibitory domain (DID) [Goode and Eck, 2007]. Furthermore, a dimerisation domain (DD) and a coiled coil (CC) domain are present to mediate the dimerisation of the N-terminal of DRFs which are active as dimers [Campellone and Welch, 2010]. The interaction between the N-terminal DID and the C-terminal DAD determinates the transition of the DRF to its inactive conformation. In addition, the first portion of DID domain is also involved in the interaction with Rho family GTPases [Otomo *et al.*, 2005; Rose *et al.*, 2005]. As a result of the partial overlap of DAD and Rho binding sites, the binding of Rho GTPases induces the release of the DAD autoinhibitory domain and stimulate actin nucleation functions [Young and Copeland, 2010] (Fig. 3.17). *DIAPH1*, *DIAPH2* and *DIAPH3* genes are the orthologs of the *Drosphila melanogaster diaphanous* (*dia*) gene, which has important roles in fertility and hearing in the fly. Indeed, mutant *dia* alleles have been associated both to sterility and impaired response to sound [Castrillon and Wasserman, 1994; Schoen *et al.*, 2010]. In humans, *DIAPH1* and *DIAPH3*, have been associated with different types of HL. In particular, heterozygous mutations in *DIAPH1* are responsible for autosomal dominant NSHL with or without thrombocytopenia (locus DFNA1, MIM #124900) [Lynch *et al.*, 1997; Stritt *et al.*, 2016], whereas overexpression of *DIAPH3* – resulting from

a mutation in the 5' UTR of the gene - causes auditory neuropathy, autosomal dominant, 1 (AUNA1, MIM #609129) [Schoen *et al.*, 2010]. Interestingly, homozygous truncating mutations of *DIAPH1* have been associated with seizures, cortical blindness, microcephaly syndrome (SCBMS, MIM #616632) [Ercan-Sencicek *et al.*, 2015]. Concerning *DIAPH2*, until now it has only been linked to premature ovarian failure 2 (POF2A, MIM #300511). Indeed, in a mother and daughter affected by POF, a balanced translocation t(X;12)(q21;p13) was found to disrupt the *DIAPH2* gene [Bione *et al.*, 1998]. Nonetheless, given its membership to DRFs and the important function of actin nucleation in hair-cell stereocilia (see Paragraph 1.1 and 1.4.1), we hypothesised a possible role of DIAPH2 as an actin regulator in sensory hair cells, thus considering *DIAPH2* a promising candidate NSHL gene to further characterise.

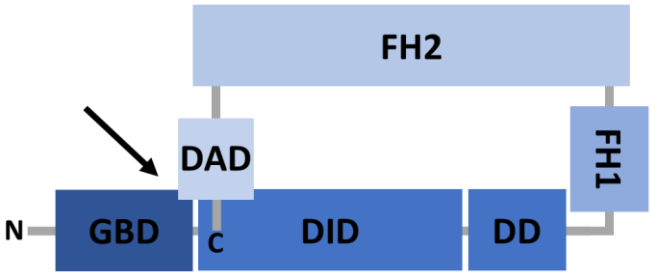


Figure 3.17. Schematic representation of *diaphanous*-related formin domains. The structural/functional domains of the DRFs are shown. The arrow indicates the binding site of Rho GTPases. GBD: GTPase-binding domain; DID: *diaphanous* inhibitory domain; DD: dimerisation domain; FH1: formin homology 1, FH2: formin homology 2; DAD: *diaphanous* autoregulatory domain; N: N-terminus; C: C-terminus.

3.6.1. Effect of the c.868A>G variant on *DIAPH2* splicing

Besides being a missense variant, the c.868A>G variant in *DIAPH2* affects a nucleotide in exon 8 located only two base pairs upstream of the donor splice site and was thus also annotated as a candidate splicing mutation. For this reason, the possible effect of the variant on *DIAPH2* splicing had been previously investigated in whole blood from all available family members. However, the RT-PCR results suggested that the variant does not alter exon 8 splicing at least in the analysed tissue. Minigene transfection experiments –

using the pBS-KS_modified hybrid minigene vector containing either wild-type or mutant exon 8 with the surrounding intronic sequences (Fig. 3.18A) – had also been performed in several human cell-lines (HeLa, HEK293, HepG2). Interestingly, in the very first experiment performed, the RT-PCR showed that the variant could induce exon 8 skipping. In addition, it was observed a different degree of exon 8 skipping in the three cell lines analysed, suggesting a tissue specificity of this splicing event. However, any attempt to replicate these results in the same cell lines failed, as no alternative splicing was observed in the cells transfected with the mutant construct in subsequent experiments (independent transfections, RNA extractions, and RT-PCR assays; data from Dr. Michela Robusto, not shown). Given the lack of reproducibility of the results, I decided to perform the same minigene transfection experiments using a cell line more suitable for inner ear studies, i.e. the UB/OC-2 cell line, derived from the prosensory epithelium of the H-2Kb-tsA58 transgenic mouse [Jat e*t al.*, 1991; Holley and Lawlor, 1997]. Nevertheless, the RT-PCR analyses did not show any difference in the splicing pattern of UB/OC-2 cells transfected with the mutant minigene vector compared with those transfected with the wild-type, both in proliferative and differentiating conditions (Fig. 3.18B, n=3 for both conditions). The results were also confirmed by a competitive fluorescent RT-PCR assay, which showed a unique peak corresponding to the amplification of the transcript including the *DIAPH2* exon 8 (data not shown). In conclusion, we could not find any evidence that the NM_006729.4:c.868A>G variant alters *DIAPH2* exon 8 splicing.

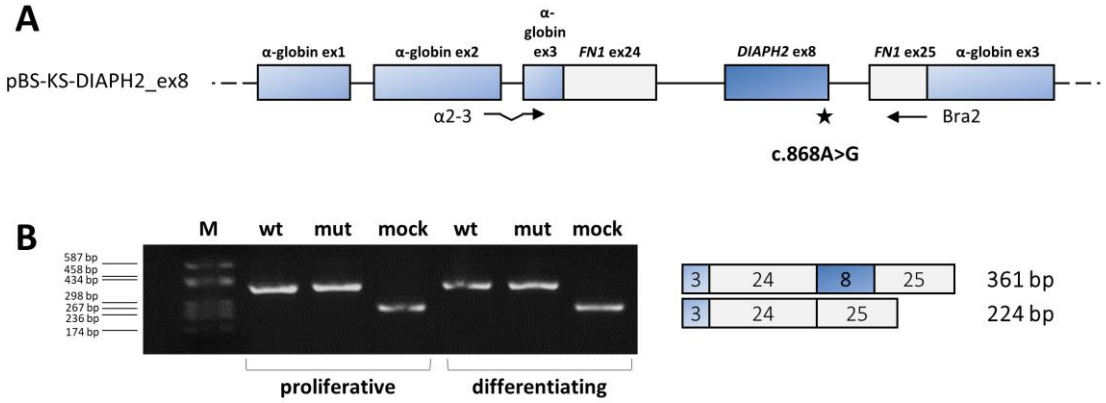


Figure 3.18. *In-vitro* analysis of the impact of c.868A>G variant on *DIAPH2* pre-mRNA splicing in UB/OC-2 cells. A. Schematic representation of the hybrid pBS-KS-DIAPH2_ex8 minigene where α-globin exons are represented by light blue boxes, fibronectin (*FN1*) exons by grey boxes, whereas introns are shown as black lines (not to scale). Exon 8 of *DIAPH2* is represented by a blue box. The c.868A>G variant in exon 8 is indicated by a black star. Black arrows represent the primers used in RT-PCR assays. **B.** On the left, agarose gel electrophoresis of RT-PCR products obtained from RNA of UB/OC-2 cells under proliferative or differentiating conditions, transfected with the wild-type (wt), mutant (mut), or empty (mock) minigene vector. M: molecular weight marker (pUC9/*HaeIII*). On the right, schematic representation of the splicing products, as verified by Sanger sequencing. The length of each fragment is shown.

3.6.2. Effect of the p.I290V variant on the DIAPH2 protein

Despite being a conservative amino-acid substitution – isoleucine to valine - the p.I290V variant is predicted to be deleterious by 7 out of 8 different software used (Fig. 3.19C). In addition, the isoleucine 290 is highly conserved across vertebrates and it localises in a region of the inhibitory domain possibly involved both in the auto-inhibition of the protein and in the interaction with Rho GTPases (Fig. 3.19A-B). To evaluate whether the variation could affect DIAPH2 localisation, I performed immunofluorescence localisation assays on HeLa and HEK293 cells 48 hours after transfection with wild-type or mutant plasmids expressing an HA-tagged DIAPH2 isoform (pCMV-HA-hDIA2B_wt, pCMV-HA-hDIA2B_mut). These studies suggested that the variant does not alter DIAPH2 cytoplasmic localisation, at least in basal conditions (Fig. 3.20). Therefore, I studied the localisation of wild-type and mutant DIAPH2 protein under activating stimuli, by

performing immunolocalisation assays on HeLa and HEK293 cells co-expressing wild-type or mutant DIAPH2 together with a constitutively active form of the activator RhoA (pCDNA3.1-Myc-RhoA-G14V). The activation of DIAPH2 by RhoA changes the localisation of DIAPH2 from the cytoplasm to the plasma membrane, where the protein co-localises with actin filaments. Co-localisation with actin can be observed both for the wild-type and the mutant DIAPH2; however, the overlap of red and green signals seems to be higher in cells expressing the wild-type protein (Fig. 3.21). In addition, the presence of RhoA promotes a morphological change of the cells, with the formation of stress fibers in HeLa cells and actin-rich membrane protrusions in HEK293 cells (Fig. 3.22). Considering base sections of the cells (at the level of their adhesion to the glass), DIAPH2 protein can be observed as red dots which, most clearly in HEK293 cells, are localised at the end of the actin bundles. Similar morphological changes are induced in cells expressing either the wild-type or the mutant DIAPH2 together with RhoA. Nevertheless, there is a statistically significant difference in the average length (measured with Fiji software) of the membrane protrusions in HEK293, being ~20% shorter in cells expressing mutant DIAPH2 (unpaired t test, $p<0.01$) (Fig. 3.23).

A

B

```
O60879 DIAP2_HUMAN            261 ARAIDPKQPNMMTEIVKILSAICIVGEENILDKLLGAITTAAERNNRERFSPIVEGLENQ 320
A0A2I3RZA9 A0A2I3RZA9_PANTR   268 ARAIDPKQPNMMTEIVKILSAICIVGEENILDKLLGAITTAAERNNRERFSPIVEGLENQ 327
F1Q1Q1 F1Q1Q1_CANLF           261 ARAIDPKQPNMMTEIVKILSAICIVGEENILDKLLGAITIAAERNNKERFSPIVEGLENH 320
F6SEZ8 F6SEZ8_HORSE           269 ARAIDPKQPSMMTEIVKILSAICIVGEENILEKLLGAITTAAERNNRERFSPIVEGLENH 328
O70566 DIAP2_MOUSE            260 ARAIDPKQQNMMTEIVKILSAICIVGEENILDKLLGGITAAAELNNRERFSPIVEGLENN 319
F6W113 F6W113_XENTR            66 ARAIDPKQANMMTEILKILSAICIIGEDNILDKLLEAITTAAEIHNKERFAPIVEALENH 125
F1RCU7 F1RCU7_DANRE           224 ARAIDPKQTNMMTEIVKILSAVCIIGEENILDKILAAMTIAAERNNKERFAPIVEGLENH 283
                                  ******** .*****:*****:**:**:***:*:* .:* *** :*:***:****.***:
```

C

	DIAPH2 variant
Predictor	**p.I290V**
SIFT	Tolerated
Polyphen 2 HumDiv	Possibly Damaging
Polyphen 2 HumVar	Possibly Damaging
MutationTaster	Disease causing
Mutation Assessor	FI M
FATHMM	Damaging
Condel	Deleterious
CADD	23.4

Figure 3.19. *In-silico* analyses of the *DIAPH2* p.I290V missense variant identified in family NSHL3. A. Schematic representation of DIAPH2 protein. The position of the p.I290V is indicated by a black star. **B.** Amino acid sequence alignments of DIAPH2 orthologs in the region surrounding the mutant residue (p.I290). Protein sequences were retrieved from UniProt, and alignments were generated with Clustal Omega. The amino acid residue affected by the mutation is shaded in red and indicated by a red arrow. Identical amino acids are marked by an asterisk, while partially conserved residues are indicated by a colon. **C.** Pathogenicity prediction of the p.I290V missense variant with 8 commonly-used software. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$).

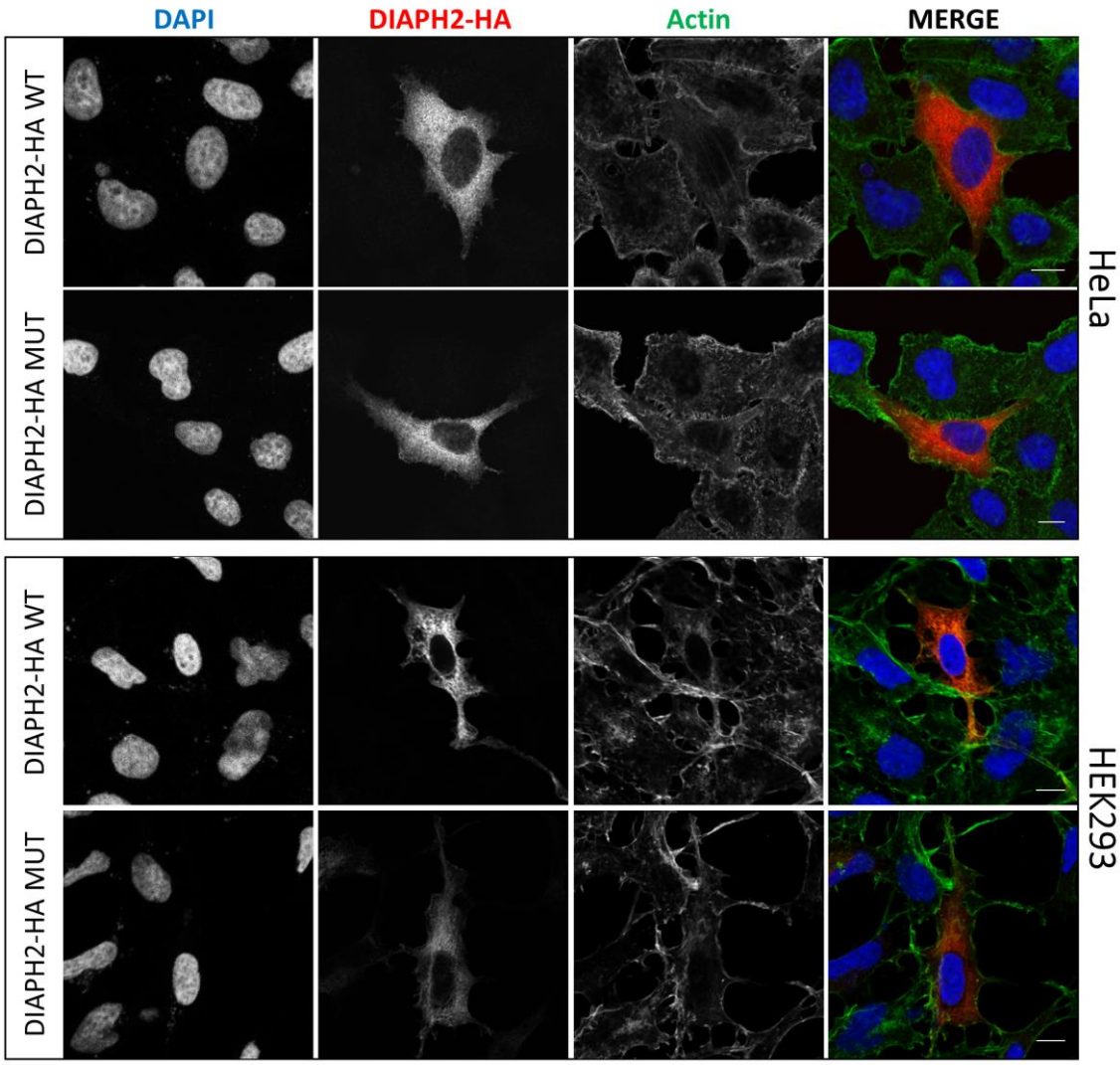


Figure 3.20. DIAPH2 immunolocalisation studies in basal conditions. DIAPH2 localisation was studied in HeLa (upper panel) and HEK293 (lower panel) cells 48 hours after transfection with plasmids coding for HA-tagged isoforms of wild-type or mutant DIAPH2 proteins (DIAPH2-HA-WT, DIAPH2-HA-MUT). Images were acquired with Leica True Confocal Scanner (TCS) SP8 and are representative of 3 experiments. Single confocal sections are shown. DAPI: 4',6-diamidino-2-phenylindole. Scale bars: 10 μm.

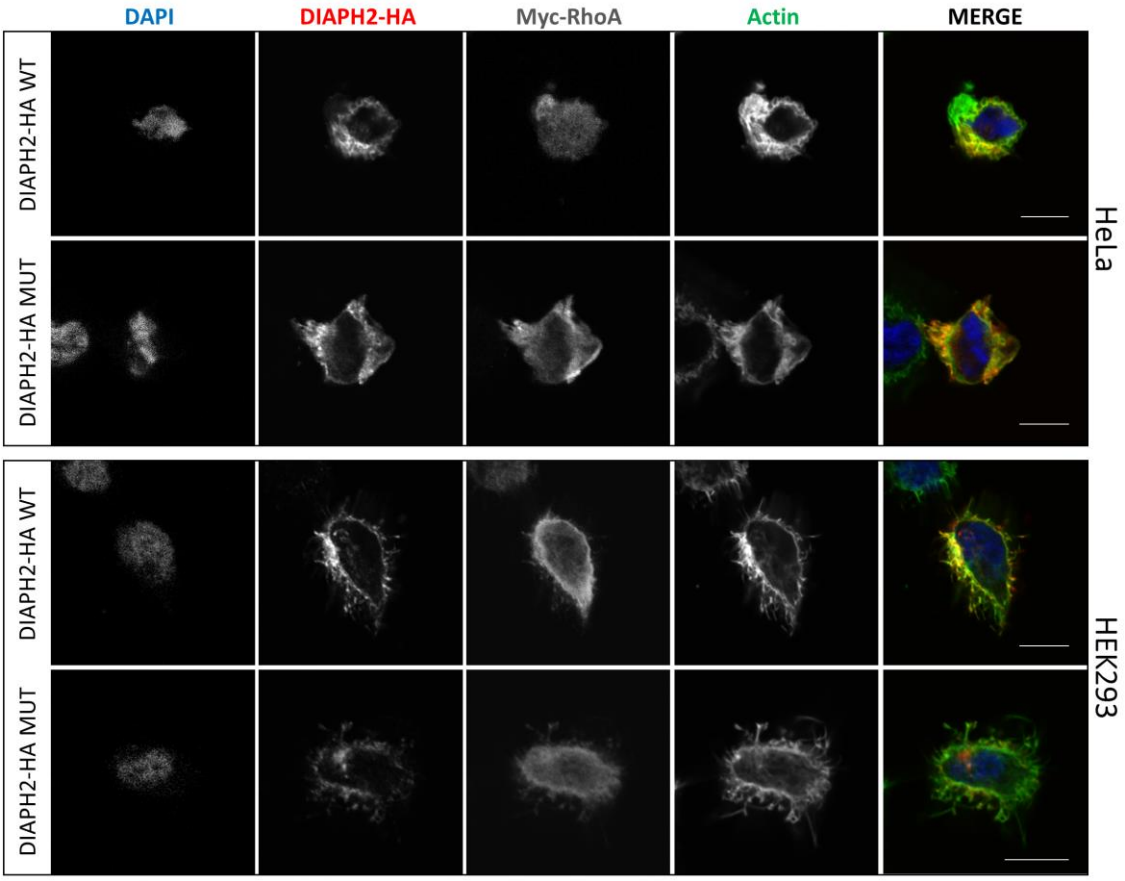


Figure 3.21. Co-localisation of DIAPH2 and actin under activating stimuli. DIAPH2 localisation was studied in HeLa (upper panel) and HEK293 (lower panel) cells 48 hours after transfection with plasmids coding for HA-tagged isoforms of wild-type or mutant DIAPH2 (DIAPH2-HA WT, DIAPH2-HA MUT), together with a Myc-tagged active form of RhoA (Myc-RhoA). On the right, merged image of the blue, red and green channels showing the partial co-localisation of DIAPH2 with actin. Images were acquired with Leica True Confocal Scanner (TCS) SP8 and are representative of 3 experiments. Single middle confocal sections are show. Scale bars: 10 μm.

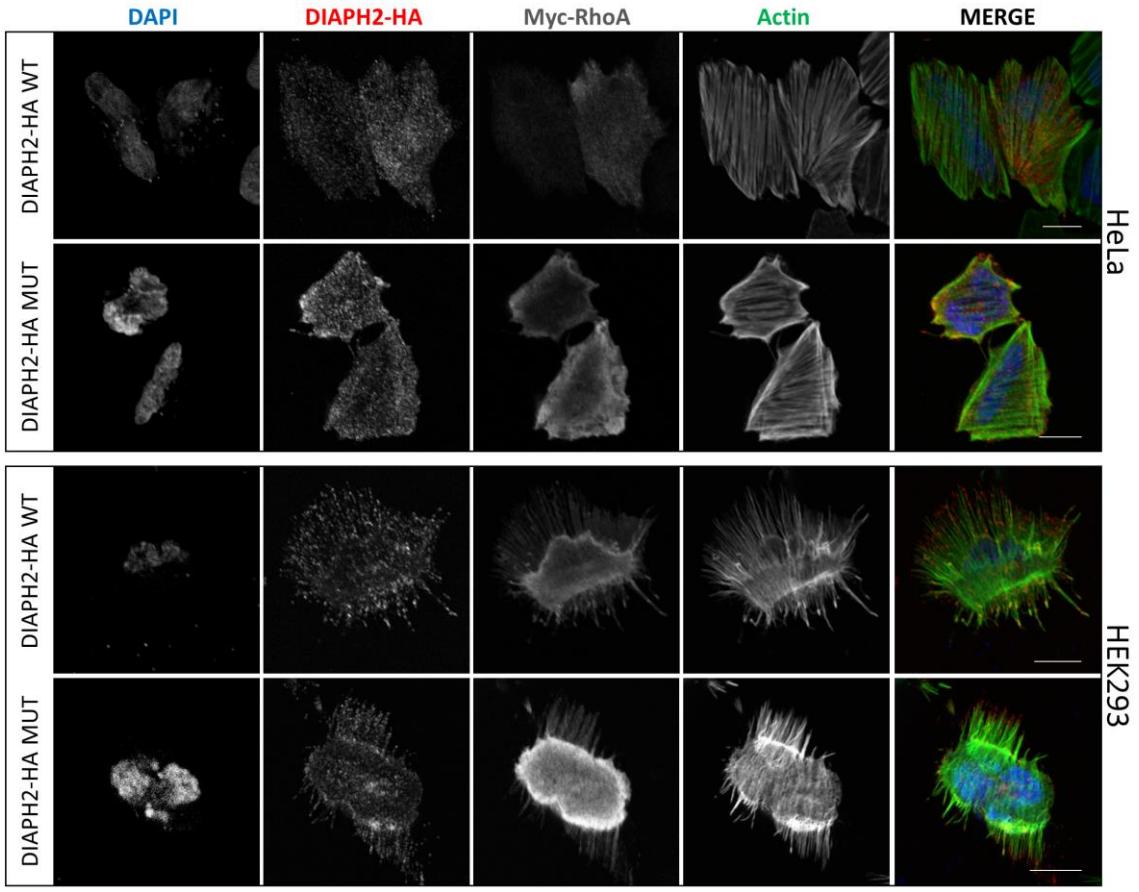


Figure 3.22. Evaluation of cell morphology and DIAPH2 localisation under activating stimuli. DIAPH2 localisation was studied in HeLa (upper panel) and HEK293 (lower panel) cells 48 hours after transfection with plasmids coding for HA-tagged isoforms of wild-type or mutant DIAPH2 (DIAPH2-HA WT, DIAPH2-HA MUT), together with a Myc-tagged active form of RhoA (Myc-RhoA). On the right, merged image of the blue, red and green channels. Images were acquired with Leica True Confocal Scanner (TCS) SP8 and are representative of 3 experiments. Single confocal sections cells at the level of their adhesion to the glass are shown. Scale bars: 10 μm.

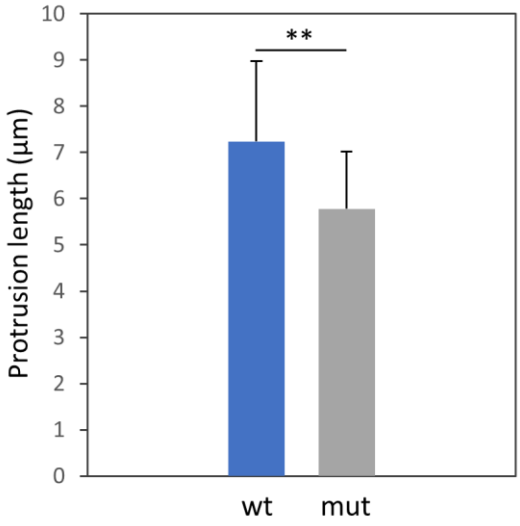


Figure 3.23. Evaluation of the length of membrane protrusions in HEK293 cells under activating stimuli. The length of membrane protrusions was measured using Fiji in HEK293 cells 48 hours after transfection with plasmids coding for HA-tagged isoforms of wild-type (wt) or mutant (mut) DIAPH2, together with a Myc-tagged active form of RhoA. The histograms represent the mean length of protrusions measured in 20 wt and 20 mut HEK293 cells, with error bars showing the standard deviations. Significance level of unpaired t-test (performed with GraphPad Prism 8.0) is shown (**: $p<0.01$).

3.6.3. Search for other HL patients with variants in *DIAPH2* gene

In order to validate *DIAPH2* as a novel NSHL gene, we aimed at identifying other HL patients carrying variants in the gene.

Since no variants were previously identified in 16 NSHL patients compatible with a X-linked pattern of inheritance from our cohort (data from Dr. Michela Robusto), *DIAPH2* gene was posted on GeneMatcher (http://www.genematcher.org), a web-based tool designed to help the identification of additional patients with variants in the same gene of interest [Sobreira *et al.*, 2015]. A total of 7 patients with hemizygous variants in the *DIAPH2* gene were identified through GeneMatcher. However, none of the patients matched with the HL phenotype (Table 3.4). Of note, the same p.R1075W variant was identified by two different groups in two patients with a different phenotype, whereas the p.R1091H variant is also present at the hemizygous state in one individual from the gnomAD database. At present, all these variations should be considered variant of

unknown significance (VUS), as no experimental evidence in support of their pathogenicity, nor of their actual involvement in producing the associated phenotype has been provided. Therefore, we still lack additional genetic data in support of *DIAPH2* role in HL, as well as in other human diseases apart from POF.

Table 3.4. GeneMatcher matches for *DIAPH2*.

	Variant				
Match	**NM_006729.4:**	**NP_006720.1:**	**Phenotype**	**gnomAD**	**CADD**
1	exon6:c.662+1G>A splicing variant	/	Intellectual disability	/	34
2	exon26:c.3223C>T	p.R1075W	Nephrotic syndrome	1.112e-5	24.4
3	exon26:c.3223C>T	p.R1075W	Cerebral palsy and epilepsy	1.112e-5	24.4
4*	exon27:c.3272G>A	p.R1091H	Congenital heart disease	1.169e-5	25.4
5**	na	na	Spasticity, motor delay and intellectual disability	na	na
6	na	na	Developmental delay, macrocephaly and epilepsy	na	na

* The variant is present at the hemizygous state in a male European non-Finnish individual from the gnomAD database.
**A de novo variant was identified in the hemizygous state in 2 male individuals. na: variant position not available.

3.6.4. Expression of the mouse ortholog protein Diap2 in the mouse inner ear

To better understand the physiologic function of DIAPH2 in the inner ear, I performed immunohistochemical studies to evaluate the expression of the mouse ortholog protein Diap2 in the mouse cochlea at different ages. Immunohistochemical studies showed a low Diap2 expression in the developing cochlear duct in E14.5 and E16.5 mice foetuses with a similar localisation of the protein in cells lining the cochlear duct. In fact, Diap2 seems to be principally expressed in the dorsal wall of cochlear duct, which will develop into the organ of Corti (Fig. 3.24 and 3.25, n=3 for each stage).

In E18.5 and P0 wild-type mice, Diap2 is highly expressed in the outer hair cells (OHCs), principally in the stereocilia. In addition, it is expressed in the Kölliker's organ and in the stria vascularis. Interestingly, the expression varies along the developing cochlea in these stages. In fact, the epithelial cells of the cochlear duct start to differentiate from the base to the apex and Diap2 is expressed in OHCs in the basal turn, while its expression towards the apex is localised in the luminal cells of the dorsal wall of the cochlear duct, similarly to the expression at earlier stages (Fig. 3.26 and 3.27, n=5 for each stage).

In P5 wild-type mice, Diap2 is expressed at low levels in the sensory epithelium, and it is mainly expressed in the stria vascularis, in the root cells and in the epithelial layer of the Reissner membrane (Fig. 3.28, n=5).

Immunohistochemical studies were also carried out in three older wild-type mice of the C57BL/6 strain (one P7 and two P14 mice), in which Diap2 was detected at very low levels in the stria vascularis only (data not shown).

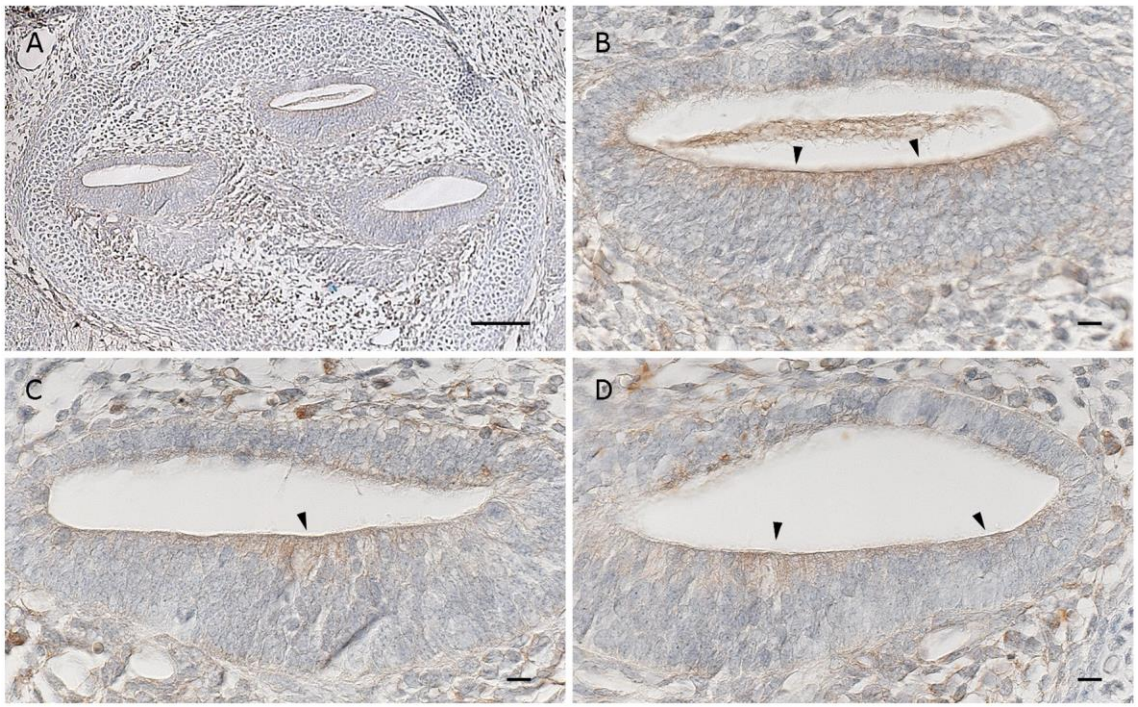


Figure 3.24. Diap2 expression in E14.5 wild-type mouse cochlea. A. Cross section of the cochlea of a E14.5 wild-type mouse embryo. At this stage, the scala vestibuli and scala tympani have not yet developed, and only one chamber – the developing scala media – is present. Also, the different cells composing the organ of Corti have not differentiated yet. Dorsal to the bottom. Scale bar: 100 μm. **B**-**D**. Higher magnification of the turns of the developing cochlear duct. Diap2 expression is indicated by arrowheads. Scale bars: 10 μm.

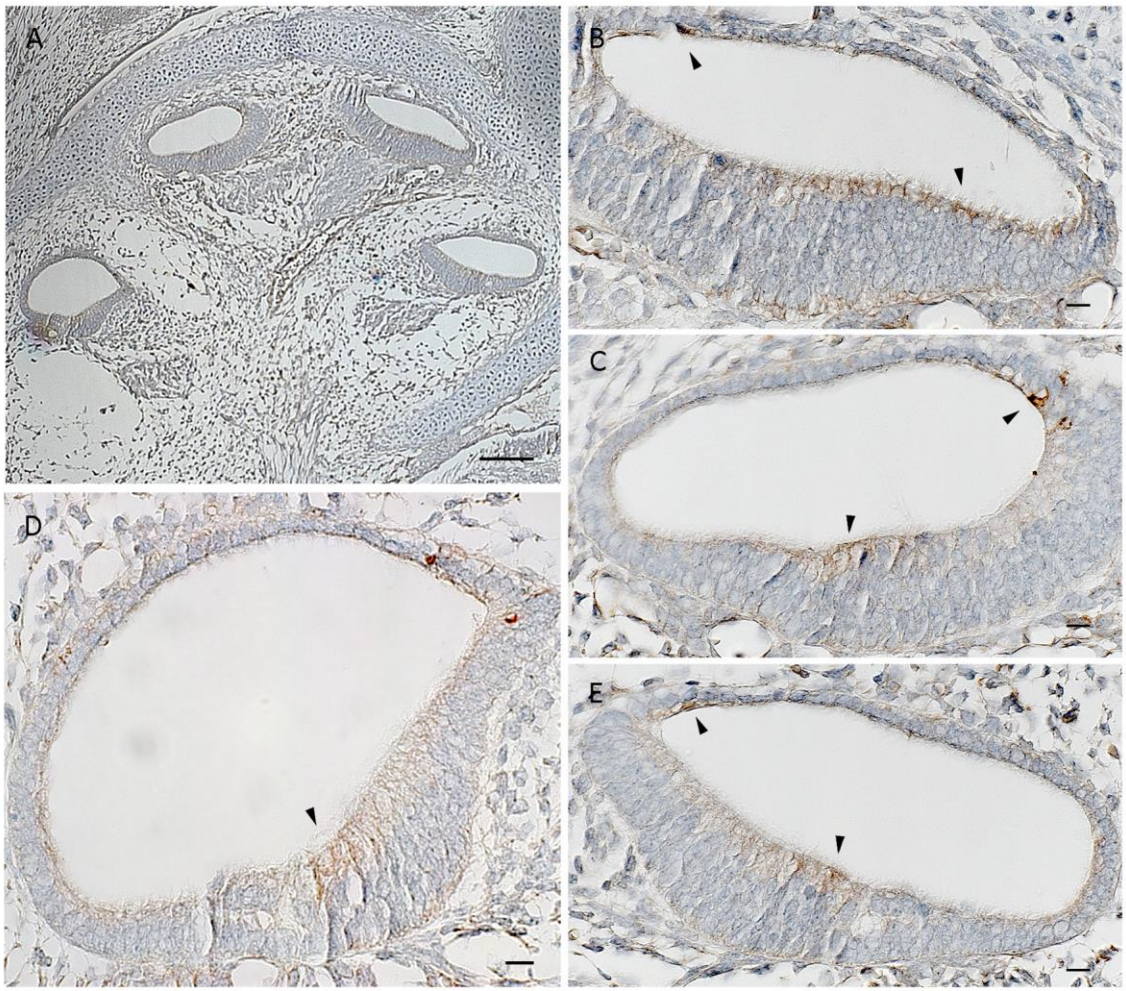


Figure 3.25. Diap2 expression in E16.5 wild-type mouse cochlea. A. Cross section of the cochlea of a E16.5 wild-type mouse foetus. At this stage, the scala vestibuli and scala tympani begin to open starting from the base of the cochlea. Dorsal to the bottom. Scale bar: 100 μm. **B-E.** Higher magnification of the apical (B), middle (C, E) and basal (D) turns of the developing cochlear duct. Arrowheads point to Diap2 expression. Scale bars: 10 μm.

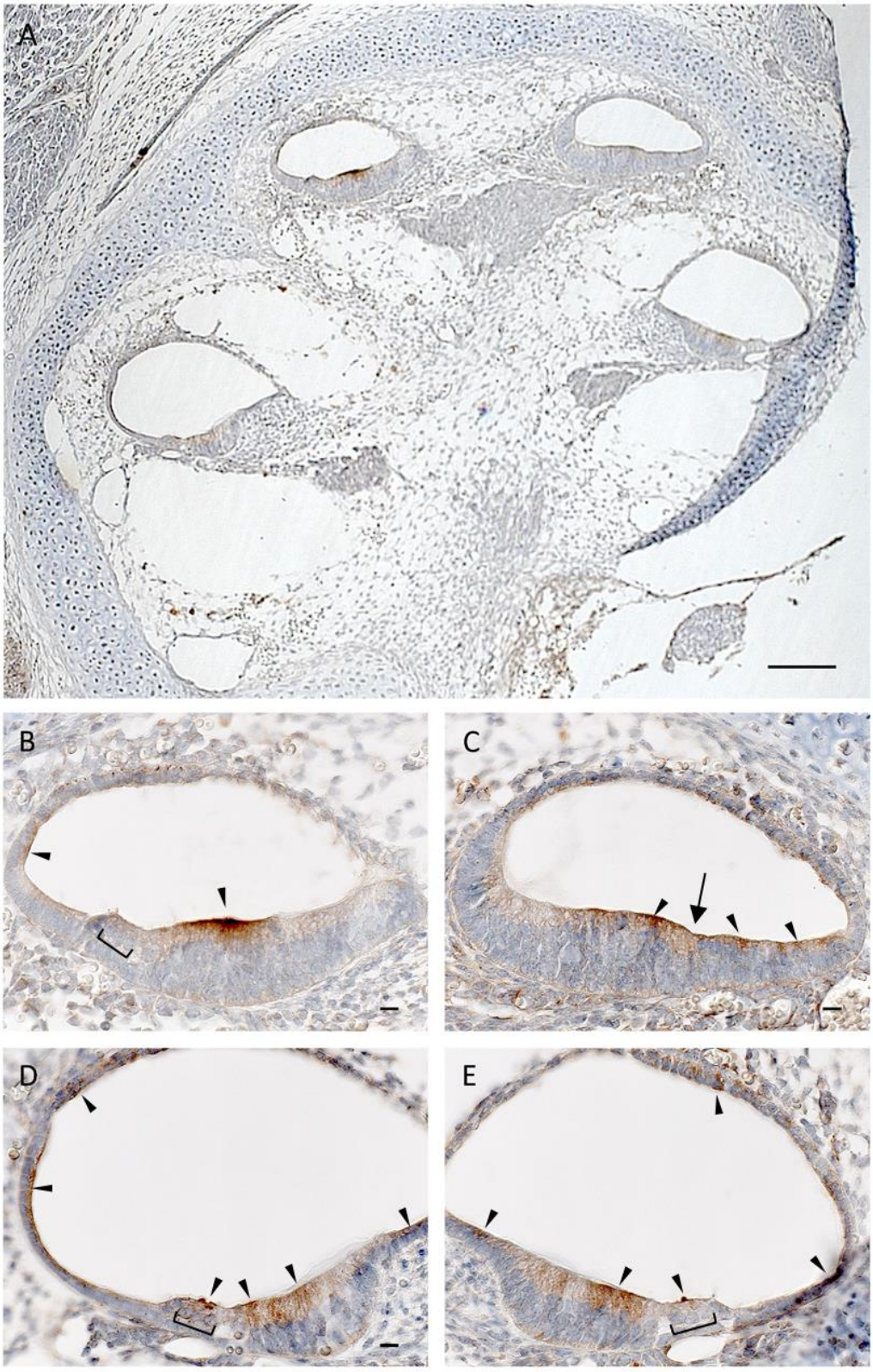


Figure 3.26. Diap2 expression in E18.5 wild-type mouse cochlea. A. Cross section of the cochlea of a E18.5 wild-type mouse foetus. At this stage, a different degree of cell differentiation can be observed from the base to the apex. Dorsal to the bottom. Scale bar: 100 μm. **B**-**E.** Higher magnification of the apical (C), middle (B, E) and basal (D) turns of the developing cochlear duct. The square brackets indicate OHCs. Diap2 expression is indicated by arrowheads. The arrow in C indicates the boundary between GER and LER. Scale bars: 10 μm.

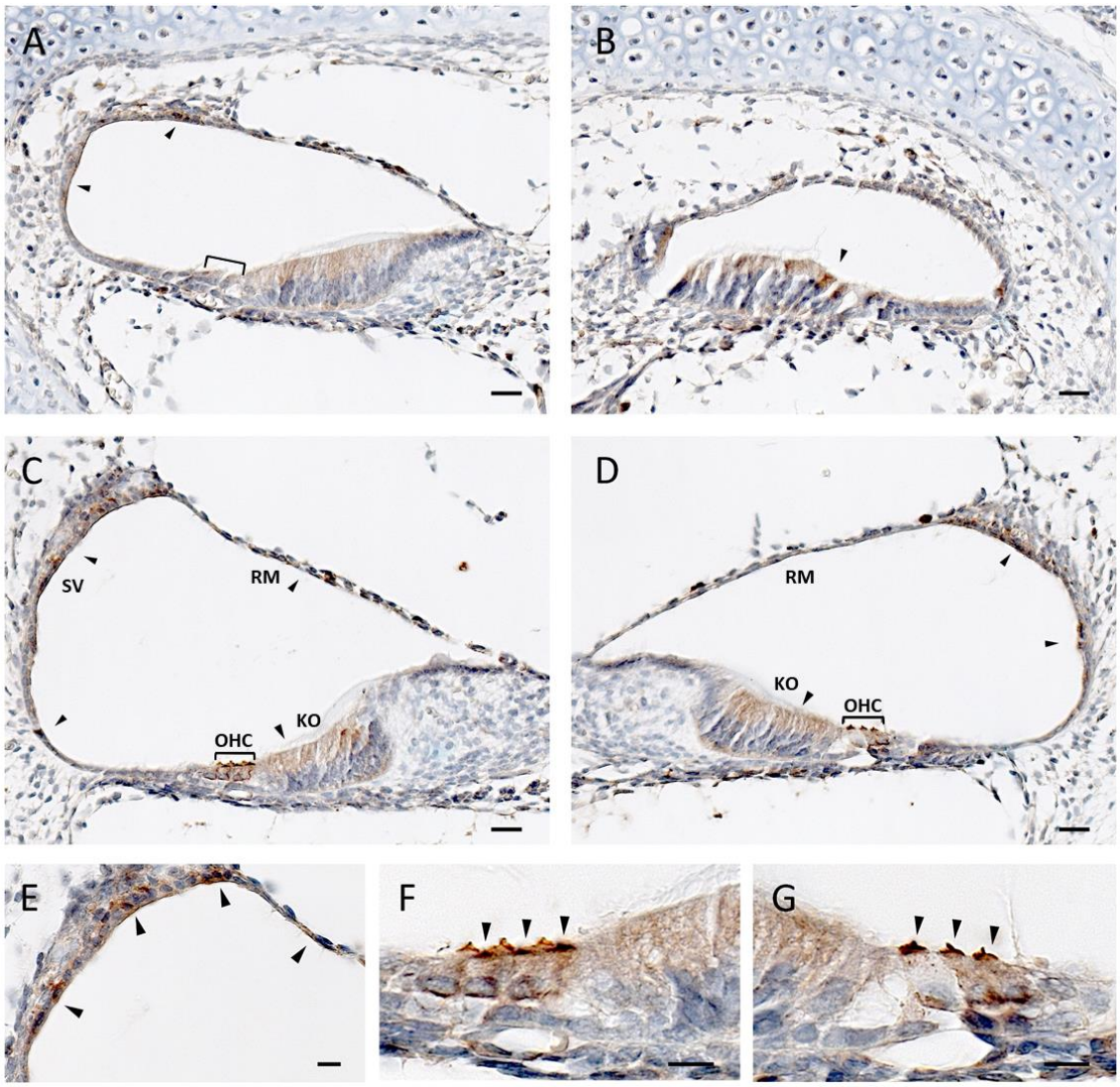


Figure 3.27. Diap2 expression in P0 wild-type mouse cochlea. A-D. Cross section of the cochlear turns (B: apical, A, D: mid, C: basal turn) of a P0 wild-type mouse. At this stage, a different degree of cell differentiation can be observed from the base to the apex. The square brackets indicate OHCs (outer hair cells). KO: Kölliker's organ, RM: Reissner membrane, SV: stria vascularis. Dorsal to the bottom. **E.** Higher magnification of the stria vascularis in C. **F-G.** Higher magnification of the OHC shown, respectively, in C and D. Diap2 expression is indicated by arrowheads. Scale bars: 10 μm.

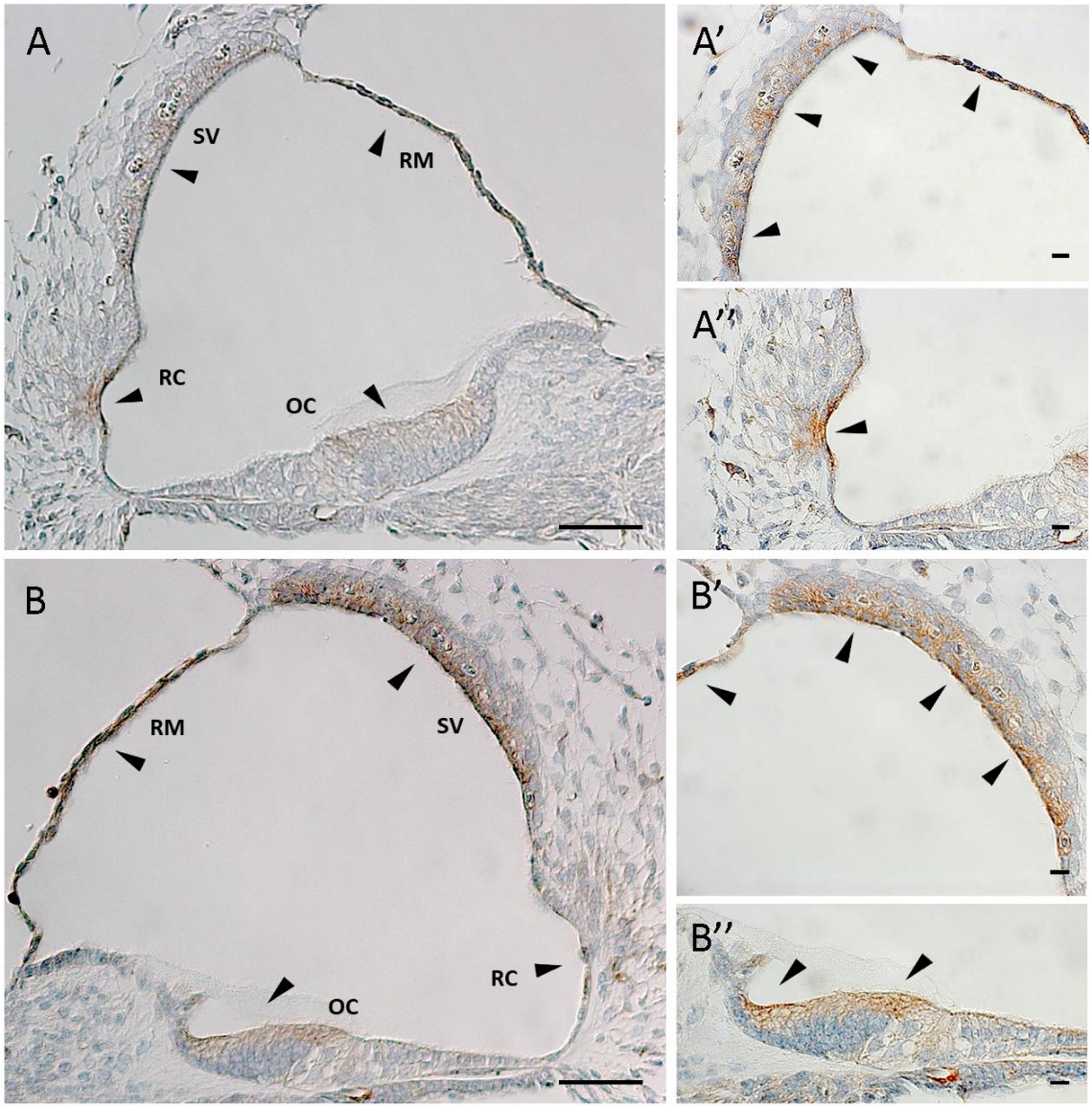


Figure 3.28. Diap2 expression in P5 wild-type mouse cochlea. A. Scala media of the middle turn of cochlea. Scale bar: 100 μm. **A'-A".** Higher magnification of the stria vascularis and of the outer sulcus shown in A. Scale bars: 10 μm. **B.** Scala media of the basal turn of cochlea. Scale bar: 100 μm. **B'-B".** Higher magnification of the stria vascularis and of the organ of Corti shown in B. Scale bars: 10 μm. Arrowheads point to Diap2 expression. OC: organ of Corti; RC: root cells; RM: Reissner membrane; SV: stria vascularis.

3.6.5. Auditory phenotype of the *Diaph2*^{em2Kcl} knock-out mouse

I evaluated the auditory function of the *Diaph2*^{em2Kcl} knock-out mice using ABR recordings at 4 and 8 weeks of age. In details, I tested 7 wild-type (4 males and 3 females), 2 heterozygous female and 6 hemizygous male mice at 4 weeks of age. Two wild-type males and 4 hemizygous males were also tested at 8 weeks of age. The remaining mice were not yet 8 weeks old at the end of my research period at Professor Steel's laboratory.

As no difference was observed in wild-type males and females, all wild-type mice were plotted together and compared to the heterozygous and hemizygous groups. A statistically significant difference in hearing thresholds between wild-type and hemizygous mice groups was only observed at 36 kHz (Fig. 3.29A); however, this small difference is probably not biologically relevant. In addition, no difference between genotypes was observed in the mice tested at 8 weeks (Fig. 3.29B). These first results on a limited number of animals suggest that lack of *Diaph2* does not impair the development of normal hearing function in mice. Nonetheless, a later onset of HL cannot be excluded, and older mice need to be tested.

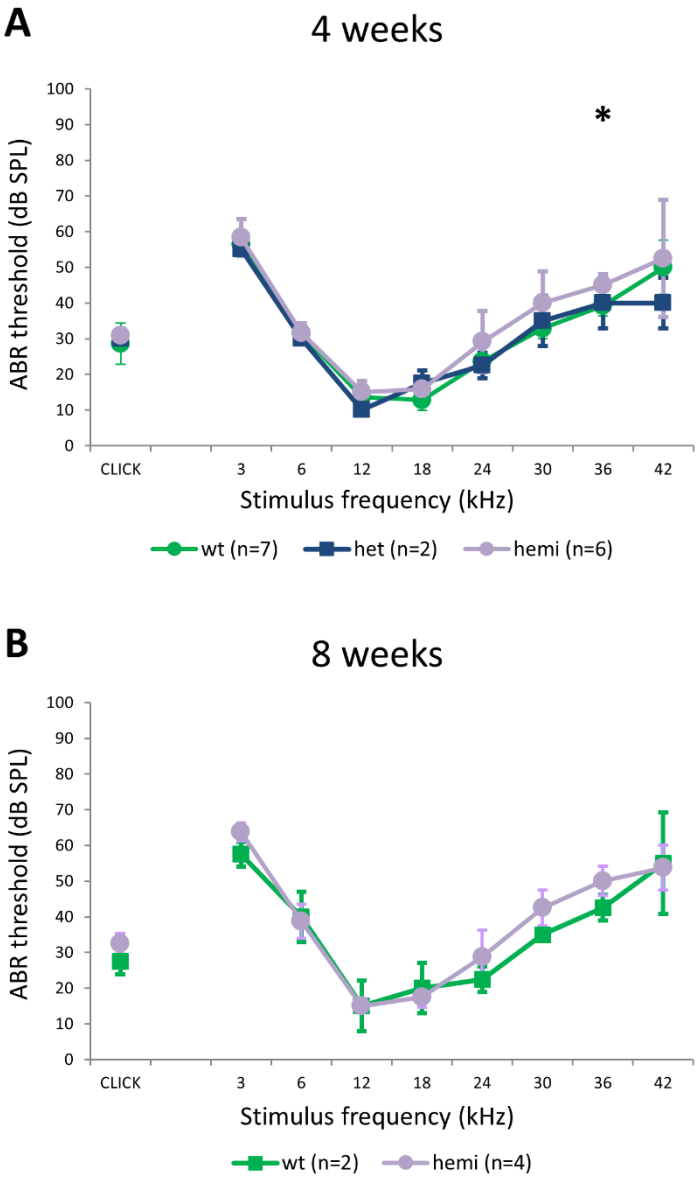


Figure 3.29. ABR recordings of *Diaph2*^{em2Kcl} knock-out mice at 4 and 8 weeks. Mean ABR thresholds (± SD) for clicks and tone pips are plotted for wild-type (wt, in green), heterozygous females (het, in blue) and hemizygous (hemi, in lilac) mice aged 4 (**A**) and 8 (**B**) weeks. **A.** Data were not normally distributed and were thus analysed by the Kruskal-Wallis test followed by the Dunn's *post-hoc* analysis to compare the three genotype groups. A statistically significant difference in hearing thresholds was only observed at 36 kHz. Asterisks indicate p-values (*: $p<0.05$). **B.** Mann-Whitney test was performed to compare ABR threshold of 8-week-old wild-type and hemizygous mice at each frequency. Only 6 (2 wild-types and 4 hemizygotes) of the 15 mice tested at 4 weeks were also tested at 8 weeks. No difference between genotypes was observed at 8 weeks. All statistical analyses were performed with GraphPad Prism 8.0. ABR measurements for each individual mouse are shown in Appendix 6.3.

3.6.6. Auditory phenotype of *the Diaph2*^{em3Kcl} knock-in mouse

In order to evaluate the effect of the specific variant identified in family NSHL3, ABR measurements were also performed in *Diaph2*^{em3Kcl} knock-in mice. A total of 5 wild-type (3 females and 2 males) and 5 hemizygous mice were tested at 4 weeks. Three wild-types (2 females and 1 male) and 3 hemizygotes were tested at 8 weeks of age. Wild-type mice of both sexes were plotted together, since no difference was observed among them. At 4 weeks, a statistically significant difference in hearing thresholds between wild-type and hemizygous groups was observed for thresholds at 30 kHz (Fig. 3.30A). This small difference was considered of no biological meaning, especially as it was not observed in 8-week-old mice (Fig. 3.30B). Even if the number of animal tested is very limited, these first recordings did not show any relevant impairment of hearing function at least at the ages tested, indicating that the point mutation in *Diaph2* gene (NM_172493.2:c.877A>G; NP_766081.1:p.I293V) does not affect the development of normal hearing in mice.

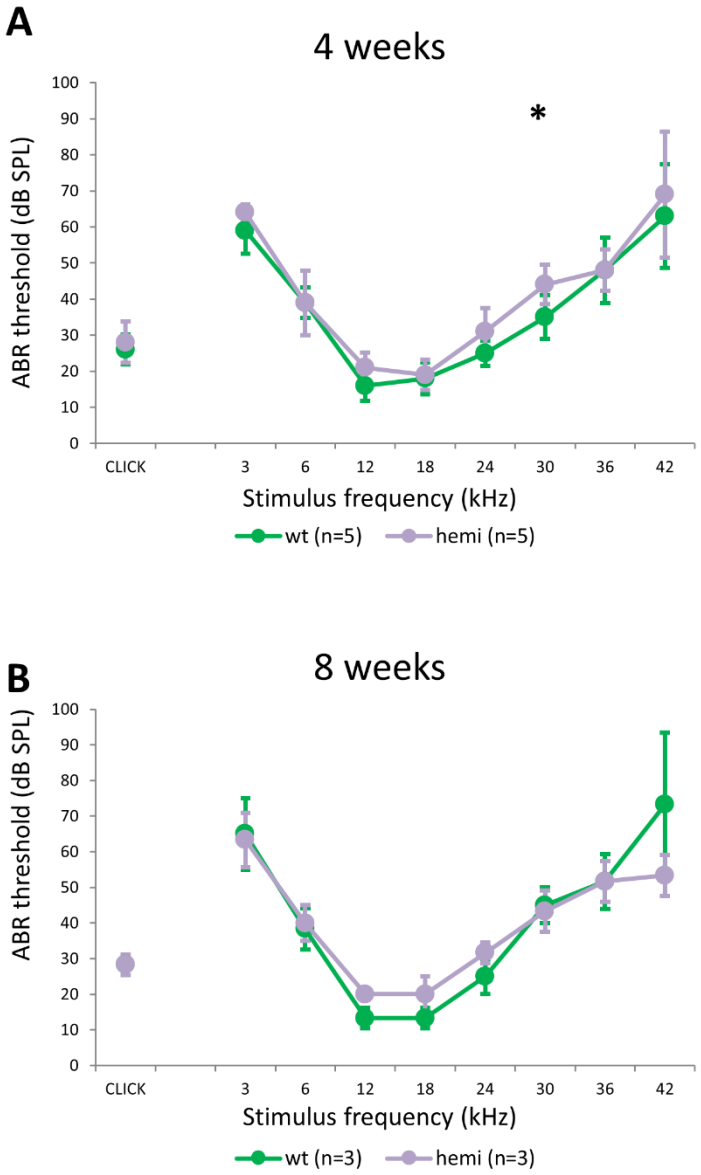


Figure 3.30. ABR recordings of *Diaph2*^{em3Kcl} knock-in mice at 4 and 8 weeks. Mean ABR thresholds (± SD) for clicks and tone pips are plotted for wild-type (wt, in green) and hemizygous (hemi, in lilac) mice aged 4 (**A**) and 8 (**B**) weeks. Mann-Whitney test was performed to compare ABR threshold of 4- (**A**) and 8- (**B**) week-old wild-type and hemizygous mice at each frequency. Only 6 (3 wild-types and 3 hemizygotes) of the 10 mice tested at 4 weeks were also tested at 8 weeks. A statistically significant difference in hearing thresholds was observed at 30 kHz at 4 weeks (**A**), but no difference was observed in 8-week-old mice (**B**). All statistical analyses were performed with GraphPad Prism 8.0. ABR measurements for each individual mouse are shown in Appendix 6.4.

3.6.7. Discussion and conclusions on NSHL3 family

Although the possibility to screen hundreds of genes by WES is a powerful tool for detecting potentially pathogenic variants underlying NSHL, the actual demonstration of the causal role of mutations - especially when found in novel candidate genes - still represents a critical and limiting step. In family NSHL3, WES variant prioritisation and segregation analyses pointed out the NM_006729.4:c.868A>G (p.I290V) variant in *DIAPH2* gene as the most likely candidate mutation for the prelingual HL of II3 and II5 individuals. Despite the progress made in this work in elucidating the function of *DIAPH2* gene, the actual pathogenic effect of the variant and the role of the gene in hearing function remain elusive. For instance, despite the predicted effect of the variant on splicing and the observation of exon 8 skipping in one single *in-vitro* assay, the irreproducibility of our results led us to exclude the c.868A>G variant as a splicing mutation.

Furthermore, the immunofluorescence studies of DIAPH2 wild-type and mutant protein did not demonstrate a major effect of the p.I290V missense variant on DIAPH2 protein localisation. Although a statistically significant reduction in the average length of the RhoA-induced membrane protrusions in HEK293 cells expressing mutant DIAPH2 was observed, it could be argued that a modest length difference might not be biologically relevant. It is important to consider that, in the cochlea, the actin filaments are essential components of the sensory hair-cell stereocilia, whose length is precisely regulated. Indeed, actin dynamics play a critical role in this regulation. As the precise architecture of the hair bundle is required for the synchronous gating of mechanotransduction channels [Narayanan *et al.*, 2015], it is likely that even a modest disruption of this finely regulated structures could have an effect on the signal transmission. In addition, it should be noted that, in HEK293 cells, RhoA activation of endogenous DIAPH2 - and possibly of other formins - might have partially compensated for the effect of the mutant protein.

Concerning the physiologic role of DIAPH2 in hearing, immunohistochemistry studies in wild-type mice gave some insights into the expression pattern of the mouse ortholog protein Diap2 in the cochlea. Indeed, Diap2 showed a low expression in the developing cochlear duct at E14.5 and E16.5 and a specific strong expression in outer hair-cell stereocilia at later stages (E18.5 and P0). Finally, in P5 mice, Diap2 was mainly restricted to structures important for the cochlear fluid homeostasis (root cells, stria vascularis). These data suggested a possible function of Diap2 in the development of stereocilia of OHCs supporting our hypothesis of a role of DIAPH2 as an actin regulator in the cochlear sensory cells. Also, the early expression in the cochlea could explain the prelingual onset of NSHL in family NSHL3. On the other hand, the expression of Diap2 – postnatally - in structures required for fluid homeostasis and endocochlear potential maintenance, suggest other possible roles of the protein in hearing function.

As mouse mutants have proven a highly effective research tool to investigate the molecular processes affected as a consequence of mutations in genes critical for hearing, we aimed at characterising the auditory phenotype of *Diaph2* knock-out and knock-in mice carrying the corresponding variant identified in NSHL3 family. However, the first ABR recordings, performed in 4- and 8-week-old mice, did not indicate a clear hearing phenotype in mutant mice, as compared to wild-type littermates. Of note, the phenotype of a knock-out model of *Diaph2* was recently reported by the IMPC (*Diaph2*$^{tm1b(EUCOMM)Hmgu}$ allele; http://www.mousephenotype.org/data/genes/MGI:1858500). Although no significant sign of hearing impairment was reported in the 14-week-old mice tested, homozygous knock-out females showed slightly higher threshold at high frequencies (30 kHz). Therefore, before ruling out *DIAPH2* as a deafness gene, it is essential to extend ABR measurements to more mice of all genotypes and to repeat the experiments in older mice.

However, given a possible functional redundancy in the *Diaph* gene family, it is conceivable that the absence of a hearing phenotype may be due to a compensation of *Diaph2* loss of

function by the other *Diaph* genes. In support of this hypothesis, it was reported that the *Diaph1* knock-out mice don't recapitulate the microcephaly phenotype described in humans with homozygous truncating mutation in *DIAPH1*. However, *Diaph1/Diaph2* double knock-outs showed several abnormalities in neural development [Toyoda *et al.*,_2013]. Indeed, while both *Diaph1* and *Diaph2* are required for the assembly of the apical actin belt in neuroepithelial cells, the loss of only *Diaph1* does not affect this structure [Ercan-Sencicek *et al.*, 2015]. A similar redundancy of *Diaph* genes might thus also exist in the cochlea.

Although we have not collected enough evidence to support *DIAPH2* causality in the pathogenesis of deafness, the c.868A>G variant is still the best candidate resulting from exome data. Therefore, it would be very important to verify the presence of pathogenic variants in the gene by gaining access to genomic data from larger cohorts of hearing-impaired individuals. Nevertheless, it is also possible to speculate that variants in other genes could have a synergic role together with *DIAPH2* variant and account for family NSHL3 phenotype. Finally, the possibility cannot be excluded of intronic/intergenic variants or other complex rearrangements not detectable using WES.

3.7. Genetically undiagnosed NSHL families: NSHL16, NSHL17, NSHL19

In three families (NSHL16, NSHL17, NSHL19), all compatible with an autosomal dominant inheritance (Fig. 3.31), the search for NSHL-causing variant is still ongoing.

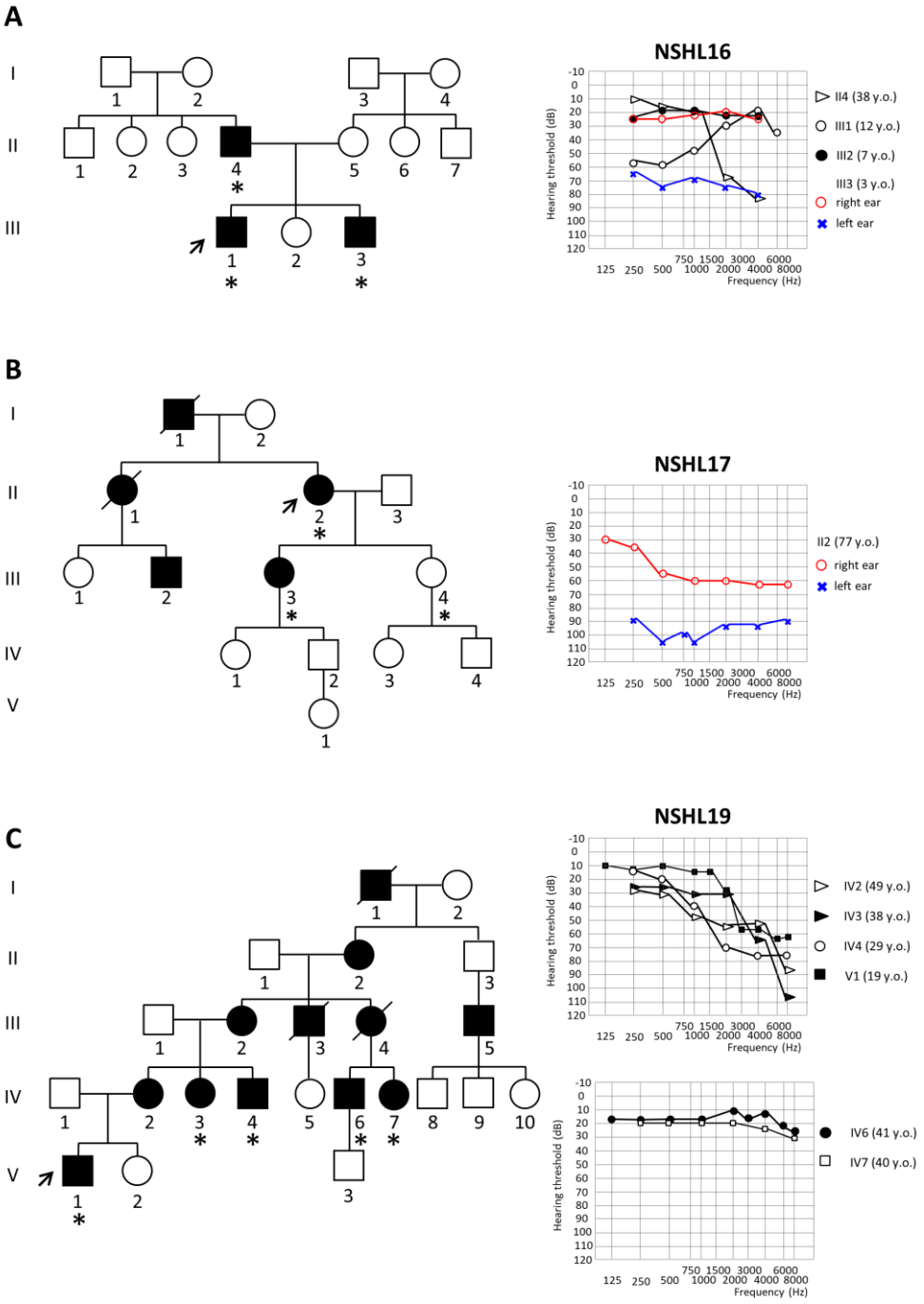


Figure 3.31. Pedigree and audiograms of genetically undiagnosed NSHL families. The pedigree and audiograms of family NSHL16 (**A**), NSHL17 (**B**) and NSHL19 (**C**) are shown. The arrows point to the probands, whereas asterisks indicate individuals selected for WES. The threshold average for the right and left ear (in black) or the distinct threshold for each ear (right in red, left in blue) and the age at audiometric evaluation for each individual are shown.

In family NSHL16, individual II4 is a 52-year-old man who developed a bilateral high-frequency HL at the age of 8 years old. He has a normal-hearing daughter (III2, 22 years old) and two deaf sons (III1, 26 years old and III3, 19 years old), who developed NSHL at 10 and 3 years old, respectively. Interestingly, the hearing impairment is different between the two siblings, being bilateral and more severe in the low frequencies in III1 and unilateral and severe in all frequencies in III3 and thus differing also from the father's phenotype (Fig. 3.31A). WES data of the three affected individuals were first analysed considering a dominant inheritance. Given the high number (> 200) of heterozygous variants shared between the affected family members, a further selection was performed. First, variants identified in genes repeatedly mutated in the exomes sequenced in our laboratory, and representing false positive of the technology, were filtered out. Then, the remaining variants were further prioritised considering their frequency in the gnomAD database and the possible role of the gene affected in hearing function (Fig. 3.32). However, the variants identified in known NSHL gene were not confirmed by Sanger sequencing, being WES false positives. Also, none of the variants in novel putative NSHL genes were segregating with the phenotype in the family, being present also in the normal hearing sister (III2).

	TECTA	*KCNQ4*	*VANGL1*	*GAD1*	*DNAH5*
	DFNA12	DFNA2A	Encodes a planar cell polarity component, involved in the hair cell polarity [Fukuda *et al.*, 2014].	Encodes the enzyme responsible for the production of GABA; it is expressed in the cochlea [Lin *et al.*, 2003].	Encodes a component of microtubule-based cilia, highly expressed in vestibular hair cells [Scheffer *et al.*, 2015].
Predictor	**p.G38E**	**p.V698M**	**p.E349G**	**p.G433E**	**p.G4202S**
SIFT	damaging	damaging	damaging	tolerated	damaging
Polyphen 2 HumDiv	probably damaging	benign	benign	benign	probably damaging
Polyphen 2 HumVar	possibly damaging	benign	benign	benign	probably damaging
MutationTaster	disease causing	polymorphism	disease causing	disease causing	disease causing
Mutation Assessor	FI M	FI L	FI M	FI L	FI H
FATHMM	damaging	tolerated	damaging	tolerated	tolerated
Condel	neutral	deleterious	deleterious	neutral	deleterious
CADD	27.4	17.06	22.9	23.5	32
Database					
gnomAD	0	0	0	0	2/250916 (7.971e-6)
in-house exomes	0	0	0	0	0

Figure 3.32. Variants shared between affected individuals screened by Sanger sequencing. Pathogenicity prediction of the variants selected for segregation analyses with 8 commonly-used software and presence in databases is shown. The variants with CADD Scaled C-score > 20 were further selected based on the possible function in hearing. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$).

Since no signs of hearing impairment were reported in the parents of the proband II4 (I1 and I2), I also explored the less likely possibility of a recessive segregation. Under this hypothesis, individual II4 would carry two mutant alleles (one inherited from the mother and one inherited from the father) and would have transmitted one of them to his sons. Therefore, III1 and III3 might have inherited the same or a different allele from the father and would at least share the mutant allele of the gene inherited from the mother. Consequently, WES data were re-analysed under a recessive transmission model, focussing on genes with two prioritised variants, of which at least one shared between affected siblings and one shared with their father. Of note, each individual carried two variants that passed the filters of our prioritisation pipeline in at least one autosomal recessive NSHL gene, but that were not annotated as pathogenic in public databases. However, I couldn't identify any shared prioritised variant fitting the hypothesised recessive mode of inheritance.

Family NSHL17 shows autosomal dominant HL with late-onset (Fig. 3.31B). Both the proband II2 (91 years old) and her daughter III3 (71 years old) developed bilateral HL - more severe in the left ear - at ~40 years; the other daughter (III4, 65 years old) shows normal hearing. WES data analysis performed on II2, III3 and III4 yielded ~400 variants passing quality and prioritisation filters in each individual. After filtering out likely false positive of WES, less than 60 variants were shared among affected family members and were not present in the normal-hearing sibling. Among these, I identified an intronic heterozygous deletion in the NSHL gene *TMC1* (NM_138691:c.1696-14_1696-13delTG, 13 bp upstream of the acceptor splice site of exon 19), which was confirmed by Sanger sequencing. The identified variant is reported in the gnomAD database in the heterozygous state in 34 out of 125640 individuals (allele frequency: 1.353e-4). On the other hand, the variant was predicted to have a possible effect on *TMC1* exon 19 acceptor site recognition by 3 out of 3 prediction software used (NNSPLICE, Human Splicing Finder and NetGene). Therefore, I further investigated the potential pathogenicity of the 2-bp deletion using an *in-vitro* assay. In details, exon 19 of the *TMC1* gene with the surrounding intronic sequences, either wild-type or carrying the identified variant, was cloned into the pBS-KS_modified hybrid minigene vector (Fig. 3.33A). The obtained constructs (pBS-KS-TMC1_ex19_wt and pBS-KS-TMC1_ex19_mut) were transiently transfected into HeLa, HEK293 or UB/OC-2 cells and *TMC1* splicing products were analysed by RT-PCR assays. No alternative splicing isoforms were observed in any cell line transfected with the mutant construct (Fig. 3.33B, n=2), as also confirmed by a competitive fluorescent RT-PCR assay (not shown). These results seem to exclude the pathogenicity of the NM_138691:c.1696-14_1696-13delTG variant. None of the other variants affecting known deafness genes were shared between the affected individuals only, and no promising novel candidate gene was identified, thus leaving NSHL17 family without a genetic diagnosis.

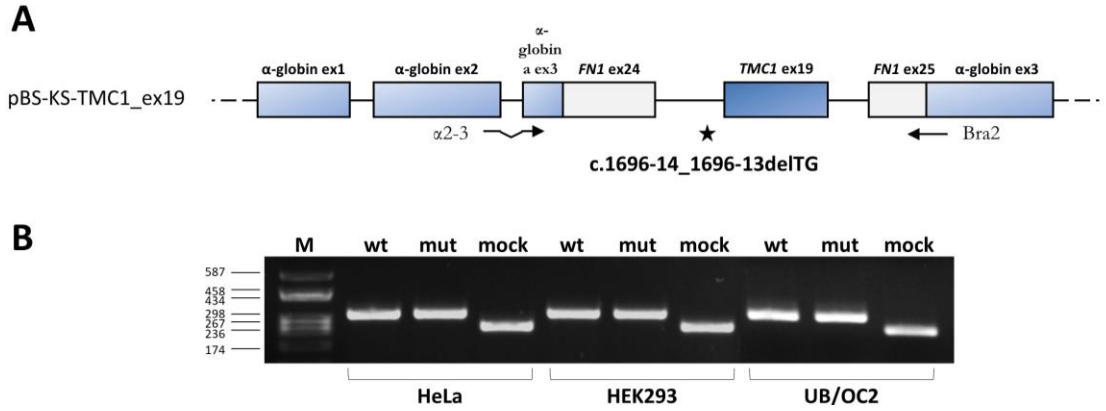


Figure 3.33. *In-vitro* analysis of the impact of c.1696-14_1696-13delTG variant on *TMC1* pre-mRNA splicing. A. Schematic representation of the hybrid pBS-KS-TMC1_ex19 minigene where α-globin exons are represented by light blue boxes, fibronectin (*FN1*) exons by grey boxes, whereas introns are shown as black lines (not to scale). Exon 19 of *TMC1* is represented by a blue box. The c.1696-14_1696-13delTG variant in intron 18 is indicated by a black star. Black arrows represent the primers used in RT-PCR assays. **B.** Agarose gel electrophoresis of RT-PCR products obtained from RNA of HeLa, HEK293 and UB/OC-2 cells under proliferative conditions, transfected with the wild-type (wt), mutant (mut), or empty (mock) minigene vector. M: molecular weight marker (pUC9/*HaeIII*).

In family NSHL19, the 25-year-old proband V1 developed moderate sensorineural HL at the high frequencies at the age of 17 years. The mother IV2 (53 years old) and her siblings IV3 (52 years old) and IV4 (43 years old) also show a similar impairment, more severe at the high frequencies. The cousins of IV2-4 (IV6, 46 years old and IV7, 45 years old) were also available to the study and show a mild hearing impairment at the high frequencies (Fig. 3.31C). Despite having analysed WES data from 5 individuals (IV3, IV4, IV6, IV7 and V1), we could not identify any candidate pathogenic variant shared between all affected family members. More than 250 variants were shared between the 5 individuals, but only 21 genes were involved. Most of these genes were excluded, as variants in these genes are known to be false positives of WES. None of the shared prioritised variants affected a deafness-associated gene, however affected individuals carried from 4 to 18 different passing-filters variants in known NSHL genes. Interestingly, individuals IV6 and IV7, showing a milder hearing impairment, carried respectively 4 and 6 variants only. It would be thus interesting to evaluate if these variants of unknown significance might contribute to the phenotype.

3.7.1. Discussion on genetically undiagnosed families

Despite the unquestionable progress promoted by NGS technologies in the detection of potentially pathogenic variants underlying heterogeneous diseases, the diagnostic yield of WES for NSHL is still under 50% [Neveling *et al.*, 2013; Likar *et al.*, 2018; Sheppard *et al.*, 2018]. In this work, the difficulties we encountered in identifying the genetic cause of deafness in three of the analysed families could originate from the WES analysis pipeline applied or from the intrinsic limitations of the technology.

For instance, synonymous variations located outside canonical splicing sites were considered non-pathogenic and filtered out. However, there is increasing evidence that synonymous variants can have a deleterious effect on the protein, for example disrupting splicing enhancer or silencer elements, affecting mRNA stability or altering the translation kinetics [Sauna and Kimchi-Sarfaty, 2011]. Therefore, the disease-causing variants might have been missed in the first step of the prioritisation pipeline. In addition, setting the MAF threshold at 0.1% for variant filtering in dominant families, generated a list of thousands of variants for each individual. The selection of the variants only shared between affected individuals considerably reduced the number of possible pathogenic variants; however, the selected variants were still too numerous, making experimental validation unfeasible without a further selection. In absence of mutations in known deafness genes, variants were further prioritised based on their predicted pathogenicity (taking CADD score into account) and the possible role of the gene affected in hearing function. This approach might have excluded the real causative variants from further validation analyses.

In the case of family NSHL16, an additional point to raise is the intrafamilial variability of the phenotype, that I attempted to explain considering a recessive inheritance pattern with affected individuals carrying different variants in the same deafness gene. Even if this hypothesis was not confirmed, it cannot be excluded that other genetic or environmental factors could play a role in this variability. In fact, the variability between ears of individual

III3 might suggest the presence of genetic modifiers, as well as the presence of stochastic effects contributing to the development of hearing loss. Therefore, it should also be considered that potential causes of acquired HL might have not been reported in the patient's medical history and that it might be necessary to modify data analyses accordingly. Another limitation of WES data analysis is the detection of CNVs. Although a CNV analysis from WES data was attempted, the interpretation of the output is particularly challenging when searching for heterozygous variants. In fact, the read count approach of the tools used for CNV detection is affected by the nonuniform efficiency of exome capture [D'aurizio *et al.*, 2016]. Therefore, it could be worth integrating the analysis of the undiagnosed families using specific technologies for the detection of CNVs and chromosomal changes, such as arrayCGH.

It is important to note that some of the limitations in the detection of pathogenic variants are intrinsic to WES technology. For instance, WES is affected by a reduction of coverage in the GC-rich regions, as a result of the PCR stage [Veal *et al.*, 2012]. Also, the detection of variants is particularly challenging in duplicated regions (e.g. genes with pseudogenes or paralogues) due to reads misalignment. Indeed, many of the rare heterozygous variants identified in the three undiagnosed families were in genes (e.g. belonging to mucin and ankyrin gene families), which are frequently found mutated and that are considered false positive of WES [Fuentes Fajardo *et al.*, 2012]. Besides false positives, the difficulties in calling variants in duplicated regions might also negatively affect the identification of mutations in deafness genes with pseudogenes (i.e. *STRC*) or located within segmental duplications (such as *OTOA*) [Shahin *et al.*, 2010; Mandelker *et al.*, 2014].

Finally, although being highly enriched in possible pathogenic variants, the exome represents only the 1-2% of the genome and variants located in deep-intronic or intergenic regions are obviously missed using WES.

3.8. SHL families: Alport syndrome

We selected 3 AS families (AS1 and AS2 with a likely autosomal dominant or X-linked inheritance and AS3 compatible with a recessive pattern of inheritance) who had been previously subjected to several genetic screenings (including targeted resequencing of the 3 AS genes *COL4A3*, *COL4A4*, and *COL4A5* by Roche 454) without obtaining a molecular diagnosis. We subjected to WES the proband of each AS family and first focused data analysis on known Alport genes only. We verified that all exons of *COL4A3*, *COL4A4*, and *COL4A5* were adequately covered by WES, with *COL4A3* exon 1 being the least well-covered exon (mean coverage of 31X, 16X, and 42X in AS1, AS2, and AS3, respectively). We identified a candidate pathogenic variant within one AS gene in each patient (Table 3.5).

Table 3.5. Variants identified by WES data analyses in AS families.

			Variants		
Family	**Inheritance**	**Gene**	**cDNA localisation**	**Protein localisation**	**Classification**
AS1	AD/XD	*COL4A5*	NM_000495.4:c.2245-40A>G	/	splicing
AS2	AD/XD	*COL4A5*	NM_000495.4:c.2822G>A	NP_000486.1:p.G941D	missense
AS3	AR	*COL4A3*	NM_000091.4:c.40_63del	NP_000082.2:p.L14_L21del	in-frame deletion

3.8.1. Family AS1: identification of a novel intronic variant in *COL4A5* gene

Proband III3 of family AS1 is a 49-year-old woman with history of hematuria and slight proteinuria; she was diagnosed with AS after kidney biopsy. The mother of the proband (II2, 71 years old) only shows microhematuria, while the proband's daughter (IV1), currently 23 years old, has hematuria, proteinuria and developed sensorineural HL in the first decade of age. We detected a novel heterozygous A-to-G transition

(NM_000495.4:c.2245-40A>G) within *COL4A5* intron 28, 40 nucleotide upstream of exon 29. The variant is absent in the proband's unaffected siblings (III1, III2) and is present in her mother and daughter (Fig. 3.34). The intronic variant is absent in the gnomAD database as well as in our in-house collection of Italian exomes. The particular location of this nucleotide substitution suggested it might affect the branch point sequence, a conserved splicing signal located 20-50 bases upstream of the 3' end of the intron, and, consequently, alter *COL4A5* mRNA splicing.

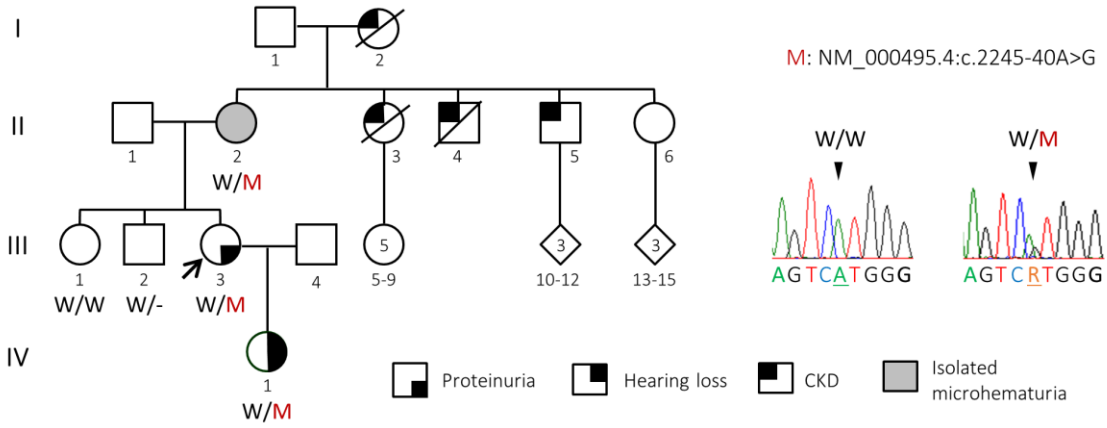


Figure 3.34. Identification of the *COL4A5* c.2245-40A>G variant in family AS1. Pedigree of family AS1 showing the segregation of the identified variant with AS. The proband, analysed by WES, is pointed by an arrow. The genotypes of all available individuals are indicated below the corresponding symbol. On the right, electropherograms show the sequence surrounding the mutated nucleotide in a wild-type relative and in the proband, carrying the variant. W: wild-type allele, M: mutant allele. R: A or G. CKD: chronic kidney disease.

The *in-silico* predictions performed with Human Splicing Finder indicate that the wild-type nucleotide resides in the best-scoring branch-point sequence, although it does not reach the software significance threshold of 67. The A-to-G substitution decreases the score of the putative branch site, suggesting that the variant could impact on exon 29 recognition (Fig. 3.35).

Given the unavailability of a suitable tissue from the patient to extract RNA, we used an *in-vitro* model to verify the effect of c.2545-40A>G variant on splicing. In details, exon 29 of

the *COL4A5* gene with the surrounding intronic sequences, either wild-type or carrying the identified variant, was cloned into the pBS-KS_modified hybrid minigene vector (Fig. 3.36A). The obtained constructs (pBS-KS-COL4A5_ex29_wt and pBS-KS-COL4A5_ex29_mut) were transiently transfected into HEK293 cells and *COL4A5* splicing products were analysed by an RT-PCR assay using primers that map in the exon upstream and downstream of the cloning site (Fig. 3.36A). All amplified products were sequenced to characterise splicing isoforms generated by either the wild-type or the mutant construct. Furthermore, relative quantitation of all splicing products was performed by competitive fluorescent RT-PCRs. Transfection with the pBS-KS-COL4A5_ex29_mut vector originated two aberrant products: the most abundant (~93.9%) and shorter amplicon derived from the skipping of the entire exon 29, whereas the other one (~6.1%) derived from the activation of a cryptic 3' acceptor splice site located downstream of the wild-type one, thus resulting in the retention of a shortened exon 29. In addition, competitive fluorescent RT-PCR of splicing isoforms also helped in confirming the absence of residual amounts of the wild-type transcript in the mutant (Fig. 3.36B, n=3). Both the skipping of the entire exon 29 and the inclusion of a shorter exon 29 are predicted to cause frameshifts leading to the introduction of a premature stop codon after 767 and 776 amino acids, respectively. Exon 29 skipping was also detected in about the 28% of splicing products derived from the pBS-KS-COL4A5_ex29_wt construct, suggesting that low amount of this aberrant transcript can be produced in physiologic conditions, probably due to "leaky" natural splice site recognition (Fig. 3.36B).

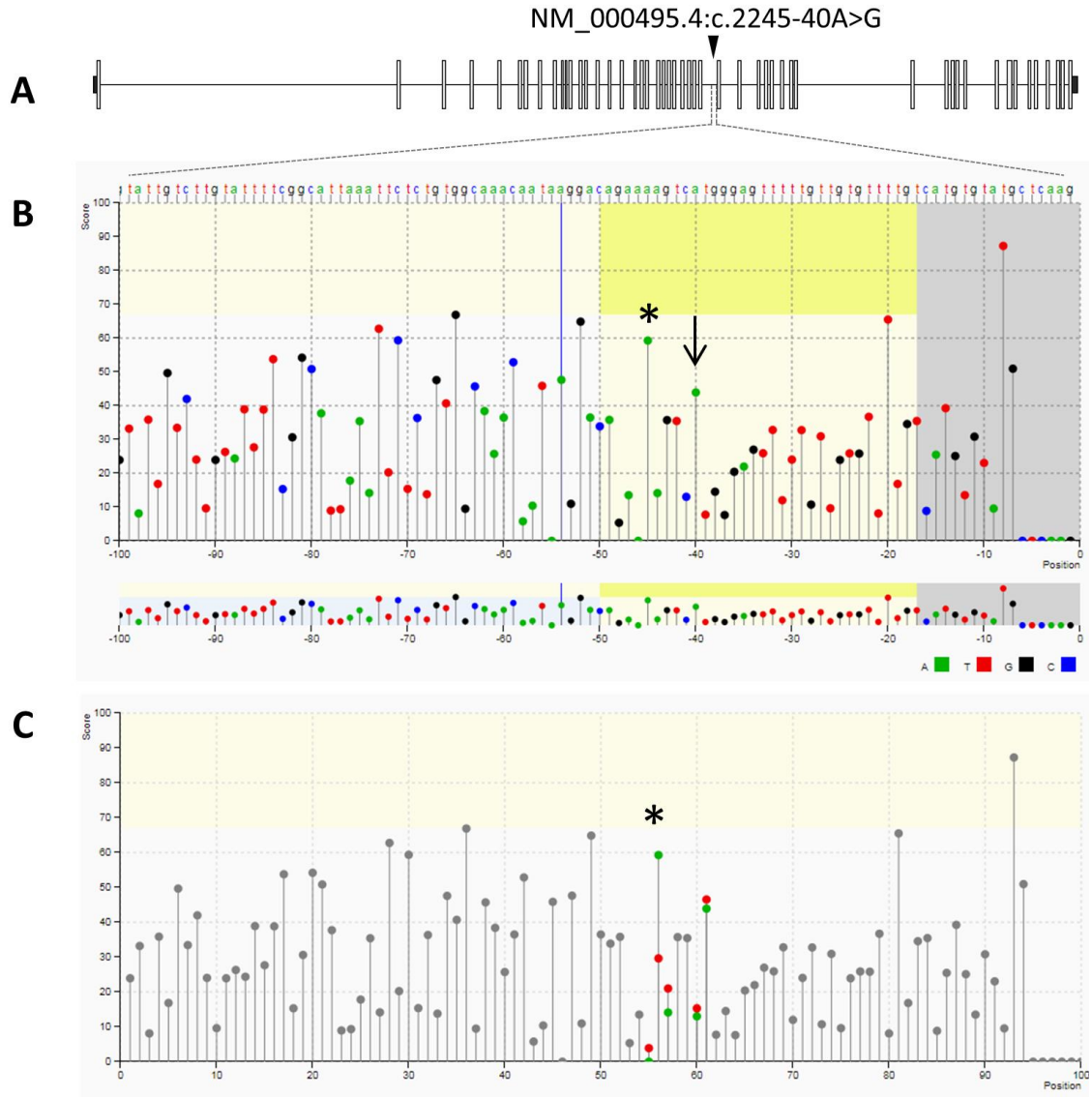


Figure 3.35. Human Splice Finder branch point sequence analysis. A. Schematic representation of *COL4A5* gene, where exons are indicated by rectangles and introns by black lines. The position of the variant in intron 28 is indicated by an arrowhead. **B.** Wild-type *COL4A5* intron 28: the score of branch point sequences (heptamers) are indicated under the first base of the heptamer. The c.2545-40 adenine (indicated by the black arrow) is included in the best-scoring branch point heptamer (whose first base is indicated by the black asterisk). The dark yellow box includes the strongest branch-point site candidates. **C.** The c.2545-40A>G substitution decreases the score of the best-scoring branch-point site (black asterisk) from 59.27 (green dot) to 29.64 (red dot).

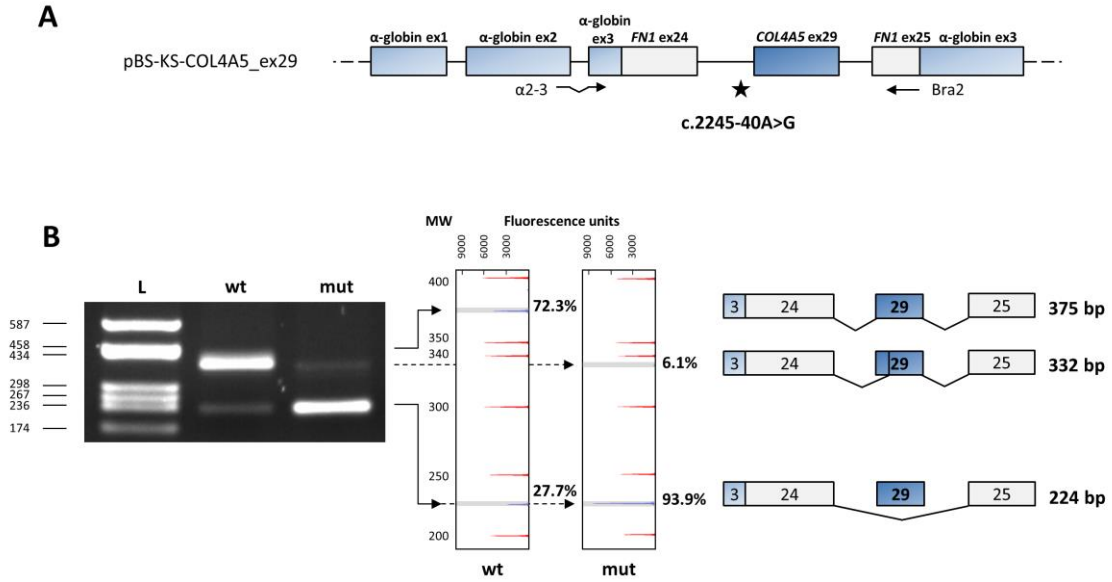


Figure 3.36. *In-vitro* analysis of the impact of c.2245-40A>G variant on *COL4A5* pre-mRNA splicing. A. Schematic representation of the hybrid pBS-KS-COL4A5_ex29 minigene where α-globin exons are represented by light blue boxes, fibronectin (*FN1*) exons by grey boxes, whereas introns are shown as black lines (not to scale). Exon 29 of *COL4A5* is represented by a blue box. The c.2245-40A>G mutation in intron 28 is indicated by a star. Primers used in RT-PCR assays are also indicated. **B.** On the left, agarose gel electrophoresis of RT-PCR products obtained from RNA of HEK293 cells transfected with the wild-type (wt) or mutant (mut) minigene vector. L: DNA ladder pUC9-*HaeIII* molecular weight marker). In the middle, GeneMapper windows show fluorescence peaks corresponding to the molecular species amplified by RT-PCR. Blue peaks correspond to the RT-PCR-labeled products, whose relative quantitation is reported on the right of the panel (%). Red peaks represent the size standard (ROX-500 HD). The *x* axis indicates fluorescence units. On the right, schematic representation of the splicing products, as verified by Sanger sequencing. The length of each fragment is shown.

Since the three heterozygous female carriers (II2, III3, IV1) of the NM_000495.4:c.2245-40A>G variant showed a different severity of phenotypic manifestations (Fig. 3.34), we investigated the possibility that these differences might be due to skewed inactivation of X chromosome, and consequent preferential expression of one of the two alleles (either the wild-type or the mutant one). Hence, we evaluated the methylation status of the two alleles, using a methylation-sensitive restriction-enzyme assay on DNA extracted from patients' blood. Subsequently, we discriminated the wild-type from the mutant allele using a polymorphic marker co-segregating with *COL4A5*, for which the tested females were heterozygous. The results showed a skewed inactivation (91%) of wild-type *COL4A5* allele

in IV1, a balanced inactivation of both alleles in III3 (45 vs 55%), and a greater inactivation (89%) of the mutant *COL4A5* allele in II2 (Fig. 3.37), nicely correlating with the progressively milder phenotypic manifestations.

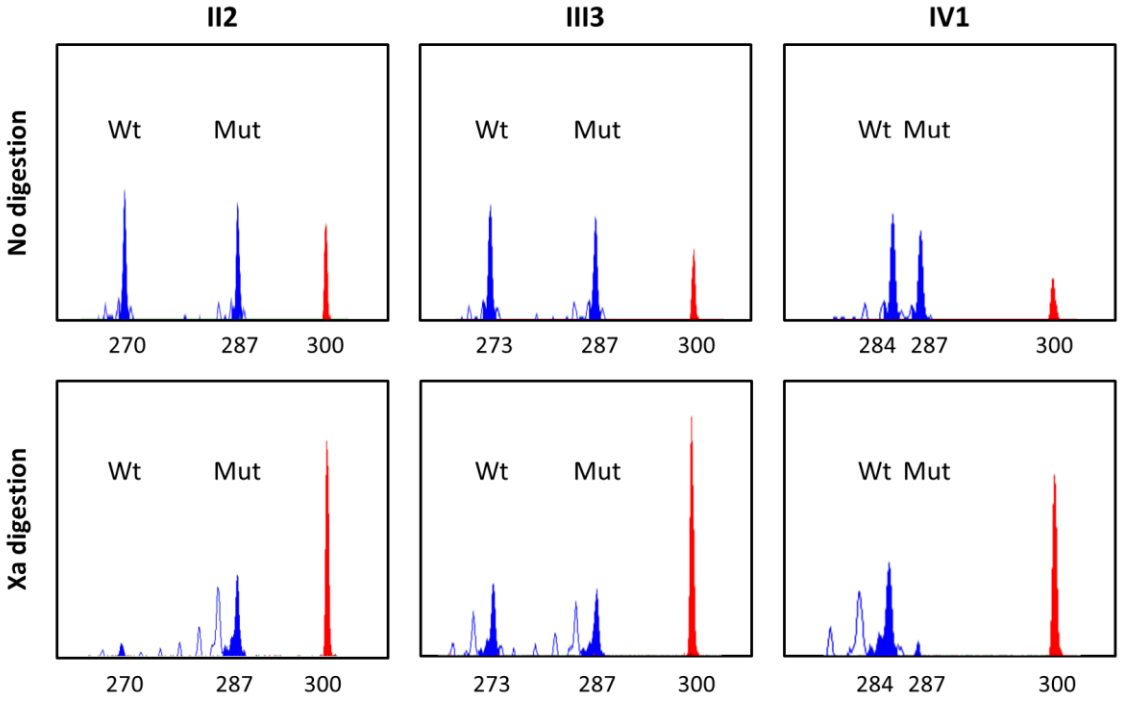


Figure 3.37. X-inactivation analysis of AS1 family. GeneMapper windows with peaks showing the X-inactivation patterns in the female carriers of the *COL4A5* c.2245-40A>G mutation. Methylation assays were performed on the *AR* CAG polymorphic region. No digestion: undigested genomic DNA; Xa digestion: DNA pre-digested with the methylation sensitive enzyme *HpaII*, which only cuts restriction sites on the unmethylated, active X (Xa). The size (bp) of the amplification products for the CAG-repeat region on the two alleles is indicated below the corresponding peak. GeneScan 500 ROX Size Standard is shown in red. Wt: wild-type allele; Mut: mutant allele.

3.8.2. Family AS2: identification of a novel missense variant in *COL4A5* gene

Proband IV3 is a 22-year-old man diagnosed with AS after renal biopsy. At the age of 3 he started to show microhaematuria and episodes of macrohaematuria with proteinuria. He has chronic kidney disease and developed sensorineural HL at the age of 10. The probands mother (III4) showed microhaematuria and proteinuria with normal renal function, while the grandmother (II3) only presented with microhaematuria (Fig. 3.38).

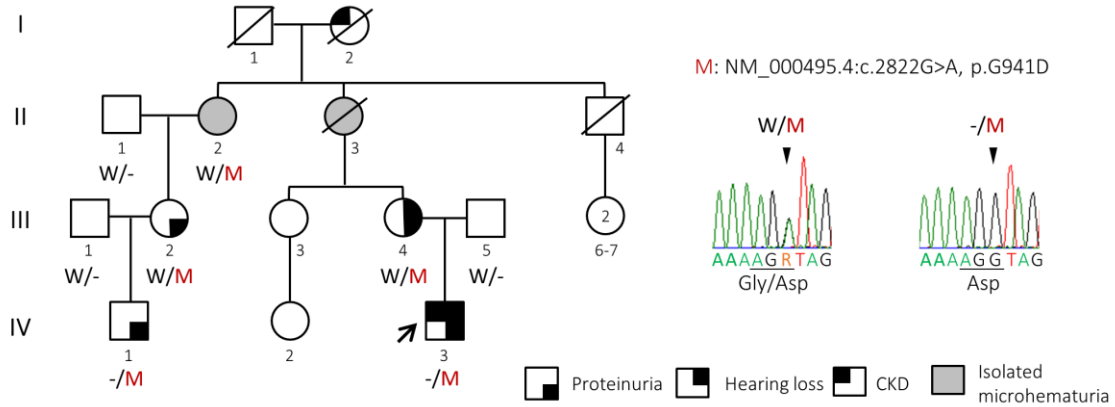
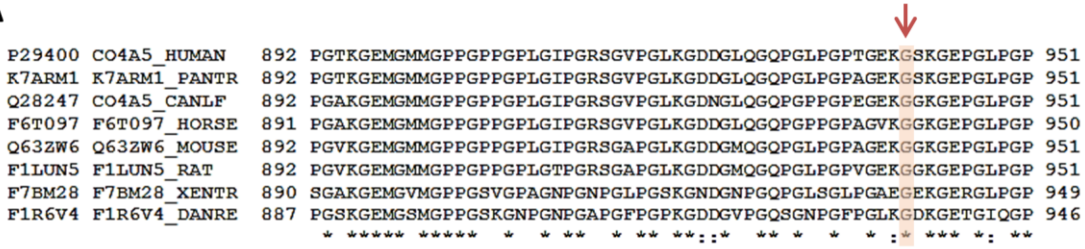


Figure 3.38. Identification of the *COL4A5* c.2822G>A variant in family AS2. Pedigree of family AS2 showing the segregation of the identified variant with AS. The proband is pointed by an arrow and the genotypes of all available individuals are indicated below the corresponding symbols. On the right, electropherograms show the sequence surrounding the mutated nucleotide in a heterozygous female carrier and in the proband, who is hemizygous for the mutant allele. W: wild-type allele, M: mutant allele. R: A or G. CKD: chronic kidney disease.

We identified a novel missense variant in *COL4A5* exon 33 (NM_000495.4:c.2822G>A), causing the p.G941D amino acid substitution. The variant is present in the hemizygous state in the proband and in his affected male cousin (IV1), and in the heterozygous state in all affected female relatives (II2, III2, III4), while it is absent in all available unaffected individuals (II1, III1, III5) (Fig. 3.38). In addition, the variant is absent both in publicly available and our in-house databases, it affects an evolutionary-conserved amino acid and is predicted to be deleterious by all the software used (Fig. 3.39). Indeed, glycine substitutions within the repetitive triplet sequence (Gly)-X-Y of the collagenous domain represent one of the most common type of missense pathogenic variant found in COL4A5, as they are suspected to introduce kinks in the molecule, thus interfering with the proper folding of the collagen triple helix [Kashtan, 2019]. Importantly, a different nucleotide substitution in the same amino acid (c.2821G>T, p.G941D) was previously reported in AS patients [Martin *et al.*, 1998; Hertz *et al.*, 2001]. Of note, exon 33 was not properly covered by the previously performed Roche 454 targeted resequencing, explaining why the variant could have been missed.

Based on the segregation analyses, *in-silico* predictions, and on the previously reported pathogenetic substitution of glycine 941, we propose that the p.G941D mutation is the cause of AS in family AS2.

A



B

	COL4A5 variant
Predictor	p.G941D
SIFT	Damaging
Polyphen 2 HumDiv	Probably Damaging
Polyphen 2 HumVar	Probably Damaging
MutationTaster	Disease causing
Mutation Assessor	FI H
FATHMM	Damaging
Condel	Deleterious
CADD	scaled C-score 28.2

Figure 3.39. *In-silico* analyses of the novel COL4A5 p.G941D missense variant identified in family AS2. A. Amino acid sequence alignments of COL4A5 orthologs in the region surrounding the mutant residue (p.G941). Protein sequences were retrieved from UniProt, and alignments were generated with Clustal Omega. The amino acid residue affected by the mutation is shaded in red and indicated by a red arrow. Identical amino acids are marked by an asterisk, while partially conserved residues are indicated by a colon. **B.** Pathogenicity prediction of the p.G941D missense variant with 8 commonly-used software. FI, Functional Impact of a variant; CADD, Combined Annotation Dependent Depletion. FI is ranked as: high (H), medium (M) for predicted functional variants and low (L) for predicted non-functional variants. CADD Scaled C-score represents the PHRED-like [-10*log10(rank/total)] score, ranking a variant relative to all possible substitutions of the human genome ($8.6x10^9$).

3.8.3. Family AS3: identification of a 24-bp in-frame deletion in *COL4A3* exon 1

Proband III4, a 27-year-old woman with a family history of microhematuria, developed microhematuria, proteinuria and sensorineural HL within the first decade of age and was diagnosed with AS after renal biopsy. A slight maculopathy was evidenced at the age of 21. WES data analysis identified an in-frame 24-bp deletion in *COL4A3* exon 1 (NM_000091.4:c.30_53del:p.V11_L18del or NM_000091.4:c.40_63del24:p.L14_L21del). The deletion is present in the homozygous state in the proband and in the heterozygous state in the mother (II5), who suffers from microhaematuria (Fig. 3.40A).

The variant is reported in the gnomAD database with a minor allele frequency of 0.0001404, being present at the heterozygous state in 17 individuals, most of whom have an Ashkenazi Jewish heritage. Moreover, given the repetitive nature of the surrounding sequence, the variant is ambiguously annotated in dbSNP147 as rs774798108 (lacking clinical annotation) or rs876657397 (OMIM:120070.0011; ClinVar: 192299) (Fig. 3.40B). We deem possible that the first annotation was produced by automated short-read alignment of NGS data, while the second could derive from operator-dependent observation of Sanger sequencing electropherograms, conforming to HGVS (Human Genome Variation Society) recommendation for variant nomenclature and annotation. Nomenclature aside, the deletion in exon 1 eliminates 8 amino acids from the signal peptide and is thus expected to alter COL4A3 protein secretion. Indeed, *in-silico* analyses, using 3 different programs to predict the presence and location of signal peptides (SignalP 4.1: http://www.cbs.dtu.dk/services/SignalP/, PrediSi: http://www.predisi.de/ and Signal-3L 2.0: http://www.csbio.sjtu.edu.cn/bioinf/Signal-3L/), indicate the disruption of the physiologic signal peptide (amino acids 1 to 28) in the mutant protein (Table 3.6). In addition, both PrediSi and WoLF PSORT (https://wolfpsort.hgc.jp/) programs suggest that the mutant protein would not be secreted.

The identified deletion has already been reported in AS patients of different ethnic origins [Longo *et al.*, 2002; Tazón Vega *et al.*, 2003; Webb *et al.*, 2014; Zhang *et al.*, 2012], although it was not functionally characterised.

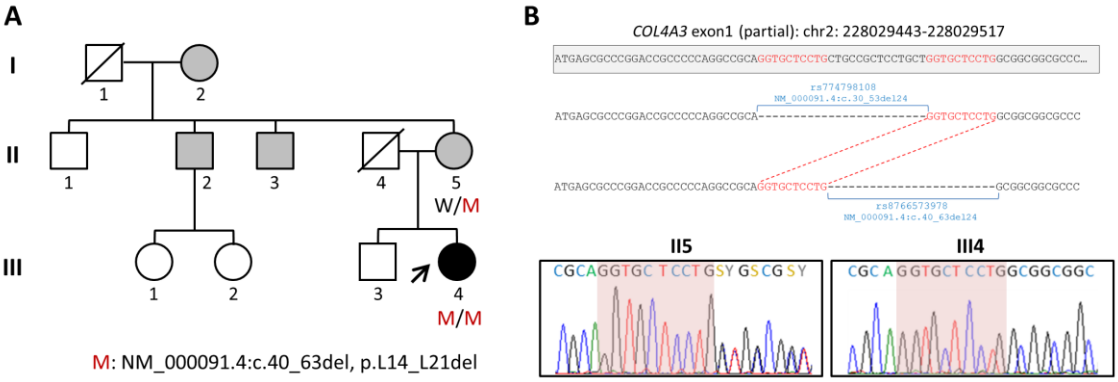


Figure 3.40. Identification of the *COL4A3* c.40_63del variant in family AS3. A. Pedigree of family AS3 showing the segregation of the identified variant with AS. The proband, analysed by WES, is pointed by an arrow. The genotypes of all available individuals are indicated below the corresponding symbols. AS phenotype is indicated by a black symbol, while grey symbols indicate isolated microhaematuria. W: wild-type allele, M: mutant allele. **B.** Given the presence of repetitions (in red), sequence carrying the 24-bp deletion can be aligned to the reference genome in two different ways, resulting in different annotations: 1) rs774798108, corresponding to deletion of bases from 228029472 to 228029495 (GGTGCTCCTGCTGCCGCTCCTGCT), and 2) rs876657397 (OMIM: 120070.0011; ClinVar: 192299), corresponding to the 228029482-228029505 deletion (CTGCCGCTCCTGCTGGTGCTCCTG). Coordinates refer to the hg19 reference genome. Electropherograms of individuals II5 and III4 (the repetitive sequence is highlighted in red) are shown. S: G or C. Y: C or T.

Table 3.6. Signal peptide *in-silico* predictions.

	SignalP 4.1			PrediSi			Signal-3L	
COL4A3	**Score**	**Cleavage position**	**Signal Peptide**	**Score**	**Cleavage position**	**Signal Peptide**	**Cleavage position**	**Signal Peptide**
Wild-type	0.79	28	Yes	1.00	28	Yes	28	Yes
p.L14_L21del	0.21	n.c.	No	0.45	22	No	n.c.	No

n.c.: not calculated

To demonstrate that the NM_000091.4:c.40_63del variant alters the physiologic COL4A3 signal peptide, potentially affecting collagen secretion, we performed localisation assays in HEK293 cells using an hybrid fluorescent reporter protein (hybEGFP, Enhanced Green Fluorescent Protein) containing at the N-terminus either the complete COL4A3 signal peptide (amino acid 1 to 29: SP-wt-hybEGFP) or the partially deleted localisation signal (p.L14_L21del: SP-del-hybEGFP). Our results show that, while the wild-type signal peptide seems to drive hybEGFP expression along the secretory pathway, the mutant peptide - lacking 8 amino acids - retains hybEGFP within the cell (Fig. 3.41). Indeed, hybEGFP fused with a deleted COL4A3 signal peptide show a diffuse and uniform localisation in the cytoplasm and in the nucleus, similarly to the soluble wild-type EGFP (pEGFP-N1; Fig. 3.41). Co-localisation studies using either a marker of the endoplasmic reticulum (the calnexin protein), or a marker of the Golgi apparatus (Trans-Golgi Network protein 38, TGN38) further support the hypothesis that the SP-del-hybEGFP is not included in the secretory pathway (Fig. 3.42).

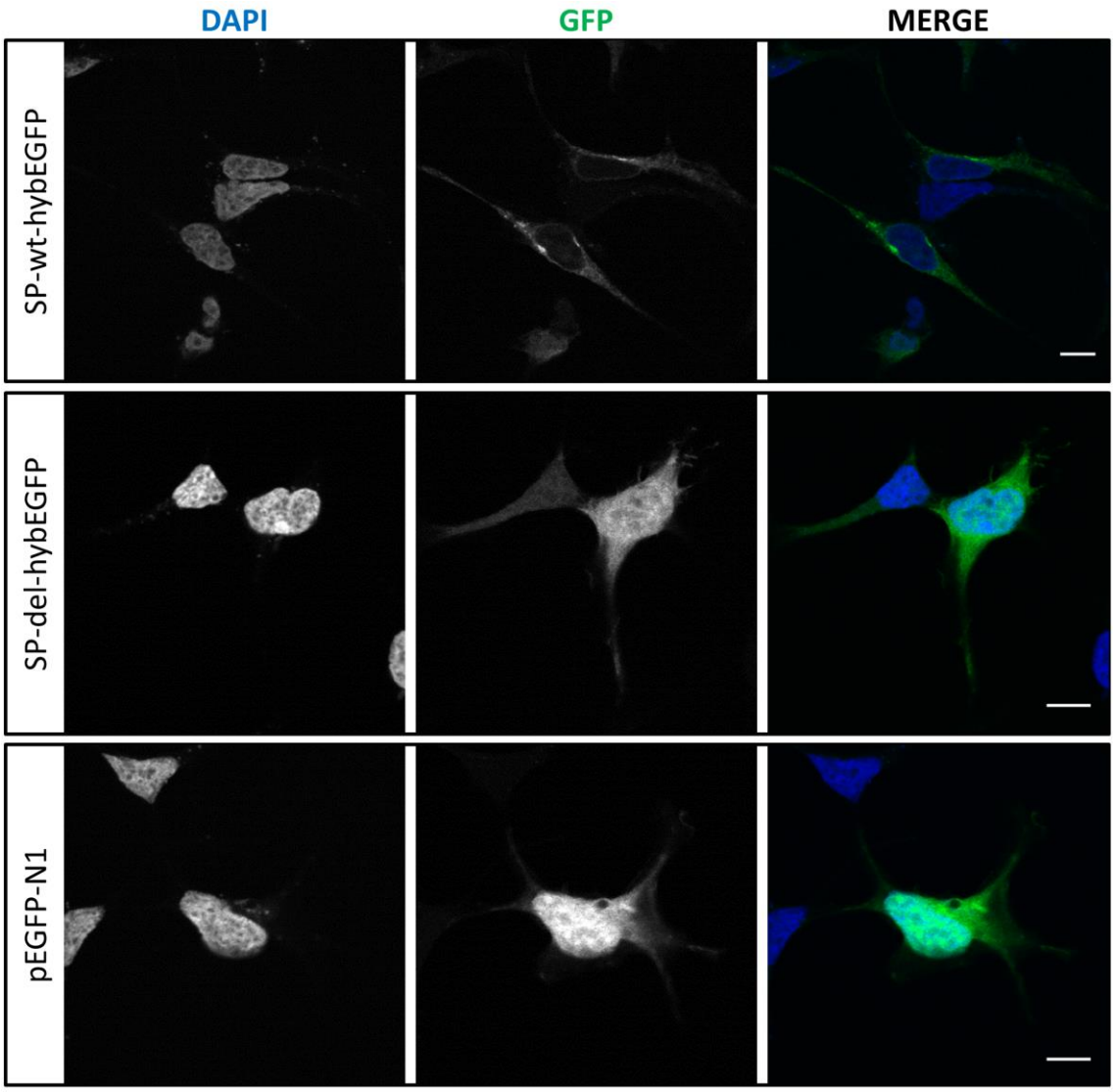


Figure 3.41. Functional characterisation of the signal peptide deletion in COL4A3. Single confocal sections of HEK293 cells expressing EGFP N-terminus fused either with the entire COL4A3 signal peptide (SP-wt-hybEGFP, top panels) or the 8-amino-acid deleted signal peptide (SP-del-hybEGFP, middle panels). Positive control cells, expressing a soluble EGFP (pEGFP-N1) are also shown (bottom panels). Cells were fixed 24 hours after transfection. Images were acquired with Olympus FV1000 Inverted Confocal IX81 Microscope and are representative of 3 independent experiments. DAPI: 4',6-diamidino-2-phenylindole; EGFP, Enhanced Green Fluorescent Protein. Scale bars: 10 μm.

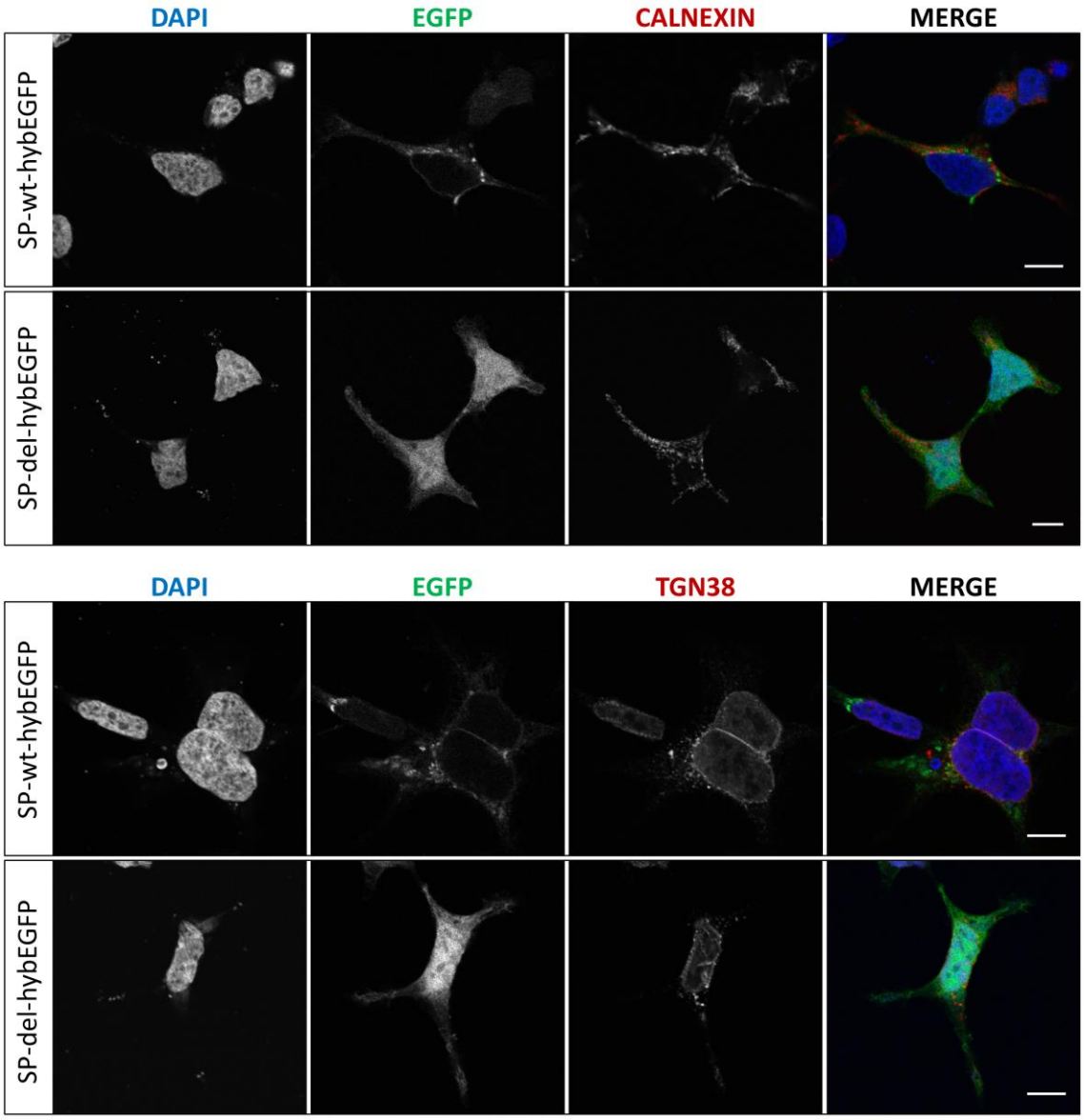


Figure 3.42. Co-localisation studies of hybEGFP with markers of the secretory pathway. Single confocal sections of HEK293 cells expressing COL4A3-SP-wt-hybEGFP or COL4A3-SP-del-hybEGFP and stained for calnexin (endoplasmic reticulum transmembrane protein, top panel) or TGN38 (trans-Golgi network protein, bottom panels). SP-del-hybEGFP does not show any co-localisation with either secretory pathway markers. Cells were fixed 24 hours after transfection. Images were acquired with Olympus FV1000 Inverted Confocal IX81 Microscope and are representative of 2 independent experiments. Scale bars: 10 μm.

3.8.4. Conclusion on AS families

Our approach, combining WES and functional studies, was fundamental to identify the pathogenic variants causing AS in 3 families with no genetic diagnosis, thus allowing us to solve elusive cases.

In family AS1, the *COL4A5* branch-point variant was probably missed in previous screenings because of its location. Indeed, only consensus splice sites and surrounding nucleotides are usually analysed, whereas branch-point site – which are more difficult to predict on the basis of nucleotide sequence alone [Houdayer *et al.*, 2012] – are not usually detected. AS1 is the first AS family in which a branch-point site mutation has been identified.

In family AS3, we identified the in-frame deletion c.40_63del, which escaped previous genetic screenings. *COL4A3* exon 1 has been reported to be difficult to sequence by NGS due to its high GC content (~77%) [Morinière *et al.*, 2014]. Indeed, it was not amplified in the previous NGS targeted re-sequencing on the proband and it was also the least covered exon among the three type IV collagen genes in our WES data. Despite being ambiguously annotated in databases, the c.40_63del variant has been repeatedly reported in AS patients. It was first described in the heterozygous state in an Italian patient with microhaematuria and mild proteinuria [Longo *et al.*, 2002]. Subsequently, it was detected in one Spanish patient [Tazón Vega *et al.*, 2003] and four Chinese patients [Zhang *et al.*, 2012] with autosomal recessive AS: in four of them, the mutation was present in compound heterozygosity with other pathogenic variants [Tazón Vega *et al.*, 2003; Zhang *et al.*, 2012], whereas in the fifth patient it was present in the homozygous state [Zhang *et al.*, 2012]. Finally, the same in-frame deletion was found as a founder mutation in Ashkenazi Jews, with an estimated carrier frequency of 1:183 [Webb *et al.*, 2014]. Indeed, in the gnomAD database the minor allele frequency of the variant in the Ashkenazi population is 0.001282

(carrier frequency ~1:390), being thus ~21 times the MAF of the variant considering all the other populations in gnomAD together.

In the heterozygous state, the c.40_63del is not associated with significant renal dysfunctions, while in the homozygous state or in compound heterozygosity with another loss-of-function variant, it causes severe AS with HL and, as in the Spanish patient [Tazón Vega *et al.*, 2003] and in the proband of family AS3, ocular abnormalities. Due to the recurrence of the variant in different populations, it would be important to specifically screen *COL4A3* exon 1 in case of diagnosis of autosomal recessive HL and/or family history of isolated microhaematuria.

The results of this work have been recently published in the Plos One journal [Chiereghin *et al.*, 2017].

4. Concluding remarks

The research reported in this thesis has provided new insights into the mutational landscape of hearing loss, showing the advantages of the use of whole-exome sequencing for the study of the molecular bases of this genetically heterogeneous disease.

In this work, WES has proven a powerful tool for the identification of possible pathogenic variants also in small families – such as NSHL22 - unsuitable for linkage analysis. However, this work has also highlighted the limitations of an approach uniquely based on exome sequencing for the definition of the genetic cause of HL. In fact, while the attribution of a genetic diagnosis is a simple task in case of known pathogenic variants in deafness genes, when a novel variation is identified its co-segregation with the phenotype within a family is not always sufficient to validate the variant as pathogenic, especially when few individuals are available for screening. Therefore, it is important to integrate WES analyses with functional studies designed to examine the specific effect of the candidate mutations identified. An approach aimed at discriminating false leads from genetic variants with clear functional effects supporting their role in deafness is critical, as many predicted pathogenic variants in HL genes are also identified in the general population [Lewis e*t al.*, 2018].

An even more cautious and critical approach is needed in case of identification of likely pathogenic variants in novel candidate deafness genes. In fact, even genes whose association to deafness was supported by the identification of likely pathogenic variants in several unrelated individuals are now being called into question [Vona *et al.*, 2015]. In particular, *MYO1A* gene, which was considered for over 10 years the gene underlying autosomal dominant DFNA48 [Donaudy *et al.*, 2003], has been disqualified as a NSHL gene. Indeed, putative pathogenic variants (including nonsense mutations) were identified both in patients and in unaffected relatives [Eisenberger *et al.*, 2014]. In the case of *DIAPH2*, the c.868A>G variant described here has not been reported in the hemizygous state in healthy individuals; however, different variants have been identified in individuals affected by diverse diseases (Paragraph 3.6.3). Nonetheless, a few deafness genes have also

been associated with quite dissimilar diseases, such as the case of *TBC1D24* gene, associated with DFNA65 and DFNB86, but also with early-infantile epileptic encephalopathy 16 and other complex diseases with epileptic features [Mucha *et al.*, 2017]. Therefore, further studies – primarily in the available *Diaph2* mouse models – are needed to confirm or disqualify *DIAPH2* variants as a cause of deafness and, eventually, of the other diseases in which hemizygous variants were identified.

Although the introduction of NGS technologies has considerably improved the genetic diagnosis of deafness, previously relying mostly on the screening of the single *GJB2* gene, the diagnostic yield of WES is still below 50% and varies based on several factors, including inheritance mode of selected patients and prescreening prior to comprehensive testing. For example, the diagnostic rate was reported to be lower in autosomal recessive or sporadic cases and in case of prescreened patients [Shearer and Smith, 2015]. Of note, families were not selected on the basis of the inheritance pathway (both recessive and dominant families without potential nongenetic causes of HL were included), and all recruited probands had been previously screened for mutations in the *GJB2* and *GJB6* genes. In this study we detected the genetic cause of the disease in at least 7 out of 11 families analysed (>60%). Although, the combination of a diligent exome-sequencing data analysis, together with an accurate functional validation, allowed us to increase the expected diagnostic outcome, some improvements might be implemented. For example, it would be very important to have an unrelated normal-hearing control cohort for screening purposes. In fact, despite the availability of 125 Italian audiologically-tested normal-hearing controls, many of these individuals were relatives of deaf patients (not included in this study). Also, the access to sequence data from large cohorts of hearing-impaired patients might help in confirming the association of prioritised variants with the disease. Finally, other approaches could be applied to further enhance our results. For instance, RNA sequencing might be a suitable approach to comprehensively detect splicing mutations in HL genes that are also

expressed in accessible tissues (e.g. blood), although the splicing pattern in the tissue analysed might not accurately represent that of the inner ear. In addition, using linkage analysis in large families to narrow the region in which to search for pathogenic variants, and including specific molecular techniques to evaluate genomic imbalances might be useful to decipher the genetic cause of HL in the undiagnosed families. Certainly, it would not be feasible - especially in terms of time - to translate the approach applied in this thesis to a high-throughput clinical service. However, functional studies aimed at discriminating between neutral polymorphisms and HL-causing variants might be collated in a publicly-available expert-curated database (e.g. the Deafness Variation Database, http://deafnessvariationdatabase.org/) to help variant interpretation. The construction of a comprehensive database would undoubtedly help clinical decision making and enhance the diagnostic yield also in a clinical setting.

5. References

Ahmed ZM, Morell RJ, Riazuddin S, Gropman A, Shaukat S, Ahmad MM, Mohiddin SA, Fananapazir L, Caruso RC, Husnain T, et al. Mutations of MYO6 are associated with recessive deafness, DFNB37. Am J Hum Genet. 2003 May;72(5):1315-22.

Ahmed ZM, Riazuddin S, Ahmad J, Bernstein SL, Guo Y, Sabar MF, Sieving P, Riazuddin S, Griffith AJ, Friedman TB, *et al.* PCDH15 is expressed in the neurosensory epithelium of the eye and ear and mutant alleles are responsible for both USH1F and DFNB23. Hum Mol Genet. 2003 Dec 15;12(24):3215-23.

Arellano B, Pera A, Ramírez-Camacho R, Villamar M, Trinidad A, García JR, Moreno F, Hernández-Chico C. Pendred's syndrome and non-syndromic DFNB4 deafness associated with the homozygous T410M mutation in the SLC26A4 gene in siblings. Clin Genet. 2005 May;67(5):438-40.

Avraham KB, Hasson T, Steel KP, Kingsley DM, Russell LB, Mooseker MS, Copeland NG, Jenkins NA. The mouse Snell's waltzer deafness gene encodes an unconventional myosin required for structural integrity of inner ear hair cells. Nat Genet. 1995 Dec;11(4):369-75.

Bademci G, Diaz-Horta O, Guo S, Duman D, Van Booven D, Foster J 2nd, Cengiz FB, Blanton S, Tekin M. Identification of copy number variants through whole-exome sequencing in autosomal recessive nonsyndromic hearing loss. Genet Test Mol Biomarkers. 2014 Sep;18(9):658-61.

Bae SH, Robertson NG, Cho HJ, Morton CC, Jung da J, Baek JI, Choi SY, Lee J, Lee KY, Kim UK. Identification of pathogenic mechanisms of COCH mutations, abolished cochlin secretion, and intracellular aggregate formation: genotype-phenotype correlations in DFNA9 deafness and vestibular disorder. Hum Mutat. 2014 Dec;35(12):1506-13.

Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011 Sep 27;12(11):745-55.

Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, Ffrench-Constant C, Baralle FE. Identification of a mutation that perturbs NF1 gene splicing using genomic DNA samples and a minigene assay. J Med Genet. 2003 Mar;40(3):220-2.

Ben Said M, Grati M, Ishimoto T, Zou B, Chakchouk I, Ma Q, Yao Q, Hammami B, Yan D, Mittal R, *et al.* A mutation in SLC22A4 encoding an organic cation transporter expressed in the cochlea strial endothelium causes human recessive non-syndromic hearing loss DFNB60. Hum Genet. 2016 May;135(5):513-24.

Bianchi LM, Fuchs PA (2010). Cochlear supporting cells. In: Fuchs PA, editor. The Oxford Handbook of Auditory Science: The Ear. Oxford: Oxford UP; 2010. pp. 329-353.

Bione S, Sala C, Manzini C, Arrigo G, Zuffardi O, Banfi S, Borsani G, Jonveaux P, Philippe C, Zuccotti M, *et al.* A human homologue of the Drosophila melanogaster diaphanous gene is disrupted in a patient with premature ovarian failure: evidence for

conserved function in oogenesis and implications for human sterility. Am J Hum Genet. 1998 Mar;62(3):533-41.

Bitner-Glindzicz M. Hereditary deafness and phenotyping in humans. Br Med Bull. 2002;63:73-94.

Bizhanova A, Kopp P. Minireview: The sodium-iodide symporter NIS and pendrin in iodide homeostasis of the thyroid. Endocrinology. 2009 Mar;150(3):1084-90.

Bork JM, Peters LM, Riazuddin S, Bernstein SL, Ahmed ZM, Ness SL, Polomeno R, Ramesh A, Schloss M, Srisailpathy CR, *et al.* Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. Am J Hum Genet. 2001 Jan;68(1):26-37.

Braunstein EM, Crenshaw EB 3rd, Morrow BE, Adams JC. Cooperative function of Tbx1 and Brn4 in the periotic mesenchyme is necessary for cochlea formation. J Assoc Res Otolaryngol. 2008 Mar;9(1):33-43.

Brownstein Z, Bhonker Y, Avraham KB. High-throughput sequencing to decipher the genetic heterogeneity of deafness. Genome Biol. 2012 May 29;13(5):245.

Burckhardt G, Wolff NA. Structure of renal organic anion and cation transporters. Am J Physiol Renal Physiol. 2000 Jun;278(6):F853-66.

Butler BE, Lomber SG. Functional and structural changes throughout the auditory system following congenital and early-onset deafness: implications for hearing restoration. Front Syst Neurosci. 2013 Nov 26;7:92.

Bykhovskaya Y, Estivill X, Taylor K, Hang T, Hamon M, Casano RA, Yang H, Rotter JI, Shohat M, Fischel-Ghodsian N. Candidate locus for a nuclear modifier gene for maternally inherited deafness. Am J Hum Genet. 2000 Jun;66(6):1905-10.

Campellone KG, Welch MD. A nucleator arms race: cellular control of actin assembly. Nat Rev Mol Cell Biol. 2010 Apr;11(4):237-51.

Cantos R, Cole LK, Acampora D, Simeone A, Wu DK. Patterning of the mammalian cochlea. Proc Natl Acad Sci U S A. 2000 Oct 24;97(22):11707-13.

Castrillon DH, Wasserman SA. Diaphanous is required for cytokinesis in Drosophila and shares domains of similarity with the products of the limb deformity gene. Development. 1994 Dec;120(12):3367-77.

Chakchouk I, Ben Said M, Jbeli F, Benmarzoug R, Loukil S, Smeti I, Chakroun A, Gibriel AA, Ghorbel A, Hadjkacem H, Masmoudi S. NADf chip, a two-color microarray for simultaneous screening of multigene mutations associated with hearing impairment in North African Mediterranean countries. J Mol Diagn. 2015 Mar;17(2):155-61.

Chatzispyrou IA, Alders M, Guerrero-Castillo S, Zapata Perez R, Haagmans MA, Mouchiroud L, Koster J, Ofman R, Baas F, Waterham HR, *et al.* A homozygous missense

mutation in ERAL1, encoding a mitochondrial rRNA chaperone, causes Perrault syndrome. Hum Mol Genet. 2017 Jul 1;26(13):2541-2550.

Chen W, Kahrizi K, Meyer NC, Riazalhosseini Y, Van Camp G, Najmabadi H, Smith RJ. Mutation of COL11A2 causes autosomal recessive non-syndromic hearing loss at the DFNB53 locus. J Med Genet. 2005 Oct;42(10):e61.

Chiereghin C, Robusto M, Mastrangelo A, Castorina P, Montini G, Giani M, Duga S, Asselta R, Soldà G. Alport syndrome cold cases: Missing mutations identified by exome sequencing and functional analysis. PLoS One. 2017 Jun 1;12(6):e0178630.

Choi BY, Kim HM, Ito T, Lee KY, Li X, Monahan K, Wen Y, Wilson E, Kurima K, Saunders TL, Petralia RS, Wangemann P, Friedman TB, Griffith AJ. Mouse model of enlarged vestibular aqueducts defines temporal requirement of Slc26a4 expression for hearing acquisition. J Clin Invest. 2011 Nov;121(11):4516-25.

Craven KB, Zagotta WN. CNG and HCN channels: two peas, one pod. Annu Rev Physiol. 2006;68:375-401.

D'Aurizio R, Pippucci T, Tattini L, Giusti B, Pellegrini M, Magi A. Enhanced copy number variants detection from whole-exome sequencing data using EXCAVATOR2. Nucleic Acids Res. 2016 Nov 16;44(20):e154.

de Kok YJ, van der Maarel SM, Bitner-Glindzicz M, Huber I, Monaco AP, Malcolm S, Pembrey ME, Ropers HH, Cremers FP. Association between X-linked mixed deafness and mutations in the POU domain gene POU3F4. Science. 1995 Feb 3;267(5198):685-8.

Del Castillo I, Moreno-Pelayo MA, Del Castillo FJ, Brownstein Z, Marlin S, Adina Q, Cockburn DJ, Pandya A, Siemering KR, Chamberlin GP, *et al.* Prevalence and evolutionary origins of the del(GJB6-D13S1830) mutation in the DFNB1 locusin hearing-impaired subjects: a multicenter study. Am J Hum Genet. 2003 Dec;73(6):1452-8.

Donaudy F, Ferrara A, Esposito L, Hertzano R, Ben-David O, Bell RE, Melchionda S, Zelante L, Avraham KB, Gasparini P. Multiple mutations of MYO1A, a cochlear-expressed gene, in sensorineural hearing loss. Am J Hum Genet. 2003 Jun;72(6):1571-7.

Dossena S, Nofziger C, Brownstein Z, Kanaan M, Avraham KB, Paulmichl M. Functional characterization of pendrin mutations found in the Israeli and Palestinian populations. Cell Physiol Biochem. 2011;28(3):477-84.

Dror AA, Avraham KB. Hearing impairment: a panoply of genes and functions. Neuron. 2010 Oct 21;68(2):293-308.

Dror AA, Brownstein Z, Avraham KB. Integration of human and mouse genetics reveals pendrin function in hearing and deafness. Cell Physiol Biochem. 2011;28(3):535-44.

Duman D, Tekin M. Autosomal recessive nonsyndromic deafness genes: a review. Front Biosci (Landmark Ed). 2012 Jun 1;17:2213-36.

Eisenberger T, Di Donato N, Baig SM, Neuhaus C, Beyer A, Decker E, Mürbe D, Decker C, Bergmann C, Bolz HJ. Targeted and genomewide NGS data disqualify mutations in MYO1A, the "DFNA48 gene", as a cause of deafness. Hum Mutat. 2014 May;35(5):565-70.

Ercan-Sencicek AG, Jambi S, Franjic D, Nishimura S, Li M, El-Fishawy P, Morgan TM, Sanders SJ, Bilguvar K, Suri M, *et al.*. Homozygous loss of DIAPH1 is a novel cause of microcephaly in humans. Eur J Hum Genet. 2015 Feb;23(2):165-72.

Everett LA, Belyantseva IA, Noben-Trauth K, Cantos R, Chen A, Thakkar SI, Hoogstraten-Miller SL, Kachar B, Wu DK, Green ED. Targeted disruption of mouse Pds provides insight about the inner-ear defects encountered in Pendred syndrome. Hum Mol Genet. 2001 Jan 15;10(2):153-61.

Fontana P, Morgutti M, Pecile V, Lenarduzzi S, Cappellani S, Falco M, Scarano F, Lonardo F. A novel OTOA mutation in an Italian family with hearing loss. Gene Reports 2017 Dec;9:111-114.

Frigeni M, Iacobazzi F, Yin X, Longo N. Wide tolerance to amino acids substitutions in the OCTN1 ergothioneine transporter. Biochim Biophys Acta. 2016 Jun;1860(6):1334-42.

Fuchs PA (2010). Introduction and overview. In: Fuchs PA, editor. The Oxford Handbook of Auditory Science: The Ear. Oxford: Oxford UP; 2010. pp. 1-13.

Fuentes Fajardo KV, Adams D, NISC Comparative Sequencing Program, Mason CE, Sincan M, Tifft C, Toro C, Boerkoel CF, Gahl W, Markello T. Detecting false-positive signals in exome sequencing. Hum Mutat. 2012 Apr;33(4):609-13.

Fukuda T, Kominami K, Wang S, Togashi H, Hirata K, Mizoguchi A, Rikitake Y, Takai Y. Aberrant cochlear hair cell attachments caused by Nectin-3 deficiency result in hair bundle abnormalities. Development. 2014 Jan;141(2):399-409.

Gale J, Jagger D (2010). Cochlear supporting cells. In: Fuchs PA, editor. The Oxford Handbook of Auditory Science: The Ear. Oxford: Oxford UP; 2010. pp. 307–327.

Gasman S, Kalaidzidis Y, Zerial M. RhoD regulates endosome dynamics through Diaphanous-related Formin and Src tyrosine kinase. Nat Cell Biol. 2003 Mar;5(3):195-204.

Goode BL, Eck MJ. Mechanism and function of formins in the control of actin assembly. Annu Rev Biochem. 2007;76:593-627.

Grati M, Yan D, Raval MH, Walsh T, Ma Q, Chakchouk I, Kannan-Sundhari A, Mittal R, Masmoudi S, Blanton SH, *et al.* MYO3A Causes Human Dominant Deafness and Interacts with Protocadherin 15-CD2 Isoform. Hum Mutat. 2016 May;37(5):481-7.

Gründemann D, Harlfinger S, Golz S, Geerts A, Lazar A, Berkels R, Jung N, Rubbert A, Schömig E. Discovery of the ergothioneine transporter. Proc Natl Acad Sci U S A. 2005 Apr 5;102(14):5256-61.

Gu X, Su W, Tang M, Guo L, Zhao L, Li H. Massively Parallel Sequencing of a Chinese Family with DFNA9 Identified a Novel Missense Mutation in the LCCL Domain of COCH. Neural Plast. 2016;2016:5310192.

Hardisty-Hughes RE, Parker A, Brown SD. A hearing and vestibular phenotyping pipeline to identify mouse mutants with hearingimpairment. Nat Protoc. 2010 Jan;5(1):177-90.

Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016 Jan;107(1):1-8.

Hertz JM, Juncker I, Persson U, Matthijs G, Schmidtke J, Petersen MB, Kjeldsen M, Gregersen N. Detection of mutations in the COL4A5 gene by SSCP in X-linked Alport syndrome. Hum Mutat. 2001 Aug;18(2):141-8.

Hertzano R, Montcouquiol M, Rashi-Elkeles S, Elkon R, Yücel R, Frankel WN, Rechavi G, Möröy T, Friedman TB, Kelley MW, Avraham KB. Transcription profiling of inner ears from Pou4f3(ddl/ddl) identifies Gfi1 as a target of the Pou4f3deafness gene. Hum Mol Genet. 2004 Sep 15;13(18):2143-53.

Hildebrand MS, DeLuca AP, Taylor KR, Hoskinson DP, Hur IA, Tack D, McMordie SJ, Huygen PL, Casavant TL, Smith RJ. AudioGene audioprofiling: a machine-based candidate gene prediction tool for autosomal dominant non-syndromic hearing loss. Laryngoscope. 2009 Nov;119(11):2211-5.

Hildebrand MS, Gandolfo L, Shearer AE, Webster JA, Jensen M, Kimberling WJ, Stephan D, Huygen PL, Smith RJ, Bahlo M. A novel mutation in COCH-implications for genotype-phenotype correlations in DFNA9 hearing loss. Laryngoscope. 2010 Dec;120(12):2489-93.

Hilgert N, Smith RJ, Van Camp G. Function and expression pattern of nonsyndromic deafness genes. Curr Mol Med. 2009 Jun;9(5):546-64.

Holley MC, Lawlor PW. Production of conditionally immortalised cell lines from a transgenic mouse. Audiol Neurootol. 1997;2:25-35.

Hosoya M, Fujioka M, Ogawa K, Okano H. Distinct Expression Patterns Of Causative Genes Responsible For Hereditary Progressive Hearing Loss In Non-Human Primate Cochlea. Sci Rep. 2016 Feb 26;6:22250.

Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 524 and BRCA2 variants. Hum Mutat. 2012;33(8):1228-38.

Hudson BG, Tryggvason K, Sundaramoorthy M, Neilson EG. Alport's syndrome, Goodpasture's syndrome, and type IV collagen. N Engl J Med. 2003 Jun 19;348(25):2543-56.

Ingham NJ, Pearson S, Steel KP. Using the Auditory Brainstem Response (ABR) to Determine Sensitivity of Hearing in Mutant Mice. Curr Protoc Mouse Biol. 2011 Jun 1;1(2):279-87.

Ison JR, Allen PD, Oertel D. Deleting the HCN1 Subunit of Hyperpolarization-Activated Ion Channels in Mice Impairs Acoustic Startle Reflexes, Gap Detection, and Spatial Localization. J Assoc Res Otolaryngol. 2017 Jun;18(3):427-440.

Jagger DJ, Nevill G, Forge A. The Membrane Properties of Cochlear Root Cells are Consistent with Roles in Potassium Recirculation and Spatial Buffering. J Assoc Res Otolaryngol. 2010 Sep;11(3):435-48.

JanssensdeVarebeke SPF, Van Camp G, Peeters N, Elinck E, Widdershoven J, Cox T, Deben K, Ketelslagers K, Crins T, Wuyts W. Bi-allelic inactivating variants in the COCH gene cause autosomal recessive prelingual hearing impairment. Eur J Hum Genet. 2018 Apr;26(4):587-591.

Jat PS, Noble MD, Ataliotis P, Tanaka Y, Yannoutsos N, Larsen L, Kioussis D. Direct derivation of conditionally immortal cell lines from an H-2Kb-tsA58 transgenic mouse. Proc Natl Acad Sci USA. 1991;88:5096-100.

Jung J, Kim HS, Lee MG, Yang EJ, Choi JY. Novel COCH p.V123E Mutation, Causative of DFNA9 Sensorineural Hearing Loss and Vestibular Disorder, Shows Impaired Cochlin Post-Translational Cleavage and Secretion. Hum Mutat. 2015 Dec;36(12):1168-75.

Kahrizi K, Mohseni M, Nishimura C, Bazazzadegan N, Fischer SM, Dehghani A, Sayfati M, Taghdiri M, Jamali P, Smith RJ, *et al.* Identification of SLC26A4 gene mutations in Iranian families with hereditary hearing impairment. Eur J Pediatr. 2009 Jun;168(6):651-3.

Kalay E, Li Y, Uzumcu A, Uyguner O, Collin RW, Caylan R, Ulubil-Emiroglu M, Kersten FF, Hafiz G, van Wijk E, *et al.* Mutations in the lipoma HMGIC fusion partner-like 5 (LHFPL5) gene cause autosomal recessive nonsyndromic hearing loss. Hum Mutat. 2006 Jul;27(7):633-9.

Kashtan CE. Alport Syndrome and Thin Basement Membrane Nephropathy. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Amemiya A, Bean LJH, Bird TD, Fong CT, Mefford HC, Smith RJH, Stephens K, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2019. 2001 Aug 28 [updated February 21, 2019].

Kato Y, Kubo Y, Iwata D, Kato S, Sudo T, Sugiura T, Kagaya T, Wakayama T, Hirayama A, Sugimoto M, *et al.* Gene knockout and metabolome analysis of carnitine/organic cation transporter OCTN1. Pharm Res. 2010 May;27(5):832-40.

Kaufman MH. The Atlas of Mouse Development. Elsevier Academic Press, 1999.

Kelley MW. Cellular commitment and differentiation in the organ of Corti. Int J Dev Biol. 2007;51(6-7):571-83.

Khan SY, Ahmed ZM, Shabbir MI, Kitajiri S, Kalsoom S, Tasneem S, Shayiq S, Ramesh A, Srisailpathy S, Khan SN, *et al.* Mutations of the RDX gene cause nonsyndromic hearing loss at the DFNB24 locus. Hum Mutat. 2007 May;28(5):417-23.

Kikkawa Y, Seki Y, Okumura K, Ohshiba Y, Miyasaka Y, Suzuki S, Ozaki M, Matsuoka K, Noguchi Y, Yonekawa H. Advantages of a mouse model for human hearing impairment. Exp Anim. 2012;61(2):85-98.

Kim YH, Holt JR. Functional contributions of HCN channels in the primary auditory neurons of the mouse inner ear. J Gen Physiol. 2013 Sep;142(3):207-23.

Koffler T, Ushakov K, Avraham KB. Genetics of Hearing Loss: Syndromic. Otolaryngol Clin North Am. 2015 Dec;48(6):1041-61.

Korver AM, Smith RJ, Van Camp G, Schleiss MR, Bitner-Glindzicz MA, Lustig LR, Usami SI, Boudewyns AN. Congenital hearing loss. Nat Rev Dis Primers. 2017 Jan 12;3:16094.

Kurima K, Peters LM, Yang Y, Riazuddin S, Ahmed ZM, Naz S, Arnaud D, Drury S, Mo J, Makishima T, *et al.*. Dominant and recessive deafness caused by mutations of a novel gene, TMC1, required for cochlear hair-cell function. Nat Genet. 2002 Mar;30(3):277-84.

Lalwani AK, Goldstein JA, Kelley MJ, Luxford W, Castelein CM, Mhatre AN. Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in nonmuscle myosin MYH9. Am J Hum Genet. 2000 Nov;67(5):1121-8.

Lamhonwah AM, Tein I. Novel localization of OCTN1, an organic cation/carnitine transporter, to mammalian mitochondria. Biochem Biophys Res Commun. 2006 Jul 14;345(4):1315-25.

Lee CH, MacKinnon R. Structures of the Human HCN1 Hyperpolarization-Activated Channel. Cell. 2017 Jan 12;168(1-2):111-120.e11.

Lemmink HH, Schröder CH, Monnens LA, Smeets HJ. The clinical spectrum of type IV collagen mutations. Hum Mutat. 1997;9:477–99.

Levy SE, Myers RM. Advancements in Next-Generation Sequencing. Annu Rev Genomics Hum Genet. 2016 Aug 31;17:95-115.

Lewis MA, Nolan LS, Cadge BA, Matthews LJ, Schulte BA, Dubno JR, Steel KP, Dawson SJ. Whole exome sequencing in adult-onset hearing loss reveals a high load of predicted pathogenic variants in known deafness-associated genes and identifies new candidate genes. BMC Med Genomics. 2018 Sep 4;11(1):77.

Lewis MA, Quint E, Glazier AM, Fuchs H, De Angelis MH, Langford C, van Dongen S, Abreu-Goodger C, Piipari M, Redshaw N, *et al.* An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. Nat Genet. 2009 May;41(5):614-8.

Li H, Jin Z, Li X, Wu L, Jin J. Associations between single-nucleotide polymorphisms and inflammatory bowel disease-associated colorectal cancers in inflammatory bowel disease patients: a meta-analysis. Clin Transl Oncol. 2017 Aug;19(8):1018-1027.

Likar T, Hasanhodžić M, Teran N, Maver A, Peterlin B, Writzl K. Diagnostic outcomes of exome sequencing in patients with syndromic or non-syndromic hearing loss. PLoS One. 2018 Jan 2;13(1):e0188578.

Lin J, Ozeki M, Javel E, Zhao Z, Pan W, Schlentz E, Levine S. Identification of gene expression profiles in rat ears with cDNA microarrays. Hear Res. 2003 Jan;175(1-2):2-13.

Lin X, Tang W, Ahmad S, Lu J, Colby CC, Zhu J, Yu Q. Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. Hear Res. 2012 Jun;288(1-2):67-76.

Lindner TH, Hoffmann K. easyLINKAGE: a PERL script for easy and automated two-/multi-point linkage analyses. Bioinformatics. 2005 Feb 1;21(3):405-7.

Liu XZ, Walsh J, Mburu P, Kendrick-Jones J, Cope MJ, Steel KP, Brown SD. Mutations in the myosin VIIA gene cause non-syndromic recessive deafness. Nat Genet. 1997 Jun;16(2):188-90.

Liu XZ, Walsh J, Tamagawa Y, Kitamura K, Nishizawa M, Steel KP, Brown SD. Autosomal dominant non-syndromic deafness caused by a mutation in the myosin VIIA gene. Nat Genet. 1997 Nov;17(3):268-9.

Liu XZ, Yuan Y, Yan D, Ding EH, Ouyang XM, Fei Y, Tang W, Yuan H, Chang Q, Du LL, *et al.* Digenic inheritance of non-syndromic deafness caused by mutations at the gap junction proteinsCx26 and Cx31. Hum Genet. 2009 Feb;125(1):53-62.

Lolicato M, Nardini M, Gazzarrini S, Möller S, Bertinetti D, Herberg FW, Bolognesi M, Martin H, Fasolini M, Bertrand JA, *et al.* Tetramerization dynamics of C-terminal domain underlies isoform-specific cAMP gating in hyperpolarization-activated cyclic nucleotide-gated channels. J Biol Chem. 2011 Dec 30;286(52):44811-20.

Longo I, Porcedda P, Mari F, Giachino D, Meloni I, Deplano C, Brusco A, Bosio M, Massella L, Lavoratti G, *et al.* COL4A3/COL4A4 mutations: from familial hematuria to autosomal-dominant or recessive Alport syndrome. Kidney Int. 2002 Jun;61(6):1947-56.

López-Bigas N, Melchionda S, de Cid R, Grifa A, Zelante L, Govea N, Arbonés ML, Gasparini P, Estivill X. Identification of five new mutations of PDS/SLC26A4 in Mediterranean families with hearing impairment. Hum Mutat. 2001 Dec;18(6):548.

Lukashkin AN, Legan PK, Weddell TD, Lukashkina VA, Goodyear RJ, Welstead LJ, Petit C, Russell IJ, Richardson GP. A mouse model for human deafness DFNB22 reveals that hearing impairment is due to a loss of inner hair cell stimulation. Proc Natl Acad Sci U S A. 2012 Nov 20;109(47):19351-6.

Lynch ED, Lee MK, Morrow JE, Welcsh PL, León PE, King MC. Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the Drosophila gene diaphanous. Science. 1997 Nov 14;278(5341):1315-8.

Mandelker D, Amr SS, Pugh T, Gowrisankar S, Shakhbatyan R, Duffy E, Bowser M, Harrison B, Lafferty K, Mahanta L, *et al.* Comprehensive diagnostic testing for stereocilin: an approach for analyzing medically important genes with high homology. J Mol Diagn. 2014 Nov;16(6):639-47.

Martin P, Heiskari N, Zhou J, Leinonen A, Tumelius T, Hertz JM, Barker D, Gregory M, Atkin C, Styrkarsdottir U *et al.* High mutation detection rate in the COL4A5 collagen gene in suspected Alport syndrome using PCR and direct DNA sequencing. J Am Soc Nephrol. 1998; 9(12):2291-301.

Martini M, Ferrara AM, Giachelia M, Panieri E, Siminovitch K, Galeotti T, Larocca LM, Pani G. Association of the OCTN1/1672T variant with increased risk for colorectal cancer in young individuals and ulcerative colitis patients. Inflamm Bowel Dis. 2012 Mar;18(3):439-48.

McGrath J, Roy P, Perrin BJ. Stereocilia morphogenesis and maintenance through regulation of actin stability. Semin Cell Dev Biol. 2017 May;65:88-95.

Melchionda S, Ahituv N, Bisceglia L, Sobe T, Glaser F, Rabionet R, Arbones ML, Notarangelo A, Di Iorio E, Carella M, *et al.* MYO6, the human homologue of the gene responsible for deafness in Snell's waltzer mice, is mutated in autosomal dominant nonsyndromic hearing loss. Am J Hum Genet. 2001 Sep;69(3):635-40.

Mencía A, Modamio-Høybjør S, Redshaw N, Morín M, Mayo-Merino F, Olavarrieta L, Aguirre LA, del Castillo I, Steel KP, Dalmay T, *et al.* Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. Nat Genet. 2009 May;41(5):609-13.

Mikaelian D, Ruben RJ. Development of Hearing in the Normal Cba-J Mouse: Correlation of Physiological Observations Anatomy. Acta Oto-Laryngologica, 1965;59:2-6, 451-461.

Morinière V, Dahan K, Hilbert P, Lison M, Lebbah S, Topa A, Bole-Feysot C, Pruvost S, Nitschke P, Plaisier E *et al.* Improving mutation screening in familial hematuric nephropathies through next generation sequencing. J Am Soc Nephrol 2014;25(12):2740-51.

Morton CC, Giersch AB. Genetics of hearing loss. In: Fuchs PA, editor. The Oxford Handbook of Auditory Science: The Ear. Oxford: Oxford UP; 2010. pp. 377–407.

Moteki H, Azaiez H, Sloan-Heggen CM, Booth K, Nishio SY, Wakui K, Yamaguchi T, Kolbe DL, Iwasa YI, Shearer AE, *et al.* Detection and Confirmation of Deafness-Causing Copy Number Variations in the STRC Gene by Massively Parallel Sequencing and Comparative Genomic Hybridization. Ann Otol Rhinol Laryngol. 2016 Nov;125(11):918-923.

Mouse Genome Sequencing Consortium, Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, *et al.* Initial sequencing and comparative analysis of the mouse genome. Nature. 2002 Dec 5;420(6915):520-62.

Mucha BE, Hennekam RCM, Sisodiya S, Campeau PM. TBC1D24-Related Disorders. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Amemiya A, Bean LJH, Bird TD, Fong CT, Mefford HC, Smith RJH, Stephens K, editors. GeneReviews® [Internet]. Seattle

(WA): University of Washington, Seattle; 1993-2018. 2015 Feb 26 [updated December 7, 2017].

Mustapha M, Weil D, Chardenoux S, Elias S, El-Zir E, Beckmann JS, Loiselet J, Petit C. An alpha-tectorin gene defect causes a newly identified autosomal recessive form of sensorineural pre-lingual non-syndromic deafness, DFNB21. Hum Mol Genet. 1999 Mar;8(3):409-12.

Nance WE, The genetics of deafness. Ment Retard Dev Disabil Res Rev. 2003;9(2):109-19.

Narayanan P, Chatterton P, Ikeda A, Ikeda S, Corey DP, Ervasti JM, Perrin BJ. Length regulation of mechanosensitive stereocilia depends on very slow actin dynamics and filament-severing proteins. Nat Commun. 2015 Apr 21;6:6855.

Nava C, Dalle C, Rastetter A, Striano P, de Kovel CG, Nabbout R, Cancès C, Ville D, Brilstra EH, Gobbi G, *et al.* De novo mutations in HCN1 cause early infantile epileptic encephalopathy. Nat Genet. 2014 Jun;46(6):640-5.

Naz S, Griffith AJ, Riazuddin S, Hampton LL, Battey JF Jr, Khan SN, Riazuddin S, Wilcox ER, Friedman TB. Mutations of ESPN cause autosomal recessive deafness and vestibular dysfunction. J Med Genet. 2004 Aug;41(8):591-5.

Naz S, Giguere CM, Kohrman DC, Mitchem KL, Riazuddin S, Morell RJ, Ramesh A, Srisailpathy S, Deshmukh D, Riazuddin S, *et al.* Mutations in a novel gene, TMIE, are associated with hearing loss linked to the DFNB6 locus. Am J Hum Genet. 2002 Sep;71(3):632-6.

Naz S, Griffith AJ, Riazuddin S, Hampton LL, Battey JF Jr, Khan SN, Riazuddin S, Wilcox ER, Friedman TB. Mutations of ESPN cause autosomal recessive deafness and vestibular dysfunction. J Med Genet. 2004 Aug;41(8):591-5.

Neveling K, Feenstra I, Gilissen C, Hoefsloot LH, Kamsteeg EJ, Mensenkamp AR, Rodenburg RJ, Yntema HG, Spruijt L, Vermeer S, *et al.* A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. Hum Mutat. 2013 Dec;34(12):1721-6.

Newman WG, Friedman TB, Conway GS, Demain LAM. Perrault Syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A, editors. GeneReviews®[Internet]. Seattle (WA): University of Washington, Seattle; 1993-2018. 2014 Sep 25 [updated September 6, 2018].

Nolan MF, Malleret G, Lee KH, Gibbs E, Dudman JT, Santoro B, Yin D, Thompson RF, Siegelbaum SA, Kandel ER, Morozov A. The hyperpolarization-activated HCN1 channel is important for motor learning and neuronal integration by cerebellar Purkinje cells. Cell. 2003 Nov 26;115(5):551-64.

Oishi A, Makita N, Sato J, Iiri T. Regulation of RhoA signaling by the cAMP-dependent phosphorylation of RhoGDIα. J Biol Chem. 2012 Nov 9;287(46):38705-15.

O'Rahilly R. The timing and sequence of events in the development of the human eye and ear during the embryonic period proper. Anat Embryol (Berl). 1983;168(1):87-99.

Otomo T, Otomo C, Tomchick DR, Machius M, Rosen MK. Structural basis of Rho GTPase-mediated activation of the formin mDia1. Mol Cell. 2005 Apr 29;18(3):273-81.

Pan B, Akyuz N, Liu XP, Asai Y, Nist-Lund C, Kurima K, Derfler BH, György B, Limapichat W, Walujkar S, *et al.* TMC1 Forms the Pore of Mechanosensory Transduction Channels in Vertebrate Inner Ear Hair Cells. Neuron. 2018 Aug 22;99(4):736-753.e6.

Parker M, Bitner-Glindzicz M. Genetic investigations in childhood deafness. Arch Dis Child. 2015 Mar;100(3):271-8.

Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X, Newman B, Van Oene M, Cescon D, Greenberg G, *et al.* Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet. 2004 May;36(5):471-5.

Petit C. Genes responsible for human hereditary deafness: symphony of a thousand. Nat Genet. 1996 Dec;14(4):385-91.

Pippucci T, Benelli M, Magi A, Martelli PL, Magini P, Torricelli F, Casadio R, Seri M, Romeo G. EX-HOM (EXome HOMozygosity): a proof of principle. Hum Hered. 2011;72(1):45-53.

Pirooznia M, Goes FS, Zandi PP. Whole-genome CNV analysis: advances in computational approaches. Front Genet. 2015 Apr 13;6:138.

Quittner AL, Leibach P, Marciel K. The impact of cochlear implants on young deaf children: new methods to assess cognitive and behavioral development. Arch Otolaryngol Head Neck Surg. 2004 May;130(5):547-54.

Ramakrishnan NA, Drescher MJ, Khan KM, Hatfield JS, Drescher DG. HCN1 and HCN2 proteins are expressed in cochlear hair cells: HCN1 can form a ternary complex with protocadherin 15 CD3 and F-actin-binding filamin A or can interact with HCN2. J Biol Chem. 2012 Nov 2;287(45):37628-46.

Raviv D, Dror AA, Avraham KB. Hearing loss: a common disorder caused by many rare alleles. Ann N Y Acad Sci. 2010 Dec;1214:168-79.

Rehman AU, Morell RJ, Belyantseva IA, Khan SY, Boger ET, Shahzad M, Ahmed ZM, Riazuddin S, Khan SN, Riazuddin S, Friedman TB. Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. Am J Hum Genet. 2010 Mar 12;86(3):378-88.

Riazuddin S, Ahmed ZM, Fanning AS, Lagziel A, Kitajiri S, Ramzan K, Khan SN, Chattaraj P, Friedman PL, Anderson JM, *et al.* Tricellulin is a tight-junction protein necessary for hearing. Am J Hum Genet. 2006 Dec;79(6):1040-51.

Riazuddin S, Castelein CM, Ahmed ZM, Lalwani AK, Mastroianni MA, Naz S, Smith TN, Liburd NA, Friedman TB, Griffith AJ, *et al.* Dominant modifier DFNM1 suppresses recessive deafness DFNB26. Nat Genet. 2000 Dec;26(4):431-4.

Richardson GP, de Monvel JB, Petit C. How the genetics of deafness illuminates auditory physiology. Annu Rev Physiol. 2011;73:311-34.

Rivolta MN, Holley MC. Cell lines in inner ear research. J Neurobiol. 2002 Nov 5;53(2):306-18.

Robertson NG, Lu L, Heller S, Merchant SN, Eavey RD, McKenna M, Nadol JB Jr, Miyamoto RT, Linthicum FH Jr, Lubianca Neto JF, *et al.* Mutations in a novel cochlear gene cause DFNA9, a human nonsyndromic deafness with vestibular dysfunction. Nat Genet. 1998 Nov;20(3):299-303.

Robertson NG, Resendes BL, Lin JS, Lee C, Aster JC, Adams JC, Morton CC. Inner ear localization of mRNA and protein products of COCH, mutated in the sensorineural deafness and vestibular disorder, DFNA9. Hum Mol Genet. 2001 Oct 15;10(22):2493-500.

Rose R, Weyand M, Lammers M, Ishizaki T, Ahmadian MR, Wittinghofer A. Structural and mechanistic insights into the interaction between Rho and mammalian Dia. Nature. 2005 May 26;435(7041):513-8.

Royaux IE, Belyantseva IA, Wu T, Kachar B, Everett LA, Marcus DC, Green ED. Localization and functional studies of pendrin in the mouse inner ear provide insight about the etiology of deafness in pendred syndrome. J Assoc Res Otolaryngol. 2003 Sep;4(3):394-404.

Rudman JR, Kabahuma RI, Bressler SE, Feng Y, Blanton SH, Yan D, Liu XZ. The genetic basis of deafness in populations of African descent. J Genet Genomics. 2017 Jun 20;44(6):285-294.

Santos-Cortez RL, Lee K, Giese AP, Ansar M, Amin-Ud-Din M, Rehn K, Wang X, Aziz A, Chiu I, Hussain Ali R, Smith JD *et al.* Adenylate cyclase 1 (ADCY1) mutations cause recessive hearing impairment in humans and defects in hair cell function and hearing in zebrafish. Hum Mol Genet. 2014 Jun 15;23(12):3289-98.

Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet. 2011 Aug 31;12(10):683-91.

Scheffer DI, Shen J, Corey DP, Chen ZY. Gene Expression by Mouse Inner Ear Hair Cells during Development. J Neurosci. 2015 Apr 22;35(16):6366-80.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, *et al.* Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012 Jun 28;9(7):676-82.

Schoen CJ, Emery SB, Thorne MC, Ammana HR, Sliwerska E, Arnett J, Hortsch M, Hannan F, Burmeister M, Lesperance MM. Increased activity of Diaphanous homolog 3

(DIAPH3)/diaphanous causes hearing defects in humans with auditory neuropathy and in Drosophila. Proc Natl Acad Sci U S A. 2010 Jul 27;107(30):13396-401.

Schraders M, Ruiz-Palmero L, Kalay E, Oostrik J, del Castillo FJ, Sezgin O, Beynon AJ, Strom TM, Pennings RJ, Zazo Seco C, *et al.* Mutations of the gene encoding otogelin are a cause of autosomal-recessive nonsyndromic moderate hearing impairment. Am J Hum Genet. 2012 Nov 2;91(5):883-9.

Schrijver I. Hereditary non-syndromic sensorineural hearing loss: transforming silence to sound. J Mol Diagn. 2004 Nov;6(4):275-84.

Scott EM, Halees A, Itan Y, Spencer EG, He Y, Azab MA, Gabriel SB, Belkadi A, Boisson B, Abel L, *et al.* Characterization of Greater Middle Eastern genetic variation for enhanced disease gene discovery. Nat Genet. 2016 Sep;48(9):1071-6.

Shahin H, Walsh T, Rayyan AA, Lee MK, Higgins J, Dickel D, Lewis K, Thompson J, Baker C, Nord AS, *et al.* Five novel loci for inherited hearing loss mapped by SNP-based homozygosity profiles in Palestinian families. Eur J Hum Genet. 2010 Apr;18(4):407-13.

Shearer AE, Hildebrand MS, Smith RJH. Hereditary Hearing Loss and Deafness Overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, Amemiya A, editors. GeneReviews® [Internet]. Seattle; 1993-2019. 1999 Feb 14 [updated July 27, 2017].

Shearer AE, Kolbe DL, Azaiez H, Sloan CM, Frees KL, Weaver AE, Clark ET, Nishimura CJ, Black-Ziegelbein EA, Smith RJ. Copy number variants are a common cause of non-syndromic hearing loss. Genome Med. 2014 May 22;6(5):37.

Shearer AE, Smith RJ. Massively Parallel Sequencing for Genetic Diagnosis of Hearing Loss: The New Standard of Care. Otolaryngol Head Neck Surg. 2015 Aug;153(2):175-82.

Shen J, Scheffer DI, Kwan KY, Corey DP. SHIELD: an integrative gene expression database for inner ear research. Database (Oxford). 2015 Jul 24;2015:bav071.

Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol. 2008 Oct;26(10):1135-45.

Sheppard S, Biswas S, Li MH, Jayaraman V, Slack I, Romasko EJ, Sasson A, Brunton J, Rajagopalan R, Sarmady M, *et al.* Utility and limitations of exome sequencing as a genetic diagnostic tool for children with hearing loss. Genet Med. 2018 Jun 15.

Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, Mujica AO, Thomas M, Harrow J, Cox T, *et al.* A conditional knockout resource for the genome-wide study of mouse gene function. Nature. 2011 Jun 15;474(7351):337-42.

Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015 Oct;36(10):928-30.

Son EJ, Wu L, Yoon H, Kim S, Choi JY, Bok J. Developmental gene expression profiling along the tonotopic axis of the mouse cochlea. PLoS One. 2012;7(7):e40735.

Stokman MF, Renkema KY, Giles RH, Schaefer F, Knoers NV, van Eerde AM. The expanding phenotypic spectra of kidney diseases: insights from genetic studies. Nat Rev Nephrol. 2016;12:472-83.

Stritt S, Nurden P, Turro E, Greene D, Jansen SB, Westbury SK, Petersen R, Astle WJ, Marlin S, Bariana TK, *et al.* A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. Blood. 2016 Jun 9;127(23):2903-14.

Taylor JP, Metcalfe RA, Watson PF, Weetman AP, Trembath RC. Mutations of the PDS gene, encoding pendrin, are associated with protein mislocalization and loss of iodide efflux: implications for thyroid dysfunction in Pendred syndrome. J Clin Endocrinol Metab. 2002 Apr;87(4):1778-84.

Tazón Vega B, Badenas C, Ars E, Lens X, Milà M, Darnell A, Torra R. Autosomal recessive Alport's syndrome and benign familial hematuria are collagen type IV diseases. Am J Kidney Dis. 2003 Nov;42(5):952-9.

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, *et al.* An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat Genet. 2003 Dec;35(4):341-8.

Torres M, Giráldez F. The development of the vertebrate inner ear. Mech Dev. 1998 Feb;71(1-2):5-21.

Toyoda Y, Shinohara R, Thumkeo D, Kamijo H, Nishimaru H, Hioki H, Kaneko T, Ishizaki T, Furuyashiki T, Narumiya S. EphA4-dependent axon retraction and midline localization of Ephrin-B3 are disrupted in the spinal cord of mice lacking mDia1 and mDia3 in combination. Genes Cells. 2013 Oct;18(10):873-85.

Tsukada K, Ichinose A, Miyagawa M, Mori K, Hattori M, Nishio SY, Naito Y, Kitajiri S, Usami S. Detailed hearing and vestibular profiles in the patients with COCH mutations. Ann Otol Rhinol Laryngol. 2015 May;124 Suppl 1:100S-10S.

Tu NC, Friedman RA. Age-related hearing loss: Unraveling the pieces. Laryngoscope Investig Otolaryngol. 2018 Feb 21;3(2):68-72.

Vahava O, Morell R, Lynch ED, Weiss S, Kagan ME, Ahituv N, Morrow JE, Lee MK, Skvorak AB, Morton CC, *et al.* Mutation in transcription factor POU4F3 associated with inherited progressive hearing loss in humans. Science. 1998 Mar 20;279(5358):1950-4.

van Wijk E, Krieger E, Kemperman MH, De Leenheer EM, Huygen PL, Cremers CW, Cremers FP, Kremer H. A mutation in the gamma actin 1 (ACTG1) gene causes autosomal dominant hearing loss (DFNA20/26). J Med Genet. 2003 Dec;40(12):879-84.

Veal CD, Freeman PJ, Jacobs K, Lancaster O, Jamain S, Leboyer M, Albanes D, Vaghela RR, Gut I, Chanock SJ, Brookes AJ. A mechanistic basis for amplification differences between samples and between genome regions. BMC Genomics. 2012 Sep 5;13:455.

Vona B, Nanda I, Hofrichter MA, Shehata-Dieler W, Haaf T. Non-syndromic hearing loss gene identification: A brief history and glimpse into the future. Mol Cell Probes. 2015 Oct;29(5):260-70.

Wainger BJ, DeGennaro M, Santoro B, Siegelbaum SA, Tibbs GR. Molecular mechanism of cAMP modulation of HCN pacemaker channels. Nature. 2001 Jun 14;411(6839):805-10.

Walsh T, Abu Rayan A, Abu Sa'ed J, Shahin H, Shepshelovich J, Lee MK, Hirschberg K, Tekin M, Salhab W, Avraham KB, *et al.* Genomic analysis of a heterogeneous Mendelian phenotype: multiple novel alleles for inherited hearing loss in the Palestinian population. Hum Genomics. 2006 Jan;2(4):203-11.

Walsh T, Walsh V, Vreugde S, Hertzano R, Shahin H, Haika S, Lee MK, Kanaan M, King MC, Avraham KB. From flies' eyes to our ears: mutations in a human class III myosin cause progressive nonsyndromic hearing loss DFNB30. Proc Natl Acad Sci U S A. 2002 May 28;99(11):7518-23.

Walsh T, Shahin H, Elkan-Miller T, Lee MK, Thornton AM, Roeb W, Abu Rayyan A, Loulus S, Avraham KB, King MC, Kanaan M. Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPSM2 as the cause of nonsyndromic hearing loss DFNB82. Am J Hum Genet. 2010 Jul 9;87(1):90-4.

Wang A, Liang Y, Fridell RA, Probst FJ, Wilcox ER, Touchman JW, Morton CC, Morell RJ, Noben-Trauth K, Camper SA, Friedman TB. Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. Science. 1998 May 29;280(5368):1447-51.

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 2013;153:910-918.

Wang QJ, Lu CY, Li N, Rao SQ, Shi YB, Han DY, Li X, Cao JY, Yu LM, Li QZ, *et al.* Y-linked inheritance of non-syndromic hearing impairment in a large Chinese family. J Med Genet. 2004 Jun;41(6):e80.

Wangemann P. K+ cycling and the endocochlear potential.Hear Res. 2002 Mar;165(1-2):1-9.

Wangemann P, Nakaya K, Wu T, Maganti RJ, Itza EM, Sanneman JD, Harbidge DG, Billings S, Marcus DC. Loss of cochlear HCO3- secretion causes deafness via endolymphatic acidification and inhibition of Ca2+ reabsorption in a Pendred syndrome mouse model. Am J Physiol Renal Physiol. 2007 May;292(5):F1345-53.

Wangemann P. Mouse models for pendrin-associated loss of cochlear and vestibular function. Cell Physiol Biochem. 2013;32(7):157-65.

Webb BD, Brandt T, Liu L, Jalas C, Liao J, Fedick A, Linderman MD, Diaz GA, Kornreich R, Trachtman H, *et al.* A founder mutation in COL4A3 causes autosomal recessive Alport syndrome in the Ashkenazi Jewish population. Clin Genet. 2014 Aug;86(2):155-60.

Weil D, Blanchard S, Kaplan J, Guilford P, Gibson F, Walsh J, Mburu P, Varela A, Levilliers J, Weston MD, *et al.* Defective myosin VIIA gene responsible for Usher syndrome type 1B. Nature. 1995 Mar 2;374(6517):60-1.

White JK, Gerdin AK, Karp NA, Ryder E, Buljan M, Bussell JN, Salisbury J, Clare S, Ingham NJ, Podrini C, *et al.* Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell. 2013 Jul 18;154(2):452-64.

Wilcox ER, Burton QL, Naz S, Riazuddin S, Smith TN, Ploplis B, Belyantseva I, Ben-Yosef T, Liburd NA, Morell RJ, *et al.* Mutations in the gene encoding tight junction claudin-14 cause autosomal recessive deafness DFNB29. Cell. 2001 Jan 12;104(1):165-72.

Yabuuchi H, Tamai I, Nezu J, Sakamoto K, Oku A, Shimane M, Sai Y, Tsuji A. Novel membrane transporter OCTN1 mediates multispecific, bidirectional, and pH-dependent transport of organic cations. J Pharmacol Exp Ther. 1999 May;289(2):768-73.

Yang T, Gurrola JG 2nd, Wu H, Chiu SM, Wangemann P, Snyder PM, Smith RJ. Mutations of KCNJ10 together with mutations of SLC26A4 cause digenic nonsyndromic hearing loss associated with enlarged vestibular aqueduct syndrome. Am J Hum Genet. 2009 May;84(5):651-7.

Yang T, Vidarsson H, Rodrigo-Blomqvist S, Rosengren SS, Enerback S, Smith RJ. Transcriptional control of SLC26A4 is involved in Pendred syndrome and nonsyndromic enlargement of vestibular aqueduct (DFNB4). Am J Hum Genet. 2007 Jun;80(6):1055-63.

Yariz KO, Duman D, Zazo Seco C, Dallman J, Huang M, Peters TA, Sirmaci A, Lu N, Schraders M, Skromne I, *et al.* Mutations in OTOGL, encoding the inner ear protein otogelin-like, cause moderate sensorineural hearing loss. Am J Hum Genet. 2012 Nov 2;91(5):872-82.

Yasuda S, Oceguera-Yanez F, Kato T, Okamoto M, Yonemura S, Terada Y, Ishizaki T, Narumiya S. Cdc42 and mDia3 regulate microtubule attachment to kinetochores. Nature. 2004 Apr 15;428(6984):767-71.

Young KG, Copeland JW. Formins in cell signaling. Biochim Biophys Acta. 2010 Feb;1803(2):183-90.

Yousaf R, Ahmed ZM, Giese AP, Morell RJ, Lagziel A, Dabdoub A, Wilcox ER, Riazuddin S, Friedman TB, Riazuddin S. Modifier variant of METTL13 suppresses human GAB1-associated profound deafness. J Clin Invest. 2018 Apr 2;128(4):1509-1522.

Yousaf R, Gu C, Ahmed ZM, Khan SN, Friedman TB, Riazuddin S, Shears SB, Riazuddin S. Mutations in Diphosphoinositol-Pentakisphosphate Kinase PPIP5K2 are associated with hearing loss in human and mouse. PLoS Genet. 2018 Mar 28;14(3):e1007297.

Zhang Y, Wang F, Ding J, Zhang H, Zhao D, Yu L, Xiao H, Yao Y, Zhong X, Wang S. Genotype-phenotype correlations in 17 Chinese patients with autosomal recessive Alport syndrome. Am J Med Genet A. 2012 Sep;158A(9):2188-93.

Zheng J, Shen W, He DZ, Long KB, Madison LD, Dallos P. Prestin is the motor protein of cochlear outer hair cells. Nature. 2000 May 11;405(6783):149-55.

Zhou X, Jen PH, Seburn KL, Frankel WN, Zheng QY. Auditory brainstem responses in 10 inbred strains of mice. Brain Res. 2006 May 26;1091(1):16-26.

Zhu M, Yang T, Wei S, DeWan AT, Morell RJ, Elfenbein JL, Fisher RA, Leal SM, Smith RJ, Friderici KH. Mutations in the gamma-actin gene (ACTG1) are associated with dominant progressive deafness (DFNA20/26). Am J Hum Genet. 2003 Nov;73(5):1082-91.

Zong X, Eckert C, Yuan H, Wahl-Schott C, Abicht H, Fang L, Li R, Mistrik P, Gerstner A, Much B, *et al.* A novel mechanism of modulation of hyperpolarization-activated cyclic nucleotide-gated channels by Src kinase. J Biol Chem. 2005 Oct 7;280(40):34224-32.

Zwaenepoel I, Mustapha M, Leibovici M, Verpy E, Goodyear R, Liu XZ, Nouaille S, Nance WE, Kanaan M, Avraham KB, *et al.* Otoancorin, an inner ear protein restricted to the interface between the apical surface of sensory epithelia and their overlying acellular gels, is defective in autosomal recessive deafness DFNB22. Proc Natl Acad Sci U S A. 2002 Apr 30;99(9):6240-5.

6. Appendix

6.1. NSHL genes and their clinical manifestations.

Gene	Locus	Onset	NSHL phenotype	Additional features (SHL)
ACTG1	DFNA20/26	Post-lingual (1st/2nd)	Downsloping; audiogram moderate-to-profound; progressive	
ADCY1	DFNB44	Prelingual	Mild-to-moderate/profound; sensorineural or mixed; stable	
BDP1	DFNB112	Post-lingual (1st)	Downsloping; mild-to-severe; progresses, also involving low frequencies	
BSND	DFNB73	Prelingual	Severe-to-profound; stable	Some variants cause Bartter syndrome (deafness and nephropathy)
CABP2	DFNB93	Prelingual	U-shaped audiogram; moderate-to-severe; stable	
CCDC50	DFNA440	Post-lingual (1st)	Low-mid frequency; moderate-to-profound; progresses, involving all frequencies	
CD164	DFNA66	Prelingual/post-lingual (1st/2nd)	Flat or U-shaped audiogram; stable/progressive	
CDC14A	DFNB32/105	Prelingual	Moderate-to-profound; progressive	Some variants cause Deafness-Infertility Syndrome in males
CDH23	DFNB12	Prelingual	Downsloping; moderate-to-profound; progressive	Also associated with Usher syndrome Type ID (profound congenital HL, retinitis pigmentosa, constant vestibular dysfunction)
CEACAM16	DFNB113	Post-lingual (2nd)	Mild-to-moderate; progressive	
	DFNA4B	Post-lingual (1st/2nd)	Severe to profound; progressive. High-frequency tinnitus at onset in some patients.	
CIB2	DFNB48	Prelingual	Severe-to-profound	
CLDN14	DFNB29	Prelingual	Severe-to-profound; stable	
CLIC5	DFNB103	Prelingual	High-mid frequency; severe-to-profound; progresses, also involving low frequencies. Vestibular areflexia in the second decade of life.	
COCH	DFNB110	Prelingual	Downsloping; moderate-to-severe	
	DFNA9	Post-lingual (2nd/3rd)	Downsloping; moderate-to-profound; progresses to anacusis. Variants in LCCL domain are associated with HL and vestibular symptoms.	
COL11A1	DFNA37	Post-lingual (1st)	Mild-to-moderate; slowly progressive	Also associated with Stickler syndrome type II and Marshall syndrome (ocular, auditory, skeletal, and orofacial abnormalities)
COL11A2	DFNB53	Prelingual	Severe-to-profound; stable	
	DFNA13	Post-lingual (2nd/4th)	Moderate-to-severe; progressive	Some variants cause otospondylomegaepiphyseal dysplasia (sensorineural deafness and short extremities with large epiphyses)
CRYM	DFNA40	Prelingual	Moderate-to-severe	
DCDC2	DFNB66	Prelingual	Severe-to-profound	
DIABLO	DFNA64	Post-lingual (2nd/3rd)	Mild-to-severe; progressive. Frequent high-frequency tinnitus at onset.	
DIAPH1	DFNA1	Post-lingual (1st)	Low-frequency; progresses to profound HL involving all frequencies	Some variants cause high-frequency progressive HL with thrombocytopenia
DIAPH3	AUNA1	Post-lingual (2nd)	Profound; auditory neuropathy	
DMXL2	DFNA71	Post-lingual (2nd)	Mild-to-profound; progressive	
ELMOD3	DFNB88	Prelingual	Severe-to-profound; mixed	

Gene	Locus	Onset	NSHL phenotype	Additional features (SHL)
EPS8	DFNB102	Prelingual	Severe-to-profound	
EPS8L2	DFNB106	Post-lingual (1st)	Downsloping; moderate-to-profound; progressive	
ESPN	DFNB36	Prelingual	Profound with or without vestibular areflexia	
ESRP1	DFNB109	Prelingual	Severe-to-profound. Vestibular dysplasia without balance problems.	
ESRRB	DFNB35	Prelingual	Severe-to-profound; stable	
EYA4	DFNA10	Post-lingual (2nd/5th)	Severe-to-profound; progressive	
GAB1	DFNB26	Prelingual	Severe-to-profound. Incomplete penetrance: a variant in *METTL13* gene prevents *GAB1*-associated deafness.	
GIPC3	DFNB15/72/95	Prelingual	Moderate-to-severe/profound; stable	
GJB2	DFNB1A	Prelingual	Mild-to-profound (variable); progressive in ~50% of cases. ~50% of cases report vestibular symptoms.	
	DFNA3A	Prelingual/post-lingual (1st/2nd)	Downsloping; mild-to-profound (variable); progressive	Associated with different syndromic forms causing HL and skin problems
GJB3	DFNA2B	Post-lingual (4th)	Downsloping; progressive	
GJB6	DFNB1B	Prelingual	Severe; stable. Vestibular symptoms in some patients.	
	DFNA3B	Prelingual	Downsloping; mild-to-profound (variable); progressive	
GRHL2	DFNA28	Post-lingual (variable, from 1st)	Mild-to-moderate; progresses to severe in high frequencies by 5th decade	
GRXCR1	DFNB25	Prelingual	Moderate-to-profound; progressive. Vestibular symptoms in some patients.	
GRXCR2	DFNB101	Prelingual	Moderate-to-profound; progressive	
GSDME	DFNA5	Post-lingual (1st)	Downsloping; progresses, involving all frequencies	
HGF	DFNB39	Prelingual	Downsloping; severe-to-profound	
HOMER2	DFNA68	Post-lingual (1st)	Downsloping; progresses, involving all frequencies	
IFNLR1	DFNA2C	Post-lingual (2nd)	Downsloping; progresses, involving all frequencies. Generally accompanied by high-frequency tinnitus.	
ILDR1	DFNB42	Prelingual	Moderate-to-profound; stable	
KARS	DFNB89	Prelingual	Moderate-to-severe; stable	
KCNQ4	DFNA2A	Post-lingual (1st/2nd)	Downsloping; progresses, involving all frequencies	
KITLG	DFNA69	Prelingual	Unilateral or bilateral asymmetric; stable. Subclinical vestibular dysfunction in some patients.	
LHFPL5	DFNB66/67	Prelingual	Severe-to-profound; stable	
LMX1A	DFNA7	Prelingual/post-lingual (1st/3rd)	Variable severity and progression rate. Vestibular symptoms in some patients.	
LOXHD1	DFNB77	Post-lingual (1st)	Downsloping; moderate-to-profound; progressive, flattening of the audiogram over time	
LRTOMT	DFNB63	Prelingual	Severe-to-profound; stable	

Gene	Locus	Onset	NSHL phenotype	Additional features (SHL)
MARVELD2	DFNB49	Prelingual	Moderate-to-profound; stable	
MCM2	DFNA70	Post-lingual (from 1st)	Flat/downsloping; mild-to-profound; progressive	
MET	DFNB97	Prelingual	Severe-to-profound	
MIR96	DFNA50	Post-lingual (1st)	Flat; mild-to-profound; progressive	
MPZL2	DFNB111	Prelingual/post-lingual (1st)	Moderate-to-severe; slowly progressive	
MSRB3	DFNB74	Prelingual	Profound	
MYH14	DFNA4A	Post-lingual (1st/2nd)	Severe-to-profound; progressive	One variant has been associated with peripheral neuropathy, myopathy, hoarseness, and hearing loss in one family
MYH9	DFNA17	Post-lingual (1st/3rd)	Downsloping; progressive	Also associated with macrothrombocytopenia and granulocyte inclusions with or without nephritis or sensorineural hearing loss
MYO15A	DFNB3	Prelingual	Severe-to-profound; stable	
MYO3A	DFNB30	Prelingual	Downsloping; moderate-to-profound; progressive	
	-	Post-lingual (1st)	Moderate-to-profound; progressive	
MYO6	DFNB37	Prelingual	Severe-to-profound; progressive	
	DFNA22	Post-lingual (1st)	Mild-to-profound; progressive	One variant has been associated with deafness and hypertrophic cardiomyopathy
MYO7A	DFNB2	Prelingual/post-lingual (1st/2nd)	Downsloping; severe-to-profound. Vestibular symptoms in some patients.	Also associated with Usher syndrome Type IB (profound congenital HL, retinitis pigmentosa, constant vestibular dysfunction)
	DFNA11	Post-lingual (1st/5th)	Flat/gently downsloping; moderate; progressive. Mild vestibular symptoms in some patients.	
NARS2	DFNB94	Prelingual	Profound; stable	
NLRP3	DFNA34	Post-lingual (1st/4th)	Variable severity; slowly progressive.	Some patients have pure NSHL, while others have features of an autoinflammatory disorder with systemic manifestations, including periodic fevers, arthralgias, and episodic urticaria. *NLRP3* gene is also associated with Muckle-Wells syndrome (episodic skin rash, arthralgias, and fever associated with late-onset sensorineural deafness and renal amyloidosis).
OSBPL2	DFNA67	Post-lingual (1st/3rd)	Downsloping; mild-to-profound; progresses to all frequencies. Generally accompanied by high-frequency tinnitus.	
OTOA	DFNB22	Prelingual	Moderate-to-profound; stable	
OTOF	DFNB9	Prelingual	Severe-to-profound; stable	
OTOG	DFNB18B	Prelingual	Downsloping; mild -to-moderate; stable	
OTOGL	DFNB84B	Prelingual	Downsloping; moderate; stable	
P2RX2	DFNA41	Post-lingual (2nd)	Downsloping in males, upsloping audiogram in females; patterns tend to converge with HL progression. Moderate-to-severe; progressive. Generally accompanied by high-frequency tinnitus.	

Gene	Locus	Onset	NSHL phenotype	Additional features (SHL)
PCDH15	DFNB23	Prelingual	Severe-to-profound; stable	Also associated with Usher syndrome Type IF (profound congenital HL, retinitis pigmentosa, constant vestibular dysfunction)
PDE1C	DFNA74	Post-lingual (1st/3rd)	Downsloping; mild-to-profound; progressive	
PDZD7	DFNB57	Prelingual	Downsloping; mild-to-profound; mildly progressive	Also associated with Usher syndrome Type IIC (moderate-to-severe congenital HL, retinitis pigmentosa)
PJVK	DFNB59	Prelingual	Severe-to-profound	
PNPT1	DFNB70	Prelingual	Severe	Also associated with combined oxidative phosphorylation deficiency-13 (multisystem disorder characterised by severe neurologic impairment, including HL)
POU4F3	DFNA15	Post-lingual (2nd/6th)	Flat/downsloping; mild-to-severe; progressive	
PPIP5K2	DFNB100	Prelingual	Severe-to-profound	
PTPRQ	DFNB84A	Prelingual	Moderate-to-profound; progressive. Vestibular dysfunction.	
	DFNA73	Post-lingual (1st/3rd)	Mild-to-severe	
RDX	DFNB24	Prelingual	Severe-to-profound; stable	
REST	DFNA27	Post-lingual (1st/3rd)	Moderate-to-profound; progressive	
RIPOR2	DFNB104	Prelingual	Severe-to-profound	
ROR1	DFNB108	Prelingual	Profound. Fusion of the cochlea and vestibule into a common cavity; no vestibular dysfunction.	
S1PR2	DFNB68	Prelingual	Profound	
SERPINB6	DFNB91	Post-lingual (2nd)	Moderate-to-severe; progressive	
SIX1	DFNA23	Prelingual	Downsloping; mild-to-profond; sensorineural or mixed	Also associated with Branchiootic syndrome 3 (HL, branchial arch defects, lacrimal duct stenosis)
SLC17A8	DFNA25	Post-lingual (2nd/6th)	Downsloping; slowly progressive	
SLC22A4	DFNB60	Prelingual	Severe-to-profound	
SLC26A4	DFNB4	Prelingual	Moderate-to-profound. Associated with cochlear abnormalities (EVA or Mondini dysplasia) and vestibular symptoms in some patients.	Also associated with Pendred syndrome (HL and goiter)
SLC26A5	DFNB61	Prelingual	Moderate-to-profound; stable	
STRC	DFNB16	Prelingual	Downsloping; mild-to-profound; stable	Deletion of *STRC* in conjunction with *CATSPER2* gene causes Deafness-Infertility syndrome
SYNE4	DFNB76	Prelingual	Downsloping; moderate-to-severe; progressive	
TBC1D24	DFNB86	Prelingual	Flat audiogram; profound	
	DFNA65	Post-lingual (2nd/3rd)	Downsloping; severe-to-profound; progresses, also involving low frequencies	
TECTA	DFNB21	Prelingual	Flat/U-shaped audiogram; severe-to-profound; stable	
	DFNA8/12	Prelingual/post-lingual (1st/2nd)	U-shaped audiogram; moderate-to-severe; stable/progressive	
TJP2	DFNA51	Post-lingual (4th)	Downsloping; severe-to-profound; progresses, involving all frequencies	

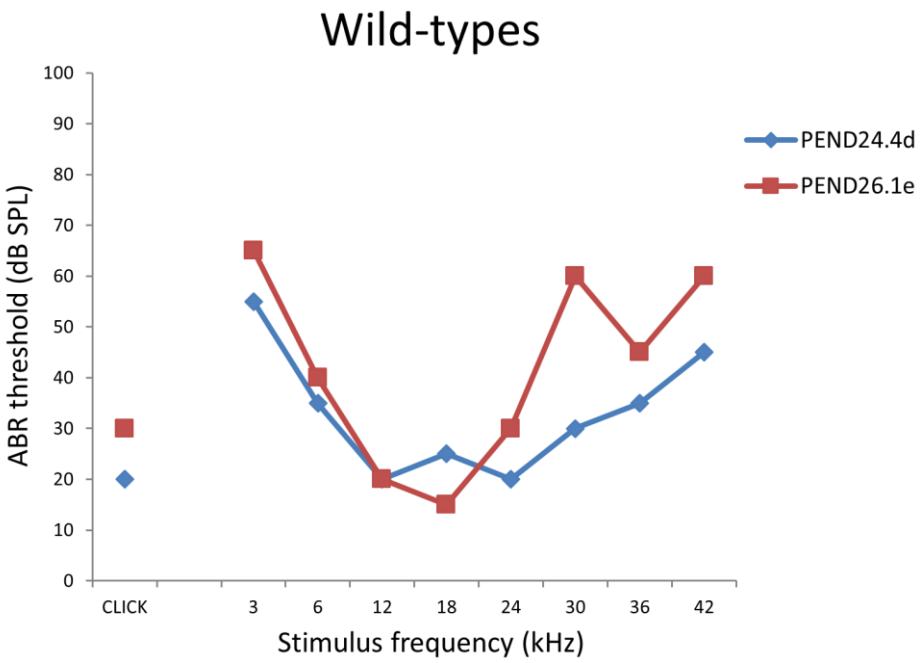
Gene	Locus	Onset	NSHL phenotype	Additional features (SHL)
TMC1	DFNB7/11	Prelingual	Profound; stable	
	DFNA36	Post-lingual (1st/3rd)	Downsloping; severe-to-profound; progresses, involving all frequencies	
TMEM132E	DFNB99	Prelingual	Severe-to-profound	
TMIE	DFNB6	Prelingual	Severe-to-profound; stable	
TMPRSS3	DFNB8/10	Prelingual/post-lingual (1st)	Moderate-to-profound	
TNC	DFNA56	Post-lingual (1st/3rd)	Low-frequency; mild-to-severe; progresses, involving all frequencies	
TPRN	DFNB79	Prelingual	Severe-profound; progressive	
TRIOBP	DFNB28	Prelingual	Severe-to-profound; stable	
TSPEAR	DFNB98	Prelingual	Flat audiogram; profound	
USH1C	DFNB18A	Prelingual	Severe-to-profound; stable	Also associated with Usher syndrome Type IC (profound congenital HL, progressive retinitis pigmentosa, vestibular hypofunction)
WBP2	DFNB107	Prelingual	Downsloping; severe-to-profound	
WFS1	DFNA6/14/38	Prelingual	Low-frequency; progressive	Also associated with Wolfram syndrome 1 (diabetes insipidus and mellitus with optic atrophy and HL)
WHRN	DFNB31	Prelingual	Profound	Also associated with Usher syndrome Type IID (moderate congenital HL, retinitis pigmentosa)

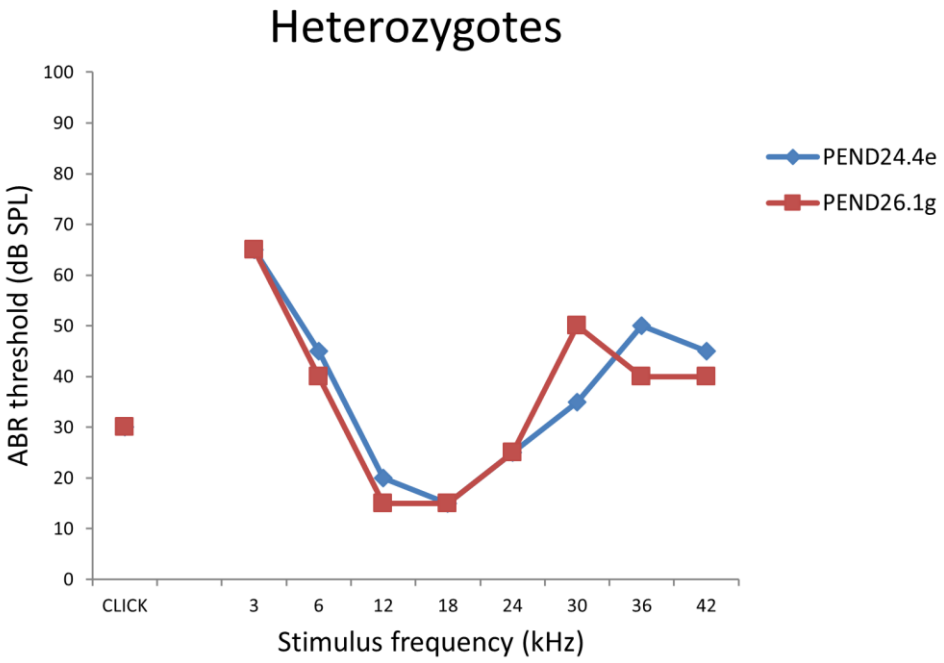
Adapted and updated from Shearer *et al.*, 2017 and Hereditary Hearing Loss Homepage, https://hereditaryhearingloss.org/.

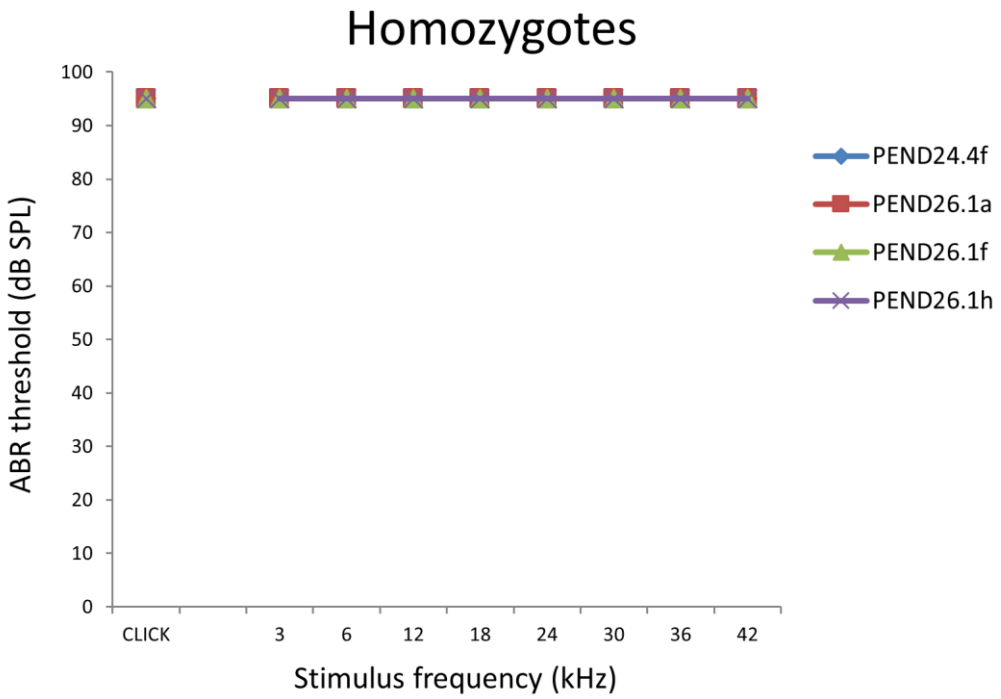
6.2. ABR recordings of individual *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice

The auditory function of *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice was evaluated using ABR recordings at 3, 4 and 8 weeks of age. The individual ABR recordings of each mouse I personally tested are reported. The personal ID of each mouse is indicated on the right.

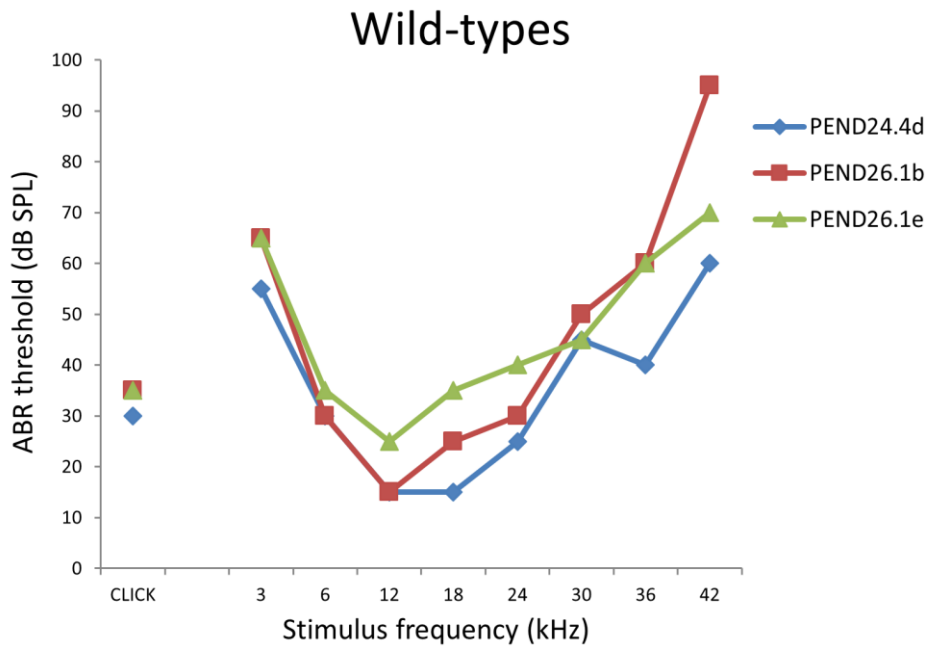
6.2.1. ABR recordings of individual 3-week-old *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice

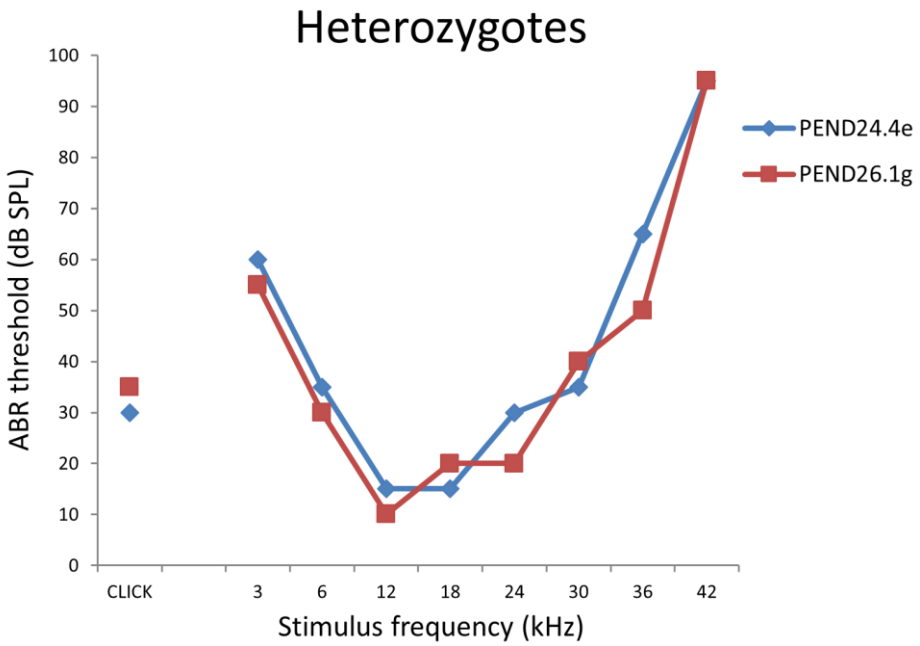


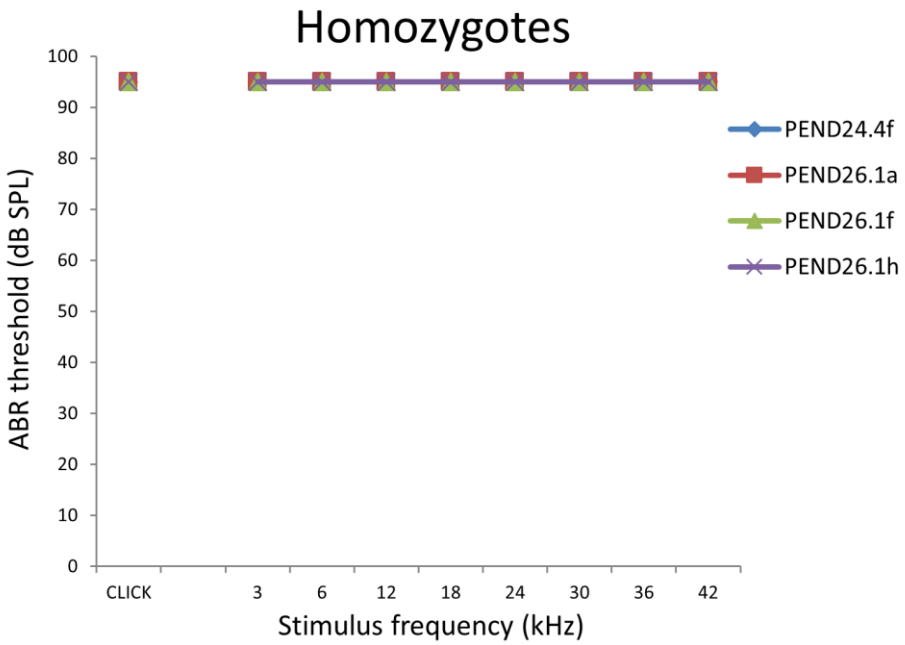




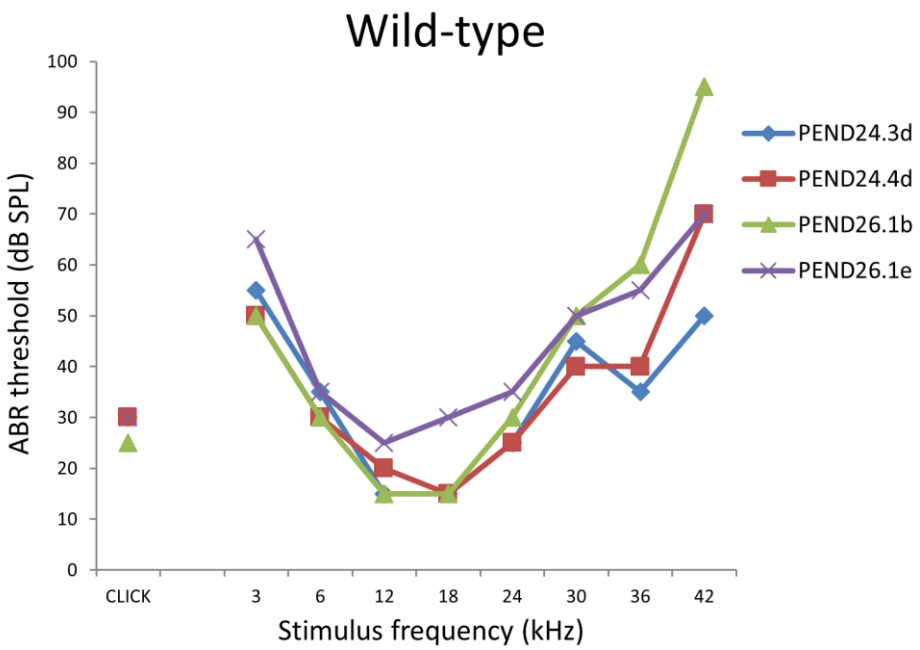
6.2.2. ABR recordings of individual 4-week-old *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice

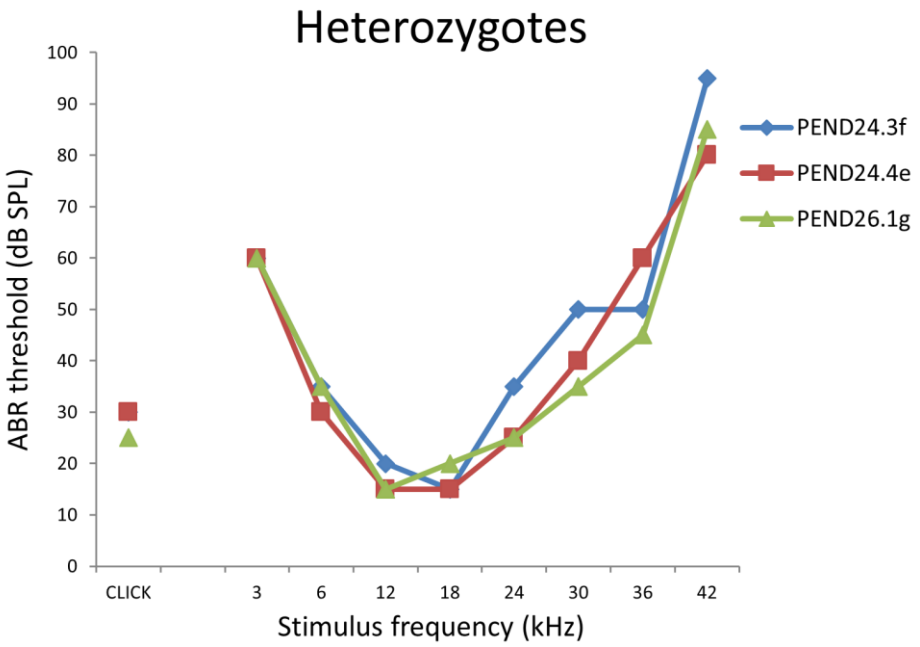


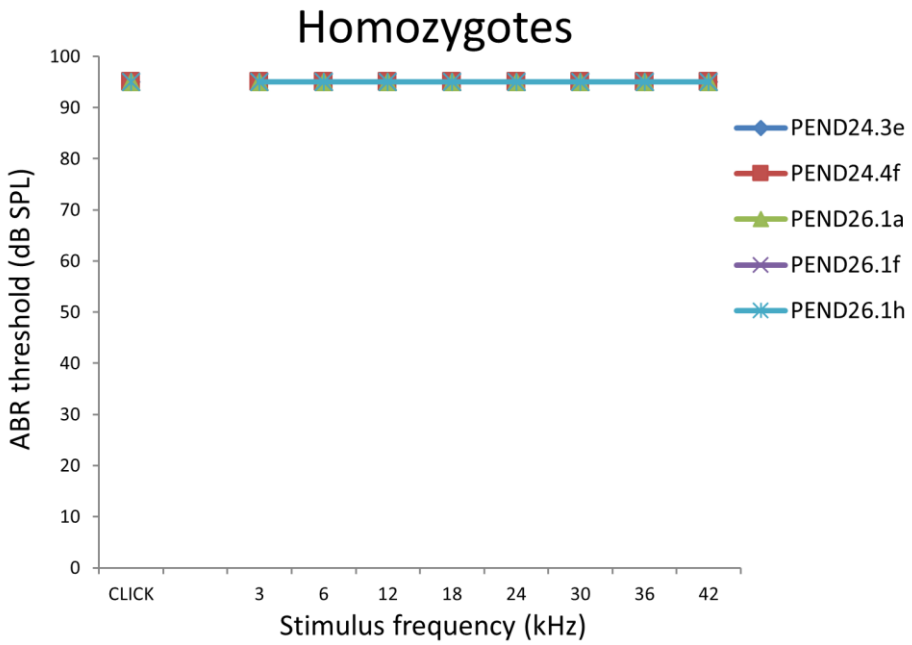

Heterozygotes
ABR threshold (dB SPL)
100
90
80
70
60
50
40
30
20
10
0
CLICK
3
6
12
18
24
30
36
42
Stimulus frequency (kHz)
PEND24.4e
PEND26.1g


Homozygotes
ABR threshold (dB SPL)
100
90
80
70
60
50
40
30
20
10
0
CLICK
3
6
12
18
24
30
36
42
Stimulus frequency (kHz)
PEND24.4f
PEND26.1a
PEND26.1f
PEND26.1h

6.2.3. ABR recordings of individual 8-week-old *Slc26a4*$^{tm1(CreERT2\_EGFP)Wtsi}$ mice



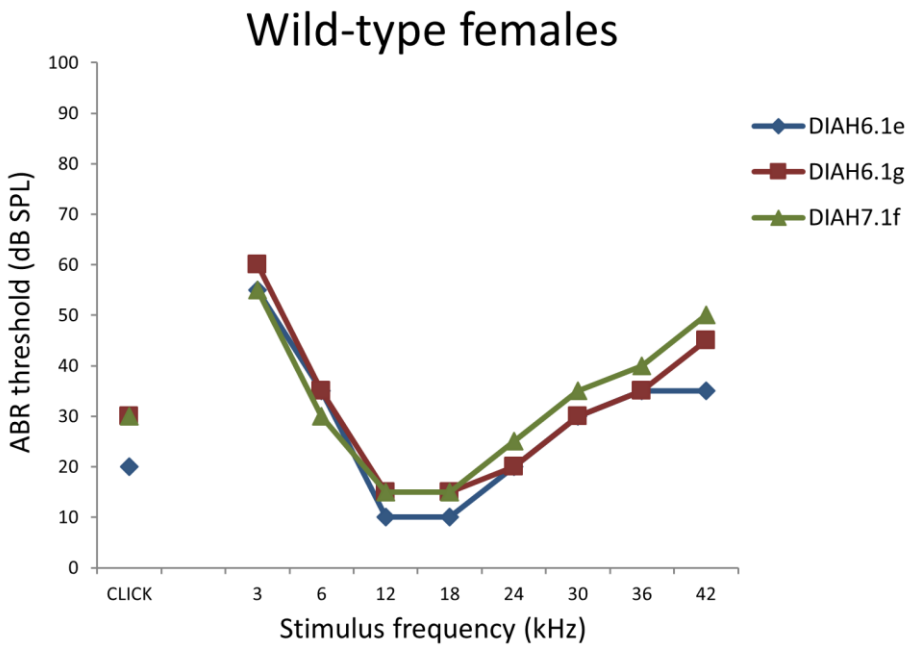


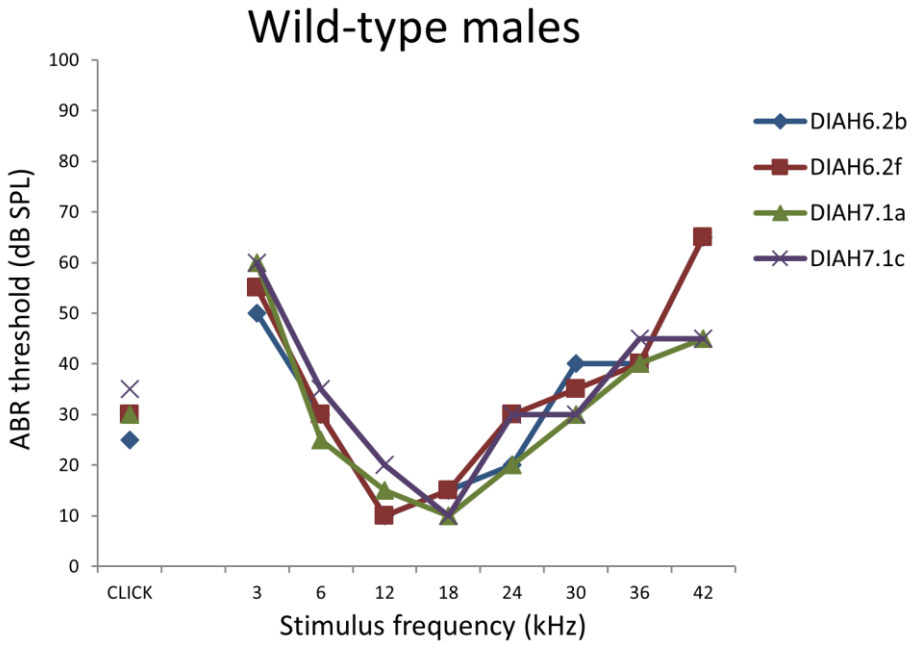


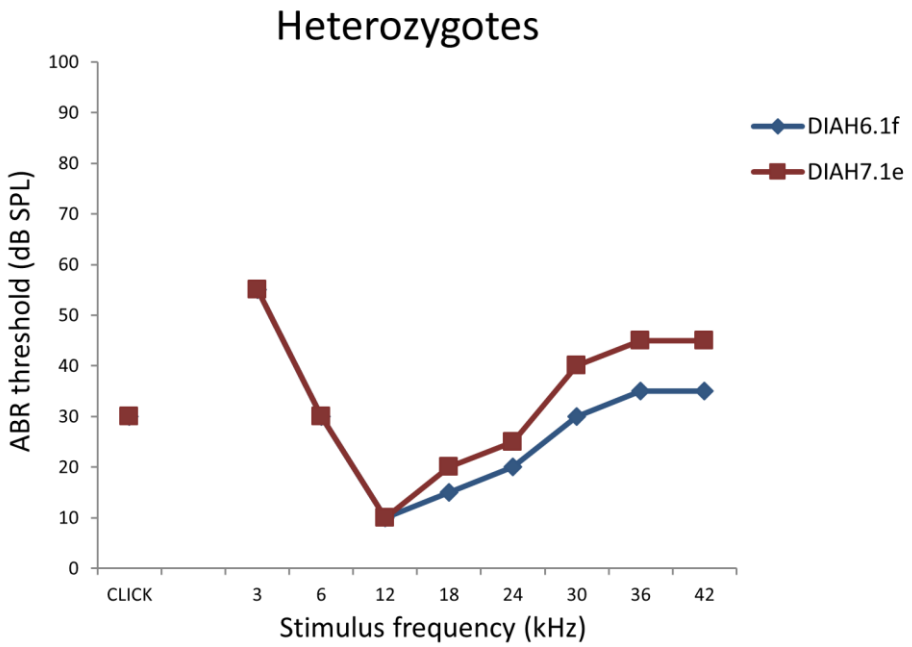
6.3. ABR recordings of individual *Diaph2*^{em2Kcl} knock-out mice

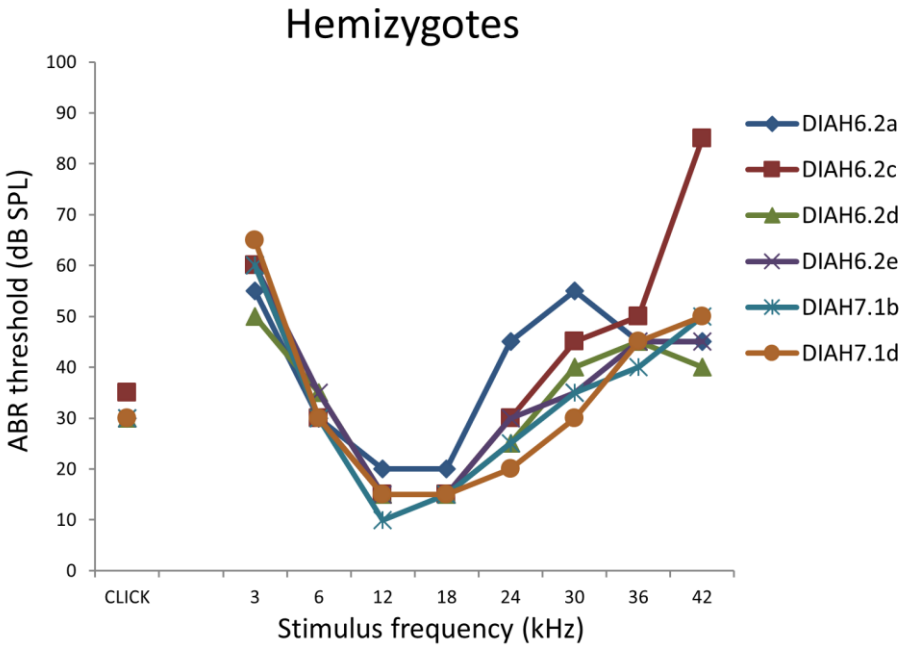
The auditory function of *Diaph2*^{em2Kcl} mice was evaluated using ABR recordings at 4 and 8 weeks of age. The individual ABR recordings of each mouse I personally tested are reported. The personal ID of each mouse is indicated on the right.

6.3.1. ABR recordings of individual 4-week-old *Diaph2*^{em2Kcl} knock-out mice

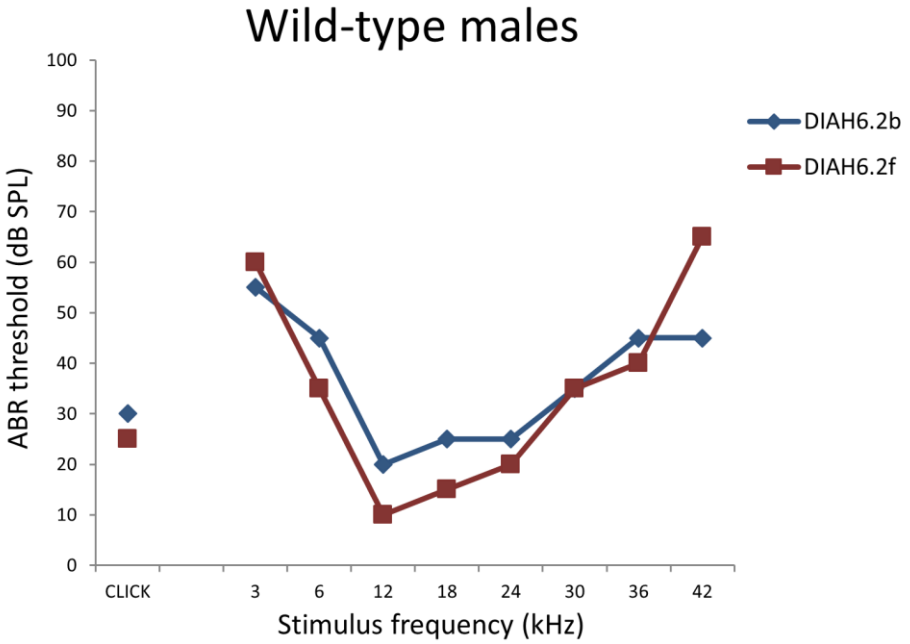


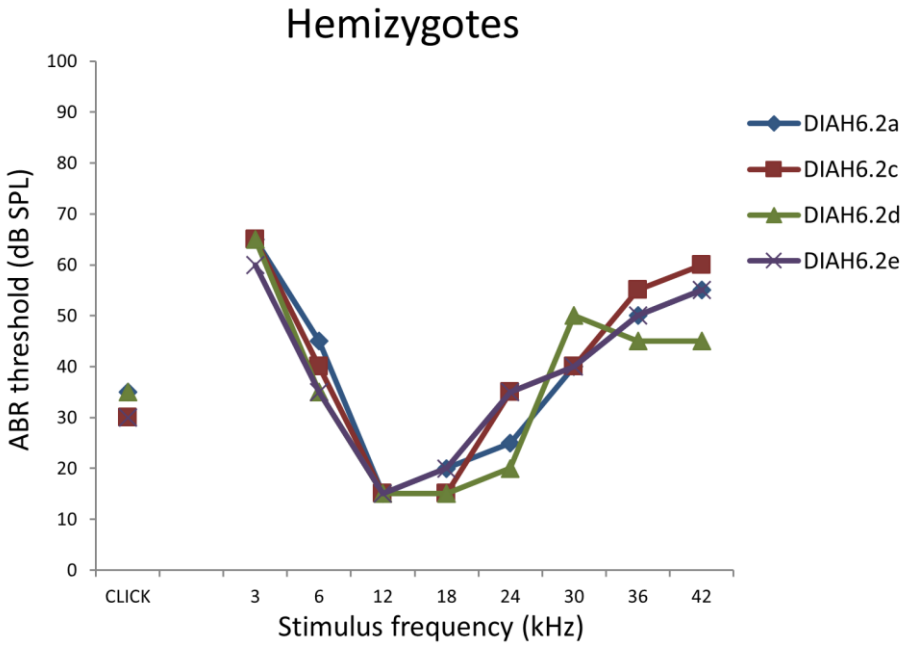

Wild-type males
ABR threshold (dB SPL)
100
90
80
70
60
50
40
30
20
10
0
CLICK
3
6
12
18
24
30
36
42
Stimulus frequency (kHz)
DIAH6.2b
DIAH6.2f
DIAH7.1a
DIAH7.1c


Heterozygotes
ABR threshold (dB SPL)
100
90
80
70
60
50
40
30
20
10
0
CLICK
3
6
12
18
24
30
36
42
Stimulus frequency (kHz)
DIAH6.1f
DIAH7.1e



6.3.2. ABR recordings of individual 8-week-old *Diaph2*^{em2Kcl} knock-out mice

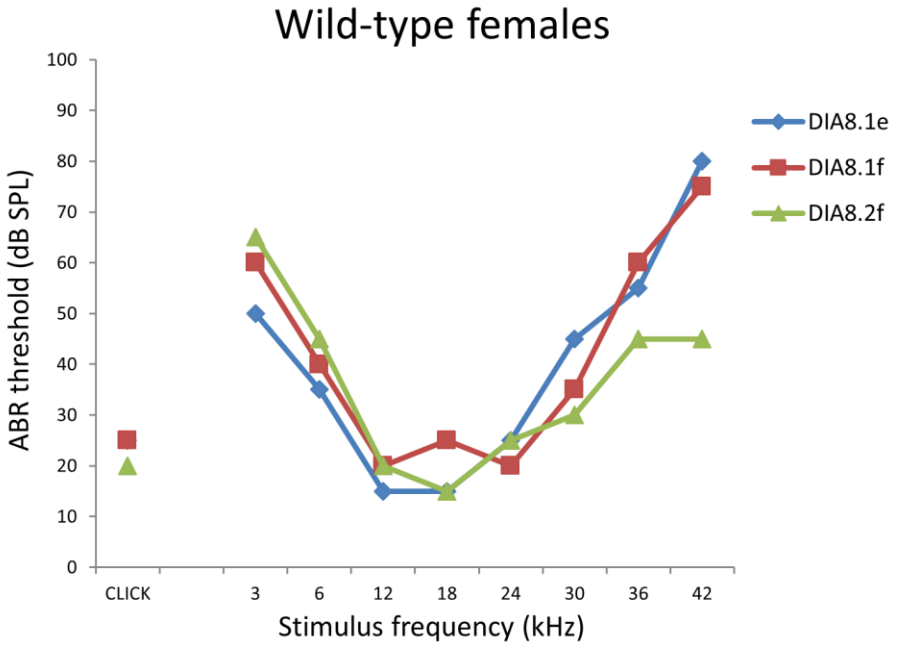


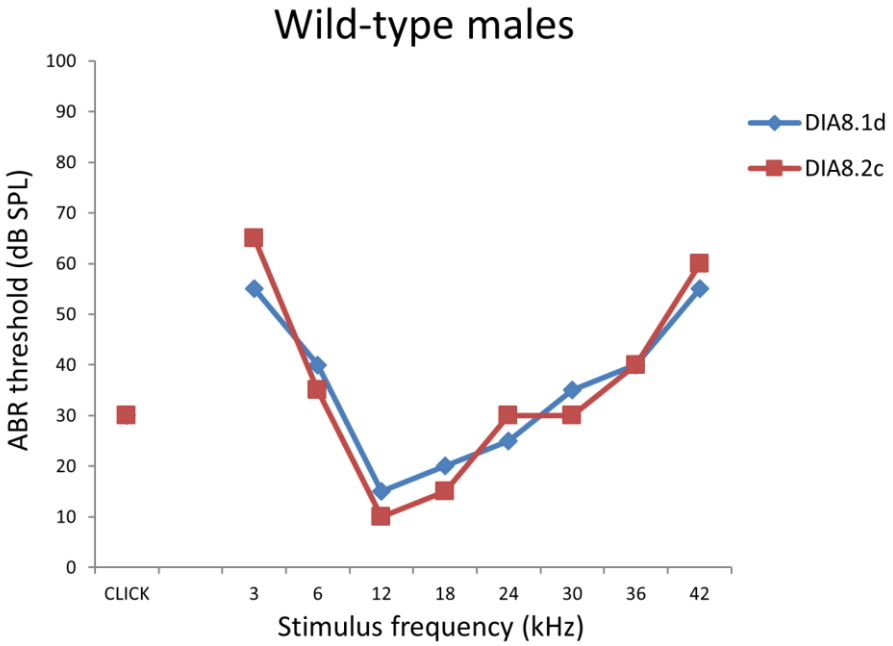


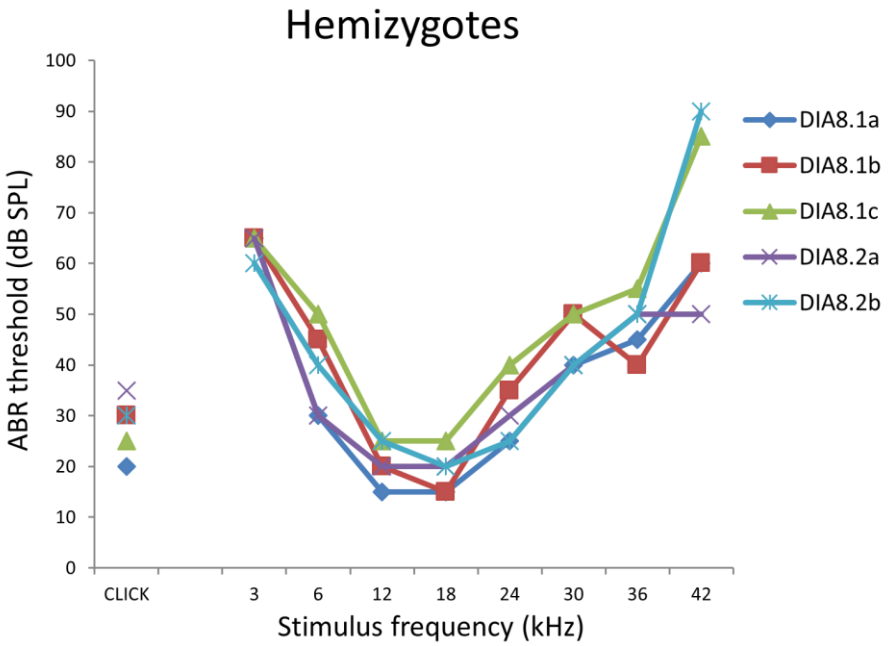
6.4. ABR recordings of individual *Diaph2*^{em3Kcl} knock-in mice

The auditory function of *Diaph2*^{em3Kcl} mice was evaluated using ABR recordings at 4 and 8 weeks of age. The individual ABR recordings of each mouse I personally tested are reported. The personal ID of each mouse is indicated on the right.

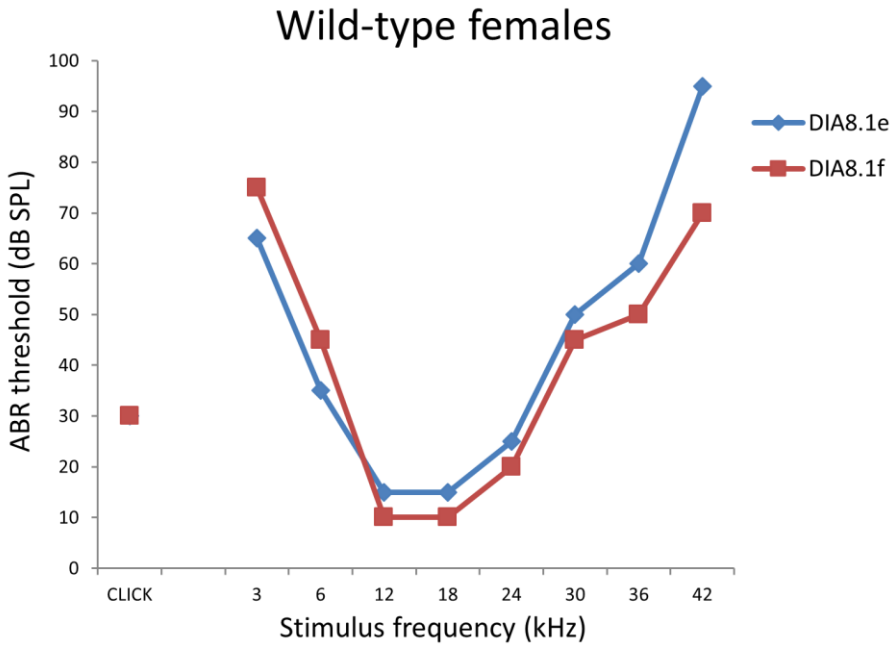
6.4.1. ABR recordings of individual 4-week-old Diaph2em3Kcl knock-in mice

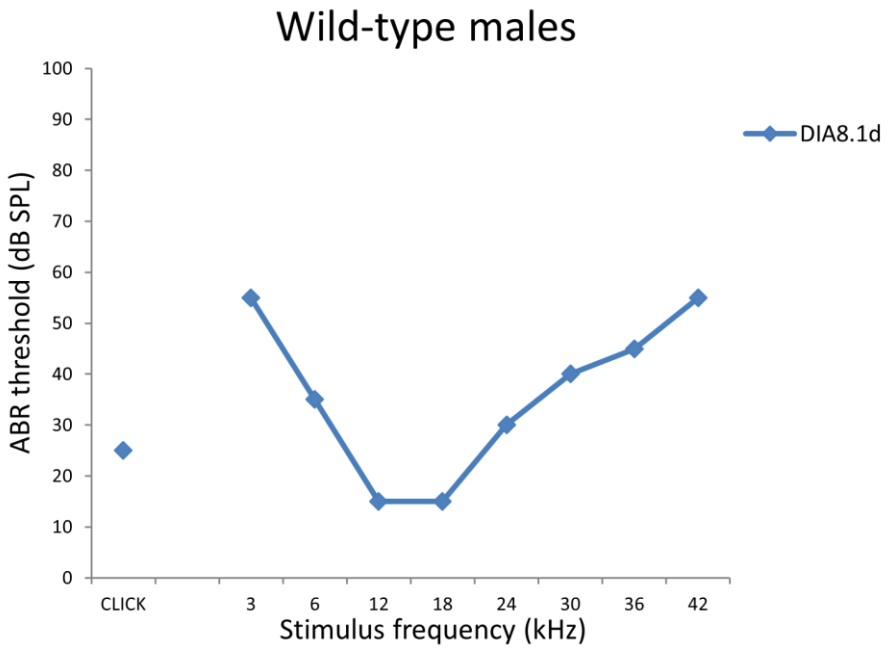


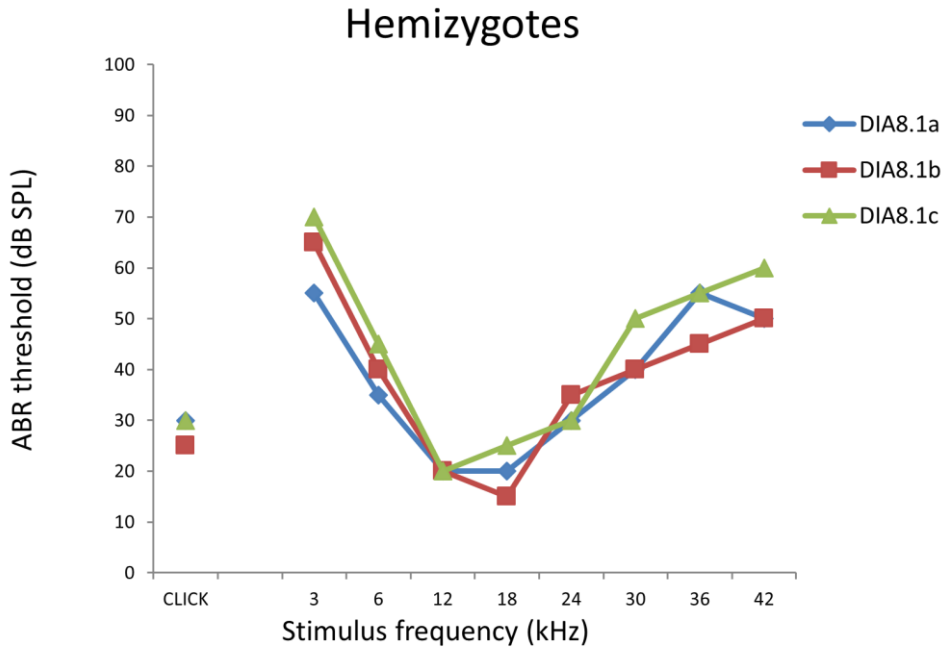

Wild-type males
ABR threshold (dB SPL)
100
90
80
70
60
50
40
30
20
10
0
CLICK
3
6
12
18
24
30
36
42
Stimulus frequency (kHz)
DIA8.1d
DIA8.2c


Hemizygotes
ABR threshold (dB SPL)
100
90
80
70
60
50
40
30
20
10
0
CLICK
3
6
12
18
24
30
36
42
Stimulus frequency (kHz)
DIA8.1a
DIA8.1b
DIA8.1c
DIA8.2a
DIA8.2b

6.4.2. ABR recordings of individual 8-week-old *Diaph2*^{em3Kcl} knock-in mice







6.5. Relevant publications

Chiereghin C, Robusto M, Mauri L, Primignani P, Castorina P, Ambrosetti U, Duga S, Asselta R, Soldà G. Role of SLC22A4 gene in hereditary nonsyndromic hearing loss: recurrence and incomplete penetrance of the p.C113Y mutation in Northwest Africa. Under submission.

Chiereghin C, Robusto M, Mastrangelo A, Castorina P, Montini G, Giani M, Duga S, Asselta R, Soldà G. Alport syndrome cold cases: missing mutations identified by exome sequencing and functional analysis. PLoS One. 2017 Jun 1;12(6).

Soldà G, Caccia S, Robusto M, Chiereghin C, Castorina P, Ambrosetti U, Duga S, Asselta R. First independent replication of the involvement of LARS2 in Perrault syndrome by whole-exome sequencing of an Italian family. J Hum Genet. 2016 Apr;61(4):295-300.