Application of Massively Parallel Sequencing for Mutation Discovery and Genetic Diagnosis of Hereditary Hearing Loss in a Chinese Population

CAO, Ye

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Obstetrics and Gynaecology

The Chinese University of Hong Kong
July 2014

Abstract of thesis entitled: Application of Massively Parallel Sequencing for Mutation Discovery and Genetic Diagnosis of Hereditary Hearing Loss in a Chinese Population

Submitted by CAO, Ye

For the degree of Doctor of Philosophy in Obstetrics and Gynaecology

At the Chinese University of Hong Kong in July 2014

Abstract

Hearing loss is one of the most common birth defects, with a prevalence rate of approximately 3 per 1000 neonates. In the latest WHO report, 360 million people worldwide have hearing loss at least greater than 30dB. In China, based on the data of the Second China National Sample Survey on Disability, it was estimated that 20 million people were diagnosed with hearing disability. This accounts for 1.5% of the total Chinese population, and represents 115 000 children under the age of 7 years with severe-to-profound deafness and 30 000 newborns (~0.15%) each year with hearing impairment.

Hereditary hearing loss is due to genetic defects. Therefore establishing a genetic diagnosis for hereditary hearing loss is favorable for explanation of the disease, determination of clinical management, and to offer appropriate genetic counseling for patients and their family members. It has a well-recognized phenotype which is relatively easy to be diagnosed clinically. Nonetheless, genetic causes contributing to hearing impairment are unparalleled heterogeneous. Implementing genetic diagnosis in clinical practice for hereditary hearing loss is challenging.

It has been reported that different ethnic groups show a unique spectrum of common mutations of hearing loss genes. Understanding the spectrum and carrier frequency of common mutations particularly in the Chinese population could facilitate the molecular diagnosis of hearing loss. According to common mutations extracted from the literatures regarding hearing loss patients, we studied the carrier frequency of 15 common mutations in *GJB2*, *SLC26A4* and the mitochondrial genome by the SNapShot method. This is the first epidemiological study of nonsyndromic hearing loss in Chinese that reveals high carrier frequencies (one per 6.3 newborns) of 15 common mutations in newborns.

Considering the limitation of mutations screened by Sanger sequencing, we developed the solution-based capture system with massively parallel sequencing (MPS) to target more than 252 genes related with hearing loss. An in-house bioinfomatic analysis pipeline was constructed to identify candidate variants with a high validation rate. We designed both Agilent SureSelect and NimbleGen SeqCap target enrichment platforms and adopted them for the mutation discovery of hearing loss genes in unrelated Chinese families with nonsyndromic hearing loss. Our targeted genome enrichment (TGE) MPS approach identified an additional 28% (2 out of 7) of unrelated Chinese families carrying novel mutations.

As more hearing loss patients receive a genetic diagnosis, there will be a foreseeable growing demand for prenatal diagnosis. Standard prenatal testing requires an invasive procedure to obtain a fetal sample, which has the risk of fetal loss. Therefore we explored the application of a noninvasive method using cell free fetal DNA in a family with a *GJB2* mutation. Using haplotype analysis of a 1.1 Mb flanking region around *GJB2*, we successfully recovered the fetal haplotype and

deduced that the fetus inherited the paternal mutant allele and maternal wide-type allele.

In summary, our carrier frequency data aid in effective risk assessment and genetic counseling for hearing loss patients in the Chinese population. The newly established target genome enrichment MPS method and noninvasive haplotype approach enhanced the success of genetic diagnosis of hereditary hearing loss.

摘　要

聽力障礙是最常見的出生缺陷之一，大約每一千個新生兒中就有三個表現為聽力障礙。世界衛生組織的一份最新報告顯示，世界範圍內有三億六千萬人有大於或等於 30 分貝聽閾值的聽損。在中國，根據第二次國家殘疾調查數據估計有兩千萬人群已經診斷為聽力障礙。約佔中國總人口的百分之一點五。其中有十一萬伍仟小於七歲的兒童診斷為重度失聰， 且每年至少有三萬名新生兒有聽力障礙。

遺傳性耳聾可根據臨床表現診斷。然而導致耳聾的遺傳背景卻千差萬別。遺傳因素約佔學語前性耳聾的百分之五十。明確導致耳聾的遺傳性致病因素有助於解釋疾病的發生發展，臨床診療決策，針對病人及家屬有效的遺傳咨詢。然而在臨床上開展針對遺傳性耳聾的分子診斷確非易事。

根據文獻報導不同種族耳聾患者尤其獨特的熱點突變類型。了解中國人群的突變種類和人群攜帶率有助於開展耳聾的分子診斷。經過針對中國耳聾患者文獻報導的查詢，我們發現 15 個常見突變。並在中國新生兒人群中採用 SNaPshot 方法鑑定這 15 個突變的基因型。由於 Sanger 測序方法所能檢測該方法所能檢測的基因型有限，考慮到遺傳性耳聾的致病基因眾多，我們將液相捕獲和二代測序相結合來實現高通量篩查。針對二百多個耳聾相關基因設計捕獲探針，雜交捕穫後同時進行二代測序。針對該流程自設的生物信息分析方法可有效定位變異位點，且 sanger 測序驗證率高。我們亦同時設計了 NimbleGen 和 Agilent 兩種平台的液相捕獲探針文庫。最終在七名患者中的兩名找到可能與耳聾相關的遺傳突變。

伴隨著分子診斷對耳聾患者的廣泛應用，產前診斷需求亦有增加。傳統產前診斷需要有創操作獲取胎兒組織。這些操作可能對胎兒不利從而導致流產。我們將母親血漿 DNA 分析胎兒基因型的無創方法应用在一例遺傳性耳聾的產前診斷。胎兒經過分析為致病突變攜帶者。該結果與孕中期有創產前基因檢查結果一致。

總而言之，該論文中所總結的耳聾熱點突變的中國人群攜帶率有助於耳聾患者的風險評估和遺傳咨詢。目標區域富集結合二代測序以及無創單倍體分型的方法建立都將增強遺傳性耳聾的基因診斷的應用。

Acknowledgements

First, I would like to express my sincerest and profoundest gratitude to my supervisor Dr. Richard Choy, who supported my Ph.D. study and research with his patience, motivation, enthusiasm and knowledge. His guidance and encouragement helped me throughout my research and writing of this thesis. His words always inspired me and brought me to a higher level. What I learned from him is not only how to do science itself, but also the self-drive to be a better person.

I have been blessed with a friendly and lovely group of team mates: Jin Huang, Kenneth Wong, Wilson Chong, Dawn Lui, Cherry Wong, Yvonne Kwok, Yongjian Yue, Christina Lin, for their valuable help and advice during my work, promoting a stimulating and welcoming academic and social environment.

My sincere appreciation goes to Prof. Cynthia Morton and Dr. Jun Shen, at Department of Obstetrics and Gynaecology and Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, USA, for offering me the exchange opportunity in their unit, allowing me to work with their exciting projects and providing great help and advice on my projects.

I am grateful to Prof. Sau Wai Cheung and Prof. Carlos Bacino, Prof. V. Reid Sutton in the Department of Molecular and Human Genetics, Baylor Medical College USA, for their precious help and advice in clinical genetics on my work.

I would like to thank colleagues and collaborators who helped in recruiting samples: Prof. Tak Yeung Leung, Dr. Yvonne Cheng, other staff in the Prenatal Diagnosis Unit, Prince of Wales Hospital and Dr. Ying Chen and Dr. Huijun Yuan from China.

It is an honor to work with my labmates Tao Zhang, Gene Man, Ariel Cheung, Kristy Cheung, Loucia Chan, Tat San Lau, Yan Tong, Dr. Joseph Kwok, Dr. Ronald Wong and staffs in my department. I cannot find words to express my gratitude to my friends Leina Lu, Yang Liu, Jiajun Zhang, Hui Li and Lizi Shi. Thanks for sharing their incredible ideas, cheering me up when I felt frustrated and taking care of my life.

I owe my deepest gratitude to my parents and family with their unconditional support all these years. Without their encouragement, I could never be here to pursue my dreams.

Abbreviations

Nucleotides:

A	Adenine
T	Thymine
C	Cytosine
G	Guanine

Amino acids:

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

ABR	Auditory Brain stem Response testing
aCGH	Array Comparative Genomic Hybridization
AD	Autosomal Dominant
AR	Autosomal Recessive
CI	Cochlear Implantation
CVS	Chorionic Villus Sampling
dB	Decibel
DFNA	Nonsyndromic hearing loss, autosomal dominant
DFNB	Nonsyndromic hearing loss, autosomal recessive
DFNX	Nonsyndromic hearing loss, X-linked
FISH	Fluorescence In Situ Hybridization
Indel	Small insertion and deletions
MAF	Minor Allele Frequency
MLPA	Multiplex Ligation Dependent Probe Amplification
MPS	Massively Parallel Sequencing
NIPT	Noninvasively Prenatal Diagnosis Testing
OAE test	Otoacoustic Emissions test
SNHL	Sensorineural Hearing Loss
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variants
UNHS	Universal Newborn Hearing Screening
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

Publications

1 Chen Y*, **Cao Y*,** Li H, Mao J, Liu M, Liu Y, Wang B, Jiang D, Zhu Q, Ding Y, Wang W, Li H, Choy KW. SNaPshot Revealed a High Mutation and Carrier Frequency of 15 Common Hearing Loss Mutants in a Chinese Newborn Cohort. Clin Genetics (PMID:24989646) (*****Co-first author)

2 Lu Y, Zhou X, Jin Z, **Cao Y**, Cheng J, Shen W, Ji F, Liu L, Zhang X, Zhang M, Han D, Choy KW, Yuan H. Resolving the Genetic Heterogeneity of Prelingual Hearing Loss within One Family: Performance Comparison and Application of Two Targeted Next Generation Sequencing Approaches. J Human Genetics (In press).

3 **Cao Y,** Li Z, Rosenfeld J, Patel A, Huang J, SunX, Leung T, Cheung S, Choy K. Identification of pathogenic CNV by arrayCGH in prenatal cases with oral clefts. American Society of Human Genetics (ASHG) 63rd Annual Meeting, Boston, Massachusetts, October 22-26, 2013. (Poster presentation)

4 **Cao Y** and Choy K. Targeted Massively Parallel Sequencing as a Molecular Test for Diagnosis of Hereditary Hearing Loss. Genomic Disorders 2013-from 60 years of DNA to human genomes in the clinic, The Wellcome Trust, Cambridge, UK, 10-12 April 2013 (Poster presentation)

List of Figures

List of Tables

Table of Contents

Section I: Introduction

Section II: Materials and methods

Section III: Results

Section I: Introduction

Chapter 1 Overview of hearing loss

1.1 Definition

Hearing loss or hearing impairment is a partial or complete inability to hear sound in one or both ears.

1.2 Epidemiology

The latest data released by World Health Organization (WHO) in 2012 showed that 360 million people, more than 5.3% of the world's population have disabling hearing loss. It consists of 328 million adults (91%) older than 15 years old with hearing loss > 40dB in the better ear (183 million males and 145 million females), and 32 million children (9%) with hearing loss >30dB in the better ear (http://www.who.int/mediacentre/factsheets/fs300/en/). The disabling hearing loss is most prevalent in south Asia, Asia Pacific, and sub-Saharan Africa (WHO 2012). The reported prevalence of permanent hearing loss was mainly identified by universal newborn screening program. In China, it is approximately 1-3 per 1000 with bilateral hearing loss, and 5 in 1000 unilaterally (WHO 2009). In the US, 1.86 per 1000 newborns is predicted to have hearing loss (>35dB), and the incidence of permanent hearing loss is increased to 2.7 per 1000 children by the age of five and to 3.5 per 1000 during adolescence (MORTON and NANCE 2006). It was estimated that approximately 10% of the whole population suffers from hearing loss during their life time (NADOL 1993). Up to one-third of people over 65 years of age are affected by disabling hearing loss. (WHO)

1.3 Clinical manifestation

1.3.1 Severity of hearing loss

The severity of hearing loss is measured in decibels (dB) on the basis of 500Hz, 1000 Hz, and 2000 Hz. Hearing threshold levels of severity are graded as following: (http://www.ncbi.nlm.nih.gov/books/NBK1434/)

Table 1-1 The severity of hearing loss in decibels

Hearing threshold in Decibels	Description
Up to 25dB	Normal
26-40	Mild
41-55	Moderate
56-70	Moderately severe
71-90	Severe
>90	Profound

1.3.2 Frequency of hearing loss

In humans, the hearing perception range is from 20 Hz to 20,000 Hz. It has highest sensitivity at 1000Hz. The frequency of hearing loss is classified into three levels, low (<500Hz), middle (501-2000Hz), and high (>2000Hz).

1.3.3 Age of onset

Age of onset is often described related to development of speech because auditory function is important for the acquisition of spoken language. Hearing impairment occuring prior to the development of speech is called prelingual, while after that is postlingual. Prelingual hearing loss is much more common than postlingual hearing loss. The period from birth to 5 years old, especially the first six months of life, is critical for normal speech and language development (SHEMESH 2010).

1.3.4 Temporal course of hearing loss

Hearing loss may develop progressively, becoming even worse from the very beginning, or present stably. Most recessively inherited hearing loss is stable. Progressive hearing loss is more often observed in with a dominantly inherited pattern. Fluctuation of hearing loss may be due to ear infection such as recurrent otitis media.

1.3.5 Type of hearing loss

Conductive hearing loss: It is associated with abnormalities of any anatomic components that transduce sound to the cochlea, which are located in the outer or middle ear (PRIEVE *et al.* 2013). It does not imply any particular etiology. It would be caused by otitis media, otosclerosis, skull fracture or other orofacial abnormality as part of a certain syndrome (PRIEVE *et al.* 2013).

Sensorineural hearing loss (SNHL): It relates to dysfunction of any components of the auditory pathway that convert sound vibration into electrical stimulus from the inner ear to the brain. Genetic mutation are an important cause for it, especially in children (MATSUNAGA 2009).

Mixed hearing loss: It is a combination of conductive hearing loss and sensorineural hearing loss. It could be caused by several pathologies, common causes such as otosclerosis or chronic otitis media and less seldom causes such as congenital malformations, trauma or temporal bone fracture (VERHAERT *et al.* 2013).

1.4 Etiology of hearing loss

Hearing loss can be caused either by genetic defects, environmental factors or a combination of both. Genetic causes are extremely heterogeneous. It is estimated that around 1% of human genes are involved in the hearing process (FRIEDMAN and

GRIFFITH 2003). Genetic causes account for at least 50% cases of prelingual hearing loss (SCHRIJVER 2004). Novel genetic causes for hearing loss are being discovered at a rapid rate with the advent of sequencing technologies (DELTENRE and VAN MALDERGEM 2013). Presbyacusis (age-related hearing loss) is the most common type of hearing loss in older people, considered as a multifactorial sensorineural hearing loss (PACALA and YUEH 2012). Common environmental factors include pre- or postnatal bacterial or viral infections such as rubella, cytomegalovirus (CMV), meningitis, prematurity, brain trauma, ototoxic drugs or noise exposure. Congenital CMV infection is the leading nongenetic cause of sensorineural hearing loss (LACKNER *et al.* 2009). Ototoxic drug induced hearing loss such as aminoglycoside ototoxicity always has genetic susceptibility (DELTENRE and VAN MALDERGEM 2013).

1.5 Hereditary hearing loss

Hereditary hearing loss results from genetic defects which are extremely heterogeneous while new genes that are implicated in hearing loss are being discovered increasingly. Most of the hereditary hearing loss cases are monogenic diseases. It accounts for 50-60% of prelingual hearing loss (MORTON and NANCE 2006). However, different mutant genes may lead to the same clinical expression while different mutations within a same gene can cause variable hearing loss. In clinical management, it is important to distinguish syndromic or nonsyndromic hearing loss.

1.5.1 Nonsyndromic hearing loss

Nonsyndromic hearing loss (NSHL) is associated with abnormalities of the inner ear without other organs or outer ear abnormalities. Seventy percentage of prelingual

hearing loss shows nonsyndromic (SHEARER and SMITH 2012). Approximately 77% of nonsyndromic hearing loss is autosomal recessive (AR) inheritance, 22% autosomal dominant(AD), and 1% X-linked and mitochondrial inheritance (2002). ARNSHL is typically prelingual while ADNSHL more often occurs postlingually. So far more than 180 loci or human genes have been mapped and cloned for nonsyndromic hearing loss (http://hereditaryhearingloss.org). In Table1-2 80 nonsyndromic hearing loss genes are identified according to the Hereditary Hearing Homepage website update on May 19, 2014. Despite the heterogeneity of NSHL, *GJB2* is responsible for 50% of nonsydromic hearing loss cases, the second most frequent cause is *SLC26A4,* which accounts for 15% (DELTENRE and VAN MALDERGEM 2013). Other relatively common mutant genes include *MYO15A, OTOF, CDH23, TMC1 and TMPRSS3*. (HILGERT *et al.* 2009a; DUMAN and TEKIN 2012).

The proteins encoded by nonsyndromic hearing genes are generally involved in the following functions of the inner ear: hair bundle morphogenesis, components of the extracellular matrix, cochlear ion homeostasis or transcription factors (HILGERT *et al.* 2009b). *ACTG1, DIAPH, ESPN, RDX, TRIOBP* and *CCDC50* encode proteins of cytoskeleton in the hair bundles which result in morphological transformation associated with actin assembly. *GJB2, GJB3, GJB6,* and *GJA1* encode the connexins, transmembrane proteins which form gap junctions coupling adjacent cells in the inner ear. *SLC26A4* and *KCNQ4* encode ion channels of anion and potassium. *POU4F3* and *POU3F4* are transcription factors which bind target DNA and contribute to the regulation of gene expression underlying the molecular mechanism of hearing. Patients with mutations in *POU3F4* may present with mixed or purely

NSHL, and stapes fixation and defective bony labyrinth of the inner ear would be observed (HILGERT *et al.* 2009b).

1.5.1.1 Nonsydromic hearing loss and the connexin gene family

Human connexin genes include a family of 23 connexin genes which encode gap junction proteins, which permit ions and small molecules space (less than 1kDa) passage between neighbouring cells (SCOTT and KELSELL 2011). Connexin genes are classified into five α, β, γ, δ and ε groups based on their structures and homologous sequence (http://www.genenames.org/genefamily/gj.php). Connexins (Cx) are widely expressed in the majority of human tissues. Gap junctions that consists of connexins are crucial for intercellular communication and homeostasis in organisms, allowing responses to external stimuli. Mutations in connexin genes are associated with a range of disorders. Mutations on *GJB2, GJB3, GJB6, GJA1* genes which are expressed in the inner ear, are associated with nonsyndromic hearing loss (http://davinci.crg.es/deafness/). In particular, *GJB2* mutations account for 50% of nonsyndromic hearing loss in many ethnic groups. *GJB2* c.35delG might account for 70% of all Caucasian *GJB2* mutations cases (Snoeckx et al. 2005). Specific mutations are more prevalent in certain ethnic groups and carrier frequencies of these mutations are relatively high, such as c.35delG in 2.55% of European ancestry (STORM *et al.* 1999), c.167delT in 4% of Ashkenazi Jewish population (MORELL *et al.* 1998), c.235delC in 1% of Japanese (ABE *et al.* 2000), and 1.3% in Chinese (LIU *et al.* 2002). Mutations in *GJB2* result in Cx26 loss of function, which is hypothesized to interrupt the recycling of $K^{+,}$ and hemostasis of cochlear endolymph. This would ultimatedly cause hair cell death and introduce the hearing loss phenotype.

Mutation in these genes can cause either recessive or dominant inheritance of monogenic hearing loss disorders. Interestingly, it was reported that digenic mutations in *GJB2* and *GJB6* also were responsible for DFNB1 (OMIM 220290) (LERER *et al.* 2001; DEL CASTILLO *et al.* 2002). In contrast to a variety of mutations detected in *GJB2*, only one dominant mutation is reported in the *GJB6* coding region (GRIFA *et al.* 1999). Three large deletions truncating the 5' end of *GJB6* were identified: del(*GJB6-D13S1830*) 309 kb, del(*GJB6-D13S1854*) 232 kb, and del(chr13:19,837,344–19,968,698) 131kb, in which there is a 95.4Kb common interval overlap (WILCH *et al.* 2010). Biopsies of patients with these deletions reveal reduced expression of both *GJB2* and *GJB6*, suggesting disruption of a *GJB2 cis*-regulatory element located within the *GJB6* deleted interval (COMMON *et al.* 2005; WILCH *et al.* 2006; RODRIGUEZ-PARIS and SCHRIJVER 2009). Completed *Gjb2* knock out mice are embryonic lethal. However, *Gjb6* knock out mice have normal hearing, with only half of Cx26 expression preserved. This might imply that *GJB6* only was dispensable in cochlear functions (BOULAY *et al.* 2013) .

1.5.1.2 Nonsydromic hearing loss and the myosin family

Myosins are a group of motor proteins that bind actin and hydrolyze adenosine triphosphate (ATP) to produce force and movement (FRIEDMAN *et al.* 1999). They regulate a variety of cellular functions: rearrangement of the actin cytoskeleton, tension of actin filaments and transport of organelles (MERMALL *et al.* 1998). This protein family was ubiquitously expressed as multiple isoforms in all eukaryotic cells. To date, seven of the big myosin genes families, *MYO1A, MYO3A, MYO6, MYO7A, MYO15A, MYH9* and *MYH14,* have been implicated in nonsyndromic hearing loss DFNA48, DFNB30, DFNA22/DFNB37, DFNB2/DFNA11, DFNB3, DFNA17, and

DFNA4 respectively. These myosins were mainly located in the hair bundles of stereocilia. Actin-rich stereocilia on outer and inner hair cells open or close ion channels and sequentially depolarize or hyperpolare the hair cell membrane. This is where the sound waves are transferred into electrical signals which are could be are processed in the brain (MERMALL *et al.* 1998). Besides causing nonsyndromic hearing loss, *MYO7A* is also associated with Usher syndrome Type 1B (OMIM, 276903), which presents with hearing loss, vestibular areflexia and adolescent onset retinitis pigmentosa.

1.5.2 Syndromic hearing loss

Syndromic hearing loss may be associated with malformations of the external ear or other organs systems. Up to 30% of prelingual hearing loss cases are syndromic. Over 400 syndromes with hearing loss have been described (Toriello et al 2004). However, the hearing loss is not a constant phenotype in every patient. The severity of hearing loss can be moderate to profound and unilateral or bilateral. SNHL is a major feature in several well-characterized syndromes, such as Pendred syndrome, Waardenburg syndrome, Usher syndrome, Jervell and Lange-Nielsen syndrome. Usher syndrome is the most frequent recessive syndromic hearing loss, characterized with progressive vision loss caused by retinitis pigmentosa (AVRAHAM 2013). Three major types are clinically classified based on severity and age of onset. Eleven genes (*MYO7A, USH1C, CDH23, PCDH15, SANS, CIB2, USH2A, VLGR1, WHRN, CLRN1, PDZD7*) have been identified and are related to different molecular types of this syndrome. Pendred syndrome is always associated with a thyroid disorder called goiter, which usually presents between late childhood and early adulthood. Patients with Pendred syndrome may also present with enlarged vestibular aqueduct which

might not be the cause of the hearing loss. Different mutations in the same gene can cause either syndromic or nonsyndromic hearing loss with different inheritance patterns. Six common syndromic hearing loss are reviewed in Table 1-3.

1.5.3 Noncoding RNA and hearing loss

MicroRNA is recognized to have a role in many biological processes in mammals. Lewis et al. and Mencia et al. first reported that a mutation in microRNA-96 (miR-96) was responsible for hair cell defects and progressive hearing loss in humans and the mouse(LEWIS *et al.* 2009; MENCIA *et al.* 2009). It was also the first report showing the dysfunctional microRNA causing a Mendelian disorder (METZLER 2009). The three gene set of miR-96/182/183 is transcribed as a polycistronic unit in ciliated neurosensory organs. This miR-96/182/183 cluster is expressed in a relative stable level in mouse inner ear which persists into adulthood, suggesting that these microRNAs were evolutionarily associated with mechanosensory cell development and function (WESTON *et al.* 2006). Therefore they were likely to mediate the hearing loss in mouse and human. Li et al. showed the miR-183 family is required for normal numbers of hair cells and neurons in the zebrafish inner ear and might play a role in cell fate determination (LI *et al.* 2010).

1.5.4 Mitochondria and hearing Loss

Mitochondrial DNA in humans is 16569 bps in length and double stranded. It encodes 13 essential polypeptides of oxidative phosphorylation as well as the two rRNAs and 22 tRNAs required for mitochondrial protein synthesis (MITOMAP, 2013). Mitochondria (mt) play an essential role in the process of cellular energy generation. Deleterious mtDNA mutations adversely affect cell function by reduction of ATP and consequently causes a wide spectrum of disease (GUAN 2004).

Mutations of mtDNA are one of the common causes of sensorineural hearing loss. The incidence of mitochondrial defects causing hearing impairment is estimated to be around 6-33% of all hearing deficiencies, with an even higher percentage for some syndromic cases (GUAN 2004). MtDNA is inherited exclusively from the mother but causes diseases equally in both sexes. The most commonly reported mutations including A1555G, C1494T in the 12S rRNA gene, A7445G, 7472insC, T7510C, T7511C in the tRNA-Ser (UCN) and A3243G in tRNA-Leu (ZHAO *et al.* 2004) are associated with nonsyndromic or aminoglycoside-induced hearing loss. However, the phenotypic manifestation may be different even within the same mtDNA mutation or within family members, demonstrating a wide range of severity and age at onset. A1555G, A7445G with different mitochondrial haplotype were also reported with different severity of hearing loss (Lu et al. 2010) (MAASZ *et al.* 2008). Some studies strongly indicate that mtDNA mutations such as A1555G are a crucial factor underlying development of hearing impairment but in itself insufficient to cause the phenotype. These mutants might interact with other modifier factors such as aminoglycosides or nuclear-modifier genes to modulate phenotypic manifestations of patients (GUAN *et al.* 2006). Besides mutations in mtDNA, it is also reported that mitochondrial DNA depletion which was related with defects in mtDNA maintenance caused by mutations in nuclear genes such as thymidine kinase 2 gene (*TK2*), are also involved with sensorineural auditory dysfunction (GUAN *et al.* 2006; EL-HATTAB and SCAGLIA 2013).

1.5.5 Pseudogenes and hearing loss

Pseudogenes are heritable but non-functional DNA genetic elements due to their presumed inability to produce functional protein or RNA (HARRISON *et al.* 2005).

They are prevalent in the human genome and extremely similar in sequence to wild-type cognate genes, which can introduce artifacts in molecular experiments targeted to the cognate genes (ZHENG *et al.* 2007). In contrast with the previous presumption of their non-function, it has been shown that pseudogenes are expressed in a surprisingly genome wide pattern both in the cancer or normal tissues (KALYANA-SUNDARAM *et al.* 2012). Pseudogenes can regulate coding genes by acting as short interfering RNAs (PINK *et al.* 2011). These findings shed light on redefining the functional properties of pseudogenes in the human genome, and also their role in disease etiology and development. Hong et al. reported that two novel missense mutations in the *GJA1* pseudogene result in a loss of function of connexin 43, and might contribute to hearing loss (HONG *et al.* 2010). With the wide application of next generation sequencing and comprehensive study of transcriptomics, functional pseudogenes should be taken into consideration as genetic causes of hearing loss.

1.5.6 Aneuploidy and hearing loss

Patients with aneuploidy have a wide spectrum of phenotype, including dysmorphysm, or developmental delay. Hearing loss is a common phenotype in the aneuploidy syndromes. Since multiple systems are affected, it is difficult to understand exactly how an extra chromosome gain or loss could impact hearing. A Down Syndrome child with trisomy 21 tends to have smaller external auditory canals, higher rates of chronic ear infection and effusion , and can represent with conductive, sensorineural or mixed hearing loss (SHOTT 2006; CHIN *et al.* 2014). In Turner syndrome patients, the type and the severity of hearing loss are related to the karyotype, age and also other metabolic conditions of the patient. Morimoto et al. summarized the prevalence of hearing loss in Turner syndrome patients based on the

previous study, in which 50% of patients with Turner syndrome (TS) have mild hearing loss: 60% sensorineural hearing loss (SNHL), 25% conductive hearing loss (CHL), and 15% mixed hearing loss (MORIMOTO *et al.* 2006). In this review, more than 50% of patients were high-frequency (8 kHz) sensorineural hearing loss (HFQ-SNHL). Hearing threshold increasing in high frequency was correlated with age and karyotyping but not with the middle ear infection (MORIMOTO *et al.* 2006). Oliveira et al. reported that patients with a full 45,X karyotype and isochromosomes with different degrees of deletion in the p-arm of the X chromosome had a greater risk of developing hearing loss than patients with 45,X mosaicism (OLIVEIRA *et al.* 2013). Otological problems also increased with decreases in body weight and serum levels of Insulin-like Growth Factor-1 (GRAVHOLT 2004). The absence of estrogen in the TS also plays a role in the hearing impairment since estrogen was considered to be hearing protective (HULTCRANTZ *et al.* 2006). However, the cause for sensorineural hearing loss is not well understood.

1.6 Establishing the diagnosis of hereditary hearing loss

A deaf patient with normal physical examination, negative medical history of aminoglycoside exposure or trauma, excluding infections such as toxoplasmosis, rubella, CMV, herpes, meningitis, with or without a family history of hearing loss, is likely to have a genetic cause. Low birth weight, hyperbilirubinemia or other complicated pregnancy should also be taken into consideration in an infant hearing loss (JCIH, 2000). Audiometric information and temporal bone imaging should be reviewed or referred to specialists. Genetic testing should be offered to screen hearing loss genes.

1.6.1 Audiologic evalutation

Audiologic evaluation is made up of a series of tests. It assesses the severity and which part of the auditory system might be damaged, and establishes the clinical diagnosis of hearing loss. There are several tests commonly used in clinics Otoacoustic emissions testing (OAE): otoacoustic emissions are sounds originating from outer hair cells that can be measured in the external auditory canal. It primarily reflects the activity of the outer hair cells and presents in ears with hearing sensitivity greater than 40-50 dB (http://www.ncbi.nlm.nih.gov/books/NBK1434/). It might miss a mild severity of hearing loss <40dB. Auditory brain stem response testing (ABR): this test acquires information about the inner ear, auditory nerve and brain stem by using a stimulus to evoke electrophysiologic responses. Pure-tone audiometry is used to determine the lowest intensity at which an individual can hear a pure tone at different frequencies (typically 250, 500, 1000, 2000, 4000, and 8000 Hz). Pure tone auditory can be used to classify hearing loss such as conductive and sensorineural or mixed. It can provide information as to the type, degree and pattern of hearing loss of each ear. There are three main patterns: sloping, rising, and flat. Some autosomal dominant hearing loss genes such as *DIAPH1* and *WFS1* have characteristic audioprofiles showing low frequency sensorineural hearing loss. Based on the special audioprofile of some nonsyndromic hearing loss, Audiogene, a machine-based gene prediction tool is developed to predict the genotype-phenotype correlation (HILDEBRAND *et al.* 2009; TAYLOR *et al.* 2013).

1.6.2 Genetic testing

Based on physical testing of the auditory system, it is not difficult to make a clinical diagnosis of hearing loss. However, the genetic causes are extremely heterogeneous including a variety of genomic changes from aneuploidy, submicroscopic copy number changes, to single nucleotide variants.

A variety of techniques can be used in clinics to detect copy number changes. A standard 450-500 G-banded karyotype detects structural variants larger than 5Mb across the whole genome on an individual cell basis. However, it needs an actively growing source of cells to collect metaphase chromosomes. Fluorescence in situ hybridization (FISH) applies to cells at any stage of the cell cycle and also to archived fixed tissues. The sizes of probes for clinical use mostly range from 100Kb to 500Kb. PCR based method such as multiplex ligation-dependent probe amplification (MLPA) (EIJK-VAN OS and SCHOUTEN 2011) assesses copy number changes at multiple loci in a single reaction. MLPA reaction generates fragments ranging between 100 and 500 bps. Nevertheless, MLPA and FISH only detect those regions complementary with the targeted regions. The implementation of array comparative genomic hybridization (aCGH) drastically changed the efficiency of molecular cytogenetic diagnosis to scan quickly through the entire genome for imbalances with a resolution as precise as 1000 bps (Theisen, A., 2008). Through hybridizing the mixture of different fluorescence-labeled reference genome and sample genome with millions of probes on a microarray, aCGH compares the intensity of the two fluorescences to differentiate the gain or loss of genomic segments. Single nucleotide polymorphic (SNP) microarray is another widely used whole genome scanning array-based technique (LAFRAMBOISE 2009). It directly

quantifies the samples SNPs hybridized to probes to investigate copy number changes. Meanwhile, it also provides genotypes of sample DNAs which can identify the stretches of homozygosity inferring uniparental disomy. Computer based analysis tools make variants calling of microarrays more unified and comparable among platforms. Nonetheless it can still miss balanced translocations even through the break points might interrupt the gene sequences.

Most hereditary hearing loss is a monogenic disorder. However, single gene screening by the Sanger approach was not sustainable. Single gene testing could primarily be provided to those patients with a recognizable phenotype, such as Waardenburg syndrome (abnormal pigmentation of iris, hair, and skin), Alport syndrome (hematuria, postlingual hearing loss), or with a characteristic audioprofile such as *WFS1*. Meanwhile, based on the well-studied global epidemiologic data, molecular testing for *GJB2/GJB6* should be prioritized. Besides singe gene testing, hot spot mutations in several prevalent genes or multiple panel testing for genes implicated in hereditary hearing loss would is also an alternative choice to yield a diagnosis.

1.6.3 Prenatal genetic testing

For those normal hearing parents with affected children, Brunger JW et al showed 96% of respondents had positive attitudes of genetic testing for hearing loss. (ARNOS *et al.* 1991). For prenatal diagnosis of hereditary hearing loss, some studies showed 21% of congenital deaf, 39 % of hard hearing, 49 % of normal hearing participants with affected family members would consider prenatal diagnosis, but most of them would not consider termination of pregnancy if the fetus was genetically diagnosed with hearing loss (MIDDLETON *et al.* 2001).

Before any prenatal testing, the genetic diagnosis of the hearing loss under investigations should be rigorous, definitive and definable. The purpose is to identify or exclude the known mutation. Two strategies are possible: indirect and direct genetic testing (WIEACKER and STEINHARD 2010). Direct testing assesses the exact mutation detected in the index patients. Indirect testing involves detecting a fetus inherited with the high risk haplotype harboring the mutation. The standard prenatal genetic testing currently is done invasively. Invasive procedures such as amniocentesis or chorionic villus sampling (CVS) are used to obtain fetal DNA for testing.

In contrast with the risk of fetal loss due to an invasive procedure, noninvasive prenatal testing shows great potential. Noninvasive procedure for fetal samples in prenatal testing has been under study for a long way. Early in 1969, Walknowska et al. found male fetal cells in maternal blood by karyotyping analysis (WALKNOWSKA *et al.* 1969). Isolation of fetal nucleated cells from maternal peripheral blood as an alternative sample source for noninvasive prenatal diagnosis was extensively studied (LO *et al.* 1989; SIMPSON and ELIAS 1993; CHEUNG *et al.* 1996). However, low numbers of fetal cells in maternal blood and the complicated manipulation of fetal cell isolation and enrichment hampered its application in practical use. Following the discovery that circulating tumor specific DNA fragment was found in plasma and serum of cancer patients, Lo et al. reported the presence of cell free fetal DNA in maternal plasma and serum (LO *et al.* 1997). Lo et al. detected unexpected high fetal DNA concentrations in maternal plasma using real time PCR, which contributed to 3.4% and 6.2% of total plasma DNA at 11–17 gestational weeks and 37-43 gestational weeks prior to delivery respectively (LO *et al.* 1998). Fetal DNA could be

detected as early as 7 gestational weeks and the concentration increased as pregnancy progressed (LO *et al.* 1998). With implementation of massively parallel sequencing, the entire fetal genome was approached with specific profiling in the maternal plasma (CHIU *et al.* 2008; FAN *et al.* 2008; FAN *et al.* 2012), which in theory made all types of fetal genetic testing feasible noninvasively. To date noninvasive prenatal testing (NIPT) has made tremendous progress in the clinic as a screening test for pregnant women at high risk for bearing a fetus with aneuploidy, especially in the US and China. Besides detecting numerical chromosome changes, NIPT is also reported to detect structural abnormalities such as microdeletions of the fetal genome (PETERS *et al.* 2011; JENSEN *et al.* 2012). For single gene disorders, considering the high background of the maternal genome, directly sequencing plasma DNA was useful in detection of fetal sex, RhD blood type and also other gene mutations paternally inherited or *de novo*.

1.6.4 Genetic counseling

Genetic counseling is defined as a communication process that deals with the human problems associated with the occurrence, or the risk of occurrence of a genetic disorder in a family (1975). It provides helps of comprehending the diagnosis, the probable course and the available management of the disorder; in appreciating the inheritance pattern and recurrence risk in the family; in understanding the alternative choices dealing with the recurrence risk and in making the best management decisions for the disease for both the affected individuals and other family members. This process should be offered by trained specialists. The genetic counselor makes sure that patients and family members understand the information being conveyed so they can make informed decisions (ARNOS *et al.* 1991). With discovery of causative

deafness genes and advances in molecular technology, genetic testing becomes more routine for hearing loss families. Appropriate genetic counseling gives patients useful information to understand the benefits, risks and limitations of the genetic testing (BRUNGER *et al.* 2000). Pre-test genetic counseling also enhances deaf and hard-of-hearing individuals' knowledge of genetics (BALDWIN *et al.* 2012).

In the genetic counseling, patients or parents of hearing loss children should fully understand the information of diagnosis, inheritance pattern, prognosis and treatment options. Recurrence risk should be estimated according to the inheritance pattern of the hereditary hearing loss, and consanguineous marriage. For hearing parents whose first child is deaf due to an etiology that cannot be determined after genetic evaluation, an empiric risk estimate of 9% is typically given for subsequent offspring (Bronya J.B. 2002).

1.7 Management

1.7.1 Universal newborn hearing screening

Universal newborn hearing screening programmes (UNHS) have been applied globally. The purpose of UNHS is to decrease the age of identification and intervention for infants with hearing loss (SININGER *et al.* 2009). It focuses on detecting moderate-to-severe bilateral congenital hearing loss. Acquired or progressive hearing loss might not be detected at birth. Although it has limitations, it could screen hearing loss effectively at birth and provide a foundation of identification of hearing loss. Before implementation of UNHS, the average age of diagnosis and intervention exceeded two years of age, which is well beyond the beginning of the critical interval for speech and language acquisition. With the advent of newborn screening, the average age at which hearing loss is confirmed has

dropped from 24 to 30 months to 2 to 3 months, which facilitates early intervention. Children with hearing loss who had newborn hearing screening also had better language outcomes than those not screened (NELSON *et al.* 2008).

The detection threshold of hearing loss ranges from 30 dB to 40 dB and is performed bilaterally. Screening methods used in most countries are transient-evoked otoacoustic emission emissions (TEOAE) testing and automated auditory brainstem response (AABR). In China, the government has recommended newborn hearing screening as a routine procedure since 1999. In 2009, the ministry of Health of China augmented UNHS, and recommended two stage OAE plus AABR as screening protocols (TOBE *et al.* 2013). In the first stage neonates are screened by OAE within two to seven days after birth before discharge from the hospital. Those who failed first stage OAE are referred to a second stage rescreening test by OAE or OAE plus AABR within 42 days of birth. Diagnosis of hearing loss for infants who fail two stage screening should be made by an ENT specialist before six months of age. The reported prevalence of permanent hearing loss identified by the newborn hearing screening program in China is: 1–3/1000 (bilateral) and ~5/1000 (unilateral) (OLUSANYA 2011) (HARRISON *et al.* 2003). Another study also showed up to 5.4 per 1000 neonates with hearing loss in rural areas of China (CHEN *et al.* 2012a). Genetic newborn screening can work complementary to UNHS (PRERA *et al.* 2014). 85% of hearing and 62% of deaf or hard-of-hearing individuals would allow genetic testing for deafness in their own babies (MARTINEZ *et al.* 2003).

1.7.2 The hearing aid and cochlear implants

The hearing aid is used for improving mild to moderate hearing loss. Criteria for cochlea implantation (CI) include bilateral severe to profound sensorineural hearing

loss and little benefit from appropriately fitted hearing aids. CI has achieved remarkable success for those individuals with severe to profound bilateral cochlear hearing loss. Cochlear implantation can be considered in children over age 12 months with severe-to-profound hearing loss. Cochlear implants bypass the hair cells and directly stimulate the spiral ganglion for sound transmission. It is also wildly employed in genetic hearing loss patients. Robert W. Eppsteiner *et al.* did a systematic review on the relationship between the genetic mutations and CI performance. This results supported the hypothesis that patients would have better CI performance if they had found mutations on genes expressed in the membranous labyrinth (*GJB2, LOXHD1, OTOF, CDH23, MYO7A, KCNQ1, TMC1, COCH, SLC26A54*), compared with those also expressed on the spiral ganglion (*CHD7,* or *DDP1*) (EPPSTEINER *et al.* 2012) . Patients with *GJB2* mutations, one of the largest genetic defect groups benefit from cochlear implants. CI can be a reasonable habilitation for patients with the mtDNA mutation A1555G since such patients have been shown to exhibit remarkable improvement in the test of auditory performance (TONO *et al.* 2001; ULUBIL *et al.* 2006). For other syndromes with hearing loss, the less severe the dysplasia of inner ear, the better the outcome. Understanding patients' genetic causes of hearing loss could provide a unique CI strategy and performance metrics for each individual. It would also help reduce the likelihood of ineffective cochlear implantation and decrease healthcare costs.

1.7.3 Gene therapy and stem cell therapy for the inner ear

Fifty percent of prelingual hearing loss have a genetic cause. Identification of the causative gene of hearing loss and pathogenic mechanism of these mutant genes make medical management of hereditary hearing loss include consideration of gene

based treatment. Hearing loss genes could be potential therapeutic targets. Applications of gene therapy in the inner ear are currently under investigation. Lentz et al. reported rescue of low and middle frequency hearing loss in the mouse model of USHIC by antisense oligonucleotides correcting the splicing defect of *USH1C* (LENTZ *et al.* 2013). Most affected individuals with hearing loss suffer from a sensory defect of inner ear. In addition, stem cell technology opens new approaches for hair cell regeneration (OKANO and KELLEY 2012). There is great anticipation regarding the use of these new therapies but it seems to be yet a long way to becoming a reality for humans (LIMB 2012).

1.8 Genetic susceptibility in other types of hearing loss

1.8.1 Noise induced hearing loss

Noise induced hearing loss usually is observed in certain occupations which are exposed to a continuous and regular noise. It is also one of the most frequent environmental factors of hearing loss. A typical pure-tone audiogram of NIHL shows a notch around 3 to 6 kHz (KONINGS *et al.* 2009). Depending on the frequency of the noise, different parts of the cochlea will be damaged. It is a complicated hearing loss caused by a combination of environmental and genetic factors. In the animal studies, the C57BL/6J mouse is more susceptible to noise exposure than other strains (ERWAY *et al.* 1996). Some knockout mice studies showed that deficits in the genes which involve response to oxidative stress, potassium cycling in the inner ear or that encode proteins as a component of the cochlear would increase the susceptibility of ear to acoustic overstimulation (ERWAY *et al.* 1996). Carlsson et al. showed there was no significant difference between the noise susceptible and noise resistant groups in terms of *GJB2* c.35delG carrier in a

Swedish population. In humans *KCNQ4* and *KCNE1, PCDH15, MYH14* and heat shock protein (HSP70) are susceptibility genes for noise induced hearing loss susceptibility genes (SLIWINSKA-KOWALSKA and PAWELCZYK 2013).

1.8.2 Age related hearing impairment

Age related hearing impairement (ARHI), or presbycusis is a complex and prevalent disease due to interactions between genetic predisposition and environmental factors such as noise exposure and ototoxic medications. ARHI shows pathological deficits in hair cells, cochlear neurons, and stria vascularis. Twin studies indicate that genetic factors play an important role in hearing loss in the later middle age or older than 70 years old (WINGFIELD *et al.* 2007) (CHRISTENSEN *et al.* 2001). Hearing acuity tends to decline faster and earlier in males than females (WILEY *et al.* 2008), and estrogen tends to preserve auditory function (KILICDAG *et al.* 2004). A genome wide association study of a large cohort of European older adults lead to one haplotype in *GRM7* (glutamate metabotrophic receptor 7) associated with ARHI (FRIEDMAN *et al.* 2009). Another study elucidated the genetic basis of ARHI in which identified only one gene, *GRHL2*, out of 70 candidate genes, had a significant p value. In the Caucasian *GJB2* c.35delG with 2% carrier frequency, no increase in susceptibility for the development of age related hearing loss was found (VAN EYKEN *et al.* 2007).

1.8.3 Otosclerosis

Otosclerosis is a disorder of the bony labyrinth and stapes resulting in progressive conductive and sensorineural hearing loss (CUREOGLU *et al.* 2006). 75% of patients show bilateral hearing impairment. Age of onset is mainly between 15-40 years with more prevalence in females than males. It shows the abnormal growth of stapes footplate preventing its free movement. The etiology of otosclerosis is still not clear.

It is common in Caucasian women within reproductive age, which suggests a hormonal component. It also appears to aggregate in families, suggesting a genetic cause of this disease. Various factors including genetic causes or infection could contribute to the diseases (EALY and SMITH 2011). Generally it is considered as an autosomal dominant disease with reduced penetrance. To date eight loci have been identified associated with otosclerosis: OTSC1 (chromosome 15q26.1-qter), OTSC2 (chromosome 7q34-36), OTSC3 (chromosome6p21.3-22.3), OTSC4 (chromosome 16q21-23.2), OTSC5 (chromosome 3q22-24), OTSC7 (chromosome 6q13-16.1), OTSC8 (chromosome 9p13.1-9q21.11), and OTSC10 (chromosome 1q41-44) (http://hereditaryhearingloss.org). However, no otosclerosis pathogenic mutations have been identified in genes within these loci. Association studies on the genetic etiology of otosclerosis with a low risk of bias showed *COL1A1* associated with otosclerosis (El Gezeery. 2012) and no linkage with *NOG* gene in Japanese patients (USAMI *et al.* 2012). Genes involved in the TGF-beta1 pathway are also likely an important factors in the pathogenesis of otosclerosis (THYS and VAN CAMP 2009). Measles virus RNA detected in patients with otosclerosis support a role for measles in the pathogenesis of some cases (KAROSI *et al.* 2004; KAROSI *et al.* 2007). Nevertheless, no etiology can be identified in most sporadic cases could not identify the etiology. Sporadic otosclerosis may be caused by multiple genetic variants as a synergistic effect or by the combined action of genetic and environmental factors.

Table 1-2 Nonsyndromic hearing loss genes

Causative Gene	OMIM No.	Locus (OMIM)	Cytogenetic Location	Genomic Coordinaties(hg19): Chr : start-end	Protein, Other notes
ACTG1	102560	DFNA20/26	17q25.3	chr17:79476996-79479891	actin, gamma 1, cytoplasmic actin found in non-muscle cells; Baraitser-Winter syndrome 2
ADCY1		DFNB44	7p12.3	chr7:45613739-45762714	adenylate cyclase 1, responsible for the synthesis of cAMP
BSND	606412	DFNB73	1p32.3	chr1:55464617-55474465	barttin, an essential beta subunit for the chloride channels; Bartter syndrome type 4a
CABP2	607314	DFNB93	11q13.2	chr11:67286418-67290899	calcium binding protein 2
CCDC50	611051	DFNA44	3q28	chr3:191046873-191116458	coiled-coil domain containing 50, a negative regulator of NF-kB signaling and an effector of epidermal growth factor (EGF)-mediated cell signaling
CDH23	605516	DFNB12	10q22.1	chr10:73156691-73575704	cadherin-23; Usher syndrome, type 1D
CEACAM16	614591	DFNA4B	19q13.32	chr19:45202420-45213985	carcinoembryonic antigen-related cell adhesion molecule 16, a adhesion protein that belongs to immunoglobulin (Ig)-related glycoproteins

CIB2	605564	DFNB48	15q25.1	chr15:78396948-78423877	calcium and integrin binding family member 2; Usher syndrome, type IJ
CLDN14	605608	DFNB29	21q22.13	chr21:37832920-37948867	claudin 14, a component for tight junctions
CLIC5	607293	DFNB102	6p21.1	chr6:45866187-46048084	chloride intracellular channel 5
COCH	603196	DFNA9	14q12	chr14:31343740-31364284	cochlin, knockout mouse has normal hearing
COL11A2	120290	DFNB53/DFNA13	6p21.32	chr6:33130469-33160245	collagen XI alpha 2; otospondylomegaepiphyseal dysplasia; Stickler syndrome
COL4A6	303631	DFNX?	Xq22.3	chrX:107398837-107681660	collagen type IV alpha-6 isoform, a component of the basement membrane
DFNA5	608798	DFNA5	7p15.3	chr7:24737973-24797638	DFNA5 protein ,unknown function knockout mouse has normal hearing
DFNB59/ PJVK	610219	DFNB59	2q31.2	chr2:179316163-179326110	Pejvakin, with auditory neuropathy, DFNB59 within the DFNB27 locus
DIAPH1	602121	DFNA1	5q31.3	chr5:140894587-140998621	drosophila diaphanous gene homologue 1 protein; regulation of actin polymerization in hair cells of the inner ear
DIAPH3	614567	AUNA1	13q14-21	chr13:60239723-60738119	drosophila diaphanous gene homologue 3 protein, Auditory neuropathy, actin polymerization

DSPP	605594	DFNA39	4q22.1	chr4:88529680-88538024	dentin sialophosphoprotein; with dentinogenesis
ELMOD3	615427	DFNB88	2p11.2	chr2:85581843-85618875	ELMO domain-containing protein 3; GTPase-activating protein
EPS8	600206	DFNB?	12p12.3	chr12:15773074-15942509	epidermal growth factor receptor pathway substrate 8 protein; MYO15A-whirlin-EPS8 complex essential for stereocilia elongation in mice studies
ESPN	606351	DFNB36	1p36.31	chr1:6484848-6521004	espin, actin-bundling protein; hearing loss with vestibular areflexia
ESRRB	602167	DFNB35	24q24.3	chr14:76837690-76968180	estrogen-related receptor beta
EYA4	603550	DFNA10	6q23.2	chr6:133562495-133853258	eyes absent homolog 4; cardiomyopathy, dilated 1J (CMD1J) [MIM:605362]
GIPC3	608792	DFNB15/72/95	19q13.3	chr19:3585569-3593539	PDZ domain-containing protein, Required for postnatal maturation of the hair bundle and long-term survival of hair cells and spiral ganglion
GJB2	121011	DFNB1A/DFNA3A	13q12.11	chr13:20761601-20768604	connexin 26 , gap junction protein, beta2, form gap junctions, 30-50% nonsyndromic hearing loss; keratitis-ichthyosis-deafness syndrome; Bart-Pumphrey syndrome; Vohwinkel syndrome; hystrix-

					like ichthyosis with deafness; palmoplantar keratoderma with deafness
GJB3	603324	DFNA2B	1p34.3	chr1:35246789-35251966	connexin 31, gap junction beta3; digenic GJB2/GJB3; Erythrokeratodermia variabilis (EKV) [MIM:133200]
GJB6	604418	DFNA1B/3B	13q12.11	chr13:20796100-20806533	connexin 30, gap junction beta 6, digenic GJB2/GJB6; Ectodermal dysplasia 2, Clouston type (ECTD2) [MIM:129500]
GPSM2	609245	DFNB32/82	1p13.3	chr1:109419603-109473044	G-protein-signaling modulator 2 ; modulate activation of G proteins; Chudley-McCullough syndrome (MIM604213)
GRHL2	608576	DFNA28	8q22.3	chr8:102504667-102681953	grainyhead-like protein 2 homolog
GRXCR1	613283	DFNB25	4p13	chr4:42895234-43032602	glutaredoxin domain-containing cysteine-rich protein 1 ; actin organization in hair cells
GRXCR2	615762	DFNB101	5q32	chr5:145239296-145252531	glutaredoxin domain-containing cysteine-rich protein 2; maintaining cochlear stereocilia bundles
HGF	142409	DFNB39	7q21.11	chr7:81331444-81399452	hepatocyte growth factor precursor

ILDR1	609739	DFNB42	3q13.33	chr3:121706170-121741127	Immunoglobulin-like domain-containing receptor 1 precursor
KARS	601421	DFNB89	16q23.1	chr16:75661622-75681585	lysyl-tRNA synthetase, catalyzes aminoacylation of tRNA-lys in the cytoplasm and mitochondria; Charcot-Marie-Tooth Disease, Recessive Intermediate B (CMTRIB)
KCNQ4	603537	DFNA2A	1p34.2	chr1:41249683-41306123	potassium voltage-gated channel subfamily KQT member 4; recycle potassium ions
LHFPL5/TMHS	609427	DFNB66/67	6p21.31	chr6:35773071-35791852	LHFP-like protein 5, Tetraspan membrane protein of hair cell stereocilia
LOXHD1	613072	DFNB77	18q21.1	chr18:44057217-44236996	lipoxygenase homology domain-containing protein 1; located in stereocilium
LRTOMT/COMT2	612414	DFNB63	11q13.4	chr11:71791377-71807938	leucine-rich repeat-containing protein 51; Transmembrane O-methyltransferase 2
MARVELD2/TRIC	610572	DFNB49	5q13.2	chr5:68710939-68737890	MARVEL domain-containing protein 2 , or tricellulin; a component of tight junctions
MIR96	611606	DFNA50	7q32.2	chr7:129414531-129414608	microRNA 96, short noncoding RNA
MSRB3	613719	DFNB74	13q14.3	chr12:65672423-65860687	methionine-R-sulfoxide reductase B3;

					precursor catalyzes the reduction of oxidized methionine residues and thus repairs oxidatively damaged proteins, might targeted to mitochondria
MYH14	608568	DFNA4A	19q13.33	chr19:50706884-50813800	myosin 14, conventional non-muscle; peripheral neuropathy, myopathy, hoarseness, and hearing loss (PNMHH)
MYH9	160775	DFNA17	22q12.3	chr22:36677322-36784106	myosin-9; macrothrombocytopenia and progressive sensorineural deafness
MYO15A	602666	DFNB3	17p11.2	chr17:18012020-18083116	myosin XVA, unconventional
MYO1A	601478	DFNA48	12q13.3	chr12:57422300-57444548	myosin IA, unconventional
MYO3A	606808	DFNB30	10p12.1	chr10:26223002-26501465	myosin-IIIa
MYO6	600970	DFNB37/DFNA22	6q14.1	chr6:76458892-76629253	unconventional myosin-VI; autosomal dominant hearing loss with hypertrophic cardiomyopathy
MYO7A	276903	DFNB2/DFNA11	11q13.5	chr11:76839310-76926286	myosin VIIA , Usher syndrome, type 1B
OTOA	607038	DFNB22	16p12.2	chr16:21689835-21772050	otoancorin, specifically expressed in the inner ear

OTOF	603681	DFNB9	2p23.3	chr2:26680071-26781566	otoferlin, associated with auditory neuropathy (AN)
OTOG	604487	DFNB18B	11p15.1	chr11:17568919-17667490	otogelin, associated with vestibular dysfunction
P2RX2	600844	DFNA41	12q24.33	chr12:133195403-133198972	P2X purinoceptor 2 , form a ligand-gated ion channel gated by extracellular ATP
PCDH15	605514	DFNB23	10q21.2	chr10:55562533-56561051	protocadherin 15, Usher syndrome type 1F; digenic Usher Syndrome Type ID/F
POU3F4	300039	DFNX2/DFN 3	Xq21.1	chrX:82763269-82764775	POU domain, class 3, transcription factor 4 ; Mixed with conductive hearing loss resulting from stapes fixation
POU4F3	602460	DFNA15	5q32	chr5:145718586-145720082	POU domain, class 4, transcription factor 3
PRPS1	311850	DFNX1/DFN 2	Xq22	chrX:106871654-106894256	phosphoribosyl pyrophosphate synthetase 1, Pseudogen, PRPS1L2; Charcot-Marie-Tooth disease X-linked 5 (CMTX5)
PTPRQ	603317	DFNB84	12q21.31	chr12:80838126-81073968	phosphatidylinositol phosphatase
RDX	179410	DFNB24	11q22.3	chr11:110045605-110167437	radixin, pseudogenes *RDXP1, RDXP2*
SERPINB6	173321	DFNB91	6p25.2	chr6:2948393-2972399	serpin B6, a serine proteinase inhibitor

SIX1	601205	DFNA23	14q23.1	chr14:61111417-61116155	homeobox protein SIX1 , transcription factors; branchiootorenal (BOR) syndrome
SLC17A8	607557	DFNA25	12q23.1	chr12:100750856-100815836	vesicular glutamate transporter 3
SLC26A4	605646	DFNB4	7q22.3	chr7:107301080-107358252	pendrin; Pendred syndrome; *SLC26A4*/KCNJ10, SLC26A4/FOXI1 digenic, with enlarged vestibular Aqueduct
SLC26A5	604943	DFNB61	7q22.1	chr7:102993177-103086624	prestin, with enlarged vestibular aqueduct
SMAC/DIABLO	605219	DFNA64	12q24.31	chr12:122692208-122712080	diablo homolog, mitochondrial protein, an inhibitor of apoptosis protein (IAP)-binding protein
SMPX	300226	DFNX4/DFN6	Xp22	chrX:21724090-21776278	small muscular protein, might maintain inner ear cells subjected to mechanical stress
STRC	606440	DFNB16	15q15.3	chr15:43891761-43910998	stereocilin, Deafness-infertility syndrome (DIS) [MIM:611102]
SYNE4	615535	DFNB76	19q13.12	chr19:36494002-36499672	nesprin-4, part of a LINC (linker of the cytoskeleton and nucleoskeleton) tethering complex, cellular positioning of organelles
TBC1D24	613577	DFNB86/DFNA?	16p13.3	chr16:2525147-2555734	TBC1 domain family member 24, a GTPase-activating proteins
TECTA	602574	DFNB21/DF	11q23.3	chr11:120973374-121062201	alpha-tectorin, major noncollagenous components of

		NA 8/12			the tectorial membrane in the inner ear
TJP2	607709	DFNA51	9q21.11	chr9:71714874-71870123	tight junction protein ZO-2, connecting hair cells and supporting cells
TMC1	606706	DFNB7/11 / DFNA36	9q21.13	chr9:75136716-75451266	transmembrane channel-like 1
TMIE	607237	DFNB6	3p21.31	chr3:46742823-46752413	transmembrane inner ear protein
TMPRSS3	605511	DFNB8/10	21q22.3	chr21:43791996-43816200	transmembrane protease serine 3
TNC	187380	DFNA56	9q33.1	chr9:117781854-117880536	tenascin, an extracellular matrix protein
TPRN	613354	DFNB79	9q34.3	chr9:140086069-140095163	taperin, localized prominently at the tapered regions of hair cell stereocilia
TRIOBP	609761	DFNB28	22q13.1	chr22:38092995-38172563	TRIO and F-actin binding protein, dense bundling of actin filaments essential for stereocilia bundle
TSPEAR	612920	DFNB98	21q22.3	chr21:45917775-46131495	thrombospondin-type laminin G domain and EAR repeat-containing protein, a secreted protein
USH1C	605242	DFNB18A	11p15.1	chr11:17515442-17565963	harmonin, maintenance of cochlear hair cell bundles; Usher syndrome, type 1C
WFS1	606201	DFNA6/14/38	4p16.1	chr4:6271576-6304991	wolframin, Wolfram syndrome

WHRN/ DFNB31		DFNB31	9q32	chr9:117164360-117267736	whirlin, organization and stabilization of sterocilia, Usher syndrome 2D (USH2D) [MIM:611383]:

Data from Hereditary Hearing Homepage (http://hereditaryhearingloss.org/), Genetic Home Reference (http://ghr.nlm.nih.gov/), Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim)

Table 1-3 Common syndromic hearing loss diseases

Syndrome		**Phenotype**	**Molecular classification Loci, gene, gene OMIM**	**Comments**
Usher syndrome	Most common type of AR syndromic hearing loss	sensorineural hearing loss and retinitis pigmentosa (RP)		3-6% of childhood hearing loss cases, 50% of deafblind adult patients.
	Type I	congenital severe-to-profound SNHL and abnormal vestibular dysfunction	USH1A , Unknown, 276900 USH1B, *MYO7A*, 276903 USH1C, *USH1C* , 276904 USH1D, *CDH23* , 601067 USH1E, Unknown, 602097 USH1F, *PCDH15* , 602083 USH1G, *SANS*, 606943 USH1H, Unknown, 612632 USH1J, *CJB*, 614869 USH1K, unknown,614990	Incidence : 4 in 100,000
	Type II	congenital mild-to-severe SNHL and normal vestibular function	USH2A , *USH2A* , 276901 USH2B, Unknown USH2C , *GPR98*, 605472 USH2D, *WHRN* , 611383	Most common type of Usher syndrome. It could be caused by *GPR98/PDZD7* digenic mutations.
	Type III	progressive hearing loss and variable vestibular function,	USH3, *CLRN1* , 276902	

Pendred syndrome	Second most common type of AR syndromic hearing loss.	Congenital SNHL severe-to-profound and goiter, with enlarged vestibular aqueduct (EVA)	PDS, *SLC26A4* , 605646	Around 10% of all hereditary hearing loss, most with normal thyroid function.
Jervell and Lange-Nielsen syndrome	Third most common type of AR syndromic hearing loss.	congenital profound SNHL and prolongation of the QT interval as detected by electrocardiography	JLNS1 , *KCNQ1*, 192500; JLNS2, *KCNE1* , 176261	90% of cases due to mutant *KCNQ1*
Waardenburg Syndrome	Most common type of AD syndromic hearing loss	sensorineural deafness, pigmentary abnormalities of the skin, hair, and eyes		Incidence 1 in 40000, 2-5% of all congenital hearing loss
	Type 1	dystopia canthorum	WS1, *PAX3* , 606597	
	Type 2	absence of dystopia canthorum	WS2A, *MITF*, 156845 WS2B, unknown WS2C, unknown WS2D, *SNAI2*, 602150	
	Type 3	no dystopia canthorum , with upper-limb abnormalities	WS3, *PAX3*, 606597	
	Type 4 (AR)	no dystopia canthorum , with Hirschsprung disease	WS4, *EDNRB* , 131244 WS4, *EDN3*, 131242 WS4, *SOX10*, 602229	
Branchiootorenal syndrome (BOR)	the second most common type of autosomal	conductive, sensorineural, or mixed hearing loss ; branchial	BOR1 , *EYA1*, 601653; BOR2, *SIX5*, 600963;	Incidence 1 in 40,000, *EYA1* could account for

	dominant syndromic hearing loss	cleft cysts or fistulae, malformations of the external ear, renal anomalies	BOS3, *SIX1*, 601205	40% of families segregate BOR phenotype
Alport syndrome	X-link (85%), or AR	progressive SNHL of varying severity, progressive loss of kidney function, variable ophthalmologic findings, hematuria	X-linked, *COL4A5* , 303630 AR, *COL4A4*, 120131; AR, *COL4A3*, 120070	1 in 50000 newborns

Data from Hereditary Hearing Homepage (http://hereditaryhearingloss.org/), Genetic Home Reference (http://ghr.nlm.nih.gov/), Genetic Testing Registry(http://www.ncbi.nlm.nih.gov/gtr/)

Chapter 2 Objectives

Understanding of the genetic etiology for hereditary hearing loss is crucial for our understanding of the physiology and pathophysiology of hearing. Genetic diagnosis is fundamental for explanation of the disease, determination of clinical management, and appropriate genetic counseling in patients and their extended family members. Early intervention of hearing loss in children is especially of great benefit for the development of speech, language, cognitive and socio-emotional behavior. The high number of genetic entities and the huge heterogeneity of hereditary hearing loss require guidelines for requesting genetic testing when desirable. We sought to establish cost effective approaches to identify this group of individuals with such diverse genetic etiology. Common mutation screening could be served as a first tier testing. However, limited data was available in the Chinese population. With the development of molecular technologies, a single comprehensive genetic test which could screen all genes implicated in the hereditary hearing loss would be worthy of pursuit.

The objectives of this thesis are:

1) Study the spectrum and carrier frequency of common hearing loss mutations in the Chinese population
2) Establish a targeted genome enrichment massively parallel sequencing (TGE-MPS) approach for genetics diagnosis of hereditary hearing loss
3) Evaluate the clinical application of targeted genome enrichment method on hereditary hearing loss patients
4) Develop a noninvasive approach in prenatal diagnosis of hereditary hearing loss

Section II: Materials and methods

Chapter 3 Materials and Methods

3.1 Human peripheral blood DNA extraction

Genomic DNA was extracted from EDTA anticoagulant-treated blood by DNeasy Blood & Tissue Kit, QIAGEN (CA, USA) according to the manufacturer's instruction. Add 20ul proteinase K into 100-200ul EDTA-whole blood or 30-80ul buffy coat, adjust final volume into 220ul with PBS. Add 200ul AL buffer, vortexing for 3-5 seconds .Incubate the mixture at 56°C for at least 15 mins. Add 200ul ethanol (96-100%), vortexing for 5 seconds. The mixture was transmitted to a DNeasy Mini spin column and centrifuged at 6000xg for 1 min at room temperature. 500ul AW1 was added into the spin column and centrifuge at 6000xg for 1 min at room temperature. 500ul AW2 was added into the spin column and centrifuge it at 20000xg for 3mins, then further centrifuge it at 20000xg for 1min. 40-50ul AE buffer or MilliQ water was added into spin column, incubating at room temperature for at least 1 min. DNA was eluted by centrifuging at 6000xg for 1min.

3.2 Plasma isolation

Plasma isolation should be finished in 30 mins-4 hours after intravenous blood collection. Collect 5ml peripheral blood into EDTA anticoagulant tube. Centrifuge the tube at 1600g for 10 mins at 4°C. The supernatant was transferred to 1.5ml polypropylene tubes carefully to ensure that the blood pellet remains intact. The supernatant was centrifuged again at 16 000g for 10 mins at 4°C and transferred into a 1.5 ml polypropylene tube. Cell free plasma should be stored at -80°C until further processing.

3.3 Plasma DNA extraction

Cell free DNA extracted from plasma was performed using a TIANamp Micro DNA Kit (Tiange, China) according to the manufacturer's instructions. 200ul plasma was mixed with 200ul buffer GB, 20ul proteinase K and 2ul 1ug/ul carrier RNA. The mixture was vortexed and incubated at 56℃ for 10 mins. 200ul ethanol was added into the mixture, mixed thoroughly, and then incubated at room temperature for 5mins. 15ul MagAttract Suspension B was added. The mixture was vortexed thoroughly, and placed on the magnet rack 30-60s till the magnetic beads form a dot. Remove the supernatant. 500u buffer GD was added into the mixture. Vortexed thoroughly, and putted on the magnet rack 30-60s till the magnetic beads form a dot. Remove the supernatant. 600ul Buffer GB was loaded into the mixture was vortexed thoroughly, and putted on the magnet rack 30-60s till the magnetic beads form a dot. Remove the supernatant. Repeat two times. Dry the magnetic beads for 10-15 mins at room temperature. 50ul Buffer TB was added into the mixture and vortexed thoroughly, followed by incubating at 56°C for 10 mins. Place the tube on the magnet rack 30-60s till the magnetic beads form a dot. Collect the supernatant and measure DNA concentration.

3.4 Paired-end DNA library preparation for Illumina sequencing platform

3ug of genomic DNA in 90-100ul AE buffer was sheared by Covaris (Covaris Inc., Woburn, MA, US). Microtube (6x16mm) with duty 10%, intensity 10, cycle per bust 200 for 5 mins. The size peak of sheared DNA fragment was 200bps.

Cell free DNA was characterized with fragments size peak at 143bps. No need other fragmentation process was required before preparing the sequencing library.

Cell free DNA and genomic DNA fragment samples were both prepared following the protocol for paired-end sequencing on the illumina sequencing platform with some modifications. The DNA end repair was to mix the DNA fragment with T4 DNA polymerase, T4 polynucleotide kinase, and Klenow DNA polymerase at 20°C for 30mins. Then an A nucleotide overhang was added to the 3 end of DNA fragment by Klenow exo at 37°C for 30mins. Index Pair end adapters were added by incubating with T4 DNA ligase at 37°C for 30mins or at 16°C for 12hours-16hours. DNA fragment was amplified by PCR with illumina supplied primers for five cycles, during which each sample was integrated with barcode for multiplex sequencing. The quality and quantity of purified library was checked by Bioanalyzer 2100 (Agilent, CA, USA) and real time PCR.

3.5 Agilent SureSelect target enrichment for Illumina paired-end sequencing library

500-750ng of DNA library in a maximum volume of 3.4ul was used as starting material. Hybridize the library according to the manufacturer's instructions (Agilent, CA, USA). Mix 2ul SureSelect probe library with 0.5ul RNase block dilution and 4.5ul distilled water to a final volume of 7ul. DNA library was denatured at 95°C for 5 mins, then incubated with 13ul buffer mix and 7ul probe mix at 65°C for 24 h. Incubate the Dynal MyOne Streptavidin T1 (Invitrogen) magnetic beads with the mix to separate the hybrid capture. The captured DNA library was finally eluted into 36.5ul distilled water. The captured product was amplified with PCR 12 cycles using Herculase polymerse. The quality and quantity of enriched fragment was assessed by qPCR and Bioanalyzer 2100 (Agilent, CA, USA).

3.6 NimbleGen SeqCap EZ target enrichment for Illumina paired-end sequencing library

One ug of DNA library was used as the starting material. Hybridize the library according to the NimbleGen SeqCap EZ manufacturer's instructions. Sample library, COT DNA and TS-HE oligo mix were dry to membrane stature at 60°C. The DNA mix was denatured in hybridization buffer at 95°C for 10mins. Then the mixture was incubated at 47°C for 65 hours. The captured DNA bounded by streptavidin beads was eluted into 50ul distilled water. The bead bounded DNA was enriched by PCR. The quantity and quality of purified PCR produced was assessed by Real time PCR and the Bioanalyzer 2100 (Agilent)

3.7 Agilent SureSelect target enrichment library design

All exons plus flanking 50 base pair and untranslated regions of 252 human genes (Table 3-2) and the whole mitochondrial genome were identified in Refseq using the University of California Santa Cruz genome browser (http://genome.ucsc.edu/). Apart from the nonsyndromic hearing loss genes published in the Hereditary Hearing Homepage, we also included candidate genes reported in animal models. The coordinate of target genes were uploaded to Agilent eArray website (https://earray.chem.agilent.com/) for cRNA baits design following the parameter setting (Table 3-1). The custom Agilent SureSelect target enrichment library was ordered in 2011, from Agilent Technologies, USA.

Table 3-1 Agilent custom design of SureSelect target enrichment library: parameter setting

Parameters	
Design Options	Centered
Baited Length	120
Bait Tilling	4X
Allowed Overlap into Avoided Regions	20
Avoid Standard Repeat Masked Regions	RepeatMasker

3.8 Roche NimbleGen SeqCap Choice EZ library design

All exons plus flanking 50 base pair and untranslated regions of 269 human genes (Table 3-3) and the whole mitochondrial genome were identified in the Refseq using the UCSC genome browser (http://genome.ucsc.edu/). The coordinate of target regions were sent to the NimbleGen Bioinformatics team using their in house algorithm for DNA bait design. According to their notification, they would modify the target regions by padding any regions shorter than 100bp to 100bp, and merging any overlapping region. The design consists of probes that may have up to five matches with the genome as determined by the SSAHA algorithm (a SSAHA match allows up to five insertions, deletions, or mismatches). Roche NimbleGen SeqCap Choice EZ library was ordered in 2013, from Roche NimbleGen, USA.

3.9 Target enrichment massively parallel sequencing data analysis

Variant calling: Image analyses, error estimation and base calling were performed using the Illumina Pipeline to generate raw data. Index adapters and barcode sequences introduced in DNA library preparation were used to identify different reads from different samples. Unqualified sequence reads were removed from the raw data, which had more than 10 percent Ns (N is equivalent to an interrupted and resumed signals from sequential flows), 50% of base calls with a quality value of less than 5, or an average quality score <10, containing more than 40 continuous identical

bases. Remaining sequences were termed as clean reads for alignment. Cleaned reads were aligned to the GRCh37/hg19 human genome build assembly using BWA (LI and DURBIN 2009). Alignments were generated in SAM format. Reads mapping to multiple sites in the genome were not analyzed. Mapped reads with the same 5' and 3' primers were also excluded. Unique reads were used to calculate the read depth of each base. SNPs and indels were called respectively by using SOAPsnp software (LI *et al.* 2009) and the GATK Indel Genotyper (http://www.broadinstitute.org/gsa/wiki/index.php/, The Genome Analysis Toolkit). Variants were selected as supported by at least 10 reads and their sequence was more than 20% of total reads on site.

Variant filtering: Sequence variation annotation was performed by the in-house pipeline, which consisted of gene annotation (RefSeq), allele frequency (compared with dbSNP136, HapMap database, 1000 Genome, and normal controls of BGI in-house database), and variant characterization (splicing, synonymous, non-synonymous, etc.). Amino acid substitution affected protein function was also predicted by Sorting Intolerant from Tolerant (SIFT, http://sift.jcvi.org/) and Polymorphism Phenotyping v2 (PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/) (WEI *et al.* 2011). After annotation, those SNPs and indels of each sample which reach the following criteria were selected and put in a new excel file: 1) allele frequency was lower than 0.05 in dbSNP136, HapMap, 1000 genomes public database, BGI in-house normal control database, 2) resulting in nonsynonymous change (except mitochondrial variants), 3) on splice site defined as 15bp flanking each exon.

Candidate variant evaluation: Each of retained variants after filtering was be scrutinized manually for further analysis: 1) reported before and whether reported

cases had similar phenotype (HGMD); 2) pathogenic supported by functional study (Pubmed); 3) other diagnostic lab interpretations variant of unknown significance or probably pathogenic (ClinVar); 4) minor allele frequency (MAF) on exome sequencing data (ESP) <0.05; 6) MAF in the matched ethnic group <0.05 (1000 genomes); 7) inheritance pattern of the mutations consistent with the pedigree. Candidate pathogenic variants were confirmed by Sanger sequencing in each sample and its family members.

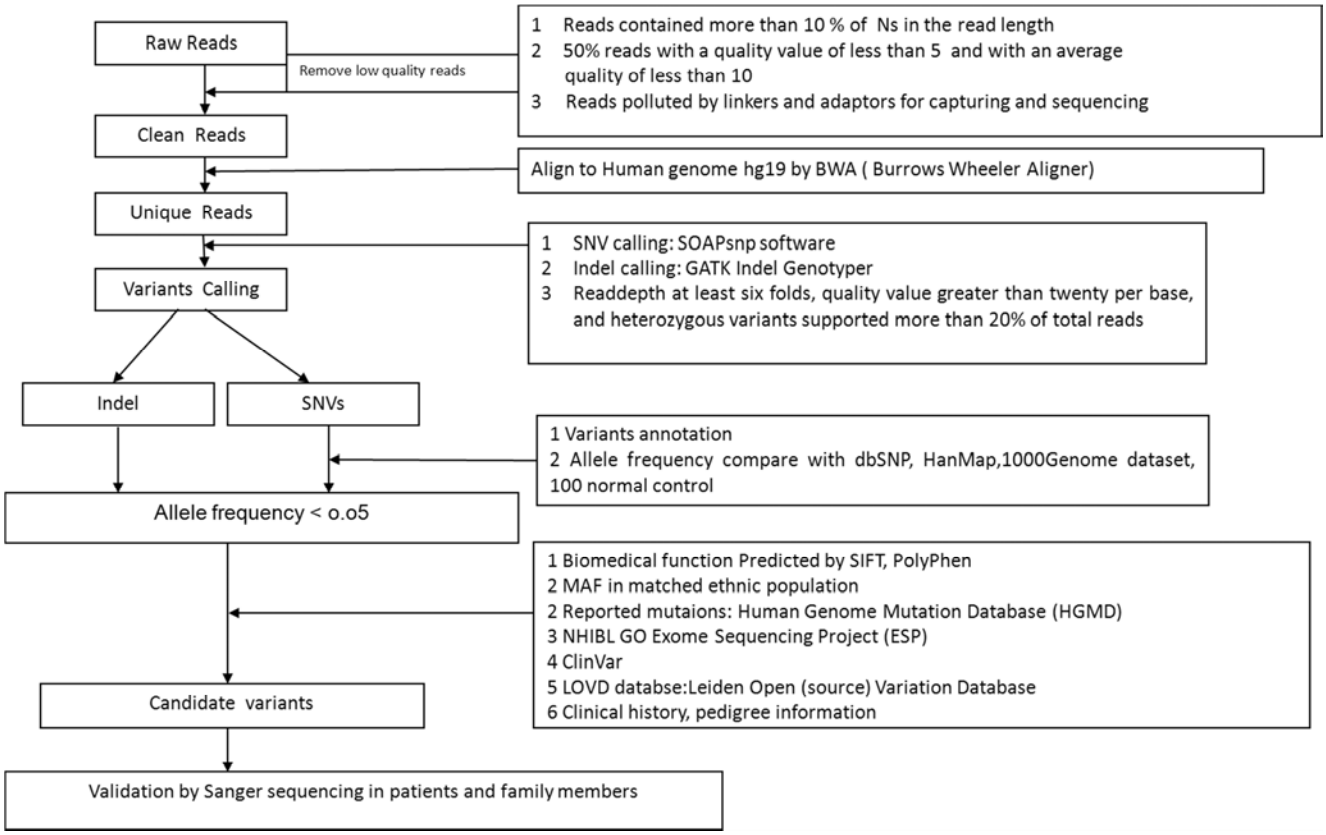


Figure 3-1 Flowchart of identifying candidate variants by target enrichment and massively parallel sequencing

3.10 Oligo primer

	F-primer	R-primer	Product size (base pair)
E2073_1 PTPRQ	TTAGCTTGCTTGCTTTCCAGA	TGCCAGTATAAGTATTTCCATCCA	251
E2073_1 MYO15A	TCCCTGTTCTGGGAGGTATG	GAAAGCAACACTGGTCGTCA	256
E2073_1 TRIOBP	TCAGCCTTGGTCAGAGAGGT	AGGTCCACTGTGTGGTGTCA	294
E2073_1 COL4A3	GAGTTCAAATTGCACCTGCTC	AGGTCCAGGGTTGGAGGTAA	212
E2073_1 MRPS16	CACCTACTGTCCTGGAGCTGA	ACTTAACCATCCGCCTTGC	296
E2073_1 ASPA	ACTGGAAACCACTGCATCCT	CCGTGTAAGATGTAAGCTGGAA	220
E2073_1 NEFL	ACTTGAAGTTGCAGGGGTTT	CGGTTTACAGACCAGCTCCT	249
E2073_1 ALMS1	TCCAACTTGAAGTCAGGCATC	TTGGTGTTGCAGTCTTCTGG	232
E2073_1 LOXHD1	TTGAATCAGGGAAGGCTGTC	GGGCCAGAGGGAATAAGAAC	208
E2073_1 TMIE	AGAAATTGGTTGGCATGAGG	CTGCCCAGGATGTCTTCATT	292
MYO1A-1	TCTCAGTCCTACCTCCCCTC	TTGCTAGTTCTGTTTGCCCC	265
ESRRB-1	TCTGAAGATACCCCTGCCTG	GTCTGTCTGTAGGTGGGCAT	395
WFS1-1	TATGGAGTGTCTGGCAGCTC	TATCCCTGAACATCCCCAGC	401
WFS1-2	ATCGAGTTCAGCACCATCCT	CCACACTGGGGAAAGGCC	267
MYO7A-1	GGAGAGCTTGTTCCCTGAGG	GCACAGAAGGAGGGAGGATC	353
SLC17A8-1	TGGGCAAGGGAATTAGGGAG	CTCAAGTGTCTGCTGAGGGA	348
HL-DFNA5- E3	TTCCTCTGCTCGTCTCCTC	CAAATCCACCTCCTGCTTC	191
HL-DFNA5- E6	AGTTCTGCCTTCTCCGAG	GAAGATATCCCATGCGCA	133

Table 3-2 Gene list of Agilent SureSelect target enrichment design (252)

ABCA4	*COL2A1*	*FOXC1*	*LOXHD1*	*OPA1*	*RLBP1*	*TNFRSF11A*
ACSL4	*COL4A3*	*FOXI1*	*LRAT*	*OPA3*	*RPGR*	*TNFRSF11B*
ACTG1	*COL4A4*	*FXN*	*LRP2*	*OTOA*	*RPS6KA3*	*TRIOBP*
ACVR1	*COL4A5*	*GALC*	*LRTOMT*	*OTOF*	*RS1*	*TULP1*
ALMS1	*COL4A6*	*GATA3*	*MARVELD2*	*PABPN1*	*RUNX2*	*TWIST1*
AMMECR1	*COL5A1*	*GBA*	*MECP2*	*PAX2*	*SAG*	*UBE3A*
ANKH	*COL5A2*	*GDAP1*	*MEN1*	*PAX3*	*SALL1*	*UBR1*
APC	*COL9A1*	*GDF5*	*MERTK*	*PAX6*	*SALL4*	*USH1C*
APTX	*COQ2*	*GDF6*	*MGP*	*PCDH15*	*SDHD*	*USH1G*
AR	*CRB1*	*GFER*	*MITF*	*PCDH18*	*SEMA3E*	*USH2A*
ASL	*CREBBP*	*GJA1*	*MPZ*	*PDE6A*	*SHOX*	*WDR36*
ASPA	*CRYBB2*	*GJB1*	*MRPS16*	*PDE6B*	*SIX1*	*WFS1*
ATP6V0A4	*CRYM*	*GJB2*	*MYH14*	*PDSS1*	*SIX5*	
ATP6V1B1	*CTSA*	*GJB3*	*MYH9*	*PDSS2*	*SLC17A8*	
ATRX	*CYP21A2*	*GJB6*	*MYO15A*	*PDZD7*	*SLC19A2*	
BCS1L	*CYP4V2*	*GLA*	*MYO1A*	*PEX1*	*SLC25A4*	
BEST1	*DFNA5*	*GPR98*	*MYO3A*	*PEX10*	*SLC26A4*	
BSND	*DFNB31/WHRN*	*GRHL2*	*MYO6*	*PEX12*	*SLC26A5*	
BTD	*DFNB59/PJVK*	*GRXCR1*	*MYO7A*	*PEX13*	*SLC29A3*	
CACNA1A	*DIABLO/SMAC*	*GRXCR2*	*MYOC*	*PEX14*	*SLC4A11*	
CALM2	*DIAPH1*	*HBB*	*NDP*	*PEX26*	*SMC1A*	
CCDC50	*DIAPH2*	*HBD*	*NDUFA7*	*PEX3*	*SMN1*	
CD151	*DSPP*	*HBG2*	*NDUFAF2*	*PEX5*	*SMN2*	
CDH23	*EDN3*	*HEXA*	*NDUFAF4*	*PEX6*	*SNAI2*	
CERKL	*EDNRB*	*HGF*	*NDUFS3*	*PEX7*	*SOST*	
CHD7	*EGR2*	*HOXA2*	*NDUFS4*	*PHEX*	*SOX10*	
CHM	*ERCC6*	*IGF1*	*NDUFS5*	*PHYH*	*SOX2*	
CLCN7	*ERCC8*	*ITM2B*	*NDUFS6*	*PLOD1*	*SOX9*	
CLCNKA	*ESPN*	*ITPR1*	*NDUFS7*	*PLOD3*	*SPTLC1*	
CLCNKB	*ESRRB*	*JAG1*	*NDUFV1*	*PMP22*	*SQSTM1*	
CLDN14	*ETFA*	*KCNE1*	*NEFL*	*POLG*	*STRC*	
CLIC5	*ETFB*	*KCNJ10*	*NF2*	*POU3F4*	*TAZ*	
CLRN1	*EYA1*	*KCNQ1*	*NIPBL*	*POU4F3*	*TCOF1*	
CNGA1	*EYA4*	*KCNQ4*	*NLRP3*	*PRPS1*	*TECTA*	
CNGB1	*FGF10*	*KIF1B*	*NOG*	*PTPN11*	*TGFB1*	
COCH	*FGF3*	*KIT*	*NPC1*	*PTPRQ*	*THRB*	
COL11A1	*FGFR1*	*LHFPL5*	*NR2E3*	*RAF1*	*TIMM8A*	
COL11A2	*FGFR2*	*LHX3*	*NRL*	*RDX*	*TMC1*	
COL1A1	*FGFR3*	*LITAF*	*OCA2*	*RGR*	*TMIE*	
COL1A2	*FLNA*	*LMNA*	*OFD1*	*RHO*	*TMPRSS3*	

Table 3-3 Gene list of NimbleGen SeqCap Choice EZ design (269)

ABCA4	*COL2A1*	*FOXC1*	*LOXHD1*	*OPA1*	*RLBP1*	*TNFRSF11A*
ACSL4	*COL4A3*	*FOXI1*	*LRAT*	*OPA3*	*RPGR*	*TNFRSF11B*
ACTG1	*COL4A4*	*FXN*	*LRP2*	*OTOA*	*RPS6KA3*	*TRIOBP*
ACVR1	*COL4A5*	*GALC*	*LRTOMT*	*OTOF*	*RS1*	*TULP1*
ALMS1	*COL4A6*	*GATA3*	*MARVELD2*	*PABPN1*	*RUNX2*	*TWIST1*
AMMECR1	*COL5A1*	*GBA*	*MECP2*	*PAX2*	*SAG*	*UBE3A*
ANKH	*COL5A2*	*GDAP1*	*MEN1*	*PAX3*	*SALL1*	*UBR1*
APC	*COL9A1*	*GDF5*	*MERTK*	*PAX6*	*SALL4*	*USH1C*
APTX	*COQ2*	*GDF6*	*MGP*	*PCDH15*	*SDHD*	*USH1G*
AR	*CRB1*	*GFER*	*MITF*	*PCDH18*	*SEMA3E*	*USH2A*
ASL	*CREBBP*	*GJA1*	*MPZ*	*PDE6A*	*SHOX*	*WDR36*
ASPA	*CRYBB2*	*GJB1*	*MRPS16*	*PDE6B*	*SIX1*	*WFS1*
ATP6V0A4	*CRYM*	*GJB2*	*MYH14*	*PDSS1*	*SIX5*	*DIAPH3*
ATP6V1B1	*CTSA*	*GJB3*	*MYH9*	*PDSS2*	*SLC17A8*	*GIPC3*
ATRX	*CYP21A2*	*GJB6*	*MYO15A*	*PDZD7*	*SLC19A2*	*GPSM2*
BCS1L	*CYP4V2*	*GLA*	*MYO1A*	*PEX1*	*SLC25A4*	*ILDR1*
BEST1	*DFNA5*	*GPR98*	*MYO3A*	*PEX10*	*SLC26A4*	*MSRB3*
BSND	*DFNB31/WHRN*	*GRHL2*	*MYO6*	*PEX12*	*SLC26A5*	*TPRN*
BTD	*DFNB59/PJVK*	*GRXCR1*	*MYO7A*	*PEX13*	*SLC29A3*	*SERPINB6*
CACNA1A	*DIABLO/SMAC*	*GRXCR2*	*MYOC*	*PEX14*	*SLC4A11*	*MIR96*
CALM2	*DIAPH1*	*HBB*	*NDP*	*PEX26*	*SMC1A*	*TJP2*
CCDC50	*DIAPH2*	*HBD*	*NDUFA7*	*PEX3*	*SMN1*	*MIR183*
CD151	*DSPP*	*HBG2*	*NDUFAF2*	*PEX5*	*SMN2*	*SMPX*
CDH23	*EDN3*	*HEXA*	*NDUFAF4*	*PEX6*	*SNAI2*	*COL9A2*
CERKL	*EDNRB*	*HGF*	*NDUFS3*	*PEX7*	*SOST*	*CLRN3*
CHD7	*EGR2*	*HOXA2*	*NDUFS4*	*PHEX*	*SOX10*	*PNPT1*
CHM	*ERCC6*	*IGF1*	*NDUFS5*	*PHYH*	*SOX2*	*OTOGL*
CLCN7	*ERCC8*	*ITM2B*	*NDUFS6*	*PLOD1*	*SOX9*	*CABP2*
CLCNKA	*ESPN*	*ITPR1*	*NDUFS7*	*PLOD3*	*SPTLC1*	*OTOG*
CLCNKB	*ESRRB*	*JAG1*	*NDUFV1*	*PMP22*	*SQSTM1*	
CLDN14	*ETFA*	*KCNE1*	*NEFL*	*POLG*	*STRC*	
CLIC5	*ETFB*	*KCNJ10*	*NF2*	*POU3F4*	*TAZ*	
CLRN1	*EYA1*	*KCNQ1*	*NIPBL*	*POU4F3*	*TCOF1*	
CNGA1	*EYA4*	*KCNQ4*	*NLRP3*	*PRPS1*	*TECTA*	
CNGB1	*FGF10*	*KIF1B*	*NOG*	*PTPN11*	*TGFB1*	
COCH	*FGF3*	*KIT*	*NPC1*	*PTPRQ*	*THRB*	
COL11A1	*FGFR1*	*LHFPL5*	*NR2E3*	*RAF1*	*TIMM8A*	

COL11A2	*FGFR2*	*LHX3*	*NRL*	*RDX*	*TMC1*	
COL1A1	*FGFR3*	*LITAF*	*OCA2*	*RGR*	*TMIE*	
COL1A2	*FLNA*	*LMNA*	*OFD1*	*RHO*	*TMPRSS3*	

Section III: Results

Chapter 4 Studying the spectrum and carrier frequency of common mutations in the Chinese population

4.1 Introduction

Hearing loss is one of the most common birth defects worldwide (ROSS *et al.* 2008). According to the universal newborn hearing screening program in US, it has been estimated that the prevalence of bilateral hearing loss in newborns is two to four per one thousand live births (OLUSANYA and NEWTON 2007). In 2006, the Second China National Survey on Disability documented that the hearing impaired population consists of 20.04 million people in China, which accounts for 1.5% of the total population (www.cdpf.org.cn). It is estimated that 115,000 children under seven years old have severe to profound hearing impairment and the number increases by 30,000 neonates with hearing loss annually (LIANG and MASON 2013). Genetic causes account for at least 50% of prelingual hearing loss (MORTON and NANCE 2006). Approximately 80 genes and 180 loci have been identified related to the nonsyndromic hearing loss (http://hereditaryhearingloss.org). Genetic heterogeneity of hearing loss makes the diagnosis quite challenging. However, some mutant genes are more frequently identified. In particular, *GJB2* is responsible for more than half of DFNB cases (HILGERT *et al.* 2009a). Based on previous studies (LI *et al.* 2011) and literature review of the inheritance on genetic etiology of nonsyndromic hearing loss patients in China (LI *et al.* 2012), at least 15 well known hearing loss mutations were commonly reported in these Chinese hearing loss patients, four of which *GJB2*-c.235delC, *GJB2*-c.299-300delAT, *SLC26A4*-c.919-2A>G, mt-1555-A>G

were summarized as the most common ones. In terms of this, a genetic screening test for common mutations would be favorable considering the large prevalence of hearing loss patients in China. Understanding the frequency of these mutation alleles in the Chinese population could facilitate design of the screening panel. However, there were limited studies on the carrier frequency of these common mutations. This study mainly aimed to estimate the frequency of these 15 well-known mutations in the Chinese population, based on the screening of a neonatal cohort in Suzhou, China.

4.2 Materials and Methods

Genomic DNA was extracted from dried blood spots of 5800 neonates, including 3077 males and 2723 females, born between October 2011 and February 2012 in the Suzhou Hospital. A well-established genotyping method, the SNaPshot® Multiplex System (ZL201110027435.1) was applied to detect the 15 mutations (Table 4-1). The genotyping analysis was performed using ABI3130 Analyzer. GeneMapper V3.0 (Applied Biosystems) was used for raw data analysis (LI *et al.* 2011). Those variants detected by this method were further confirmed by Sanger sequencing. OAEs were measured in a quiet environment and OAE recordings were evaluated by standard OAE parameters.

4.3 Results

4.3.1 Carrier frequency of the common mutations

A total of 5800 newborns were screened by the SNaPshot method in this study. All 15 mutations were found in this cohort (Table 4-1). Interestingly, 15.9% (923/5,800) of this cohort carried at least one of these 15 mutantions, indicating that one in 6.3 newborns had at least one mutant allele of a hearing loss gene. Mutations on *GJB2, SLC26A4*, mitochondrial genome accounted for 12.7% (735/5800), 2.19% (127/5800) and 1.07% (62/5800) of the recruited newborns. *GJB2* c.109G>A had the most

prevalent mutated allele with a frequency up to 5.26% (610/11600), followed by *GJB2* c.235delC (0.94%, 109/11600), and *SLC26A4 c.919-2A>G* (0.84%, 98/11600), which accounted for carrier frequencies of approximately 10.29% (597/5800), 1.88% (109/5800), and 1.62% (94/5800), respectively. For the mitochondrial genome, mt-7444G>A, mt-3243A>G and mt-1555A>G were the top three mutant alleles, with carrier frequencies of 1/141, 1/645 and 1/725, respectively, determined in 59 cases of homoplasmic mitochondrial mutations. Forty of 59 cases involved mt-7444G>A. We also identified three cases of heteroplasmy for mt-1555A>G. In addition, our data showed that the Chinese population has rather low carrier frequencies of *GJB2* c.35delG, *GJB2* c.176191del16, *SLC26A4* c.1229C>T, *SLC26A4* c.2027T>A, and mt-7445A>G, with only one to three carriers among the 5800 newborns screened (Table 4-1).

4.3.2 Clinical evaluation of genetically diagnosed newborns

Within this cohort, 0.48% of newborns (28/5800) were genetically diagnosed with hearing loss. Nineteen out of 28 newborns carried homozygous mutations in *GJB2* or *SLC26A4*. The other nine carried compound heterozygous *GJB2* mutations. Only seven of these 28 newborns failed in a universal newborn hearing screening program OAE test for at least one ear. Four of these seven cases only failed in one ear. For those thirteen cases with *GJB2 c.109G>A* mutations, one case failed an OAE test unilaterally, three cases failed bilaterally, and the rest nine cases passed the OAE test. No newborn, carrying only one mutant allele, failed both OAE tests.

Table 4-1 Comparison of carrier frequency in three studies of mutations associated with hearing loss

	Carrier frequency			Estimated Carrier frequency in China	Global minor allele frequency (MAF)†	Carrier frequency in other population
Genes and mutations	**Our study**	**Gansu[12]**	**Guangdong[10]**			
Type of Subjects	**Newborn**	**Newborn**	**Women in children bearing age**			
No. of Subjects	**5800**	**10043**	**7263**	**23106**		
***GJB2* c.35delG**	1 (0.017%)	-	1 (0.014%)	0.015%	-	2-4%‡
***GJB2* c.109G>A**	597 (10.290%)	-	-	10.290%	A=0.012	
***GJB2* c.176-191del16bp**	11 (0.190%)	-	7 (0.096%)	0.137%	-	
GJB2* c.235delC	109 (1.879%)	119 (1.185%)	128 (1.762%)	1.54%	-	1-2%‡
***GJB2* c.299-300delAT**	17 (0.293%)	-	28 (0.386%)	0.344%	-	
SLC26A4 c.919-2A>G* *	94 (1.621%)	93 (0.9%)	90 (1.239%)	1.199%	-	
***SLC26A4* c.1174A>T**	9 (0.156%)	-	-	0.156%	T=0.000	
***SLC26A4* c.1229C>T**	1 (0,017%)	-	-	0.017%	-	
***SLC26A4* c.2027T>A**	2 (0.034%)	-	-	0.034%	-	
***SLC26A4* c.2168A>G**	20 (0.350%)	-	19 (0.262%)	0.299%	G=0.001	
mt-1494C>T	1 (0,017%)	3 (0.029%)	3 (0.041%)	0.030%	-	
mt-1555-A>G	8 (0.138%)	16 (0.159%)	18 (0.25%)	0.182%	-	
mt-3243A>G	9 (0.155%)	-	-	0.155%	-	
mt-7444G>A	41 (0.707%)	-	-	0.707%	-	
mt-7445A>G	3 (0.052%)	-	-	0.052%	-	
Total	923/5800 (15.914%)	230/10043 (2.290%)	294/7263 (4.049%)			

†: 1000Genome phase 1 genotype data from 1094 worldwide individuals, released in the May 2011 dataset

*&**: The carrier frequency of the two mutations among these three regions has significant difference. (Chi-square test, $p<0.01$)

‡: Carrier frequency of *GJB2* c.235delC in the Japanese population, 1–2%; carrier frequency of *GJB2 c*.35delG in Caucasian, 2–4% (PUTCHA *et al.* 2007)

Table 4-2 Summary of genetically diagnosed cases and OAE test results

Mutation Identified	Genotype	No. of cases	No. of cases failed OAE test	OAE test Evaluation
GJB2 c.109G>A	Homozygous	13	3 (23%)	Bi. Fail
			1 (7.7%)	Uni. Fail
GJB2 c.109G>A+ *GJB2* c.176-191del16	Compound Heterozygous	1	0 (0%)	Pass
GJB2 c.109G>A + *GJB2* c.235delC	Compound Heterozygous	6	2 (33.3%)	Uni. Fail
GJB2 c.109G>A+ *GJB2* c.299-300delAT	Compound Heterozygous	1	0 (0%)	Pass
GJB2 c.35delG+ *GJB2* c.235delC	Compound Heterozygous	1	0 (0%)	Pass
SLC26A4 c.919-2A>G	Homozygous	4	1 (25%)	Uni. Fail
SLC26A4 c.1174A>T	Homozygous	2	0 (0%)	Pass
Total no. of Cases		**28**	**7 (25%)**	

OAE test: otoacoustic emissions test; Bi Fail: Bilaterally failed OAE test; Uni Fail: Unilaterally failed OAE test

4.4 Discussion

In this study, we calculated the Chinese population-based carrier frequencies of 15 common hearing loss mutations. Distinct from *GJB2 c*.235delC common in Japanese (1–2%), *GJB2 c*.35delG in Caucasian (2–4%) and *GJB2 c*.167delT in the Ashkenazi Jewish population (around 4%) (PUTCHA *et al.* 2007), our study revealed that in the Chinese cohort the *GJB2* c.109G>A had the highest carrier frequency up to 10.29%, which was compatible with 11.6% in a Taiwanese population (HWA *et al.* 2003). The second and third most frequent mutations *GJB2* c.235delC and *SLC26A4* c.919-2A>G account for 1.88% and 1.62% of our cohort respectively, which are similar in other studies of different Asian populations (HWA *et al.* 2003; SAGONG *et al.* 2013). In our neonatal Chinese cohort the carrier frequency of *GJB2* c.235delC was 1.54%, which was approximately 100 times more frequent than that of *GJB2* c.35delG. Other two Chinese hearing genetic epidemiology studies examined a smaller set of mutations in the different regions of Chinese population, Guangzhou and Gansu (ZHANG *et al.* 2012; YIN *et al.* 2013) (Table 4-1). On comparison with the data for the eight shared mutations screened in this and the Guangzhou study, carrier frequencies were found to be quite similar, 4.26% (our study) vs. 4.05% (Guangzhou) (YIN *et al.* 2013). There were two shared mutants in these three studies, *GJB2* c.235delC and *SLC26A4* c.919-2A>G, the carrier frequencies of which varied statistically significantly in different regions. A possible explanation is that different proportions of minority ethnic groups in different regions impact the carrier frequency. *GJB2* c.235delC allele frequency among different ethnic deaf groups in China also showed a statistical significant difference (DAI *et al.* 2007). We combined data from these three studies in order to have a better estimation of carrier frequency of each mutation in the Chinese population (Table 4-1). The Chinese carrier

frequencies of *GJB2* c.109G>A, *SLC26A4* c.1174A>T, and *SLC26A4* c.2168A>G differed dramatically from the global minor allele frequency (MAF) calculated in the 1000 genomes dataset (Table4-1). For *GJB2* c.109G>A CHB+JPT_low_coverage_panel showed the A allele frequency of 0.05. The other two did not have ethnic-specific frequency data. It suggests that referring to the 1000 genome data based on matched ethnic populations would be more consultative.

Our data showed the total heterozygous and homozygous carrier frequencies of *GJB2* c.109G>A were as high as 9.74% and 0.22% respectively in this study. The pathogenicity of V37I is debatable across the literature. Biological study of *GJB2* c.109G>A elucidated that *in vitro* it was devoid of function the same as *GJB2* c.235delC in paired *Xenopus* oocytes and transfected HeLa cells (BRUZZONE *et al.* 2003). Several studies showed it could even be detected in normal hearing individuals. However, it was significantly overrepresented in the mild to moderate sensorineural hearing loss patients, especially from Asia (SNOECKX *et al.* 2005; DAHL *et al.* 2006; HUCULAK *et al.* 2006; GALLANT *et al.* 2013; KIM *et al.* 2013). Huculak et al. found that 43.75% and 11.5% of Chinese and Caucasian patients respectively had *GJB2* c.109G>A (HUCULAK *et al.* 2006). Pollak et al. reported that in Polish patients, this genotype was significantly over-represented, but their penetrance rates were only 10% (Pollak et al. 2007). This mutation has either low penetrance or functions as a risk factor with other mutations with undisputed pathogenicity (SNOECKX *et al.* 2005). *GJB6*-D13S1830 deletion might contribute little since it was rarely observed in the Chinese hearing loss patients (CHEN *et al.* 2012b). The reported incomplete penetrance and interaction with a second mutation

might explain that only 4 of the 13 subjects in our study with the *GJB2*-109G>A homozygous mutation failed in the OAE test bilaterally or unilaterally.

Twenty-eight cases of our newborn cohort were genetically diagnosed with hearing loss. However, only seven of them failed the newborn OAE testing. The same genotype of GJB2 c.109G>A can have a distinct phenotype. One explanation is these neonates might present with mild hearing loss at birth but would not be detected, as normal OAE responses would be absent if hearing loss thresholds are approximately greater than 30 to 40 dB (HARLOR and BOWER 2009) . Or, they have normal hearing at birth when screened and might present late onset hearing loss which also could be missed by UNHS. Our results indicate these genetic diagnosed cases would be worthy of further investigation and long term follow up. UNHS could detect bilateral and unilateral hearing loss with any possible causes in hearing impairment by one month of age (GHOGOMU *et al.* 2014). It also showed a 2.5% to 8% false positive rate and 4.0% to 12% positive predictive values (CLEMENS and DAVIS 2001). Some cases failing UNHS were actually due to temporary conditions such as presence of debris in the external canal or amniotic fluid in the middle ear. In contrast with UNHS, our data indicated the genetic tests of these 15 mutations offers potential benefits to clarify the etiology of hearing loss and to identify later onset and mildly affected cases.

Our finding of high carrier frequencies of these 15 mutants in the Chinese population is important and provides a strong argument and need for developing a genetic hearing screening program for newborns. It could play an integral role and complement UNHS in the clinics. This combination could better achieve the goal of early detection of hearing loss and prompt intervention through an integrated,

interdisciplinary and family-centered approach. It could facilitate explanation of the cases failing hearing screening, identify those late onset cases early in life, and detect pre-symptomatic cases for long-term follow-up. It could also indicate the parents as at risk couples for genetic counseling for a subsequent pregnancy. For carriers with mitochondrial mutations such as mt-1555-A>G and mt-7444G>A, guidance of avoiding aminoglycosides medication could be an effective way to lessen their risk of hearing impairment (ZHU *et al.* 2009).

To date, genetic heterogeneity of hearing loss is still a challenging issue, and the SNaPshot method can speed up the detection of up to 15 targeted hearing loss mutants. We believe the genetic screening approach could be remarkably enhanced by targeted enrichment and massively parallel sequencing (MPS) technology. This approach has recently been proven as an effective technology for targeting more than 65 well-known hearing loss genes in order to address hereditary hearing loss (SHEARER *et al.* 2013). It is foreseeable that some variants with unknown clinical significance would be detected at the same time, for which ethical issues may arise. Considering its high throughput detection capability, a targeted genomic enrichment MPS approach would be desirable as a second tier test after hot spot mutation screening in the newborns or other patients. Genetic counseling is indicated before and after the genetic testing.

In summary, our data show a high carrier frequency of these 15 common hearing loss mutations as one per 6.3 newborns carried at least one of them. It suggests that the hearing genetic screening has great potential of clinical application. One in 207 newborns (0.48%) of our cohort screened could be genetically diagnosed as having

hearing loss, which might facilitate follow-up and management after universal newborn hearing screening.

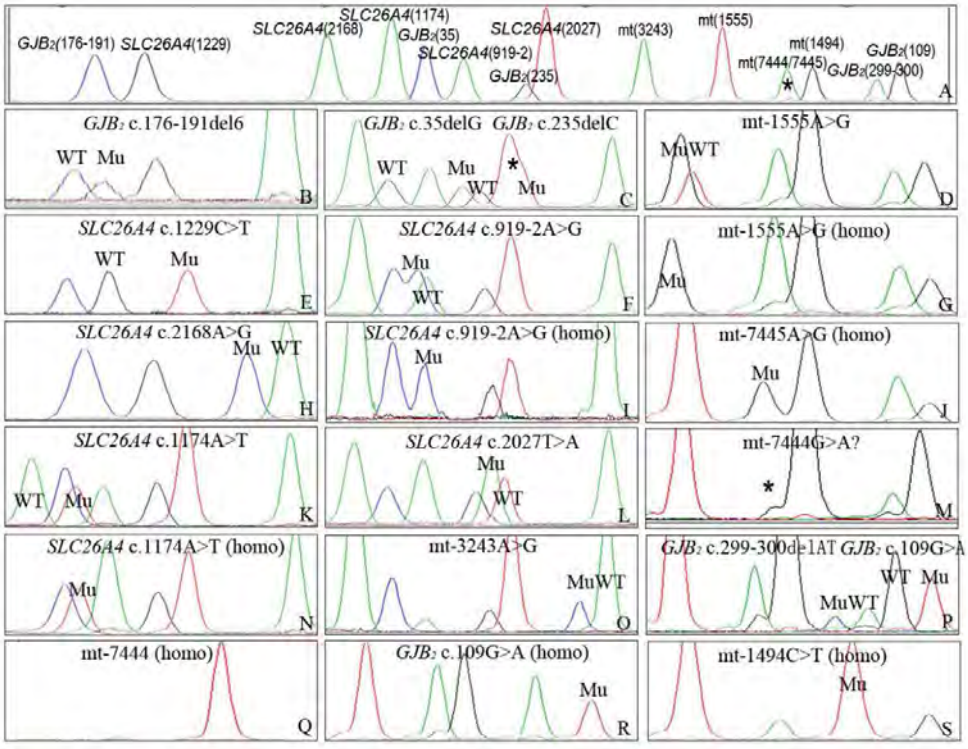


Figure 4-1 SNaPshot electropherogram plot of 15 hearing loss variants. A) shows the wild-type alleles of each variant appearing as different colored peaks at different sizes in the electropherogram plot. B) to S) each figures shows the wild-type or mutant genotyping results of these 15 SNPs. WT: Wild type allele; Mu: Mutant allele.

*(A) this peak can present both mt-7444 and 7445 genotyping: the wild-type of mt-7445A, green peak (A); mutant mt-7445G, the black peak (J); If there's no peak, it may indicate mt-7444 mutation in this site (M). For this scenario, we would further confirm it separately. (Q) Single red peak indicated this site was a homoplasmic mitochondrial mutations mt-7444G>A.

(C) *GJB2* c.235C wild-type shows a single black peak. *GJB2* c.235delC would produce a red peak which overlapped with the *SLC26A4* c.2027T wild-type red peak, eventually creating a wider red peak.

Table 4-3 Frequency of 15 common mutations detected in 5,800 Chinese newborns

Genes and Mutations	Genotype	No. of Subjects	Carrier Frequency
GJB2 c.35delG	total mutated alleles	1/1	0.017%
	WT/WT	5799	99.983%
	WT/35delG	0	0.000%
	35delG/35delG	0	0.000%
	35delG+235delC	1	0.017%
GJB2 c.109G>A	total mutated alleles/cases	610/597	10.290%
	G/G	5203	89.707%
	G/A	565	9.741%
	A/A	13*	0.224%
	109G>A+*GJB2* c.176-191del16	1*	0.017%
	109G>A+*GJB2* c.235delC	6*	0.103%
	109G>A+*GJB2* c.299-300delAT	1*	0.017%
	109G>A+*SLC26A4* c.919-2A>G	6	0.103%
	109G>A+*SLC26A4* c.2168A>G	2	0.034%
	109G>A+mt-3243A>G	1	0.017%
	109G>A+mt-7444G>A	1	0.017%
	109G>A+mt-7445A>G	1	0.017%
GJB2 c.176-191del16	total mutated alleles/cases	11/11	0.190%
	WT	5789	99.810%
	WT/del16	10	0.172%
	del16/del16	0	0.000%
	176-191del16+*GJB2* c.109G>A	1*	0.017%
GJB2 c.235delC	total mutated alleles/cases	109/109	1.879%
	WT/WT	5691	98.121%
	WT/delC	101	1.741%
	delC/delC	0	0.000%
	235delG+*GJB2* c.35delG	1*	0.017%
	235delC+*GJB2* c.109G>A	6*	0.103%

	235delC+*SLC26A4* c.919-2>G	1	0.017%
GJB2 c.299-300delAT	total mutated alleles/cases	17/17	0.293%
	WT	5783	99.707%
	WT/delAT	16	0.276%
	delAT/delAT	0	0.000%
	299-300delAT+*GJB2 c.*109G>A	1*	0.017%
SLC26A4 c.919-2A>G	total mutated alleles/cases	98/94	1.621%
	A/A	5706	98.379%
	A/G	83	1.431%
	G/G	4*	0.069%
	c.919-2A>G+*GJB2* c.109G>A	6	0.103%
	c.919-2A>G+*GJB2* c.235delC	1	0.017%
SLC26A4 c.1174A>T	total mutated alleles/cases	11/9	0.156%
	A/A	5791	99.845%
	A/T	7	0.121%
	T/T	2*	0.034%
SLC26A4 c.1229C>T	total mutated alleles/cases	1/1	0.017%
	C/C	5799	99.983%
	C/T	1	0.017%
	T/T	0	0.000%
SLC26A4 c.2027T>A	total mutated alleles/cases	2/2	0.034%
	T/T	5798	99.966%
	T/A	2	0.034%
	A/A	0	0.000%
SLC26A4 c.2168A>G	total mutated alleles/cases	20/20	0.35%
	A/A	5780	99.655%
	A/G	18	0.310%
	G/G	0	0.000%
	2168A>G+*GJB2 c.*109G>A	2	0.034%
mt-1494C>T	total mutated cases	1	0.017%

	C/C	5799	99.983%
	C/T	0	0.000%
	T/T	1	0.017%
mt-1555-A>G	total mutated cases	8	0.138%
	A/A	5792	99.862%
	A/G	3	0.052%
	G/G	5	0.086%
mt-3243A>G	total mutated cases	9	0.155%
	A/A	5791	99.845%
	A/G	8	0.138%
	G/G	0	0.000%
	m.3243A>G+*GJB2 c.*109G>A	1	0.017%
mt-7444G>A	total mutated cases	41	0.707%
	G/G	5758	99.276%
	G/A	0	0.000%
	A/A	40	0.690%
	m.7444G>A+*GJB2* c.109G>A	1	0.017%
mt-7445A>G	total mutated cases	3	0.052%
	A/A	5797	99.948%
	A/G	0	0.000%
	G/G	2	0.034%
	m-7445A>G+*GJB2 c.*109G>A	1	0.017%
Total		942/923	15.914%

*: subjects genetically diagnosed as having hearing loss screened by this SNaPShot genotyping method.

Chapter 5 Establishment of a targeted genome enrichment and massively parallel sequencing (TGE-MPS) platform for hereditary hearing loss

5.1 Introduction

Innovation of sequencing technology speeds up our understanding on human genomic biology. It is accompanied by the development of new perspectives in studying diagnosis, treatment and intervention of human diseases (LANDER 2011). Genetic diagnostics has strongly accelerated since human genome sequence was released in 2001 (LANDER *et al.* 2001). Technological breakthrough allows the application of whole genome sequencing (WGS) to analyze individual genomes. The major challenge in genetic diagnosis has shifted from sequencing data generation to the computational analysis and data interpretation (GREEN and GUYER 2011). Even though the cost of sequencing is dropping dramatically, currently WGS is not affordable for most clinical applications. Sequencing a target subset of the human genome is a compromise, such as coding region sequencing (whole exome sequencing, WES), or a limited set of genes related to certain group of diseases. WES has become a powerful tool to discover new genes and even achieve a rapid diagnostic method to differentiate genetic disorders, which favors fewer unnecessary empirical treatments, specific genetic counseling, and potentially a reasonable prediction of disease prognosis. Yang et al. demonstrated a 25% positive diagnosis in clinical exome sequencing of 250 individuals even with a broad range of phenotypic presentations (YANG *et al.* 2013). These emerging data support WES as promising for patients with a wide range of clinical features (NEED *et al.* 2012).

For certain heterogeneous genetic diseases with relatively specific, well defined phenotypes such as hereditary hearing loss, sequencing of a selected candidate gene panel is proven to be a more cost effective approach in the clinical setting (SHEARER *et al.* 2010; AMSTUTZ *et al.* 2011; REDIN *et al.* 2012). Meanwhile, the interpretation of variants would be more phenotype specific and straightforward. However, most studies only targeted limited causative hearing loss genes. In this section, we established a comprehensive test namely the targeted genome enrichment sequencing approach for hereditary hearing loss by screening 252 candidate genes and the mitochondrial genome.

5.2 Materials and Methods

Seventeen nonsyndromic hearing loss patients were recruited in this study. Three of the 17 patients were already identified with known pathogenic mutations and served as positive controls. The other fourteen patients were excluded from any known mutations on the coding region of *GJB2* and *GJB6*. The genomic DNA paired-end sequencing libraries of these samples were prepared following the protocols described in the 3.4. We employed the Agilent SureSelect target enrichment library to capture targeted genomic sequence (3.7). Captured DNA library was sequenced 90bps paired-end on the Illumina Genome Analyzer IIx (Illumina, San Diego, CA). Data Analysis was described in 3.9.

5.3 Results

5.3.1 Coverage and read depth of sequencing data

Raw sequencing data of these 17 samples were generated from 130Mb to 1004Mb (Table 5-1). Since the targeted genome of our library design included both nuclear and mitochondrial genomes, we calculated mitochondrial (mt) genome and targeted

nuclear genome regions separately. After removing duplicate reads that represent potential PCR artifacts, the average read depth for targeted nuclear genome regions was from 24x to 158x and average mitochondrial read depth 2312x-24560x (Table 5-1). In general, 92% of the targeted bases were covered at least once. We also observed that once the output reached 600Mb, coverage of the samples consistently reached 96% of the targeted region (Table 5-1). More than 20%-60% of raw data mapped to the mitochondrial genome, and its coverage is 100% in all samples. By plotting the relationship between the average read depth and coverage of target nuclear genome region (Figure 5-2), we noticed that once the average depth reached 100 fold, 96% of targeted bases were covered at least once, and 85% were covered >= 10 times, which is normally considered as being sufficient for variant calling (NG *et al.* 2009; CLARK *et al.* 2011).

5.3.2 Accuracy of variant calling

Accuracy of the variant calling is essential to assess the performance for this high throughput method. According to variant calling of sequencing data described in 3.9, around 1472-2543 SNVs, 153-400 indels were called among the 17 samples (Table 5-2). Following the variant filtering step, 10-23 variants were retained. Considering the requirement of computational power and bio-availability of the DNA resource, it would not be realistic to validate all variants for each sample to assess the accuracy of variant calling. So, we applied Sanger sequencing to validate firstly ten variants with medical interest from one sample (E2073_1) with medium average read depth (61x), according to the following criteria: (1) at

least 10 reads; (2) a variant sequence supported by >=20% of the total reads considered as a heterozygous genotype, <20% or 80% considered as homozygous

genotype; and (3) variants with quality score >=20. These ten variants (100%) were confirmed by Sanger sequencing

5.3.3 Reproducibility

DNA samples from three patients with four known mutations were sequenced as positive controls. Based on the analysis pipeline described in 3.10, all four known mutations were detected and also prioritized as candidate pathogenic variants (Table 5-4). Therefore, the reproducibility of this TGE-MPS method we developed is 100% reproducible.

Table 5-1 Summary of sequencing data (ranked in ascending raw data yield)

Sample ID	Raw Data Yield (Mb)	Data Mapped to mitochondrial(mt) genome		Data Mapped to targeted nuclear genome region		Average Depth (excluding MT data)	MT Average Depth	Coverage of Targeted Region (1x)	Coverage of Targeted Region (10x)	Coverage of Targeted Region (20x)
		Megabases	% of total raw data	Megabases	% of total raw data					
E2082_3	137.82	37.25	27.0	38.36	27.8	24.03	2312	92.71	67.8	52.08
E2073_3	151.98	48.78	32.1	44.53	29.3	28.48	3027	92.41	69.07	56.03
E2082_1	243.25	121.55	50.0	52.37	21.5	27.85	7543	93.13	70.71	56.3
E2082_9	316.15	85.93	27.2	92.08	29.1	61.41	5333	95.38	80.53	71.41
E2073_1	348.29	113.14	32.5	99.16	28.5	61.77	7022	94.85	80.37	71.67
E2073_2	353.74	104.24	29.5	108.06	30.5	66.14	6469	95.01	80.37	71.67
E2082_7	364.02	86.14	23.7	109.66	30.1	74.26	5346	95.47	82.11	74.15
E2073_5	369.31	75.46	20.4	109.26	29.6	72.52	4683	95.06	81.9	73.72
E2082_8	384.6	89.2	23.2	112.7	29.3	76.17	5536	95.54	83.15	75.76
E2082_2	444	268.6	60.5	81.12	18.3	37.69	16669	93.80	75.02	63.54
E2073_6	498.8	169.09	33.9	137.97	27.7	85.84	10494	95.34	83.15	75.76
E2082_4	624.54	172.58	27.6	193.05	30.9	122.15	10710	96.13	86.83	81.14
E2073_4	696.42	201.4	28.9	264.42	38.0	138.59	12499	95.99	86.19	80.62

E2195_5	767.49	245.55	32.0	218.65	28.5	137.81	15239	96.04	86.16	80.43
E2195_7	845.77	367.69	43.5	190.17	22.5	109.74	22819	96.05	85.49	79.49
E2082_10	973.04	382.38	39.3	231.4	23.8	154.24	23731	96.10	84.01	76.76
E2195_6	1007.46	395.74	39.3	263.62	26.2	158.5	10845	96.31	87.2	82.14

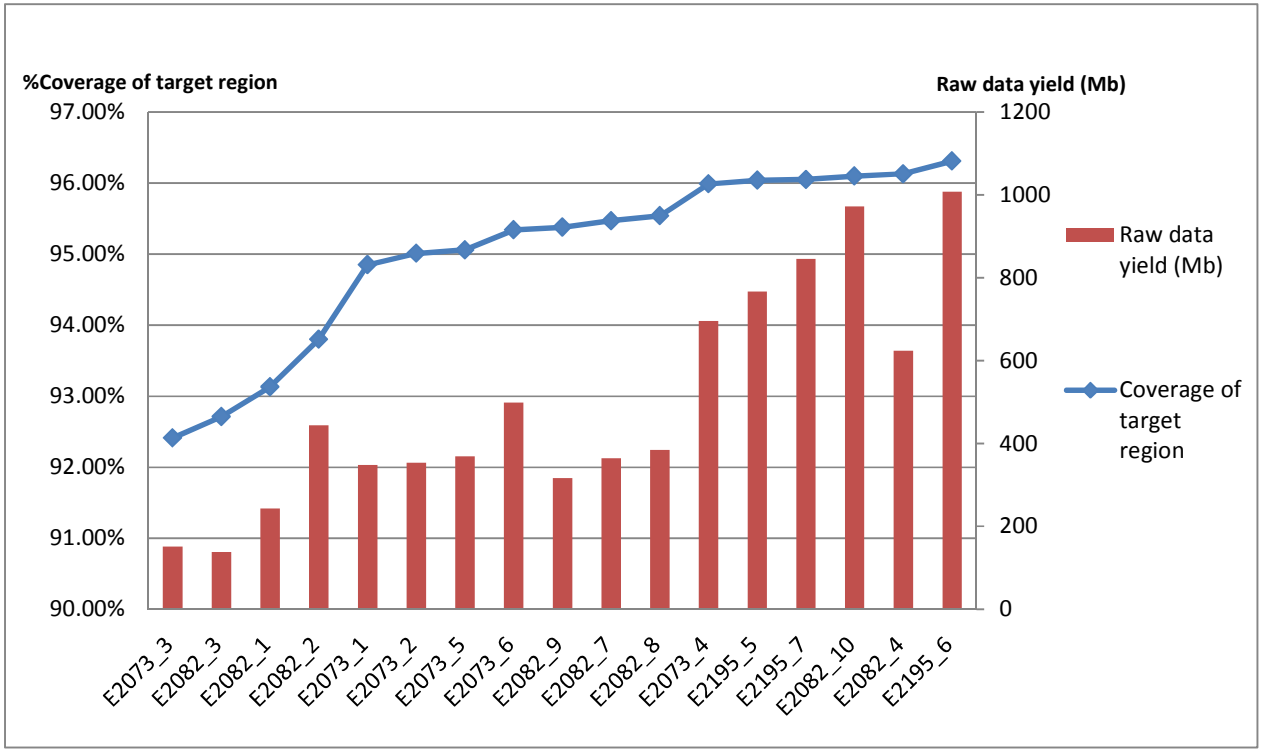


Figure 5-1 Relationship between the coverage of targeted region and output raw data for each sample

X axis: sample ID, Left Y axis: %Coverage of targeted region, Right Y axis: raw data (Megabases).

Red bars represented output raw data for each sample. Blue line represents the coverage of each sample at read depth 1X of targeted base.

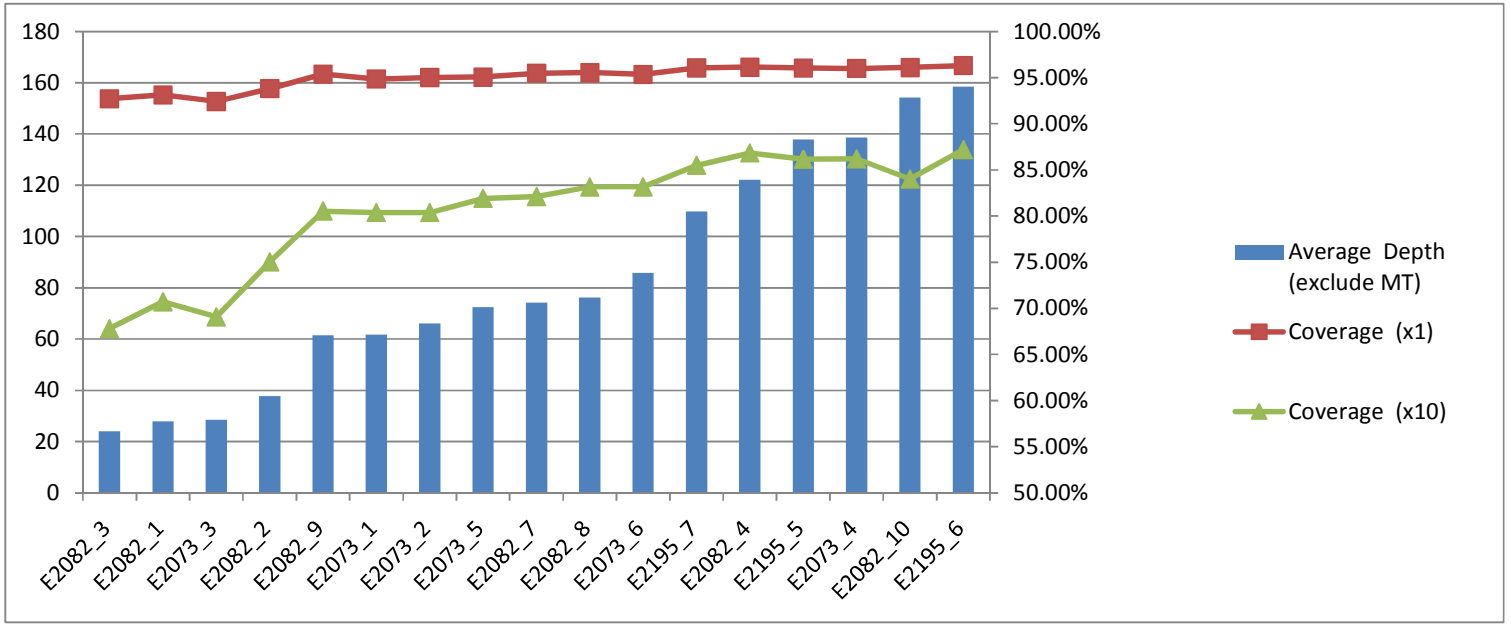


Figure 5-2 Relationship between the average read depth and the coverage of targeted nuclear genome region

X axis: sample ID. Left Y axis: read depth of nuclear genome. Right Y axis: %Coverage of nuclear genome.

Blue bars represented average read depth of each samples. Colored lines represented the coverage of each sample at read depth 1X, and 10x for targeted base.

Table 5-2 Summary of variant calling

Sample ID	total no. of SNPs	Candidate SNVs	total no. of Indel	Candidate Indel	Novel SNVs	% of Novel SNVs	Novel SNVs in candidate SNVs	Mitochondrial Variant Calling SNVs+Indel
E2082_3	1498	11	255	1	75	5.0%	9	13+0
E2073_3	1472	9	232	1	76	5.2%	6	36+0
E2082_1	1614	19	242	2	100	6.2%	15	0+0
E2082_9	2154	10	316	2	150	7.0%	9	51+0
E2073_1	1984	17	296	1	156	7.9%	13	10+0
E2073_2	1984	16	153	1	156	7.9%	13	8+0
E2082_7	2300	11	373	2	171	7.4%	9	47+0
E2073_5	2200	18	334	0	194	8.8%	13	32+0
E2082_8	2255	19	347	1	198	8.8%	15	42+0
E2082_2	1865	13	399	0	101	5.4%	12	0+0
E2073_6	2145	17	300	5	202	9.4%	14	2+0
E2082_4	2452	12	367	0	234	9.5%	10	0+0

E2073_4	2525	21	400	2	281	11.1%	17	1+0
E2195_5	2520	12	392	2	251	10.0%	9	62+0
E2195_7	2345	15	321	0	205	8.7%	13	37+0
E2082_10	2210	15	326	1	199	9.0%	11	3+0
E2195_6	2543	18	354	1	261	10.3%	14	66+0

. Novel SNVs: single nucleotide variants not reported in dbSN136

Table 5-3 Validation of candidate variants of E2073_1

GENE	Coordinate (hg19)	Genotype (Ref /Alt)	Reads of each allele	Reads ratio of alternative allele	cDNA Change	Sanger Validation
ALMS1	chr2: 73675227..73675228	insCTC	Ref 65/Alt52	44.44%, hetero	c.1570_1571insCTC	Yes
ASPA	chr17: 3402272	G/A	G46/A48	51.06%, hetero	C832G>A	Yes
COL4A3	chr2:228121101	G/T	G40/37T	48.05%, hetero	c.976G>T	Yes
LOXHD1	chr18:44181250..44181252	delCCT	Ref 21/Alt8	27.59%, hetero	c.1062_1064delCCT	Yes
MPRS16	chr10: 75011557	C/T	C36/18T	33.33%, hetero	c.238C>T	Yes
MYO15A	chr17:18041451	T/C	T11/C8	42.11%, hetero	c.4898T>C	Yes
NEFL	chr8: 24811065	delC	Ref16/Alt2	88.89%, homo	c.1414delC	Yes
PTPRQ	chr12:80838564	C/T	C8/ T11	57.89%, hetero	c.98C>T	Yes
TMIE	chr3:46751074..46751076	delAAG	Ref 12/Alt6	33.33%, hetero	c. 367_369delAAG	Yes
TRIOBP	chr22:38119213	G/A	G10/7A	41.18%, hetero	c.650G>A	Yes

Table 5-4 Characteristics of the three positive controls with four known variants detected by TGE-MPS method

Sample ID	**Gene with Known Variants**	**Detected by TGE-MPS**	**Read counting (wild allele/ mutant allele) , mutant reads ratio**
E2195_5	*GJB2* c.299_300delAT, p.H100RfsX14, het	Yes	W78/V59, 43%
	GJB2 c.235delC, p. L79CfsX3, het	Yes	W87/V46, 35%
E2195_6	Mt-1494_C/T	Yes	C0/T255, 100%
E2195_7	*SLC26A4* c.919-2 A->G, Homo	Yes	A0/G203, 100%

5.4 Discussion

In this study, we established a TGE-MPS approach for hereditary hearing loss, which covered 252 HHL related genes plus the mitochondrial genome. The accuracy of variant calling by the in-house established bioinformatics analysis pipeline was estimated to be 100% based on validation of 10 candidate variants in one sample E2073_1.

An ideal sequencing platform would cover the targeted region completely and detect all polymorphic alleles (SIMS *et al.* 2014). However, the sequencing method could introduce sequence errors and some targeted regions lack coverage because of genome structure (such as repetitive regions, GC-rich regions), or biases introduced by sample preparation for sequencing. Nevertheless, sequencing errors could be overcome by increasing the read depth (SIMS *et al.* 2014). Based on the validation work in this study, variant sites supported by at least 10 reads and its sequence supported>= 20% reads might effectively distinguish the sequence error and true variants. In order to address the false negative or false positive rate of the variant calling in our TGE-MPS approach, we will apply this established method to two well genotyped DNA samples.

For our data, we demonstrated that by increasing read depth, we could improve coverage based on the Figure 5-1 and Figure 5-2. We proposed that 600 Mb output raw data or 100X average read depths should be considered as minimum to ensure the 85% of target genome region is covered at least 10 times. However, generating short reads could not cover all targeted regions since short reads cannot be unambiguously aligned to those long repetitive regions in the human genome. Meanwhile, the more raw data generated, the more data mapped to the mitochondrial

genome. The copy number of the mitochondrial genome was 100-1000 times the diploid nuclear genome. Our data indicated that eArray could not adjust the bait titer for the mitochondrial genome accordingly. This leads to extremely high coverage of the mitochondrial genome which did not have a benefit for variant calling. Therefore, we are aiming at achieving a more even distribution of read depth between the mitochondria genome and nuclear genome for future optimization of our TEG-MPS design.

Chapter 6 Evaluate the clinical application of the TGE-MPS approach

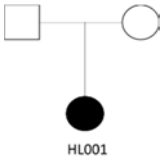
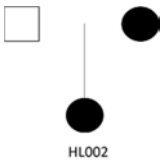
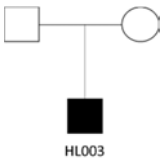
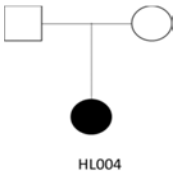
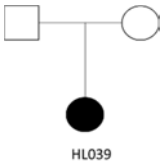
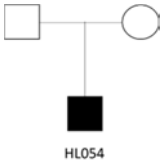
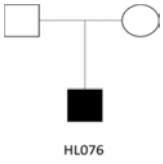
6.1 Introduction

Hereditary hearing loss is a group of disorders with a well-defined phenotype characterized with unparalleled genetic heterogeneity. Besides common mutations studied in chapter 3, most pathogenic genes do not have hot spot mutations or founder mutations, suggesting unrelated patients possess different mutations. In this study, we applied our TGE-MPS method on nonsyndromic hearing loss patients who screened negative for 15 common mutations by the SNapShot method we established. Meanwhile, to address the overrepresented mitochondrial data in the Agilent SureSelect enrichment method, in this study we also designed another in-solution target capture method from NimbleGen, USA. During the two years from the first Agilent design in 2011 to the new NimbleGen design in 2013, 17 more new genes were identified as hereditary hearing loss genes. We included these new genes into our NimbleGen design.

6.2 Materials and methods

Seven nonsyndromic patients were recruited. All patients denied exposure to ototoxic drugs. No other congenital abnormality was reported. Family pedigree and clinical information for each patient were depicted as following (Table 6-1). Two in-solution capture methods were applied in this study (Table 6-2). Targeted genomes were sequenced paired-end 90bps, Illumina 2500, rapid run mode.

Table 6-1 Patients' clinical information

Patient ID	Clinical information	Pedigree of patient family
Patient HL001	female, born in Jan 2011, bilateral severe to profound sensorineural hearing impairment (SNHL), prelingual onset, MRI brain/cochlea reported normal	
Patient HL002	female, 27 years old, bilateral severe SNHL, congenital onset, no DM/DI features (*diabetes* mellitus/*diabetes* insipidus), mother of patient also affected	
Patient HL003	male, born in Jun 2004, congenital onset, bilateral severe SNHL	
Patient HL004	female, born in Nov 2010, bilateral profound SNHL, congenital onset	
Patients HL039	female, born in 2000, presented bilateral severe to profound SNHL, onset age 5 years old	
Patients HL054	male, born in 2002, congenital bilateral severe to profound SNHL	
Patients HL076	male, born in 2000, bilateral severe sensorineural hearing impairment, around 2 years old onset	

6.3 Results

6.3.1 Summary of sequencing data

According to our previous study and other targeted sequencing studies (Table 6-7), we increased the raw data for each samples, and the average read depth of targeted nuclear genome regions ranged from 300-800 fold. The increased data output enhanced the target region coverage impressively to 98%, and 93.4%-97.4% of target bases were covered at least 10 times. Compared with the Agilent capture system, the sequencing data mapping to the mitochondrial genome following using the NimbleGen system targeted enrichment method were much reduced. The ratio of average depth (excluding MT data) to MT average depth in HL039, HL054, HL076 suggested that the NimbleGen target enrichment method achieved a more even distribution of read depth between mitochondria and nuclear genomes.

Table 6-2 Comparison of parameters in Agilent and NimbleGen in solution capture system

	Agilent SureSelect Target Enrichment	**NimbleGen SeqCap EZ Choice Library**
Baits property	120bps, cRNA, in-solution	50-105bps, DNA , in-solution
Design strategy	online eArray	proprietary tool design
Targeted genes	252 genes + whole mitochondria covering 54 nonsyndromic hearing loss genes	269 genes + whole mitochondria, covering 65 nonsyndromic hearing loss genes
Sample ID	HL001, HL002, HL003, HL004	HL039, HL054, HL076

Table 6-3 Summary of sequence data

Sample ID	Raw Data Yield (Mb)	Data Mapped to MT Genome		Data Mapped to targeted nuclear genome region		Average Depth (excluding MT data)	MT Average Depth	%Coverage of Targeted Region (1x)	%Coverage of Targeted Region (10x)
		Megabases	% of total raw data	Megabases	% of total raw data				
HL001	3109.45	904.2	29.1	607.86	19.5	430.74	54808.03	98.03	94.64
HL002	2521.63	752.4	29.8	494.39	19.6	349.88	45610.76	97.92	93.6
HL003	2319.78	589.05	25.4	420.88	18.1	296.98	35675.93	97.90	93.4
HL004	3369.47	983.07	29.2	490.48	14.6	347.66	59581.93	98.16	95.53
HL039	1843.45	15.12	0.82	719.38	39.02	602.07	914.57	98.2	97.06
HL054	1854.26	25.95	1.40	708.9	38.23	593.3	1569.99	98.1	97.05
HL076	2591.95	15.5	0.60	962.42	37.13	805.47	937.88	98.3	97.39

6.3.2 Discovery of novel variants

6.3.2.1 *WFS1*

HL002 family (Figure 6-1) has both an affected daughter and mother, but no other family members with hearing loss reported. The hearing loss was consistent with autosomal dominant or a maternally inherited pattern. Patients presented with congenital bilateral severe sensorineural hearing impairment. No DM (diabetes mellitus)/DI (diabetes insipidus) features or other optical phenotype. Previous genetic tests excluded any mutations in the coding regions of *GJB2* and *GJB6* and mitochondrial 1555G>A. In genomic DNA from the index patient HL002, two variants were observed in the *WFS1*, Wolfram syndrome type 1 gene, corresponding to a known mutation c.173C>T, p. Ala58Val (GOMEZ-ZAERA *et al.* 2001) and a novel variant c.2527C>T, p.Lys843Gln (Table 6-4). Variant reads were 56% and 52% of total reads, suggesting heterozygosity for both alleles in the patient. Sanger sequencing further confirmed the mother of the patient who was also affected and carrying these two heterozygous mutations. These two variants could be on the same allele of the maternal genome.

WFS1-related disorders included Wolfram syndrome (WFS, OMIM 222300) and *WFS1*-related low frequency sensorineural hearing loss (LFSNHL). WFS is a progressive neurodegenerative disorder characterized by onset of diabetes mellitus and optic atrophy before age 16 years, and typically associated with sensorineural hearing loss and progressive neurologic abnormalities (http://www.ncbi.nlm.nih.gov/books/NBK4144/). Median age of death is 30 years old. *WSF1* encodes wolframin, a hydrophobic transmembrane protein that is related to Wolfram syndrome or LFSNHL associated with DFNA6/14/38. *WSF1* c.173C>T,

p. Ala58Val was reported before in a Spanish WFS patient who also carried another homozygous nonsense mutation, c. 1558C > T, Q520X. This nonsense mutation truncated at protein residue 520 when a glutamine codon (CAG) is changed to a stop codon (UAG). It implied that c.173C>T, p. Ala58Val might represent a minor alteration in wolframin (GOMEZ-ZAERA *et al.* 2001). SIFT and Polyphen also predicted this as tolerated or benign. *WSF1* c.2527C>T caused amino acid substitution, changed p.Lys843Gln. This amino acid was highly conserved across vertebrates (Figure 6-2). SFIT score was 0.04, deleterious. It has not been reported in 1000 Genomes, Exome Variant Server, HGMD or dbSNP databases or LOVD before. An adjacent missense mutation (A844T) was detected in a Japanese family with bilateral symmetrical LFSNHL (NOGUCHI *et al.* 2005). This amino acid was also located in the C terminal domain of the protein which played a key role in calcium regulation of the inner ear. Considering the patients' clinical manifestation, this novel heterozygous variant c.2527C>T (p.Lys843Gln) might be associated with the hearing loss in this family. This variant should be validated in other extended family members if possible to see whether it was a *de novo* mutation in the mother's genome, or segregated in the family. Screening *WFS1* in patients with low frequency hearing loss might help us to find other patients with the same variant. If biological functional study of the mutant protein's proved to interrupt the protein normal function, it might further confirm its clinical significance.

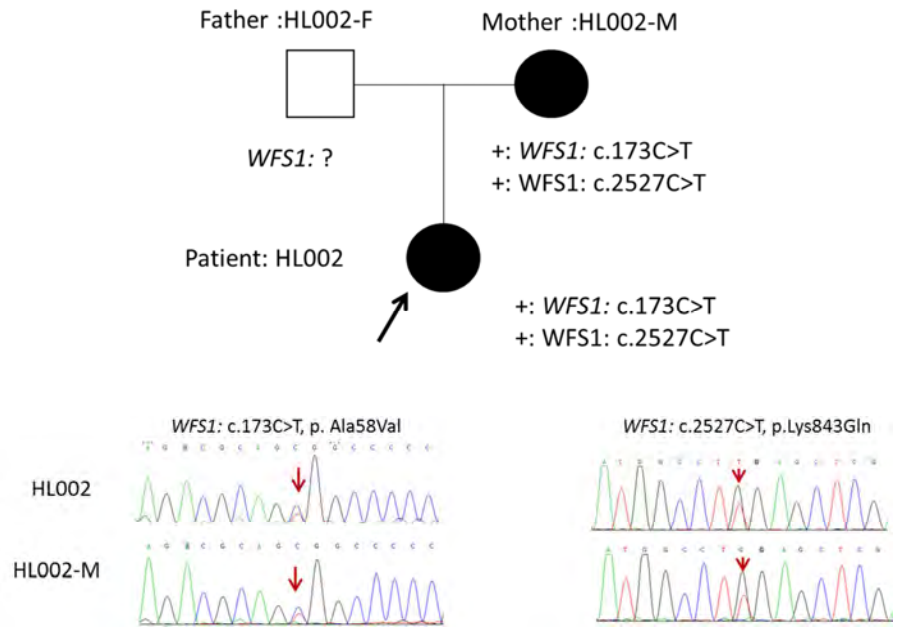


Figure 6-1 Pedigree of HL002 family and variant validation by Sanger seuqencing

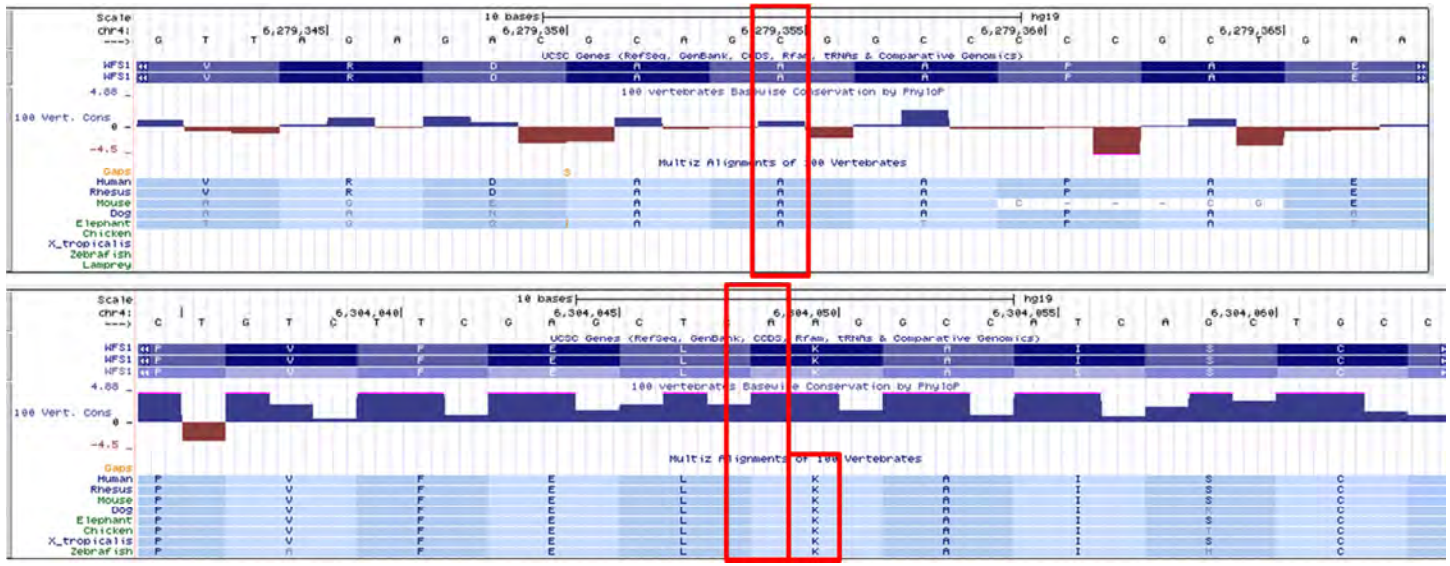
Table 6-4 The candidate variants of HL002

Sample	Gene	Locus (OMIM)	Nucleotide change	Protein Change	Read depth (Ref/Alt)	MAF	HAPMAP-CHB	GO-ESP MAF (%) (EA/AA)	ClinVar	Validation
HL002	*WFS1*	DFNA6	Chr4:6279355 C>T · c.173C>T	p. Ala58Val	55C/T70	-	-	-	-	yes
	WFS1	DFNA6	Chr4:6304049 A>C c.2527A>C	p.Lys843Gln	A109/100C	-	-	-	-	yes

Figure 6-2 Evolutionary conservation analysis of the candidate variants detected in patient HL002

Chr4:6279355
c.173C

Chr4:6304049
c.2527A



6.3.2.2 *MYO15A*

Patient HL004 (Figure 6-3), female, was born in Nov 2010, presented with bilateral profound sensorineural hearing impairment, congenital onset. No other congenital abnormality was reported. Both parents had normal hearing. It could be either autosomal recessive nonsyndromic hearing loss or autosomal dominant HL caused by *de novo* mutation. Previous genetic tests excluded any known mutations in the coding region of *GJB2* or *GJB6* and mitochondrial 1555G>A. After sequencing the capture targets of this case, two novel variants, c.104G>A and c.5917G>A were found on *MYO15A* which is associated with DFNB3 hearing loss (Table 6-5).

MYO15A encodes an unconventional myosin, myosin XVa. Mouse studies showed this protein is necessary for actin organization in the hair cells. These two mutations were not reported in 1000 Genomes, Exome Variant Server, HGMD or dbSNP databases or LOVD. The C.104G>A substitution occurs in exon 2 and results in an amino acid substitution Arg to Gln (p.Arg35Gln) on N-terminal extension of myosin15a. The Arg35 is conserved across mammals and most vertebrate species, including human, chimpanzee, dog, mouse, chicken, and X.tropicalis (Figure 6-4). The c.5917G>A also caused a single amino acid substitution p.Ala1973Thr in the neck or regulatory domain of the protein. Even though SIFT predicted this amino acid substitution was tolerated, the Ala1973 was highly conserved in vertebrate species (Figure 6-4). Sanger sequencing validated these two variants and also demonstrated that the two parents were heterozygous carriers of each variant. These data, together with the clinical presentation of the affected children indicated that the compound heterozygous *MYO15A* mutations might be the cause of the ARNSHL in

this family. However, further study is needed to establish the clinical significance of these variants.

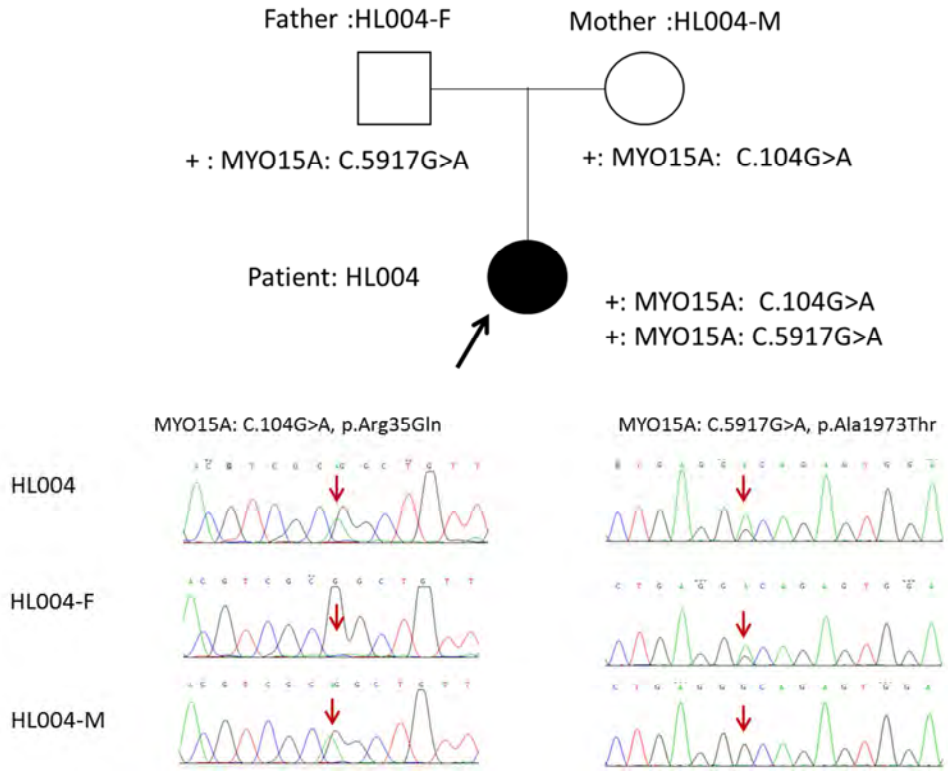


Figure 6-3 Pedigree of HL004 family and variant validation by Sanger seuqencing

Table 6-5 The candidate variants of HL004

SampleID	Gene	Locus (OMIM)	Nucleotide change	Protein Change	Read depth (Ref|Alt)	MAF	HAPMAP-CHB	GO-ESP MAF (%) (EA/AA)	ClinVar	Validation
HL004	*MYO15A*	*DFNB3*	*chr17: 18022218 c.104G>A*	*p.Arg35Gln*	*63G/ A64*	-	-	-	-	*Yes*
	MYO15A	DFNB3	chr17: 18046886 c.5917G>A	p.Ala1973Thr	32G /A57	-	-	-	-	Yes

chr17: 18022218
c.104G

chr17: 18046886
c.5917G

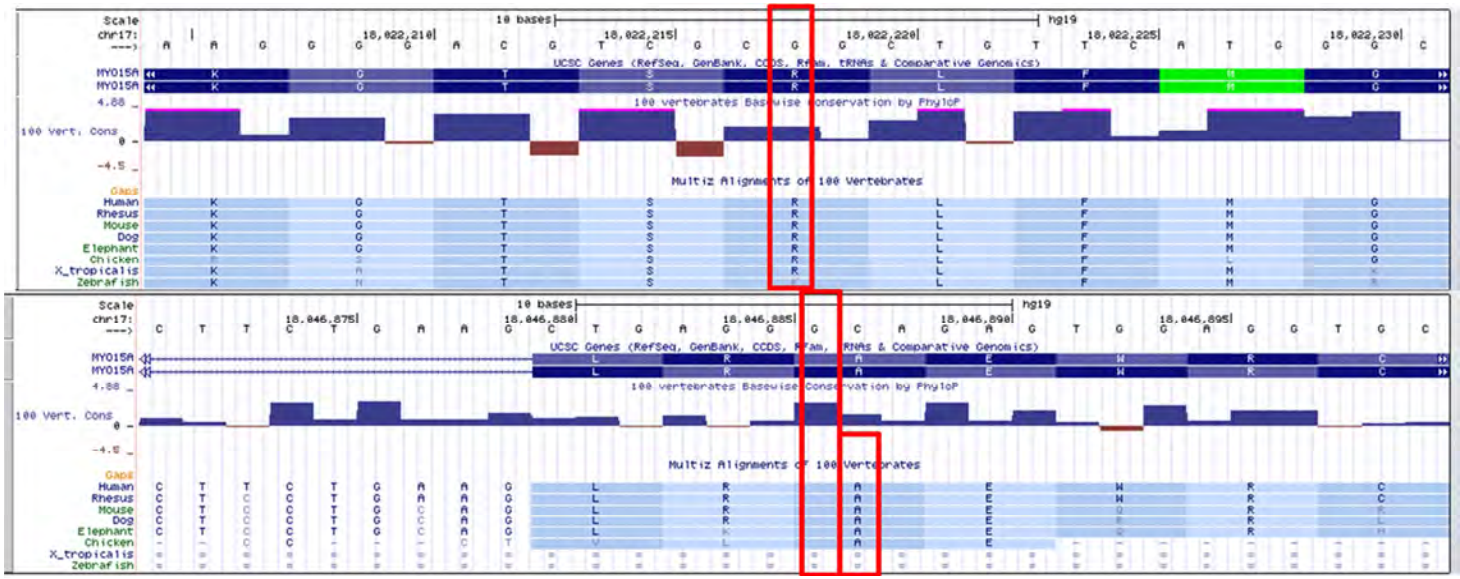


Figure 6-4 Evolutionary conservation analysis of candidate variants detected in patient HL004

6.3.2.3 *DFNA5* deletion

Patient HL039 was a female, 9 yrs, Han-Chinese, nonsyndromic hearing loss, onset age at 5 years old (Figure 6-5). Her language was reported as normal development. No family history or ototoxic medication history was reported. There were four candidate variants filtered out and shown in Table 6-5. Two variants c.2353 A>C, c.2359 G>T on *OTOA* were close to each other and on the same exon. Variant reads ratios were 20.9%, and 20.8% suggesting both were heterozygous alleles. Because the mother was identified as a heterozygous for these two variants by SNapShot genotyping method, these two *OTOA* variants were considered to have no relationship with the hearing loss. We also noticed that *OTOA* has one pseudogene LOC653786, located on chr16:22,557,019-22,588,186. Using the BLAST method to align the 201 nucleotides including these two variants with the same exon, we found that these two variants were actually aligned to the pseudogene sequence (Figure 6-6). We further investigated whether these two variants also would be detected in two other samples (HL054, HL076) of the same batch. One variant chr16:21747639G>T was also detected in HL076, with alternative allele read ratio of only 16%, which was less than 20%, and not considered as a true variant.

The other two variants on *OTOF* c.145 C>T and *DFNA5* c.619 G>A might not be defined as pathogenic mutations for the patients because of their carrier frequency in the population or already reported as benign variants in a public database. When we reviewed all of the variants detected on the *DFNA5*, they were all unexpected to be homozygous. The loss of heterozygosity implied that there might be a copy loss in this genome region. We therefore further used SYBR Green quantitative PCR

method for detect the copy number of *DFNA5*. Compared with an internal control, we found there was one copy loss in this patient while the mother still had two copies (Figure 6-7). Because the father's genome was not available for testing, we could not conclude that deletion of this gene was related to the hearing loss of the patient. Van Laer reported an insertin/deletion in intron 7 causing skipping of exon 8 of *DFNA5* mRNA and resulting in hearing loss in a Dutch family (VAN LAER *et al.* 1998). In their study, Dfna5 knock out mice with deletion of exon 8 (Dfna5-/-) were established to mimic the genotype of their patients. The result showed that outer hair cells of Dfna5-/- mice were significantly different than outer hair cells of wild type mice (Dfna5+/+). However, frequency specific Auditory Brainstem Response showed no difference between wild-type and knockout mice (VAN LAER *et al.* 2005). Two patients with 7p14 deletion encompassing *DFNA5* were also reported (DEVRIENDT *et al.* 1999; DUNO *et al.* 2004). In these two studies, reports of one patient at five years old (DEVRIENDT *et al.* 1999) and the other at 21 years (DUNO *et al.* 2004), did not specifically mention whether these patients had normal hearing or subtle hearing loss. Considering the onset age of DFNA5 mutation in the family varied between 5 and 15 years (VAN LAER *et al.* 1998), or presumably congenital (BISCHOFF *et al.* 2004), we could not exclude hearing loss in those two patients with 7p14 deletions. Currently we still cannot make any direct relationship between the *DFNA5* deletion and hearing impairment in HL039.

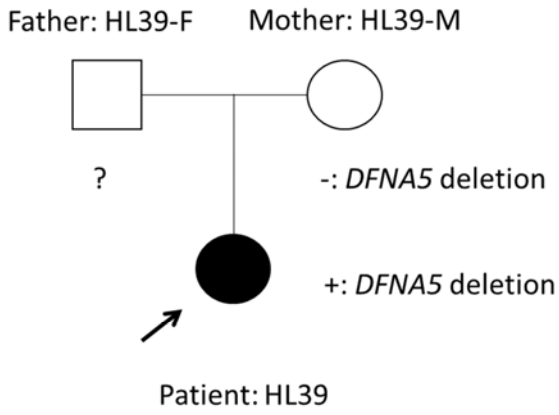


Figure 6-5 Pedigree of HL039 family

A

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHRO	STRAND	START	END	SPAN
browser details	YourSeq	201	1	201	201	100.0%	16	+	21747539	21747739	201
browser details	YourSeq	192	1	200	201	98.0%	16	+	22563662	22563861	200

B

Side by Side Alignment

```
00000001 ggttttaattattattcctcttctttcatcttctttcacacagtttcctg 00000050
>>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>>
22563662 ggttttaattattattcctcttctttcatcttctttcacacagtttcctg 22563711

00000051 agatccttctgcaagcagcttccaagatggccaggaccctgccactaaa 00000100
>>>>>>>> |||||||||||||||||||||||||||||||||||||||||||| ||||| >>>>>>>>
22563712 agatccttctgcaagcagcttccaagatggccaggaccctgcccctaaa 22563761

00000101 gattcctctgggctgtctttcagtctgttcggaacagcagtgataagat 00000150
>>>>>>>>  |||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>>
22563762 tattcctctgggctgtctttcagtctgttcggaacagcagtgataagat 22563811

00000151 ccccagctatgaccctatgcctggtgagtgttttcagggtatctgagcca 00000200
>>>>>>>> |||||||| ||||||||||||||||||||||| ||||||||||||||| >>>>>>>>
22563812 ccccagctctgaccctatgcctggtgagtgtttccagggtatctgagcca 22563861
```

Figure 6-6 BLAT alignment of 201 bps sequence (chr16:21747539-21747739)

A: BLAT search results of the a 201 bps sequence (chr16:21747539-21747739). The first line of results was the original region of this 201 bps sequence on chr16:21747539-21747739. The second line showed another region (chr16:22563662-22563861) highly identical (98.0%) with the orignial sequences. B: Side by side alignment of chr16:22563662-22563861. chr16:22563662-22563861 was within *OTOA* pseudogene LOC653786 (chr16:22,557,019-22,588,186). Orange boxes highlight the two sites had consistent variant sequences with the two candidate variants on *OTOA* chr16:21747633A>C, chr16:21747639 G>T.

Table 6-6 The candidate variants of HL039

HL039	**Locus (OMIM)**	**Nucleotide change**	**Protein Change**	**Read depth (Ref/Alt)**	**MAF (1000Genome)**	**HAPMAP -CHB**	**GO-ESP** MAF (%) (EA/AA)	**ClinVar**	**Validation**
OTOA	DFNB22 chr16:21747633	c.2353 A>C	p.Thr785Pro	197A/52C, 20.9%	-	-	-	-	Interruppted by OTOA pseudogene (PGOHUM00000249020)
OTOA	DFNB22 chr16:21747639	c.2359 G>T	p.Glu787*	198G/52T, 20.8%	-	-	-	-	
OTOF	DFNB9, AUNB1 chr2:26750782	c.145 C>T	p.Arg49Trp	126C/124T, 49.6%	A=0.008	-	0.28%/0.73%	Benign, observed in 5 families	-
DFNA5	DFNA5, chr7:24756951	c.619 G>A	p.Val207Met	0G/194A, 100%	T=0.097	T=0.159	9.3%/4.9%	Benign, observed in 19 families	-

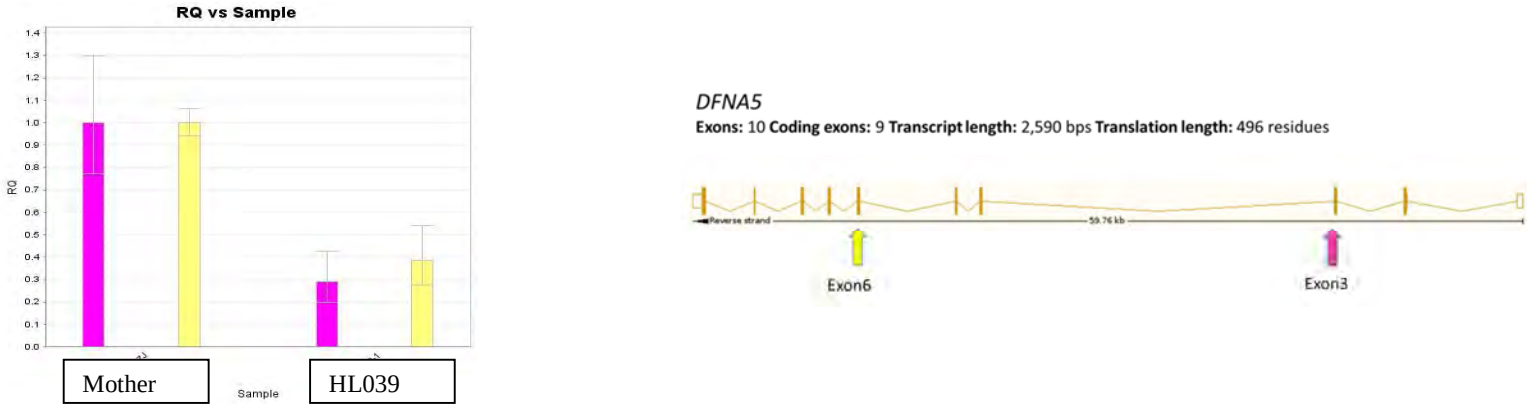


Figure 6-7 Quantitative analysis of copy number of DFNA5 on HL039

Primers were designed to amplify exon 3 and exon 6 of DFNA5 (right). SYGR PCR was applied to quantify relatively the fluorescence signals of exons 3 and 6 PCR products which were normalized with an internal control. This result showed normalized signals intensity of PCR products of *DFNA5* exon 3 and exon 6 in patient were half of that in the mother, which suggested one copy loss of *DFNA5* in the patient genome.

6.4 Discussion

In the first part of this study, we established a TGE-MPS platform for hereditary hearing loss. The accuracy of variant calling by the in-house bioinformatics analysis pipeline was estimated to be 100% based on the 10 candidate variants in one sample E2073_1. However, coverage at 10 fold read depth also showed 15% of targeted regions were missed. In the second part, based on other group of studies and our previous study, we increased the raw data to achieve a better coverage. As reaching 300 fold average read depth, 93.4% of target bases could be covered at least ten times in Agilent SureSelect target enrichment method, while the problem of overrepresention of mitochondrial data still existed. In three other samples which were captured by the NimbleGen SeqCap library, the read depth of the mitochondrial DNA genome was much closer to that of the nuclear genome target region in the NimbleGen capture system. Comparison of these two capture systems should be assessed by testing the same samples in the future.

In our study, two out of seven patients (28%) were identified with candidate variants related with hereditary hearing loss. Three novel variants, *WFS1* c.2527C>T, *MYO15A* c.104G>A, *MYO15A* c.5917G>A were first discovered in this study. Further studies such as biological function studies of these candidate variants may help to determine further their pathogenicity. For the remaining undiagnosed cases, the following possibilities might account for failure of mutation detection: 1) novel deafness genes which were not targeted or identified yet; 2) causative mutations missed by targeted sequencing because of incomplete coverage; 3) pathogenic mutations not on exons or splice sites which were not targeted; or 4) copy number variations.

Table 6-7 Published genetic testing for nonsyndromic hearing loss platforms have been released on a clinical or commercial basis

Reference	**No. of Genes targeted**	**Method**	**Coverage**	**Average read depth**	**Samples**	**Results**
(SHEARER *et al.* 2010)	54 NSHL genes	cRNA solution-phase TGE/ Illumina GAIIx	1X, 95.1%; 40×, 91.5%	903X	10, 4 positive controls	Identified causative mutation in positive controls and in five of six unknowns
(BROWNSTEIN *et al.* 2011)	246 genes responsible for either human or mouse deafness.	cRNA solution-phase TGE/ Illumina HiSeq	>10X,95%; >30X,92%	757×-2,080×	11	Identified in 6 of the 11 original probands
(DE KEULENAER *et al.* 2012)	15 genes responsible for ARNSHL :all the coding sequences and most of the UTRs	semi-automated PCR amplification/ Roche 454 Genome Sequencer FLX	>5X, 94.6%-99.0%; >30X, 86.8%-97.2%	73-221X	5 samples including one positive control. 3 of these patients are members of families in which a region of interest has previously been characterized by linkage studies.	Identified causative mutation in positive control and three of four unknowns.
(CHOI *et al.* 2013)	80 NSHL genes	Otogenetics Corporation, cDNA	>10X, 88.9+/-3.7%	218.2+/-56.1 X	20 probands from 20 multiplex families	identified in 13(65%) of the 20 probands

		solution TGE				
(MUTAI *et al.* 2013)	84 genes (61/23, nonsyndro/ sydro)	custom-designed SureSelect Target Enrichment System/Illumina GAIIx		>100X	58 subjects from 15 families, 22 subjects with normal hearing	identified in 7 of the 15 families
(WU *et al.* 2013)	80 NSHL genes	Otogenetics Corporation	>30x, 99.3%	490x	Twelve multiplex families with idiopathic nonsyndromic SNHL	4 families diagnosed
(MIYAGAWA *et al.* 2013)	58 target candidate genes	emulsion PCR/Ion Torrent	>20X, 94.2%	326.5X	8 (4 early-onset, 4 late-onset) Japanese CI/EAS patients,	four rare causative mutations in the *MYO15A, TECTA, TMPRSS3,* and *ACTG1* genes in four patients
(BAEK *et al.* 2012)	80 genes	nucleotide probes./Illumina HiSeq 2000 paired-end read	>10x, 78.26%-92.55%	43-337x	8 AD NSHL Korean families, 31 individuals , including 13 unaffected	confirmed in 5 of the 8 families
(SHAHZAD *et al.* 2013)	24 well-studied SNHL genes	OtoSeq, microdroplet PCR-			34 consanguineous Pakistani families	28 of the 34 families were identified with mutations.

		based/ HiSeq 2000			segregating prelingual hearing Loss (co-segregating recessive hearing loss and STR markers linked to MYO7A CDH23 or SLC26A4	
(SCHRAUWEN *et al.* 2013)	34 ARNHSL GENES	Microdroplet PCR-Based/Illumina Hiseq2000	>20x, 96% >30x, 95%	1585X	Twenty-four patients	identify the genetic basis of hearing loss in 9 of 24 patients

Chapter 7 Noninvasive prenatal testing in hereditary hearing loss

7.1 Introduction

As genetic testing is widely available for hereditary hearing loss, more patients and families could receive definitive genetic diagnoses for their hearing impairment. The growth of prenatal diagnosis of hereditary hearing loss is accompanied with the parents' intentions (DAGAN *et al.* 2002; WITHROW *et al.* 2009). Pregnant couples with a family history of hearing loss always request for genetic consultation of the risk of having a child with hearing loss. If they already were to have a specific genetic diagnosis of carrier status, couples could undergo prenatal genetic diagnosis of the fetus to test for the precise mutation. Now the standard prenatal diagnostic test is done invasively for fetal tissue sampling. These invasive procedures have a low yet non-negligible risk of fetal loss which is the biggest concern for the parents. Noninvasive prenatal testing such as sequencing the fetal genome in maternal plasma DNA might be an alternative choice for those high risk couples. However, only paternal inherited or de novo mutations in the fetal genome could be directly sequenced considering the high background of the maternal genome in the plasma DNA. For autosomal recessive hereditary hearing loss, maternal inherited mutations could be indirectly tested by examining whether fetus inherited the maternal-derived haplotype including the mutation. However, the cost of whole genome sequencing the cell free DNA to build up the haplotype is not affordable clinically, especially when 20x depth is needed to ensure confident variant calling (CHEN *et al.* 2013). The parental inherited haplotype would not interfere with maternal inherited haplotype as long as the haplotype region is longer than 1 Mb (FAN *et al.* 2012). It is suggested

that a 1 Mb genome region around the certain mutation should be sufficient to define two haplotypes of each parent. Meanwhile it has already been demonstrated that target enrichment method could evenly enrich the fetal and maternal genome fragments, and target sequencing could unbiased detect the fetal genome (LIAO *et al.* 2011). Therefore in this study, by using a haplotype-assisted method, we performed NIPT in a pregnant couple with a hearing loss child carrying a GJB2 c.109G>A homozygous mutation.

7.2 Materials and Methods

7.2.1 Subject information

We recruited a pregnant couple with a hearing loss child (family pedigree as Figure 7-2). 5 mls of peripheral blood were collected from the father, mother and proband in EDTA tubes. Maternal blood samples were separately collected in EDTA tubes on both11w+4d and 17W+4d gestation age. The details of plasma isolation and DNA extraction from peripheral blood and plasma were described in chapter 3.2 and 3.3.

7.2.2 Target enrichment system

Two customer-designed NimbleGen EZ capture systems were combined at 1:1 ratio of volume per reaction for DNA capture. One was designed for single nucleotide polymorphism with high minor allele frequency (MAF≈0.5) around genes of common monogenic disorders, including *GJB2* for hearing loss and the 39 other genes for other studies. SNPs were specifically selected based on the 1000 genome data of Chinese population. The other capture system was NimbleGen SeqCap EZ target enrichment library tested in chapter 6 (details also in chapter 3.8).

7.2.3 Sequencing library preparation and target genome enrichment

Library preparation was described in chapter 3.4. Genomic DNA library was pooled together and hybridized with one reaction of 1:1 merged baits from two customer designed in-solution capture systems following the protocol described in 3.6. Around 10ng - 20ng cell free DNA extracted from 1.5ml maternal plasma was separated into three equal parts. Each aliquot of plasma DNA was prepared with different indexes. Three libraries from one plasma DNA sample were pooled together and also hybridized with one reaction of 1:1 merged libraries following the protocol described in 3.6.

7.2.4 Massively parallel sequencing and analysis

Captured genomic regions were sequenced in paired-end at 90 bps on Illumina HiseqTM 2500, using rapid run mode. Sequencing reads were aligned to reference human genome (hg19) using SOAP2. Two mismatches were allowed for alignment for each read. After a filtration of multiple locations mapped reads and duplicated reads generated in PCR amplification, SNP calling was performed by SOAPsnp (LI *et al.* 2009). Low-quality SNPs with read depth<4 and a value of Q20 <90% reads were excluded (YOU *et al.* 2014).

SNPs were classified into two categories (CHEN *et al.* 2013): 1) homozygous in both parental genomes, of which different genotypes between parents were selected to calculate the fetal DNA concentration in the maternal plasma, and of which the same genotypes were used to evaluate the sequencing error rate; 2) heterozygous on either parental genome, which were used for deducing fetal haplotype. The work Flow of Fetal Haplotype was described as Figure 7-1.

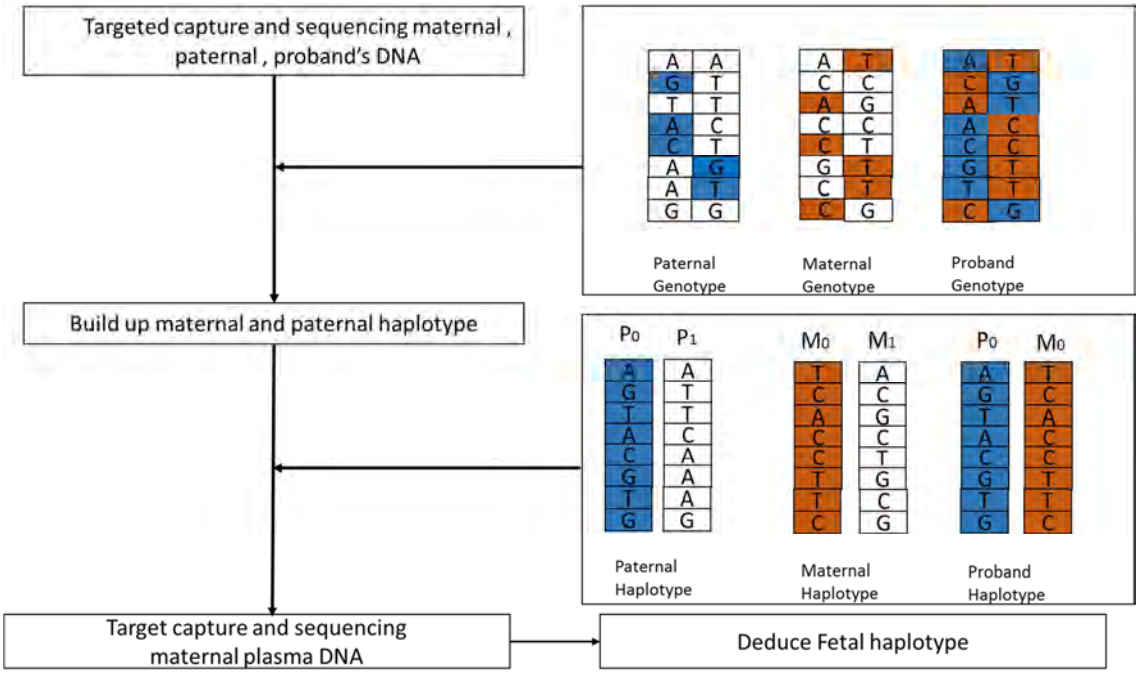


Figure 7-1 Schematic diagram of work flow on haplotype construction

7.3 Results

7.3.1 Clinical information

In our study, we recruited a pregnant couple with a hearing loss child (family pedigree in Fig 7-1). The proband, female, born in 2010, was found to have mild to moderate hearing loss, postlingual onset. The proband passed the newborn hearing screening. No pregnancy complication, no history of CMV, rubella infection or ototoxic drug medication history were reported. The patient was diagnosed with homozygous *GJB2* c.109G>A. Both parents were further confirmed as heterozygous carriers of *GJB2* c.109G>A. According to its autosomal recessive inheritance pattern, this couple has a 25% recurrence risk of have a baby with homozygous *GJB2* c.109G>A. The patient's mother requested prenatal testing of hearing loss for the current pregnancy.

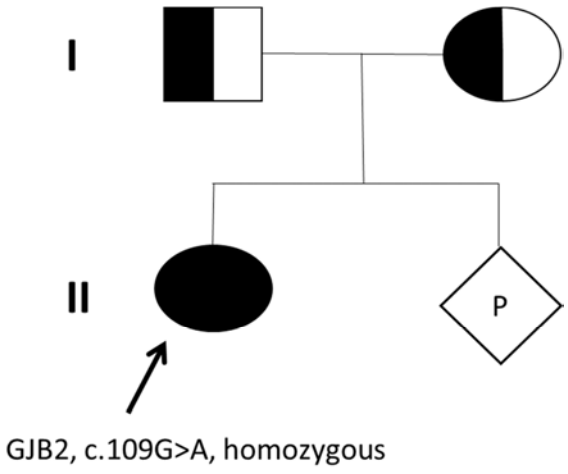


Figure7-2: Pedigree of the *GJB2* family

Table 7-1 Targeted regions of the two NimbleGen target enrichment libraries

Target enrichment libraries	**Targeted analysis region**	**Coordinate of targeted region (hg19)**	**Size**	**No. of SNPs on haplotype construction region**
NimbleGen SeqCap SZ library (details in 3.8)	*GJB2* coding region	chr13:20,761,604-20,767,114	5.5Kb	4
NimbleGen SeqCap SZ library for SNPs	*GJB2* coding region + flanking region	chr13:20,503,793-21,621,558	1.1Mb	54+255
				313

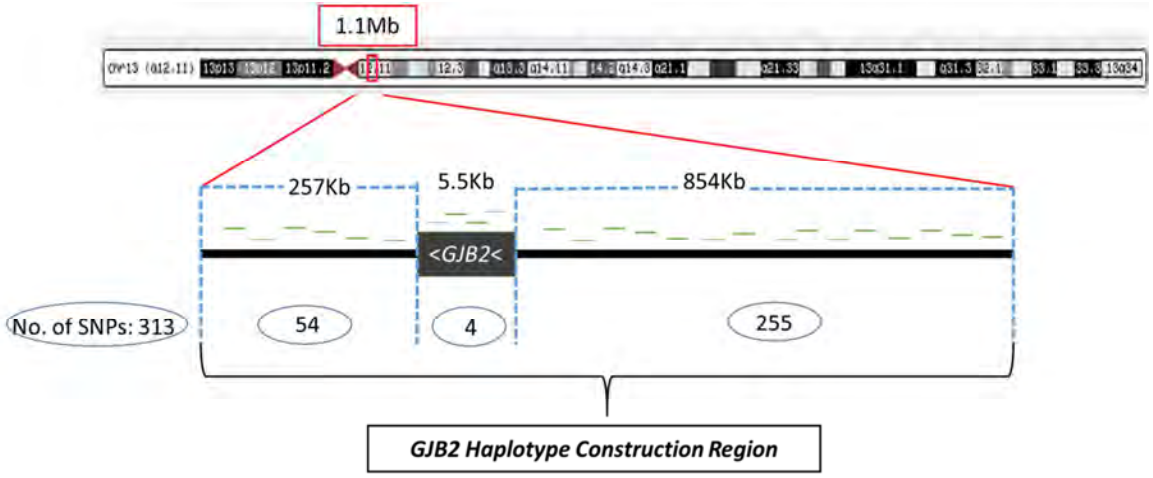


Figure 7-3 Schematic diagram of *GJB2* haplotype construction region.

7.3.2 Sequencing data summary

Three genomic DNA libraries and two cell free DNA libraries were target sequenced. Total output data were 12.4 Gigabases. After filtration of the duplicated paired-end reads, we calculated the sequencing depth of the targeted region by dividing the total number of sequenced bases within the targeted region by the size of the two merged capture systems. A minimum read depth of 116 fold was achieved (Table 7-2).

For the trio family, a total of 6,652,733 SNPs filtered were shared among these three genomes. 603 SNPs did not follow the Mendelian inheritance, meaning that the nucleotide on this loci of the proband was not inherited from the parental genome. This indicated that the 603 SNPs might be generated by *de novo* mutations occurring in fetal genome or a sequencing error, which accounted for $9.06X10^{-5}$ in the sequencing data. 6,649,073 SNPs could be phased into parental haplotype construction.

7.3.3 Fetal DNA fraction calculation

For massively parallel sequencing of the cell-free DNA, fetal DNA faction was required to be assessed. Fetal DNA fraction was calculated based on the informative SNPs around the whole genome. Informative SNPs were on those loci which both parental genotypes were homozygous but of different types. Fetal DNA fraction equals (2X read depth of paternal inherited nucleotide) / (total read depth of this locus). Early in the first trimester fetal DNA fraction was up to 17.16%, which increased slightly to 17.66% in the second trimester in this case.

Table 7-2 Sequencing data summary

Sample source	Sample ID	Raw data	Percentage of Unique reads alignment	Percentage of Paired-end Reads alignment	% Duplication Rate	% Coverage of targeted region (1X depth)	Average Read Depth	% Fetal DNA fraction
Paternal DNA	HL-6198	1.63G	96.40	92.36	2.39	83.20	122x	
Maternal DNA	HL-6197	1.47G	96.48	92.31	2.21	81.57	116x	
Proband DNA	HL-16189	1.64G	96.24	92.00	2.45	82.63	124x	
11w+4d Maternal Plasma	HL-731	3.15G	93.20	88.29	9.80	86.86%	135x	17.16
17w+4d Maternal Plasma	HL-911	4.53G	93.60	88.74	20.50	89.16	135x	17.66

7.3.4 Fetal haplotype recovery

Since the purpose of this study was to tell whether the fetus has inherited the *GJB2* mutations, we mainly focused on analysis of the targeted *GJB2* region (Figure 7-3). We constructed the *GJB2* haplotype block based on 313 informative SNPs spanning chr13:20,503,793-21,621,558 (1.1Mb) (Table 7-1). Firstly, we constructed the trio haplotype with phasing the 313 SNP markers following Mendelian inheritance. Considering the proband inherited both parental haplotype carrying *GJB2* c.109G>A mutation, each parental haplotype was assigned as P_0=paternal haplotype with the mutation, P_1=paternal haplotype with the wild-type allele, M_0= maternal haplotype with the mutation, or M_1= maternal haplotype with the wild-type allele. Secondly, recover fetal haplotype. On these 313 SNPs loci of cell-free DNA sequencing data, it should be a mixture of maternal genotype and fetal genotype. Based on Mendel's law, the fetal genotype could be one of four possible combinations of parental haplotypes: P0M0, P1M0, P0M1, and P1M1 (Figure 7-3). Different combinations would lead to different allelic ratios at each locus. Thus, with the assistance of the parental haplotype, we can deduce the fetal haplotype by constructing a Hidden Markov Model (HMM) with consideration of the single locus probability obtained from allelic ratios and the linkage relationship obtained from the parental haplotype (CHEN *et al.* 2013). Finally, compared with the proband haplotype, the fetus was only sharing one allele with the proband haplotype (P_0) which was paternally inherited. The fetal *GJB2* haplotype was deduced as M_1P_0. It indicated that the fetus is a carrier of *GJB2* c.109G>A (Figure 7-4). Fetal haplotypes were consistent with cell free DNA samples isolated from two different gestational ages.

We further preformed invasive prenatal diagnosis to confirm the result of this NIPT study. Our result confirmed that the fetus was a heterozygous carrier of *GJB2* c.109G>A. The baby was born on 2014 and passed OAE testing.

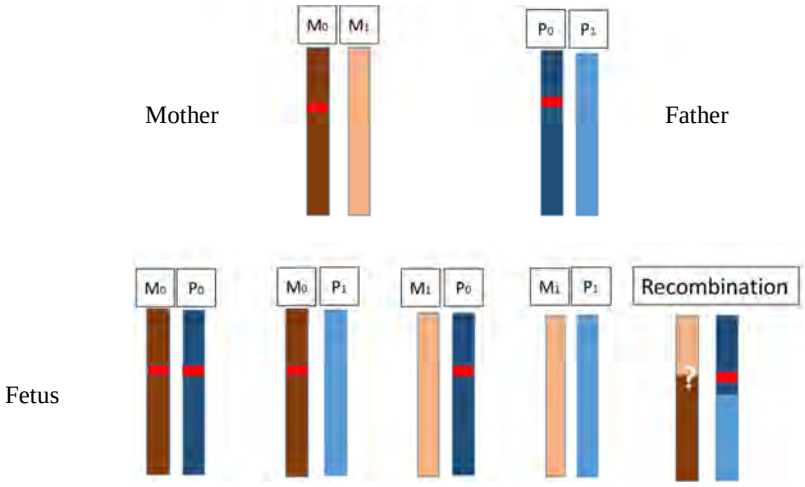


Figure 7-4 Four possible combination of fetal haplotype: M0P0, M0P1, M1P0, M1P1. Red bar: mutation

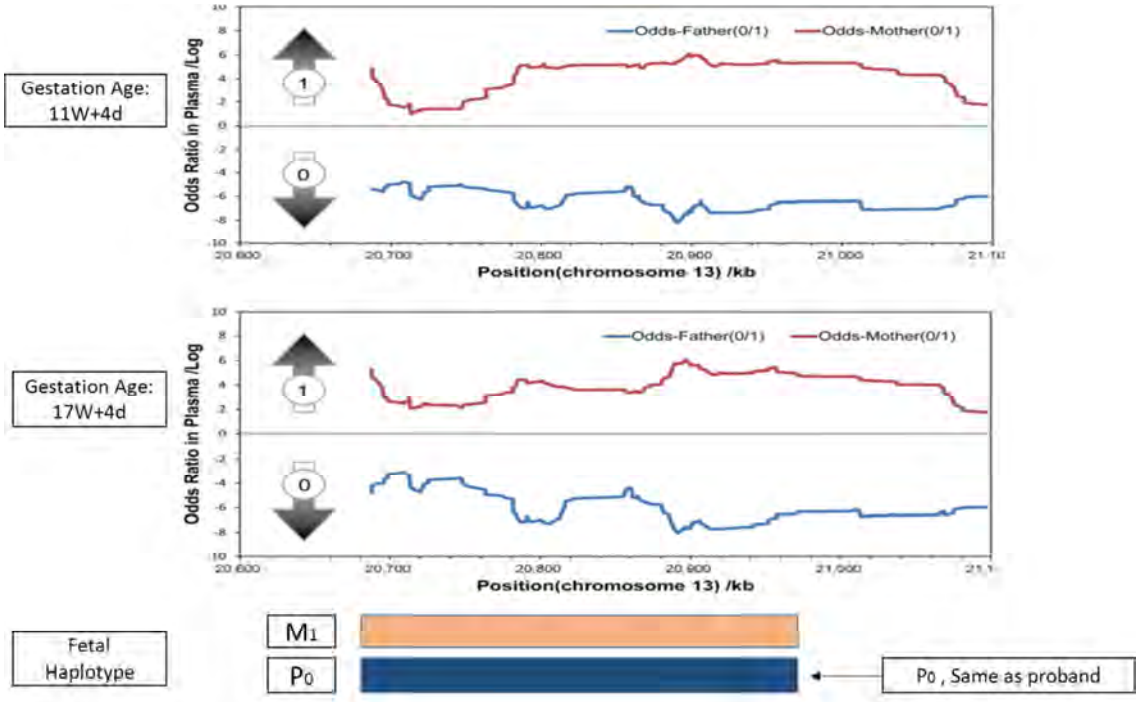


Figure 7-5 Schematic diagram of fetal haplotype in the *GJB2* target region

The fetal haplotype predicted at two different gestational ages was consistent.

'0' means the same haplotype as proband, '1' means opposite.

X axis was coordinate of targeted regions on chromosome 13. Y axis indicated the possibility of each fetal allele same as which proband haplotype. Odd ratio used zero as cut off: >0, this fetal allele was different from proband haplotype; <0, this fetal allele was same as proband haplotype. Red curve: maternal-inherited allele; Blue curve: paternal-inherited allele. Fetus haplotype was deduced as M1P0.

7.4 Discussion

In this pilot study, we demonstrated the feasibility of targeted massively parallel sequencing in noninvasive prenatal test for a hearing loss family with a proband carrying a homozygous mutation *GJB2* c.109G>A. Using a haplotype-assisted method, we recovered the fetal haplotype with the guide of the family trio genotype, and deduced that the fetus inherited the same paternal haplotype (P1) as the proband, and the other normal maternal haplotype (M0). It indicated that the fetus was a heterozygous carrier of *GJB2* c.109G>A. This haplotype assistance method could differentiate the fetal maternal or paternal inherited haplotype which implies its comprehensive application in hereditary disorders with different inheritance pattern.

Compared with recovering the fetal haplotype based on whole genome sequencing, targeted genome region sequencing could be much more affordable if considering application in clinics. Since it needs to generate much less data, the rapid run mode of the Hiseq2500 could complete the run in two days. Even taking into account one to three more days for capture hybridization, analysis results would be available one week earlier than whole genome sequencing. Digital PCR, which is another commonly used technology in NITP, could also address this cost and time issue. Nevertheless, the number of SNP markers covered in each PCR assay was much limited, and each mutation would require specific primer design and optimization of protocol. In contrast, the customer design capture system was more flexible and comprehensive. The NIPT studies by digital PCR depend on counting limited mutant/wild-type allele subtle changes caused by fetus and maternal genotypes (LUN *et al.* 2008; GU *et al.* 2014). Thus it requires a higher fetal DNA fraction to increase this difference to distinguish the fetal genotype (GU *et al.* 2014).

In our study, besides targeting SNPs for building up the haplotype, we employed a target enrichment library covering the coding region of *GJB2*. It could also detect the de novo mutations of the fetal genome in the *GJB2*. However, this haplotype assistance method could only indirectly predict the fetal disease status by deducing the fetal haplotype. If recombination occurs within the haplotype region it would interrupt the interpretation of fetal haplotype. The schematic diagram of hypothetical recombination crossing over in the *GJB2* haplotype region is shown in Figure 7-3.

GJB2 c.109G>A is the most common mutation in the Chinese population with a carrier rate as high as 10.9% based on our previous population screening study. Gallant et al. proposed a haplotype consisting of thirteen SNPs that was specific in Asian hearing loss patients, and in particular in Chinese, with homozygous *GJB2* c.109G>A (GALLANT *et al.* 2013). According to this specific *GJB2* c.109G>A haplotype (Figure 7-5), we found ten SNPs were also targeted in our *GJB2* haplotype (Table 7-3). The proband also carried this specific homozygous *GJB2* c.109G>A haplotype. A digital PCR assay could be designed particular for this mutation.

Figure 7-6 The GJB2 V37I specific haplotype (GALLANT *et al.* 2013)

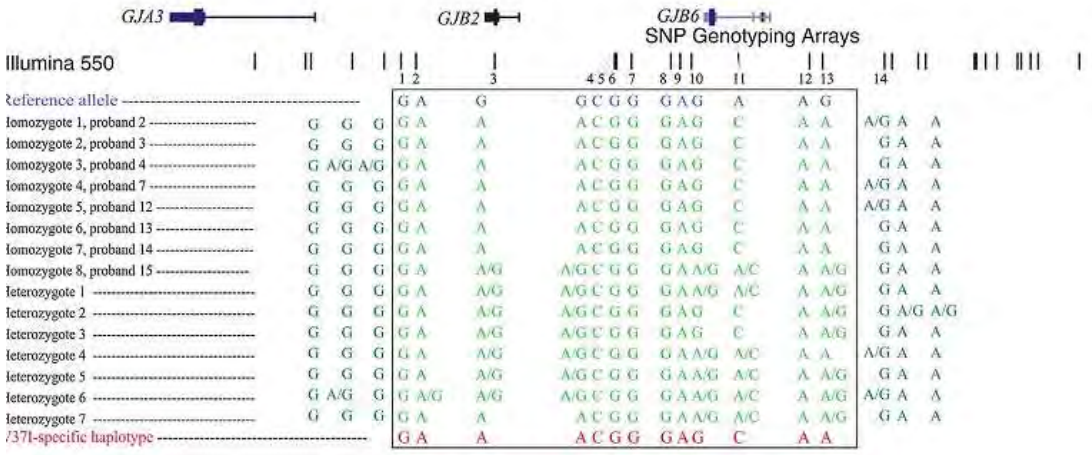


Table 7-3 Genotype of our subjects on *GJB2* V37I specific haplotype

SNP ID	Genotype	proband	father	Mother
rs9578257	G	G,G	G,G	G,G
rs2313475	A	/	/	/
rs3751385	A	A,A	A,A	A,G
rs870729	A	A,A	A,A	A,G
rs11620460	C	C,C	C,C	C,C
rs1889784	G	C,C	C,C	C,C
rs6490527	G	/	/	/
rs7981756	G	C,C	C,T	C,C
rs7330206	A	A,A	A,G	A,A
rs2065796	G	G,G	G,G	G,G
rs945369	C	G,G	G,T	G,G
rs7328044	A	/	/	/
rs945373	A	A,A	A,G	A,A

Section IV: Concluding remarks

Chapter 8 Conclusion and Future work

By screening the carrier frequency of those 15 common hearing loss mutations in a newborn cohort of the Chinese population, we surprisingly found that one in 6.3 neonates carried at least one pathogenic allele. This high carrier frequency suggested that the genetic screening test of these 15 common mutations could be an effective test for hereditary hearing loss patients as a first tier approach. Among the 5800 neonates screened, 28 cases (0.48%) could be genetically diagnosed as hearing loss, which indicated a higher prevalence rate of hearing loss in neonates than previously estimated as 0.3% by newborn screening. As 75% of the 28 genetically diagnosed cases passed the OAE testing of newborn hearing screening, this also implies that the genetic newborn screening might be implemented complementary to the universal newborn hearing screening program in the future to achieve the purpose of early genetic detection of hearing loss. However, a proper large population follow up study should be performed to justify the effectiveness of genetic newborn screening.

Besides the common mutation screening, there is still a substantial portion of hearing loss patient due to rare mutations, and single gene screening by the Sanger approach is difficult. We established a TGE-MPS approach which provides a much more comprehensive screening for hereditary hearing loss with high accuracy (100%). Besides those known hearing loss genes, our capture systems also included human homologous genes associated with a mouse model or other unidentified genes in the gene families which play an important role in the inner ear. The approach would also be applied to discover novel hearing loss genes. The extra 28% diagnostic yield

among seven common mutations screened negative cases also highlighted the potential and advantage of this TGE-MPS approach in molecular diagnosis. It also discovered three novel variants which might be pathogenic for hearing loss. Functional studies on these variants will help understand the molecular pathology of hearing loss. In the next step we should systematically assess the performance of this approach and difference between the two target enrichment methods by testing well genotyped DNA samples.

As more patients and families seek genetic diagnoses for their hearing impairment, we anticipate that there will be a growing demand for prenatal diagnosis for this condition. In contrast to the risk of fetal loss caused by invasive testing, noninvasive prenatal testing using cell free DNA could be considered as an alternative choice. We applied targeted sequencing NIPT to a hearing loss family with a proband with a *GJB2* c.109G>A homozygous mutation. With analysis of the targeted haplotype region, we accurately predict the fetus as a *GJB2* c.109G>A carrier as early as first trimester. Our study indicated this approach has great potential application for hereditary hearing loss and also other single gene disorders with autosomal dominant, recessive or sex-linked inheritance patterns. Since the haplotype assistance method indirectly tests the fetal inheritance of the mutation, it suggests that NIPT should work as a screening test to exclude those fetuses that do not inherit any mutation, or are only mutation carriers. Presently, confirmation of mutations should still be performed by standard invasive prenatal testing.

References

Abe, S., S. Usami, H. Shinkawa, P. M. Kelley and W. J. Kimberling, 2000 Prevalent connexin 26 gene (GJB2) mutations in Japanese. J Med Genet 37**:** 41-43.

Amstutz, U., G. Andrey-Zurcher, D. Suciu, R. Jaggi, J. Haberle *et al.*, 2011 Sequence capture and next-generation resequencing of multiple tagged nucleic acid samples for mutation screening of urea cycle disorders. Clin Chem 57**:** 102-111.

Arnos, K. S., J. Israel and M. Cunningham, 1991 Genetic counseling of the deaf. Medical and cultural considerations. Ann N Y Acad Sci 630**:** 212-222.

Avraham, K. B., 2013 Rescue from hearing loss in Usher's syndrome. N Engl J Med 369**:** 1758-1760.

Baek, J. I., S. K. Oh, D. B. Kim, S. Y. Choi, U. K. Kim *et al.*, 2012 Targeted massive parallel sequencing: the effective detection of novel causative mutations associated with hearing loss in small families. Orphanet J Rare Dis 7**:** 60.

Baldwin, E. E., P. Boudreault, M. Fox, J. S. Sinsheimer and C. G. Palmer, 2012 Effect of pre-test genetic counseling for deaf adults on knowledge of genetic testing. J Genet Couns 21**:** 256-272.

Baxter, M., and Y. Agrawal, 2014 Vestibular dysfunction in Turner syndrome: a case report. Otol Neurotol 35**:** 294-296.

Bischoff, A. M., M. W. Luijendijk, P. L. Huygen, G. van Duijnhoven, E. M. De Leenheer *et al.*, 2004 A novel mutation identified in the DFNA5 gene in a Dutch family: a clinical and genetic evaluation. Audiol Neurootol 9**:** 34-46.

Boulay, A. C., F. J. del Castillo, F. Giraudet, G. Hamard, C. Giaume *et al.*, 2013 Hearing is normal without connexin30. J Neurosci 33**:** 430-434.

Bronya J.B. Keats, Arthur N. Popper, Richard R. Fay - 2002, Genetics and Auditory Disorders)

Brownstein, Z., L. M. Friedman, H. Shahin, V. Oron-Karni, N. Kol *et al.*, 2011 Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. Genome Biol 12**:** R89.

Brunger, J. W., G. S. Murray, M. O'Riordan, A. L. Matthews, R. J. Smith *et al.*, 2000 Parental attitudes toward genetic testing for pediatric deafness. Am J Hum Genet 67**:** 1621-1625.

Bruzzone, R., V. Veronesi, D. Gomes, M. Bicego, N. Duval *et al.*, 2003 Loss-of-function and residual channel activity of connexin26 mutations associated with non-syndromic deafness. FEBS Lett 533**:** 79-88.

Chen, G., X. Yi, P. Chen, J. Dong, G. Yang *et al.*, 2012a A large-scale newborn hearing screening in rural areas in China. Int J Pediatr Otorhinolaryngol 76**:** 1771-1774.

Chen, P., H. Chen, S. Fu, G. Chen and J. Dong, 2012b Prevalence of GJB6 mutations in Chinese patients with non-syndromic hearing loss. Int J Pediatr Otorhinolaryngol 76**:** 265-267.

Chen, S., H. Ge, X. Wang, X. Pan, X. Yao *et al.*, 2013 Haplotype-assisted accurate non-invasive fetal whole genome recovery through maternal plasma sequencing. Genome Med 5**:** 18.

Chen, W., V. Kalscheuer, A. Tzschach, C. Menzel, R. Ullmann *et al.*, 2008 Mapping translocation breakpoints by next-generation sequencing. Genome Res 18**:** 1143-1149.

Cheung, M. C., J. D. Goldberg and Y. W. Kan, 1996 Prenatal diagnosis of sickle cell anaemia and thalassaemia by analysis of fetal cells in maternal blood. Nat Genet 14**:** 264-268.

Chin, C. J., M. M. Khami and M. Husein, 2014 A general review of the otolaryngologic manifestations of Down Syndrome. Int J Pediatr Otorhinolaryngol 78**:** 899-904.

Chiu, R. W., K. C. Chan, Y. Gao, V. Y. Lau, W. Zheng *et al.*, 2008 Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proc Natl Acad Sci U S A 105**:** 20458-20463.

Choi, B. Y., G. Park, J. Gim, A. R. Kim, B. J. Kim *et al.*, 2013 Diagnostic application of targeted resequencing for familial nonsyndromic hearing loss. PLoS One 8**:** e68692.

Christensen, K., H. Frederiksen and H. J. Hoffman, 2001 Genetic and environmental influences on self-reported reduced hearing in the old and oldest old. J Am Geriatr Soc 49**:** 1512-1517.

Clark, M. J., R. Chen, H. Y. Lam, K. J. Karczewski, R. Chen *et al.*, 2011 Performance comparison of exome DNA sequencing technologies. Nat Biotechnol 29**:** 908-914.

Clemens, C. J., and S. A. Davis, 2001 Minimizing false-positives in universal newborn hearing screening: a simple solution. Pediatrics 107**:** E29.

Common, J. E., M. Bitner-Glindzicz, E. A. O'Toole, M. R. Barnes, L. Jenkins *et al.*, 2005 Specific loss of connexin 26 expression in ductal sweat gland epithelium associated with the deletion mutation del(GJB6-D13S1830). Clin Exp Dermatol 30**:** 688-693.

Cureoglu, S., P. A. Schachern, A. Ferlito, A. Rinaldo, V. Tsuprun *et al.*, 2006 Otosclerosis: etiopathogenesis and histopathology. Am J Otolaryngol 27**:** 334-340.

Dagan, O., H. Hochner, H. Levi, A. Raas-Rothschild and M. Sagi, 2002 Genetic testing for hearing loss: different motivations for the same outcome. Am J Med Genet 113**:** 137-143.

Dahl, H. H., S. E. Tobin, Z. Poulakis, F. W. Rickards, X. Xu *et al.*, 2006 The contribution of GJB2 mutations to slight or mild hearing loss in Australian elementary school children. J Med Genet 43**:** 850-855.

Dai, P., F. Yu, B. Han, Y. Yuan, Q. Li *et al.*, 2007 The prevalence of the 235delC GJB2 mutation in a Chinese deaf population. Genet Med 9**:** 283-289.

De Keulenaer, S., J. Hellemans, S. Lefever, J. P. Renard, J. De Schrijver *et al.*, 2012 Molecular diagnostics for congenital hearing loss including 15 deafness genes using a next generation sequencing platform. BMC Med Genomics 5**:** 17.

del Castillo, I., M. Villamar, M. A. Moreno-Pelayo, F. J. del Castillo, A. Alvarez *et al.*, 2002 A deletion involving the connexin 30 gene in nonsyndromic hearing impairment. N Engl J Med 346**:** 243-249.

De Keulenaer, S., J. Hellemans, S. Lefever, J. P. Renard, J. De Schrijver *et al.*, 2012 Molecular diagnostics for congenital hearing loss including 15 deafness genes using a next generation sequencing platform. BMC Med Genomics 5**:** 17.

Deltenre, P., and L. Van Maldergem, 2013 Hearing loss and deafness in the pediatric population: causes, diagnosis, and rehabilitation. Handb Clin Neurol 113**:** 1527-1538.

Devriendt, K., J. Jaeken, G. Matthijs, H. Van Esch, P. Debeer *et al.*, 1999 Haploinsufficiency of the HOXA gene cluster, in a patient with hand-foot-genital syndrome, velopharyngeal insufficiency, and persistent patent Ductus botalli. Am J Hum Genet 65**:** 249-251.

Duman, D., and M. Tekin, 2012 Autosomal recessive nonsyndromic deafness genes: a review. Front Biosci (Landmark Ed) 17**:** 2213-2236.

Duno, M., H. Hove, M. Kirchhoff, K. Devriendt and M. Schwartz, 2004 Mapping genomic deletions down to the base: a quantitative copy number scanning approach used to characterise and clone the breakpoints of a recurrent 7p14.2p15.3 deletion. Hum Genet 115**:** 459-467.

Ealy, M., and R. J. Smith, 2011 Otosclerosis. Adv Otorhinolaryngol 70**:** 122-129.

Eijk-Van Os, P. G., and J. P. Schouten, 2011 Multiplex Ligation-dependent Probe Amplification (MLPA(R)) for the detection of copy number variation in genomic sequences. Methods Mol Biol 688**:** 97-126.

El-Hattab, A. W., and F. Scaglia, 2013 Mitochondrial DNA depletion syndromes: review and updates of genetic basis, manifestations, and therapeutic options. Neurotherapeutics 10**:** 186-198.

El Gezeery A. Otosclerosis: association of COLIA1 Sp1 binding site polymorphism in Alexandria, Egypt. Asian Biomed.2012;6:747-752

Eppsteiner, R. W., A. E. Shearer, M. S. Hildebrand, A. P. Deluca, H. Ji *et al.*, 2012 Prediction of cochlear implant performance by genetic mutation: the spiral ganglion hypothesis. Hear Res 292**:** 51-58.

Erway, L. C., Y. W. Shiau, R. R. Davis and E. F. Krieg, 1996 Genetics of age-related hearing loss in mice. III. Susceptibility of inbred and F1 hybrid strains to noise-induced hearing loss. Hear Res 93**:** 181-187.

Fan, H. C., Y. J. Blumenfeld, U. Chitkara, L. Hudgins and S. R. Quake, 2008 Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proc Natl Acad Sci U S A 105**:** 16266-16271.

Fan, H. C., W. Gu, J. Wang, Y. J. Blumenfeld, Y. Y. El-Sayed *et al.*, 2012 Non-invasive prenatal measurement of the fetal genome. Nature 487**:** 320-324.

<Feasibility of implementing a universal neonatal hearing screening programme using distortion product otoacoustic emission detection at a university hospital in Hong Kong_2004_HK MED J.PDF>.

Friedman, R. A., L. Van Laer, M. J. Huentelman, S. S. Sheth, E. Van Eyken *et al.*, 2009 GRM7 variants confer susceptibility to age-related hearing impairment. Hum Mol Genet 18**:** 785-796.

Friedman, T. B., and A. J. Griffith, 2003 Human nonsyndromic sensorineural deafness. Annu Rev Genomics Hum Genet 4**:** 341-402.

Friedman, T. B., J. R. Sellers and K. B. Avraham, 1999 Unconventional myosins and the genetics of hearing loss. Am J Med Genet 89**:** 147-157.

Gallant, E., L. Francey, E. A. Tsai, M. Berman, Y. Zhao *et al.*, 2013 Homozygosity for the V37I GJB2 mutation in fifteen probands with mild to moderate sensorineural hearing impairment: further confirmation of pathogenicity and haplotype analysis in Asian populations. Am J Med Genet A 161**:** 2148-2157.

Genetic counseling. 1975Am J Hum Genet 27**:** 240-242

Genetics Evaluation Guidelines for the Etiologic Diagnosis of Congenital Hearing Loss. Genetic Evaluation of Congenital Hearing Loss Expert Panel. 2002. ACMG statement. Genet Med 4**:** 162-171.

Ghogomu, N., A. Umansky and J. E. Lieu, 2014 Epidemiology of unilateral sensorineural hearing loss with universal newborn hearing screening. Laryngoscope 124**:** 295-300.

Gomez-Zaera, M., T. M. Strom, B. Rodriguez, X. Estivill, T. Meitinger *et al.*, 2001 Presence of a major WFS1 mutation in Spanish Wolfram syndrome pedigrees. Mol Genet Metab 72**:** 72-81.

Gravholt, C. H., 2004 Epidemiological, endocrine and metabolic features in Turner syndrome. Eur J Endocrinol 151**:** 657-687.

Green, E. D., and M. S. Guyer, 2011 Charting a course for genomic medicine from base pairs to bedside. Nature 470**:** 204-213.

Grifa, A., C. A. Wagner, L. D'Ambrosio, S. Melchionda, F. Bernardi *et al.*, 1999 Mutations in GJB6 cause nonsyndromic autosomal dominant deafness at DFNA3 locus. Nat Genet 23**:** 16-18.

Gu, W., W. Koh, Y. J. Blumenfeld, Y. Y. El-Sayed, L. Hudgins *et al.*, 2014 Noninvasive prenatal diagnosis in a fetus at risk for methylmalonic acidemia. Genet Med.

Guan, M. X., 2004 Molecular pathogenetic mechanism of maternally inherited deafness. Ann N Y Acad Sci 1011**:** 259-271.

Guan, M. X., Q. Yan, X. Li, Y. Bykhovskaya, J. Gallo-Teran *et al.*, 2006 Mutation in TRMU related to transfer RNA modification modulates the phenotypic expression of the deafness-associated mitochondrial 12S ribosomal RNA mutations. Am J Hum Genet 79**:** 291-302.

Harlor, A. D., Jr., and C. Bower, 2009 Hearing assessment in infants and children: recommendations beyond neonatal screening. Pediatrics 124**:** 1252-1263.

Harrison, M., J. Roush and J. Wallace, 2003 Trends in age of identification and intervention in infants with hearing loss. Ear Hear 24**:** 89-95.

Harrison, P. M., D. Zheng, Z. Zhang, N. Carriero and M. Gerstein, 2005 Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. Nucleic Acids Res 33**:** 2374-2383.

Hildebrand, M. S., A. P. DeLuca, K. R. Taylor, D. P. Hoskinson, I. A. Hur *et al.*, 2009 A contemporary review of AudioGene audioprofiling: a machine-based candidate gene prediction tool for autosomal dominant nonsyndromic hearing loss. Laryngoscope 119**:** 2211-2215.

Hilgert, N., R. J. Smith and G. Van Camp, 2009a Forty-six genes causing nonsyndromic hearing impairment: which ones should be analyzed in DNA diagnostics? Mutat Res 681**:** 189-196.

Hilgert, N., R. J. Smith and G. Van Camp, 2009b Function and expression pattern of nonsyndromic deafness genes. Curr Mol Med 9**:** 546-564.

Hong, H. M., J. J. Yang, J. C. Shieh, M. L. Lin and S. Y. Li, 2010 Novel mutations in the connexin43 (GJA1) and GJA1 pseudogene may contribute to nonsyndromic hearing loss. Hum Genet 127**:** 545-551.

Huculak, C., H. Bruyere, T. N. Nelson, F. K. Kozak and S. Langlois, 2006 V37I connexin 26 allele in patients with sensorineural hearing loss: evidence of its pathogenicity. Am J Med Genet A 140**:** 2394-2400.

Hultcrantz, M., R. Simonoska and A. E. Stenberg, 2006 Estrogen and hearing: a summary of recent investigations. Acta Otolaryngol 126**:** 10-14.

Hwa, H. L., T. M. Ko, C. J. Hsu, C. H. Huang, Y. L. Chiang *et al.*, 2003 Mutation spectrum of the connexin 26 (GJB2) gene in Taiwanese patients with prelingual deafness. Genet Med 5**:** 161-165.

Jensen, T. J., Z. Dzakula, C. Deciu, D. van den Boom and M. Ehrich, 2012 Detection of microdeletion 22q11.2 in a fetus by next-generation sequencing of maternal plasma. Clin Chem 58**:** 1148-1151.

JCIH, 2000. Joint Committee on Infant Hearing, 2000 Position statement.

Kalyana-Sundaram, S., C. Kumar-Sinha, S. Shankar, D. R. Robinson, Y. M. Wu *et al.*, 2012 Expressed pseudogenes in the transcriptional landscape of human cancers. Cell 149**:** 1622-1634.

Karosi, T., J. Konya, L. Z. Szabo and I. Sziklai, 2004 Measles virus prevalence in otosclerotic stapes footplate samples. Otol Neurotol 25**:** 451-456.

Karosi, T., J. Konya, L. Z. Szabo and I. Sziklai, 2007 Measles virus prevalence in otosclerotic foci. Adv Otorhinolaryngol 65**:** 93-106.

Kilicdag, E. B., H. Yavuz, T. Bagis, E. Tarim, A. N. Erkan *et al.*, 2004 Effects of estrogen therapy on hearing in postmenopausal women. Am J Obstet Gynecol 190**:** 77-82.

Kim, S. Y., G. Park, K. H. Han, A. Kim, J. W. Koo *et al.*, 2013 Prevalence of p.V37I variant of GJB2 in mild or moderate hearing loss in a pediatric population and the interpretation of its pathogenicity. PLoS One 8**:** e61592.

Konings, A., L. Van Laer and G. Van Camp, 2009 Genetic studies on noise-induced hearing loss: a review. Ear Hear 30**:** 151-159.

Korf, B. R., and H. L. Rehm, 2013 New approaches to molecular diagnosis. JAMA 309**:** 1511-1521.

Lackner, A., A. Acham, T. Alborno, M. Moser, H. Engele *et al.*, 2009 Effect on hearing of ganciclovir therapy for asymptomatic congenital cytomegalovirus infection: four to 10 year follow up. J Laryngol Otol 123**:** 391-396.

LaFramboise, T., 2009 Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. Nucleic Acids Res 37**:** 4181-4193.

Lander, E. S., 2011 Initial impact of the sequencing of the human genome. Nature 470**:** 187-197.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody *et al.*, 2001 Initial sequencing and analysis of the human genome. Nature 409**:** 860-921.

Lentz, J. J., F. M. Jodelka, A. J. Hinrich, K. E. McCaffrey, H. E. Farris *et al.*, 2013 Rescue of hearing and vestibular function by antisense oligonucleotides in a mouse model of human deafness. Nat Med 19**:** 345-350.

Lerer, I., M. Sagi, Z. Ben-Neriah, T. Wang, H. Levi *et al.*, 2001 A deletion mutation in GJB6 cooperating with a GJB2 mutation in trans in non-syndromic deafness: A novel founder mutation in Ashkenazi Jews. Hum Mutat 18**:** 460.

Leung, T. N., J. Zhang, T. K. Lau, L. Y. Chan and Y. M. Lo, 2001 Increased maternal plasma fetal DNA concentrations in women who eventually develop preeclampsia. Clin Chem 47**:** 137-139.

Lewis, M. A., E. Quint, A. M. Glazier, H. Fuchs, M. H. De Angelis *et al.*, 2009 An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. Nat Genet 41**:** 614-618.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25**:** 1754-1760.

Li, H., H. B. Li, J. Mao, M. J. Liu, H. B. Cheng *et al.*, 2011 [Genetic screening of mutation hot-spots for nonsyndromic hearing loss in southern Jiangsu province with SNaPshot]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 28**:** 383-386.

Li, H., Q. Li, H. Li and Y. Chen, 2012 [A literature review of epidemiological studies on mutation hot spots of Chinese population with non-syndromic hearing loss]. Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi 26**:** 589-594.

Li, H., W. Kloosterman and D. M. Fekete, 2010 MicroRNA-183 family members regulate sensorineural fates in the inner ear. J Neurosci 30: 3254-3263.

Li, R., Y. Li, X. Fang, H. Yang, J. Wang *et al.*, 2009 SNP detection for massively parallel whole-genome resequencing. Genome Res 19**:** 1124-1132.

Li, R., C. Yu, Y. Li, T. W. Lam, S. M. Yiu *et al.*, 2009b SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25**:** 1966-1967.

Liang, Q., and B. Mason, 2013 Enter the dragon--China's journey to the hearing world. Cochlear Implants Int 14 Suppl 1**:** S26-31.

Liao, G. J., F. M. Lun, Y. W. Zheng, K. C. Chan, T. Y. Leung *et al.*, 2011 Targeted massively parallel sequencing of maternal plasma DNA permits efficient and unbiased detection of fetal alleles. Clin Chem 57**:** 92-101.

Limb, C. J., 2012 When inner ear stem cell therapy becomes a reality. Trends Amplif 16**:** 3.

Liu, X. Z., X. J. Xia, X. M. Ke, X. M. Ouyang, L. L. Du *et al.*, 2002 The prevalence of connexin 26 (GJB2) mutations in the Chinese population. Hum Genet 111**:** 394-397.

Lo, Y. M., N. Corbetta, P. F. Chamberlain, V. Rai, I. L. Sargent *et al.*, 1997 Presence of fetal DNA in maternal plasma and serum. Lancet 350**:** 485-487.

Lo, Y. M., P. Patel, J. S. Wainscoat, M. Sampietro, M. D. Gillmer *et al.*, 1989 Prenatal sex determination by DNA amplification from maternal peripheral blood. Lancet 2**:** 1363-1365.

Lo, Y. M., M. S. Tein, T. K. Lau, C. J. Haines, T. N. Leung *et al.*, 1998 Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. Am J Hum Genet 62**:** 768-775.

Lun, F. M., N. B. Tsui, K. C. Chan, T. Y. Leung, T. K. Lau *et al.*, 2008 Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. Proc Natl Acad Sci U S A 105**:** 19920-19925.

Maasz, A., K. Komlosi, K. Hadzsiev, Z. Szabo, P. J. Willems *et al.*, 2008 Phenotypic variants of the deafness-associated mitochondrial DNA A7445G mutation. Curr Med Chem 15**:** 1257-1262.

Martinez, A., J. Linden, L. A. Schimmenti and C. G. Palmer, 2003 Attitudes of the broader hearing, deaf, and hard-of-hearing community toward genetic testing for deafness. Genet Med 5**:** 106-112.

Matsunaga, T., 2009 Value of genetic testing in the otological approach for sensorineural hearing loss. Keio J Med 58**:** 216-222.

Mencia, A., S. Modamio-Hoybjor, N. Redshaw, M. Morin, F. Mayo-Merino *et al.*, 2009 Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. Nat Genet 41**:** 609-613.

Meng, M., X. Li, H. Ge, F. Chen, M. Han *et al.*, 2014 Noninvasive prenatal testing for autosomal recessive conditions by maternal plasma sequencing in a case of congenital deafness. Genet Med.

Mermall, V., P. L. Post and M. S. Mooseker, 1998 Unconventional myosins in cell movement, membrane traffic, and signal transduction. Science 279**:** 527-533.

Metzler, M., 2009 MicroRNAs become large in hearing loss. Clin Genet 76**:** 343-345.

Middleton, A., J. Hewison and R. Mueller, 2001 Prenatal diagnosis for inherited deafness--what is the potential demand? J Genet Couns 10**:** 121-131.

MITOMAP: A Human Mitochondrial Genome Database. http://www.mitomap.org.

Miyagawa, M., S. Y. Nishio, T. Ikeda, K. Fukushima and S. Usami, 2013 Massively parallel DNA sequencing successfully identifies new causative mutations in deafness genes in patients with cochlear implantation and EAS. PLoS One 8**:** e75793.

Morell, R. J., H. J. Kim, L. J. Hood, L. Goforth, K. Friderici *et al.*, 1998 Mutations in the connexin 26 gene (GJB2) among Ashkenazi Jews with nonsyndromic recessive deafness. N Engl J Med 339**:** 1500-1505.

Morimoto, N., T. Tanaka, H. Taiji, R. Horikawa, Y. Naiki *et al.*, 2006 Hearing loss in Turner syndrome. J Pediatr 149**:** 697-701.

Morton, C. C., and W. E. Nance, 2006 Newborn hearing screening--a silent revolution. N Engl J Med 354**:** 2151-2164.

Mutai, H., N. Suzuki, A. Shimizu, C. Torii, K. Namba *et al.*, 2013 Diverse spectrum of rare deafness genes underlies early-childhood hearing loss in Japanese patients: a cross-sectional, multi-center next-generation sequencing study. Orphanet J Rare Dis 8**:** 172.

Nadol, J. B., Jr., 1993 Hearing loss. N Engl J Med 329**:** 1092-1102.

Need, A. C., V. Shashi, Y. Hitomi, K. Schoch, K. V. Shianna *et al.*, 2012 Clinical application of exome sequencing in undiagnosed genetic conditions. J Med Genet 49**:** 353-361.

Nelson, H. D., C. Bougatsos and P. Nygren, 2008 Universal newborn hearing screening: systematic review to update the 2001 US Preventive Services Task Force Recommendation. Pediatrics 122**:** e266-276.

Ng, E. K., N. B. Tsui, T. K. Lau, T. N. Leung, R. W. Chiu *et al.*, 2003 mRNA of placental origin is readily detectable in maternal plasma. Proc Natl Acad Sci U S A 100**:** 4748-4753.

Ng, S. B., E. H. Turner, P. D. Robertson, S. D. Flygare, A. W. Bigham *et al.*, 2009 Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461**:** 272-276.

Noguchi, Y., T. Yashima, A. Hatanaka, M. Uzawa, M. Yasunami *et al.*, 2005 A mutation in Wolfram syndrome type 1 gene in a Japanese family with autosomal dominant low-frequency sensorineural hearing loss. Acta Otolaryngol 125**:** 1189-1194.

Okano, T., and M. W. Kelley, 2012 Stem cell therapy for the inner ear: recent advances and future directions. Trends Amplif 16**:** 4-18.

Oliveira, C. S., F. M. Ribeiro, R. Lago and C. Alves, 2013 Audiological abnormalities in patients with turner syndrome. Am J Audiol 22**:** 226-232.

Olusanya, B. O., 2011 Highlights of the new WHO Report on Newborn and Infant Hearing Screening and implications for developing countries. Int J Pediatr Otorhinolaryngol 75**:** 745-748.

Olusanya, B. O., and V. E. Newton, 2007 Global burden of childhood hearing impairment and disease control priorities for developing countries. Lancet 369**:** 1314-1317.

Olusanya, B. O., 2011 Highlights of the new WHO Report on Newborn and Infant Hearing Screening and implications for developing countries. Int J Pediatr Otorhinolaryngol 75**:** 745-748.

Ordulu, Z., K. E. Wong, B. B. Currall, A. R. Ivanov, S. Pereira *et al.*, 2014 Describing sequencing results of structural chromosome rearrangements with a suggested next-generation cytogenetic nomenclature. Am J Hum Genet 94**:** 695-709.

Ozguc, M., 2011 Genetic testing: predictive value of genotyping for diagnosis and management of disease. EPMA J 2**:** 173-179.

Pacala, J. T., and B. Yueh, 2012 Hearing deficits in the older patient: "I didn't notice anything". Jama 307**:** 1185-1194.

Peters, D., T. Chu, S. A. Yatsenko, N. Hendrix, W. A. Hogge *et al.*, 2011 Noninvasive prenatal diagnosis of a fetal microdeletion syndrome. N Engl J Med 365**:** 1847-1848.

Pink, R. C., K. Wicks, D. P. Caley, E. K. Punch, L. Jacobs *et al.*, 2011 Pseudogenes: pseudo-functional or key regulators in health and disease? Rna 17**:** 792-798.

Prera, N., E. Lohle and R. Birkenhager, 2014 [Progressive hearing impairment with deletion in GJB2 gene despite normal newborn hearing screening]. Laryngorhinootologie 93**:** 244-248.

Prieve, B. A., M. P. Feeney, S. Stenfelt and N. Shahnaz, 2013 Prediction of conductive hearing loss using wideband acoustic immittance. Ear Hear 34 Suppl 1**:** 54S-59S.

Putcha, G. V., B. A. Bejjani, S. Bleoo, J. K. Booker, J. C. Carey *et al.*, 2007 A multicenter study of the frequency and distribution of GJB2 and GJB6 mutations in a large North American cohort. Genet Med 9**:** 413-426.

Redin, C., S. Le Gras, O. Mhamdi, V. Geoffroy, C. Stoetzel *et al.*, 2012 Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alstrom syndromes. J Med Genet 49**:** 502-512.

Rodriguez-Paris, J., and I. Schrijver, 2009 The digenic hypothesis unraveled: the GJB6 del(GJB6-D13S1830) mutation causes allele-specific loss of GJB2 expression in cis. Biochem Biophys Res Commun 389**:** 354-359.

Ronchi, E., D. Harper, A. Taylor and A. G. Halsberger, 2000 Genetic testing: policy issues for the new millennium. Community Genet 3**:** 161-163.

Ross, D. S., W. J. Holstrum, M. Gaffney, D. Green, R. F. Oyler *et al.*, 2008 Hearing screening and diagnostic evaluation of children with unilateral and mild bilateral hearing loss. Trends Amplif 12**:** 27-34.

Sagong, B., J. I. Baek, S. K. Oh, K. J. Na, J. W. Bae *et al.*, 2013 A rapid method for simultaneous screening of multi-gene mutations associated with hearing loss in the Korean population. PLoS One 8**:** e57237.

Schrauwen, I., M. Sommen, J. J. Corneveaux, R. A. Reiman, N. J. Hackett *et al.*, 2013 A sensitive and specific diagnostic test for hearing loss using a microdroplet PCR-based approach and next generation sequencing. Am J Med Genet A 161a**:** 145-152.

Schrijver, I., 2004 Hereditary non-syndromic sensorineural hearing loss: transforming silence to sound. J Mol Diagn 6**:** 275-284.

Scott, C. A., and D. P. Kelsell, 2011 Key functions for gap junctions in skin and hearing. Biochem J 438**:** 245-254.

Shahzad, M., T. A. Sivakumaran, T. A. Qaiser, J. M. Schultz, Z. Hussain *et al.*, 2013 Genetic analysis through OtoSeq of Pakistani families segregating prelingual hearing loss. Otolaryngol Head Neck Surg 149**:** 478-487.

Shearer, A. E., E. A. Black-Ziegelbein, M. S. Hildebrand, R. W. Eppsteiner, H. Ravi *et al.*, 2013 Advancing genetic testing for deafness with genomic technology. J Med Genet 50**:** 627-634.

Shearer, A. E., A. P. DeLuca, M. S. Hildebrand, K. R. Taylor, J. Gurrola, 2nd *et al.*, 2010 Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. Proc Natl Acad Sci U S A 107**:** 21104-21109.

Shearer, A. E., and R. J. Smith, 2012 Genetics: advances in genetic testing for deafness. Curr Opin Pediatr 24**:** 679-686.

Shemesh R. 2010. Hearing Impairment: Definitions, Assessment and Management. In: JH Stone, M Blouin, editors. International Encyclopedia of Rehabilitation. Available online: http://cirrie.buffalo.edu/encyclopedia/en/article/272/

Shott, S. R., 2006 Down syndrome: common otolaryngologic manifestations. Am J Med Genet C Semin Med Genet 142c**:** 131-140.

Simpson, J. L., and S. Elias, 1993 Isolating fetal cells from maternal blood. Advances in prenatal diagnosis through molecular technology. JAMA 270**:** 2357-2361.

Sims, D., N. E. Ilott, S. N. Sansom, I. M. Sudbery, J. S. Johnson *et al.*, 2014 CGAT: computational genomics analysis toolkit. Bioinformatics 30**:** 1290-1291.

Sininger, Y. S., A. Martinez, L. Eisenberg, E. Christensen, A. Grimes *et al.*, 2009 Newborn hearing screening speeds diagnosis and access to intervention by 20-25 months. J Am Acad Audiol 20**:** 49-57.

Sliwinska-Kowalska, M., and M. Pawelczyk, 2013 Contribution of genetic factors to noise-induced hearing loss: a human studies review. Mutat Res 752**:** 61-65.

Snoeckx, R. L., P. L. Huygen, D. Feldmann, S. Marlin, F. Denoyelle *et al.*, 2005 GJB2 mutations and degree of hearing loss: a multicenter study. Am J Hum Genet 77**:** 945-957.

Storm, K., S. Willocx, K. Flothmann and G. Van Camp, 1999 Determination of the carrier frequency of the common GJB2 (connexin-26) 35delG mutation in the Belgian population using an easy and reliable screening method. Hum Mutat 14**:** 263-266.

Talkowski, M. E., C. Ernst, A. Heilbut, C. Chiang, C. Hanscom *et al.*, 2011 Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. Am J Hum Genet 88**:** 469-481.

Taylor, K. R., A. P. Deluca, A. E. Shearer, M. S. Hildebrand, E. A. Black-Ziegelbein *et al.*, 2013 AudioGene: predicting hearing loss genotypes from phenotypes to guide genetic screening. Hum Mutat 34**:** 539-545.

Theisen, A. (2008) Microarray-based comparative genomic hybridization (aCGH). Nature Education 1(1):45

Thys, M., and G. Van Camp, 2009 Genetics of otosclerosis. Otol Neurotol 30**:** 1021-1032.

Tobe, R. G., R. Mori, L. Huang, L. Xu, D. Han *et al.*, 2013 Cost-effectiveness analysis of a national neonatal hearing screening program in China: conditions for the scale-up. PLoS One 8**:** e51990.

Tono, T., K. Kiyomizu, K. Matsuda, S. Komune, S. Usami *et al.*, 2001 Different clinical characteristics of aminoglycoside-induced profound deafness with and without the 1555 A-->G mitochondrial mutation. ORL J Otorhinolaryngol Relat Spec 63**:** 25-30.

Toriello HV, Reardon W, Gorlin RJ, eds. *Hereditary Hearing Loss and Its Syndromes.* New York: Oxford University Press; 2004.

Ulubil, S. A., A. D. Furze and S. I. Angeli, 2006 Cochlear implantation in a patient with profound hearing loss with the A1555G mitochondrial DNA mutation and no history of aminoglycoside exposure. J Laryngol Otol 120**:** 230-232.

Usami, S., S. Abe, S. Nishio, Y. Sakurai, H. Kojima *et al.*, 2012 Mutations in the NOG gene are commonly found in congenital stapes ankylosis with symphalangism, but not in otosclerosis. Clin Genet 82**:** 514-520.

Van Eyken, E., L. Van Laer, E. Fransen, V. Topsakal, J. J. Hendrickx *et al.*, 2007 The contribution of GJB2 (Connexin 26) 35delG to age-related hearing impairment and noise-induced hearing loss. Otol Neurotol 28**:** 970-975.

Van Laer, L., E. H. Huizing, M. Verstreken, D. van Zuijlen, J. G. Wauters *et al.*, 1998 Nonsyndromic hearing impairment is associated with a mutation in DFNA5. Nat Genet 20**:** 194-197.

Van Laer, L., M. Pfister, S. Thys, K. Vrijens, M. Mueller *et al.*, 2005 Mice lacking Dfna5 show a diverging number of cochlear fourth row outer hair cells. Neurobiol Dis 19**:** 386-399.

Verhaert, N., C. Desloovere and J. Wouters, 2013 Acoustic hearing implants for mixed hearing loss: a systematic review. Otol Neurotol 34**:** 1201-1209.

Walknowska, J., F. A. Conte and M. M. Grumbach, 1969 Practical and theoretical implications of fetal-maternal lymphocyte transfer. Lancet 1**:** 1119-1122.

Wei, X., X. Ju, X. Yi, Q. Zhu, N. Qu *et al.*, 2011 Identification of sequence variants in genetic disease-causing genes using targeted next-generation sequencing. PLoS One 6**:** e29500.

Weston, M. D., M. L. Pierce, S. Rocha-Sanchez, K. W. Beisel and G. A. Soukup, 2006 MicroRNA gene expression in the mouse inner ear. Brain Res 1111**:** 95-104.

Wieacker, P., and J. Steinhard, 2010 The prenatal diagnosis of genetic diseases. Dtsch Arztebl Int 107**:** 857-862.

Wilch, E., H. Azaiez, R. A. Fisher, J. Elfenbein, A. Murgia *et al.*, 2010 A novel DFNB1 deletion allele supports the existence of a distant cis-regulatory region that controls GJB2 and GJB6 expression. Clin Genet 78**:** 267-274.

Wilch, E., M. Zhu, K. B. Burkhart, M. Regier, J. L. Elfenbein *et al.*, 2006 Expression of GJB2 and GJB6 is reduced in a novel DFNB1 allele. Am J Hum Genet 79**:** 174-179.

Wiley, T. L., R. Chappell, L. Carmichael, D. M. Nondahl and K. J. Cruickshanks, 2008 Changes in hearing thresholds over 10 years in older adults. J Am Acad Audiol 19**:** 281-292; quiz 371.

Wingfield, A., M. Panizzon, M. D. Grant, R. Toomey, W. S. Kremen *et al.*, 2007 A twin-study of genetic contributions to hearing acuity in late middle age. J Gerontol A Biol Sci Med Sci 62**:** 1294-1299.

Withrow, K. A., K. A. Tracy, S. K. Burton, V. W. Norris, H. H. Maes *et al.*, 2009 Impact of genetic advances and testing for hearing loss: results from a national consumer survey. Am J Med Genet A 149a**:** 1159-1168.

World Health Organization. Global Estimations on prevalence of heairng loss. Mortality and Burden of Diseases and Prevention of Blindness and Deafness 2012. http://www.who.int/pbd/deafness/WHO_GE_HL.pdf

World Health Organization. Newborn and infant hearing screening. Current Issues and Guiding Principles For Action. Geneva, Switzerland,09-10 Nov 2009 http://www.who.int/blindness/publications/Newborn_and_Infant_Hearing_Screening_Report.pdf.

Wu, C. C., Y. H. Lin, Y. C. Lu, P. J. Chen, W. S. Yang *et al.*, 2013 Application of massively parallel sequencing to genetic diagnosis in multiplex families with idiopathic sensorineural hearing impairment. PLoS One 8**:** e57369.

Yang, Y., D. M. Muzny, J. G. Reid, M. N. Bainbridge, A. Willis *et al.*, 2013 Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 369**:** 1502-1511.

Yin, A., C. Liu, Y. Zhang, J. Wu, M. Mai *et al.*, 2013 The carrier rate and mutation spectrum of genes associated with hearing loss in South China hearing female population of childbearing age. BMC Med Genet 14**:** 57.

You, Y., Y. Sun, X. Li, Y. Li, X. Wei *et al.*, 2014 Integration of targeted sequencing and NIPT into clinical practice in a Chinese family with maple syrup urine disease. Genet Med.

Zhang, Z., W. Ding, X. Liu, B. Xu, W. Du *et al.*, 2012 Auditory screening concurrent deafness predisposing genes screening in 10,043 neonates in Gansu province, China. Int J Pediatr Otorhinolaryngol 76**:** 984-988.

Zhao, H., R. Li, Q. Wang, Q. Yan, J. H. Deng *et al.*, 2004 Maternally inherited aminoglycoside-induced and nonsyndromic deafness is associated with the novel C1494T mutation in the mitochondrial 12S rRNA gene in a large Chinese family. Am J Hum Genet 74**:** 139-152.

Zhao, M., Q. Wang, Q. Wang, P. Jia and Z. Zhao, 2013 Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. BMC Bioinformatics 14 Suppl 11**:** S1.

Zheng, D., A. Frankish, R. Baertsch, P. Kapranov, A. Reymond *et al.*, 2007 Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. Genome Res 17**:** 839-851.

Zhu, Y., Q. Li, Z. Chen, Y. Kun, L. Liu *et al.*, 2009 Mitochondrial haplotype and phenotype of 13 Chinese families may suggest multi-original evolution of mitochondrial C1494T mutation. Mitochondrion 9**:** 418-428.