

Next-Generation Sequencing in Genetic Hearing Loss

Denise Yan,¹ Mustafa Tekin,² Susan H. Blanton,² and Xue Zhong Liu^{1,2}

The advent of the \$1000 genome has the potential to revolutionize the identification of genes and their mutations underlying genetic disorders. This is especially true for extremely heterogeneous Mendelian conditions such as deafness, where the mutation, and indeed the gene, may be private. The recent technological advances in target-enrichment methods and next generation sequencing offer a unique opportunity to break through the barriers of limitations imposed by gene arrays. These approaches now allow for the complete analysis of all known deafness-causing genes and will result in a new wave of discoveries of the remaining genes for Mendelian disorders. In this review, we describe commonly used genomic technologies as well as the application of these technologies to the genetic diagnosis of hearing loss (HL) and to the discovery of novel genes for syndromic and nonsyndromic HL.

Introduction

HEARING LOSS (HL) IS relatively common among all human populations, with profound congenital deafness present in about 1 in 1000 births. The causes of HL may be genetic, environmental, or multifactorial. However, at least 50% of prelingual HL in developed countries is of genetic origin (Vele and Schrijver, 2008). In addition, many cases of later onset progressive HL are genetic in origin, and genes also play an important role in progressive HL associated with ageing (presbycusis). To date, ~64 nonsyndromic HL genes, with more than 1000 mutations, have been implicated in hearing impairment (<http://deafnessvariationdatabase.org>), making it one of the most genetically heterogeneous traits. This extreme heterogeneity, as well as the unique mating structure (assortative mating) of deaf individuals, has hindered the search for HL-related genes by researchers and medical geneticists.

Great progress has been made in our understanding of the genetic basis and molecular mechanisms of hereditary deafness since the cloning of the first human nuclear deafness gene, X-linked *POU3F4* in 1995 (de Kok *et al.*, 1995). Until now, traditional methods of genetic identification have depended on Sanger sequencing of genes in candidate regions identified through linkage analysis. The advent of sequence capture enrichment strategies and massively parallel next generation sequencing (NGS) are expected to increase the rate of gene discovery for genetically heterogeneous condition such as HL and transform the practice of medical genetics (ten Bosch and Grody, 2008; Voelkerding *et al.*, 2009; Tucker *et al.*, 2009). With high-throughput sequencing technologies that are intended to lower the cost of DNA sequencing, it will be

possible to sequence hundreds or even thousands of exons and other genomic sequences in an individual with a suspected genetic disorder. Current genetic testing for HL proceeds by sequencing a series of genes, individually or in small sets, based on the relative causative frequency of the gene and recognizable features to guide selection of appropriate genes for testing (such as enlarged vestibular aqueducts, low-frequency HL, or auditory neuropathy). If the underlying cause involves a rare gene, the causative mutation often remains unknown even after very extensive and expensive molecular testing. Now the availability of targeted NGS-based gene panels, such as OtoSCOPE (Otolological Sequence Capture Of Pathogenic Exons), developed by the University of Iowa (www.healthcare.uiowa.edu/labs/morl/) and the OtoGenome Test (Harvard Medical School; <http://pcpgm.partners.org/Imm/tests/hearing-loss/OtoGenome>) makes it possible to detect variants in known human genes associated with HL efficiently and at low cost.

NGS: Platforms and Technologies

Since its first introduction in 1977, targeted Sanger sequencing has been the gold standard for screening in the research community and clinical molecular diagnostics. It was revolutionized by the advent of polymerase chain reaction in 1983. In an industrial high-throughput setting, Sanger technology was used for sequencing the first human genome, completed in 2003. This was the result of a 13-year effort with an estimated cost of \$2.7 billion as part of the Human Genome Project (Lander *et al.*, 2001; McPherson *et al.*, 2001; Sachidanandam *et al.*, 2001). This draft of the human genome consists of roughly 3.3 billion base pairs; the size presents a challenge

¹Department of Otolaryngology and ²Dr. John T. Macdonald Department of Human Genetics, John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, Florida.

in terms of both speed and cost requirements. In 2008, the sequencing of a human genome was completed in 5 months, for approximately \$1.5 million (Wheeler *et al.*, 2008). This monumental accomplishment was enabled by the commercial launch of the first massively parallel pyrosequencing platform in 2005, which has ushered in a new era for high-throughput NGS strategies. These new sequencing technologies replace the one small-template-at-a-time paradigm of Sanger sequencing with massively parallel sequencing of millions of small fragments covering the whole genome as well as individual regions with great redundancy. The characteristics of the existing NGS platforms include methods that are grouped broadly as template preparation, sequencing and imaging, and data analysis.

The development of instrumentation for DNA sequencing initially intended for research purposes resulted in high-throughput sequencing technologies such as the 454 GS FLX (Roche), SOLiD (Life Technologies) and the HiSeq series (Illumina). However, because of the large amount of throughput and long duration of runs clearly incompatible with routine clinical use, companies have developed bench-top DNA sequencing instrumentation such as 454 GS Jr (Roche), MiSeq (Illumina), and IonTorrent Personal Genome Machine from Life technologies. The first two platforms are based on pyrosequencing or sequencing by DNA synthesis, the Ion Torrent's technology, on the other hand, uses a hydrogen ion sensing semiconductor chip (Rothberg *et al.*, 2011). Although the sequencing throughput is far less, the bench-top sequencers are more economical for smaller sample volumes in addition to the rapid turn-around time available on these newer instruments. Thus, in combination with target enrichment, they are more suitable for clinically oriented applications, where in-depth analysis of selected target genes is performed, enabling the detection of rare causal variants or mutants in a heterogeneous sample, such as cancer samples. DNA sequencing technologies and platforms are being updated and developed at a blistering pace. With the availability of a multitude of platforms and dramatically lower costs of sequencing, NGS technologies are expected to have a major impact on the way we practice medicine in the near future.

Sequencing

Targeted genomic capture

Targeted genomic capture and massively parallel sequencing technologies are revolutionizing genomics by making it possible to sequence complete genomes of model organisms. However, the cost and capacity required are still high, especially considering that the functional significance of intronic and intergenic noncoding DNA sequences is still largely unknown. One application that these technologies are well suited for is the resequencing of many selected parts of a genome such as all exons from a large set of genes. This requires that the targeted parts of the genome are highly enriched in the sample. Recent technological changes, such as genome capture, genome enrichment, genome partitioning, have successfully been used to enrich for large parts of the genome (Summerer, 2009; Turner *et al.*, 2009; Mamanova *et al.*, 2010). The targeted fragments can subsequently be captured using solid- or liquid-phase hybridization assays (Okou *et al.*, 2007; Gnirke *et al.*, 2009).

Exome capture

Through linkage mapping and candidate gene resequencing, loci of ~3000 of all known or suspected Mendelian diseases have been identified (Online Mendelian Inheritance in Man, www.ncbi.nlm.nih.gov/omim). For these loci, ~85% of the disease-causing mutations are located in the coding region or in canonical splice sites (Choi *et al.*, 2009).

The total size of the human exome is ~30 Mb, which comprises ~180,000 exons and constitute about 1% of the entire human genome. Currently, sequencing whole genomes is still a substantial undertaking, which is not a routine procedure that can be done on hundreds of samples. At present, exome sequencing represents an alternative in which, approximately 30–70 Mb sequences encompassing exons and splice sites are targeted, enriched, and sequenced using commercially available sequence capture methods. Several Human Exome Sequence Capture kits are now commercially available. These include the Agilent SureSelect Human All Exon Kit, the Illumina TruSeq Exome Enrichment Kit, the TargetSeq in-solution target enrichment kit from Life Tech/Applied Biosystems and SeqCap EZ Exome from Roche NimbleGen. Although these commercial kits have expanded the coverage beyond the exome, unknown or yet-to-be-annotated exons, evolutionary conserved noncoding regions, and regulatory sequences, such as enhancers or promoters, are not typically captured. Another limitation is that the capture efficiency varies considerably, and some regions may be poorly represented. Another consideration is that NGS technologies have higher raw base calling error rates in comparison with traditional Sanger sequencing. This makes the resequencing of mutant or variant genes using conventional sequencing techniques important for validation.

Data Analysis of NGS

The large amount of data derived from NGS platforms imposes increasing demands on statistical methods and bioinformatic tools for the analysis. Although the NGS platforms rely on different principles and differ in how the array is made, their work flows are conceptually very similar. All of them generate millions or billions of short sequencing reads simultaneously. Several layers of analysis are necessary to convert these raw sequence data into understanding of functional biology. These include alignment of sequence reads to a reference, base-calling and/or polymorphism detection, *de novo* assembly from paired or unpaired reads, and structural variant detection. To date, a variety of software tools are available for analyzing NGS data (<http://seqanswers.com/>). Although tremendous progress has been achieved over the last several years in the development of computational methods for the analysis of high-throughput sequencing data, there are still many algorithmic and informatics challenges remaining. For example, even if a plethora of alignment tools have been adapted or developed for the reconstruction of full human genotypes from short reads, this task remains an extremely challenging problem. Also, when a high-throughput technology is used to sequence an individual (the donor), any genetic difference between the sequenced genomes and a reference human genome—typically the genome maintained at NCBI, is called the variant. Although this reference genome was built as a mosaic of several individuals, it is haploid, and may not contain a number of genomic segments present in

other individuals. By simply mapping reads to the reference genome, it is impossible to identify these segments. Thus, *de novo* assembly procedures should be used instead. Nonetheless, NGS technologies continue to change the landscape of human genetics. The resulting information has both enhanced our knowledge and expanded the impact of the genome on biomedical research, medical diagnostics and treatment, and has accelerated the pace of gene discovery.

Deafness Genes Discovery Through Massively Parallel Sequencing

Traditional strategies for disease gene discovery (Botstein and Risch, 2003) include those based on linkage analysis (Lander and Botstein, 1986) as well as homozygosity mapping (Lander and Botstein, 1987) in which, genetic markers are used to reveal genomic regions, which are shared by affected individuals. This would then be followed up by positional cloning and candidate gene studies within the genomic regions to identify the causal variants. Gene discovery efforts for hearing disorders are complicated by the extreme heterogeneity. Consanguineous kindreds have proven extremely useful in mapping genes for autosomal recessive forms of HL. However, even in pedigrees in which a locus is identified, there is often an insufficient number of recombination events to narrow down the candidate interval, resulting in a large region often consisting of several megabases of genomic DNA under the linkage peak(s). Assortative matings are common within the deaf community, such that multiple genes may be segregating in a single pedigree. Moreover, a wide spectrum of genetic and environmental factors is expected to influence age-related HL, or presbycusis, which further underscores our lack of understanding of genetic variants leading to deafness.

Recent advances in sequencing technology, including enrichment by either solid-phase or in-solution targeted capture (Mamanova *et al.*, 2010) allows for targeting custom genomic regions of interest for up to hundreds of kilobases in size or capture the entire protein-coding sequence of an individual (the exome, over 30 Mb) for sequencing. This sequence cap-

ture method coupled with the high-throughput sequencing data produced by NGS technologies ensures an adequate depth of sequencing coverage to accurately detect the variants in the exome or targeted regions (Mamanova *et al.*, 2010; Metzker, 2010). However, the major challenge is to distinguish between background polymorphisms and potentially disease-causing mutations (Sirmaci *et al.*, 2012)

For nonsyndromic HL, ~125 autosomal loci have been mapped (54 for dominant and 71 for recessive deafness), with 64 genes identified to date (24 for dominant and 40 for recessive deafness). Of the five X-linked loci, genes for three of them have been cloned (<http://hereditaryhearingloss.org>). Hundreds of syndromic forms of deafness have been described, and for many of them, the underlying genes still await discovery. Since the introduction of the first NGS technology in 2004, more than 1000 NGS-related manuscripts have been published. Until now, approximately a dozen of genes for HL have been successfully determined using NGS (Shearer *et al.*, 2011; Lin *et al.*, 2012; Table 1). Most likely, the identification of other as yet unidentified deafness genes will soon follow.

Nonsyndromic HL

C9orf75, encoding Taperin (*TPRN*), was the first deafness gene identified using comprehensive targeted genomic enrichment and NGS to analyze 2.9 Mb of the *DFNB79* interval on chromosome 9q34.3 defined in a consanguineous Pakistani family (Rehman *et al.*, 2010). After filtering for quality of data obtained from the use of custom probes designed for enrichment from Roche NimbleGen and 454 sequencing, 2231 variants were identified; eight of which were homozygous and confirmed by Sanger sequencing. Screening of ethnically matched controls and segregation analysis led to the identification of a nonsense mutation in the *TPRN* gene. Analysis of three additional *DFNB79*-linked families previously reported by Khan *et al.* (2010), identified three other frameshift mutations. Li *et al.* (2010) also identified a homozygous loss of function mutations in the *TPRN* gene in affected members of a large consanguineous Moroccan family and a Dutch family

TABLE 1. DEAFNESS GENES IDENTIFIED USING GENOMIC CAPTURE AND MASSIVELY PARALLEL SEQUENCING

Hearing status	Locus/disorder	Gene	Strategy	References
Nonsyndromic Recessive	DFNB79	<i>TPRN</i>	Targeted enrichment of genomic locus	Rehman <i>et al.</i> (2010)
	DFNB82	<i>GPSM2</i>	Whole exome	Walsh <i>et al.</i> (2010)
Dominant	DFNB84	<i>OTOGL</i>	Whole exome	Yariz <i>et al.</i> (2012)
	DFNA4	<i>CEACAM16</i>	Whole exome	Zheng <i>et al.</i> (2011)
	DFNA41	<i>P2RX2</i>	Targeted enrichment of genomic locus	Yan <i>et al.</i> (2013)
X-Linked	DFNX4	<i>SMPX</i>	Targeted enrichment of genomic locus	Schraders <i>et al.</i> (2011)
Syndromic	Perrault syndrome	<i>HSD17B4</i>	Whole exome	Pierce <i>et al.</i> (2010)
	Perrault syndrome	<i>HARS2</i>	Targeted enrichment of genomic locus	Pierce <i>et al.</i> (2011)
	Carnevale, Malpuech, Michels, and oculo-skeletal-abdominal syndromes	<i>MASP1</i>	Whole exome	Sirmaci <i>et al.</i> (2010)
	Hereditary sensory and autonomic neuropathy type 1 (HSAN1) with dementia and hearing loss	<i>DNMT1</i>	Whole exome	Klein <i>et al.</i> (2011)

with DFNB79. The taperin protein has similarity to phosphatase, an actin filament pointed-end-capping protein reported to modulate the actin cytoskeleton (Kao *et al.*, 2007; Lai *et al.*, 2009).

Walsh *et al.* (2010) reported how homozygosity mapping in conjunction with exome sequencing led to the rapid identification of the causative allele for the DFNB82 locus at 1p13.3 in a consanguineous Palestinian family. The Agilent SureSelect 38 Mb All Exon Kit was used. The exome design covers sequences corresponding to the exons and flanking intronic regions of 23,739 genes and also encompasses 700 miRNAs and 300 noncoding RNAs. Alignment of the sequence reads in the linkage region revealed that sequencing of 93% of protein-coding sequences in the DFNB82 interval at 10-fold was achieved using the Illumina Analyzer Iix. After data filtering, seven unreported variants were identified, of which, a nonsense mutation located in the *GPSM2* gene was found to be responsible for DFNB82. Studies in a knockout mouse model of *Gpsm2* suggest that *GPSM2* may be essential for planar spindle orientation during apical divisions (Konno *et al.*, 2008).

Defects in *MYH14* have been previously identified in DFNA4 families (Chen *et al.*, 1995; Donaudy *et al.*, 2004). However, screening of an American family that originally defined the DFNA4 locus (Chen *et al.*, 1995) revealed no mutations in the coding region of the *MYH14* gene. Genotyping of SNPs close to the *MYH14* gene excluded it from the candidate region (Yang *et al.*, 2005). Recently, the Agilent SureSelect Human All Exon 50 Mb Kit and the SOLiD v4 sequencing system were used to confirm that a mutation of *CEACAM16* is responsible for DFNA4. The sequence alignment showed that 98.2% of the DFNA4 interval was targeted, with 70.2% of the region covered by at least 10 high-quality reads. The *CEACAM16* protein is thought to have a role in maintaining the integrity of the tectorial membrane (TM) as well as in the connection between the outer hair cell stereocilia and TM that is essential for mechanical amplification (Zheng *et al.*, 2011).

Schraders *et al.* (2011) identify *SMPX* as the causative gene at the DFNX4 locus in a large Dutch family with progressive, nonsyndromic deafness. Targeted enrichment using the Agilent SureSelect Human X Chromosome Kit and NGS on the Illumina GAII sequencer were performed. A total of 95.1% of the targeted bases was covered at least 10-fold. Bioinformatics analysis led to the identification of two novel gene variants, but only a nonsense mutation in the *SMPX* gene was located within the critical region. This mutation was found to segregate with HL in the DFNX4 family and was absent from control individuals. Screening of 26 additional small X-linked families with HL resulted in identification of another family with a single base pair deletion in the *SMPX* gene, confirming mutations of this gene as causative of DFNX4 deafness. *SMPX* encodes a proline-rich protein that may function in the inner ear development and/or maintenance in the IGF-1/integrin pathways (Huebner *et al.*, 2011; Schraders *et al.*, 2011).

Yariz *et al.* (2012) identified *OTOGL*, encoding otogelin-like, as the gene responsible for DFNB84 in a large consanguineous family. By linkage analysis, an autozygous region spanning 15 Mb on chromosome 12 and including the DFNB84 locus was found to be shared by affected members of the family. The Agilent SureSelect Human All Exon 50 MB kit was used with an Illumina HiSeq 2000 instrument. Applying a mini-

um depth of 43, 84% of the candidate region was covered with an average read depth of 100×. After multiple filters, whole exome data revealed two novel homozygous variants both located in *OTOGL* (hg19): a one base pair deletion, c.1430 delT (NM_173591.3), predicted to lead to a reading frame shift and a premature termination of protein synthesis, p.Val477-Glufs*25 (NP_775862.3), and c.4132T>C (NM_173591.3) predicted to result in p.Cys1378Arg (NP_775862.3). Both variants cosegregated with the phenotype in the family and were not detected in 373 ethnicity-matched control samples. Neither change has been found in the NHLBI Exome Sequencing Project (Exome Variant Server) or in dbSNP databases. However, because p.Cys1378Arg is located at the C-terminal relative to p.Val477Glufs*25, it is unlikely to contribute to the phenotype. Two compound heterozygous mutations, c.547C>T (p.Arg183*) and c.5238p5G>A, were subsequently identified in a second multiplex family, which provided further genetic support for the causative role of *OTOGL* mutations in HL.

We recently report that a defect of the ATP-gated P2X2 (ligand-gated ion channel, purinergic receptor 2) receptor is responsible for DFNA 41 (Yan *et al.*, 2013). To identify the gene, we designed a total of 3636 cRNA 120-mer overlapping probes to capture the 4.80-MB region spanning chr12:129,051,849–133,851,895 (hg19) that defines the DFNA41 locus. We captured and sequenced all 427 documented RefSeq, UniProt, and CCDS exons and flanking intronic splice sites (<http://genome.ucsc.edu>) in three affected family members, using the Agilent SureSelect Target Enrichment system and an Illumina HiSeq2000 analyzer to a median read depth of 350×, with 99.2% of the targeted regions covered by ≥10 reads. DNA variants were filtered against common polymorphisms documented by dbSNP135 or the Institute's Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>). Only a nonsynonymous change was found to alter a conserved nucleotide site corresponding to *P2RX2* c.178G>T (p.V60L NM_174873). This variant was subsequently confirmed by Sanger sequencing and shown to cosegregate with HL in the family defining DFNA41. Sequencing of the entire *P2RX2* coding sequence in 65 other families with autosomal dominant nonsyndromic sensorineural HL yielded one other family, also Chinese, carrying *P2RX2* c.178G>T. Finally, results from our human and mouse studies suggested that dysfunction of the ATP-gated P2X2 receptor leads to progressive HL and increased susceptibility to noise.

Syndromic deafness

Using NGS, mutations in different genes have been found to cause the Perrault syndrome. It is a rare recessive disorder characterized by ovarian dysgenesis in females and sensorineural HL in females and males (Perrault *et al.*, 1951).

In a nonconsanguineous large family of mixed European ancestry, previously described by Pallister and Opitz (1979), genome-wide linkage analysis mapped the locus at 5q31 (Pierce *et al.*, 2011). Screening of all the 58 genes in the candidate interval by Sanger sequencing led to the detection of two variants in the *HARS2* gene that cosegregated with the Perrault syndrome phenotype and were not found in control subjects. *HARS2* encodes a histidyl tRNA synthetase that is predicted to function in mitochondria (O'Hanlon *et al.*, 1995). To confirm deafness-causing variants in *HARS2*, the entire

4.142-Mb linkage region was targeted for enrichment using Agilent SureSelect technology and sequenced on an Illumina GAIIX yielding 97% of targeted bases having at least 20-fold coverage. *HARS2* was the only gene found in this region with two variants that affected conserved residues. However, one of the nucleotide substitutions also created an alternate splice leading to deletion of 12 codons from the *HARS2* message. Affected family members thus carried three mutant *HARS2* transcripts. In *Caenorhabditis elegans*, reduction of expression of the histidyl tRNA synthetase *hars-1* severely compromised fertility, explaining its implication in gonadal dysgenesis.

In another family with Perrault syndrome, the pedigree is too small for linkage mapping (Pierce *et al.*, 2010). Instead, the Agilent SureSelect Kit and sequencing on an Illumina GAIIX yielding 93.1% coverage by at least 10 high-quality reads were performed to identify the gene responsible. All the genes harboring the 207 rare variants were analyzed for cosegregation with the disease. Only one gene, *HSD17B4* on chromosome 5q23.1, contained two variants (one nonsense and one missense). Sanger sequencing subsequently confirmed that the affected family members carried both mutations. The protein encoded by *HSD17B4*, 17 β -hydroxysteroid dehydrogenase type 4 (HSD17B4), also known as the D-bifunctional protein (DBP) and multifunctional protein 2 (MFP-2), is a multifunctional peroxisomal enzyme involved in fatty acid β -oxidation and steroid metabolism (de Launoit and Adamski, 1999). Deficiency in the HSD17B4 protein likely explains the symptoms of Perrault syndrome observed in this family (Pierce *et al.*, 2010).

MASPI responsible for the Malpuech-Michels-Mingarelli-Carnevale (3MC) syndrome was identified by targeted genomic capture and NGS. Patients with 3MC syndrome present with facial, umbilical, coccygeal, and auditory findings of Carnevale, Malpuech, Michels, and oculo-skeletal-abdominal syndromes. In two consanguineous Turkish families with findings characteristic of these syndromes, homozygosity mapping yielded an autozygous region on chromosome 3q27 (Sirmaci *et al.*, 2010). In one family, whole exome sequencing using the Agilent SureSelect Kit and Illumina GAIIX sequencing achieved $>50\times$ median coverage for this region. Four rare variants were identified in the region. Only a missense substitution in the *MASPI* gene, segregated with the syndrome in the family. Sanger sequencing of the complete *MASPI* gene in the second family revealed a nonsense mutation. Neither mutation in *MASPI* was detected in controls supporting a causal link with the disorder. The *MASPI* protein is a mannan-binding lectin serine protease known to activate the complement pathway by binding lectin. The two mutations in a *MASPI* isoform are believed to play a critical role in insulin growth factor availability during craniofacial and muscle development (Sirmaci *et al.*, 2010).

Klein *et al.* (2011) identified the genetic cause of one form of autosomal dominant hereditary sensory neuropathy with dementia and HL using exome sequencing. Four kindreds, two of which were American, one Japanese, and one European with early onset dementia and sensorineural HL associated with sensory neuropathy, were investigated. A genome-wide linkage analysis in the largest kindred mapped the locus to 19p13.2. To increase the sequencing coverage and depth, two different combined methodologies were used—Agilent SureSelect 38 Mb All Exon Kit and Illumina GAIIX sequencing, and NimbleGen 2.1M Human Exome and Roche454 sequencing. Filtering and bioinformatics analysis

revealed an unreported nonsynonymous heterozygous mutation in the DNA methyltransferase 1 (*DNMT1*) gene that segregated with disease status. Subsequent Sanger sequencing of all 41 exons of *DNMT1* led to the identification of heterozygous missense mutations in the remaining three kindreds that segregated with disease and were not found in controls. These mutations were shown to cause premature degradation of mutant proteins, reduce the methyltransferase enzyme activity, and lead to global hypomethylation and site-specific hypermethylation (Klein *et al.*, 2011).

Conclusion

During the past several decades, continuous improvements in genomic technology have greatly improved our understanding of health and disease. The development of new DNA target enrichment technologies coupled with NGS technologies have already accelerated the pace of novel deafness gene discovery. This latter achievement is an essential starting point for both uncovering the molecular pathogenesis of HL and for providing clues to therapeutic approaches. The capacity to analyze thousands of genes simultaneously provides a powerful tool for detecting pathogenic mutations in disorders with genetic and phenotypic heterogeneity such as deafness. The availability of targeted NGS-based gene panels such as OtoSCOPE and OtoGenome Test makes indeed now, comprehensive genetic testing for all genes known to cause nonsyndromic HL and syndromes that can present as nonsyndromic such as Usher and Pendred possible.

As genomic information becomes more affordable and readily available, it will have a broad impact in the medical world. The inclusion of genetic information in healthcare has the potential to provide patients with valuable risk assessments based on their family history and genetic profile, and to carve a niche for personalized medicine. Over the next decade, most of the variant genes responsible for deafness will be identified and, such knowledge will lead to the development of practical treatments.

Acknowledgments

This work was supported by the National Institutes of Health grants R01DC05575, 1 R01 DC012546-01, translational R01DC012115, and the Hurong Scholar award at the Central-South University in Hunan, China to X.Z.L and R01DC009645 to M.T.

Author Disclosure Statement

No competing financial interests exist.

References

- Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat Genet* 33:228–237.
- Chen AH, Ni L, Fukushima K, *et al.* (1995) Linkage of a gene for dominant non-syndromic deafness to chromosome 19. *Hum Mol Genet* 4:1073e1076.
- Choi M, Scholl UI, Ji W, *et al.* (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* 106:19096–19101.
- de Kok YJ, van der Maarel SM, Bitner-Glindzicz M, *et al.* (1995) Association between X-linked mixed deafness and mutations in the POU domain gene POU3F4. *Science* 267:685–688.

- Donaudy F, Snoeckx R, Pfister M, *et al.* (2004) Nonmuscle myosin heavy-chain gene MYH14 is expressed in cochlea and mutated in patients affected by autosomal dominant hearing impairment (DFNA4). *Am J Hum Genet* 74:770–776.
- de Launoit Y, Adamski J (1999) Unique multifunctional HSD17B4 gene product: 17 β -hydroxysteroid dehydrogenase 4 and D-3-hydroxyacyl-coenzyme A dehydrogenase:hydratase involved in Zellweger syndrome. *J Mol Endocrinol* 22: 227–240.
- Gnirke A, Melnikov A, Maguire J, *et al.* (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27:182–189.
- Huebner AK, Gandia M, Frommolt P, *et al.* (2011) Nonsense mutations in SMPX, encoding a protein responsive to physical force, result in X-chromosomal hearing loss. *Am J Hum Genet* 88:621–627.
- Kao SC, Chen CY, Wang SL, *et al.* (2007) Identification of phostensin, a PP1 F-actin cytoskeleton targeting subunit. *Biochem Biophys Res Commun* 356:94–598.
- Khan SY, Riazuddin S, Shahzad M, *et al.* (2010) DFNB79: reincarnation of a nonsyndromic deafness locus on chromosome 9q34.3. *Eur J Hum Genet* 18:125–129.
- Klein CJ, Botuyan MV, Wu Y, *et al.* (2011) Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat Genet* 43:595–600.
- Konno D, Shioi G, Shitamukai A, *et al.* (2008) Neuroepithelial progenitors undergo LGN-dependent planar divisions to maintain self-renewability during mammalian neurogenesis. *Nat Cell Biol* 10:93–101.
- Lai NS, Wang TF, Wang SL, *et al.* (2009) Phostensin caps to the pointed end of actin filaments and modulates actin dynamics. *Biochem Biophys Res Commun* 387:676–681.
- Lander ES, Linton LM, Birren B, *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 2001;409: 860e921
- Lander ES, Botstein D. (1986) Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphism. *Proc Nat Acad Sci U S A* 83:7353–7735.
- Lander ES, Botstein D (1987) Homozygosity mapping. a way to map human recessive traits with the DNA of inbred children. *Science* 236:1567–1570.
- Li Y, Pohl E, Boulouiz R, *et al.* (2010) Mutations in TPRN cause a progressive form of autosomal recessive nonsyndromic hearing loss. *Am J Hum Genet* 86:479–484.
- Lin X, Tang W, Ahmad S, *et al.* (2012) Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. *Hear Res* 288:67–76.
- Mamanova L, Coffey AJ, Scott CE, *et al.* (2010) Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7:111–118.
- McPherson JD, Marra M, Hillier L, *et al.* (2001) A physical map of the human genome. *Nature* 409:934e41.
- Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev Genet* 11:31–46.
- O’Hanlon TP, Raben N, Miller FW (1995) A novel gene oriented in a head-to-head configuration with the human histidyl-tRNA synthetase (HRS) gene encodes an mRNA that predicts a polypeptide homologous to HRS. *Biochem Biophys Res Commun* 210:556–566.
- Okou DT, Steinberg KM, Middle C, *et al.* (2007) Microarray-based genomic selection for high-throughput resequencing. *Nat Methods* 4:907–909.
- Pallister PD, Opitz JM (1979) The Perrault syndrome: Autosomal recessive ovarian dysgenesis with facultative, non-sex-limited sensorineural deafness. *Am J Med Genet* 4:239–246.
- Perrault M, Klotz B, Housset E (1951) [Two cases of Turner syndrome with deaf mutism in two sisters]. *Bull Mem Soc Med Hop Paris* 67:79–84.
- Pierce SB, Walsh T, Chisholm KM, *et al.* (2010) Mutations in the DBP-deficiency protein HSD17B4 cause ovarian dysgenesis, hearing loss, and ataxia of Perrault syndrome. *Am J Hum Genet* 87:282–288.
- Pierce SB, Chisholm KM, Lynch ED, *et al.* (2011) Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. *Proc Natl Acad Sci U S A* 108:6543–6548.
- Rehman AU, Morell RJ, Belyantseva IA, *et al.* (2010) Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. *Am J Hum Genet* 86:378–388.
- Rothberg JM, Hinz W, Rearick TM, *et al.* (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475:348–352.
- Sachidanandam R, Weissman D, Schmidt SC, *et al.* (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409: 928e33.
- Schraders M, Haas SA, Weegerink NJD, *et al.* (2011) Next-generation sequencing identifies mutations of SMPX, which encodes the small muscle protein, X-linked, as a cause of progressive hearing impairment. *Am J Hum Genet* 88: 628–634.
- Shearer AE, Hildebrand MS, Sloan CM, *et al.* (2011) Deafness in the genomics era. *Hear Res* 282:1–9.
- Sirmaci A, Walsh T, Akay H, *et al.* (2010) MASP1 mutations in patients with facial, umbilical, coccygeal, and auditory findings of Carnevale, Malpuech, OSA, and Michels syndromes. *Am J Hum Genet* 87:679–686.
- Sirmaci A, Edwards YJ, Akay H, Tekin M (2012) Challenges in whole exome sequencing: an example from hereditary deafness. *PLoS One* 7:e32000.
- Summerer D (2009) Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. *Genomics* 94:363–368.
- ten Bosch JR, Grody WW (2008) Keeping up with the next generation: massively parallel sequencing in clinical diagnostics. *J Mol Diagn* 10:484–492.
- Tucker T, Marra M, Friedman JM (2009) Massively parallel sequencing: the next big thing in genetic medicine. *Am J Hum Genet* 85:142–154.
- Turner EH, Ng SB, Nickerson DA, *et al.* (2009) Methods for genomic partitioning. *Annu Rev Genomics Hum Genet* 10:263–284.
- Vele O, Schrijver I (2008) Inherited hearing loss: molecular genetics and diagnostic testing. *Expert Opin Med Diagn* 2: 231–248.
- Voelkerding KV, Dames SA, Durtschi JD (2009) Next-generation sequencing: from basic research to diagnostics. *Clin Chem* 55:641–658.
- Walsh T, Shahin H, Elkan-Miller T, *et al.* (2010) Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPSM2 as the cause of nonsyndromic hearing loss DFNB82. *Am J Hum Genet* 87: 90–94.
- Wheeler DA, Srinivasan M, Egholm M, *et al.* (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452:872–876.

- Yariz KO, Duman D, Seco CZ, *et al.* (2012) Mutations in OTOGL, encoding the inner ear protein otogelin-like, cause moderate sensorineural hearing loss. *Am J Hum Genet* 91: 872–882.
- Yan D, Zhu Y, Walsh T, *et al.* (2013) Mutation of the ATP-gated P2X2 receptor leads to progressive hearing loss and increased susceptibility to noise. *Proc Natl Acad Sci U S A* 110:2228–2233.
- Yang T, Pfister M, Blin N, *et al.* (2005) Genetic heterogeneity of deafness phenotypes linked to DFNA4. *Am J Med Genet A* 139:9–12.
- Zheng J, Miller KK, Yang T, *et al.* (2011) Carcinoembryonic antigen-related cell adhesion molecule 16 interacts with alpha-

tectorin and is mutated in autosomal dominant hearing loss (DFNA4). *Proc Natl Acad Sci U S A* 108:4218–4223.

Address correspondence to:
Xue Zhong Liu, MD, PhD
Department of Otolaryngology
Miller School of Medicine
University of Miami
1666 NW 12th Avenue
Miami, FL 33136
E-mail: xliu@med.Miami.edu