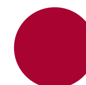
UNIVERSITY OF COPENHAGEN



Panel-based NGS reveals disease-causing mutations in hearing loss patients using BGISEQ-500 platform

Sun, Yan; Yuan, Jing; Wu, Limin; Li, Min; Cui, Xiaoli; Yan, Chengbin; Du, Lique; Mao, Liangwei; Man, Jianfen; Li, Wei; Kristiansen, Karsten; Wu, Xuan; Pan, Weijun; Yang, Yun; Yang, Yun

Published in:
Medicine

DOI:
10.1097/MD.0000000000014860

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC

Citation for published version (APA):
Sun, Y., Yuan, J., Wu, L., Li, M., Cui, X., Yan, C., ... Yang, Y. (2019). Panel-based NGS reveals disease-causing mutations in hearing loss patients using BGISEQ-500 platform. *Medicine*, *98*(12), 1-7. [e14860]. https://doi.org/10.1097/MD.0000000000014860

Download date: 09. apr.. 2020

Observational Study

Medicine®

OPEN

Panel-based NGS reveals disease-causing mutations in hearing loss patients using BGISEQ-500 platform

Yan Sun, PhD[a,b,c], Jing Yuan, PhD[d], Limin Wu, PhD[e], Min Li, MS[f], Xiaoli Cui, BS[a], Chengbin Yan, MS[a], Lique Du, MS[c], Liangwei Mao, PhD[a,g], Jianfen Man, MS[a], Wei Li, PhD[a], Karsten Kristiansen, PhD[b], Xuan Wu, MS[a], Weijun Pan, PhD[h,*], Yun Yang, PhD[a,c,*]

Abstract

Hearing loss is a highly heterogeneous disease presented with various phenotypes. Genetic testing of disease-causing mutations plays an important role in precise diagnosis and fertility guidance of heredity hearing loss. Here we reported an effective method employing target enrichment and BGISEQ-500 platform to detect clinically relevant alterations for heredity hearing patients in a single assay.

In this study, we designed an array based chip, containing 127 genes related to hearing loss. Then we conducted targeted next-generation sequencing toward 58 patients to make a precise diagnosis using BGISEQ-500 platform.

We successfully detected disease-causing mutations in 77.59% (45/58) of the patients with hearing loss. Finally, a total of 62 disease-causing mutations were identified, including 31 missense, 17 Indel, 11 splicing, 2 synonymous, and 1 copy number variant. 58.06% (36/62) of which has never been reported before.

To our knowledge, this is the first report using BGISEQ-500 platform to investigate both syndromic and nonsyndromic hearing loss in the Chinese population. The results showed that this method can greatly assist and enhance hearing loss diagnosis and improve molecular diagnostics outcome.

Abbreviations: CNVs = copy number variants, DNBs = DNA NanoBalls, HL = hearing loss, NGS = next-generation sequencing, SHI = syndromic hearing impairment.

Keywords: BGISEQ-500, hearing loss, mutation detection, targeted NGS

1. Introduction

Hearing loss (HL) is a highly heterogeneous disease presenting with various phenotypes, mainly caused by environmental or genetic factors.[1] Large scale genetic screenings have been reported in several studies.[2,3] Early identification and management play a crucial role in improving the communication and language of the patients.[4] So effective, exact and timely genetic testing is really important.

HL is a highly heterogeneous disorder. According to the phenotypes of HL patients, hereditary HL can be classified as nonsyndromic HL and syndromic HL. There are no abnormali-

Editor: Anish Thachangattuthodi.

YS, JY, and LW contributed equally to this work.

The data reported in this study are available in the CNGB Nucleotide Sequence Archive https://db.cngb.org/cnsa; accession number CNP0000280). The datasets generated and/or analyzed during the current study are not publicly available due to the ethical principles of BGI-IRB, but are available from the corresponding author on reasonable request.

This study was approved by the ethics committee of BGI (BGI-IRB15083).

This work was supported by the National Natural Science Foundation of China (grant number 81502593 to Yan Sun). This work was in part supported by Shenzhen Peacock Plan (No. KQTD20150330171505310).

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

[a] BGI Genomics, BGI-Shenzhen, Shenzhen, China, [b] Department of Biology, University of Copenhagen, Copenhagen, Denmark, [c] BGI-Wuhan, BGI-Shenzhen, Wuhan, China, [d] Reproductive Medicine Center, Department of Obstetrics and Gynecology, The First Affiliated Hospital of Anhui Medical University, Anhui Province Key Laboratory of Reproductive Health and Genetics, Biopreservation and Artificial Organs, Anhui Provincial Engineering Research Center, Anhui Medical University, [e] Center for Reproductive Medicine and Prenatal Diagnosis, Anhui Provincial Hospital, The First Affiliated Hospital of University of Science and Technology of China, Hefei, [f] Prenatal Diagnosis Center, Yancheng Maternity and Child Health Care Hospital, Yancheng, Jiangsu Province, [g] College of Life Sciences, Hubei University, Wuhan, [h] The Center for Reproductive Medicine, Ma'anshan Maternal and Child Health Hospital, Ma'anshan, China.

[*] Correspondence: Weijun Pan, The Center for Reproductive Medicine, Ma'anshan Maternal and Child Health Hospital, Ma'anshan 243000, China (e-mail: 1092832863@qq.com); Yun Yang, BGI-Wuhan, Wuhan 430075, China (e-mail: yangyun@bgi.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc.
This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

Medicine (2019) 98:12(e14860)

Received: 29 September 2018 / Received in final form: 11 February 2019 / Accepted: 15 February 2019

http://dx.doi.org/10.1097/MD.0000000000014860

ties in other organs in nonsyndromic HL patients, which account for 70% of cases affected by hereditary HL. Syndromic HL, on the other hand, accounts for about 30% of the patients with hereditary HL. The inherit patterns of HL is various, including autosomal recessive (80%), autosomal dominant (20%), X-linked, Y-linked, and mitochondrial.[5,6]

The most frequently mutated genes/mutations for HL patients might be different across ethnicities. In the Chinese patients with hereditary HL, the most commonly identified disease-causing genes are *GJB2* (21%), *SLC26A4* (14.5%), and *MT-RNR1* (4%).[7,8] Genetic testing has been applied in newborn and carrier screening of hereditary HL.[9–11] Some researchers performed Sanger sequencing in the screening of frequent mutations. Wu et al performed genetic hearing screening in 5173 newborns from Taiwan using 4 recurrent deafness mutations, including 2 mutations in gene *GJB2* (c.235delC and p.V37I), 1 mutation in *SLC26A4* (c.919-2A>G), and on mitochondrial mutation (m.1555A>G).[12] Hao et al screened 142,417 neonates from Wuhan using 4 recurrent deafness mutations, including 1 mutation in gene *GJB2* (c.235delC), 1 mutation in gene *SLC26A4* (c.919-2A>G), and 2 mitochondrial mutations (m.1494C>T and m.1555A>G).[13] However, genetic hearing screening using a small number of recurrent mutations is insufficient. To date, there are more than a 100 genes described to be related to HL.[14–16] What is more, population-scale screening and diagnosis of recurrent mutations using Sanger sequencing are laborious, costly and time-consuming. To detect recurrent mutations in large cohort, tens of thousands of polymerase chain reaction (PCR) reactions need to be performed. With the development of next-generation sequencing (NGS), it is gradually starting to play an important role to sequence clinical relevant genes in various genetic disorders. The diagnostic yield in hereditary HL can be greatly improved compared with Sanger sequencing. What is more, NGS technology is more suitable in the diagnosis of diseases with high genetic heterogeneity (such as hereditary HL).[17–20]

In this study, in order to make an accurate genetic diagnosis for patients with HL, we developed an NGS method targeting 127 genes related to HL. We tested its performance in detecting clinically relevant mutations using BGISEQ-500 platform. We also applied this panel to 58 clinical HL cases, and identified disease-causing mutations in 45 patients. This method can greatly assist and enhance HL diagnosis and help to make a precise diagnosis for the patients. To the best of our knowledge, this is the first report to investigate both syndromic and nonsyndromic HL using BGISEQ-500 platform in the Chinese population.

2. Material and methods

2.1. Design of the hereditary HL panel

In this study, we designed a custom human array spanning 620,604 bp using Roche NimbleGen. The array targets exons and 10 bp flanking intronic sequences of a total of 127 HL related genes, which were collected based on 4 well-known databases (Deafness Variation Database, The Hereditary Hearing Loss Homepage, GeneReviews, Orphanet and published papers). This array not only contained reported genes related to nonsyndromic HL, but can also be used to diagnose a total of 39 syndromic hearing impairment (SHI) (data not shown). The 127 genes can be classified into 5 categories, including 45 genes related to autosomal recessive nonsyndromic hearing loss, 3 genes related to X-link nonsyndromic hereditary hearing loss, 29 genes related to autosomal dominant nonsyndromic hearing loss, 2 genes related to maternally inherited hearing loss, and 66 genes related to SHI (Fig. 1).

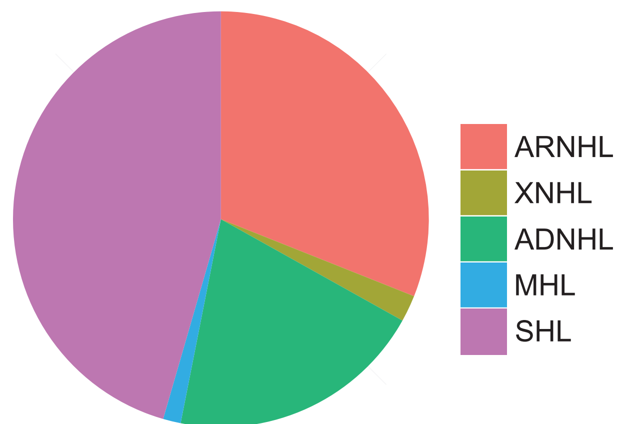


Figure 1. Gene classification of the HL panel. HL = hearing loss.

2.2. Sample collection, DNA extraction, library construction, and sequencing

In the present study, a total of 58 suspicious and confirmed HL patients were enrolled between April 2016 and May 2018 in Anhui Provincial Hospital, Ma'anshan Maternal and Child Health Hospital, and Yancheng Maternity and Child Health Care Hospital. Table 1 showed the age of the 58 patients (from 2 to 43 years old). In order to exclude HL patients due to nongenetic reasons, (such as injury and infection) and to characterize the severity of HL, clinical information of the patients was summarized by experienced clinicians.[21] Written informed consent was obtained from all the participants or their parents before sample collection. The ethics committee of BGI approved this study and all the protocols used in the study (BGI-IRB15083).

To the best of our knowledge, this is the first report to investigate both syndromic and nonsyndromic HL using BGISEQ-500 platform in the Chinese population. Figure 2 shows the procedures of library construction and sequencing using the BGISEQ-500 platform. First, genomic DNA of all the participants was isolated by QIAamp DNA Blood Midi Kit (Qiagen, Hilden, Germany). Library preparation was conducted at BGI following the guide of the standard protocol. Library preparation (BGISEQ-500RS High-throughput sequencing kit, PE50, V3.0, MGI Tech Co, Ltd, Shenzhen, China), hybridization and sequencing were performed according to the manufacturer's standard procedure provided by BGI (BGI-Shenzhen). In brief, enzymatic method including digestion by segmentase (MGI Tech Co, Ltd, BGI) was performed for DNA library construction of all the samples. Digestion reaction, end repair, and A-tailing were performed at the same time (37°C for 20 minutes, 65°C for 15 minutes, and hold at 4°C). DNA fragments were then ligated to adapter sequence after fragment selection. The ligation products were purified and amplified for 7 cycles with DNA fragments between 170 bp and 230 bp, and then subjected to the following hybridization process. Each library was pooled, and then hybridized to the HL chip for 16 hours at 65°C. After sequence capture, the libraries were washed, and then eluted according to the manufacturer's standard procedure (MGI Tech Co, Ltd). The

Table 1

Detection results of the 58 HL patients.

Sample name	Gender/age, yr old	Mutation	Mutation type	Mutation status	Reference	Validation
HL1	Female/31	OTOG (NM_001277269): c.1190G>A (p.Arg397Gln)	Het	Novel	–	–
HL2	Female/4	SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	–
		SLC26A4 (NM_000441): c.2000T>C (p.Phe667Ser);	Het;	Reported;	PMID: 28576516;	
		GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Het	Reported	PMID: 21329993	
HL3	Female/30	GJB2 (NM_004004): c.109G>A (p.Val37Ile)	Hom	Reported	PMID: 12121355	Yes
HL4	Male/9	SLC26A4 (NM_000441): c.1544+9C>T;	Het;	Reported;	PMID: 21704276;	Yes
		SLC26A4 (NM_000441): c.1519delT (p.Leu507Ter);	Het;	Novel;	–	
		WFS1 (NM_006005): c.2596G>A (p.Asp866Asn)	Het	Reported	PMID: 12955714	
HL5	Male/7	SLC26A4 (NM_000441): c.249G>A (p.Trp83Ter);	Het;	Novel;	-	Yes
		SLC26A4 (NM_000441): c.919-2A>G	Het	Reported	PMID: 18641518	
HL6	Female/43	MYO15A (NM_016239): c.3026C>A (p.Pro1009His);	Het;	Reported;	PMID: 23767834;	Yes
		MYO15A (NM_016239): c.5134-10C>G	Het	Novel	–	
HL7	Male/5	SLC26A4 (NM_000441): c.2027T>A (p.Leu676Gln);	Het;	Reported;	PMID: 24007330;	Yes
		SLC26A4 (NM_000441): c.915_916insG (p.Ile305IlefsX25)	Het	Reported	PMID: 26252218	
HL8	Male/18	SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	Yes
		SLC26A4 (NM_000441): c.1519 delT (p.Leu507TerfsX5)	Het	Novel	–	
HL9	Male/7	PTPRQ (NM_001145026): c.5942+1G>A;	Het;	Novel;	–	Yes
		PTPRQ (NM_001145026): c.6024G>A (p.Ser2008Ser)	Het	Novel	–	
HL10	Male/4	GJB2 (NM_004004): c.109G>A (p.Val37Ile);	Het;	Reported;	PMID: 12121355;	Yes
		MYO15A (NM_016239): c.3026C>A (p.Pro1009His);	Het;	Reported;	PMID: 23767834;	
		MYO15A (NM_016239): c.10420A>G (p.Ser3474Gly)	Het	Novel	–	
HL11	Female/9	GJB3 (NM_024009): c.538C>T (p.Arg180Ter);	Het;	Reported;	PMID: 22154049;	Yes
		ESPN (NM_031475): c.1464+2T>A	Het	Novel	–	
HL12	Male/24	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Het	Reported	PMID: 21329993	Yes
HL13	Male/43	MT-RNR1 (NC_012920): 908 A>G/m.1555A>G;	Hom;	Reported;	PMID: 24092330;	Yes
		MYO6 (NM_004999): c.118-2A>G	Het	Novel	–	
HL14	Male/26	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Hom	Reported	PMID: 21329993	Yes
HL15	Female/20	SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	–
		SLC26A4 (NM_000441): c.1686_1687insA (p.Phe562PhefsX12);	Het;	Reported;	PMID: 27792752;	
		GJB2 (NM_004004): c.109G>A (p.Val37Ile)				
			Het	Reported	PMID: 12121355	
HL16	Female/12	SLC26A4 (NM_000441): c.2168A>G (p.His723Arg);	Het;	Reported;	PMID: 23705809;	Yes
		SLC26A4 (NM_000441): c.2107C>G (p.Leu703Val)	Het	Reported	PMID: 26100058	
HL17	Female/7	TRIOBP (NM_001039141): c.2321delG (p.Arg774HisfsX105)	Het	Novel	–	Yes
HL18	Female/29	DIAPH3 (NM_001042517): c.3431delC (p.Thr1144SerfsX17);	Het;	Novel;	–	Yes
		MYO15A (NM_016239): c.10245_10247delCTC	Hom	Novel	–	
HL19	Male/25	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Hom	Reported	PMID: 21329993	Yes
HL20	Male/16	MT-RNR1 (NC_012920): m.1119T>C	Hom	Novel	–	–
HL21	Female/37	SLC26A4 (NM_000441): c.1229C>T (p.Thr410Met);	Het;	Reported;	PMID: 21961810;	Yes
		TRIOBP (NM_001039141): c.3256C>A (p.Pro1086Thr)	Hom	Novel	–	
HL22	Female/7	DSPP (NM_014208): c.1057delA (p.Lys353AsnfsX3)	Het	Novel	–	Yes
HL23	Male/28	DIAPH1 (NM_001079812): c.1982C>T (p.Pro661Leu)	Het	Novel	–	–
HL24	Female/5	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3);	Het;	Reported;	PMID: 21329993;	Yes
		MITF (NM_198159): c.1025_1032delGGAACAAG (p.Trp342TrpfsX25)	Het	Novel	–	
HL25	Female/28	SLC26A4 (NM_000441): EX5_6 DEL;	Het;	Reported;	PMID: 17443271;	Yes
		SLC26A4 (NM_000441): c.1339delA (p.Lys447SerfsX8);	Het;	Novel;	–;	
		TMC1 (NM_138691): c.1586_1587delTC (p.Val529ValfsX2)	Het	Reported	PMID: 21250555	
HL26	Female/28	CDH23 (NM_022124): c.9640C>A (p.Leu3214Met);	Het;	Novel;	–	–
		PCDH15 (NM_033056): c.3807G>T (p.Glu1269Asp)	Het	Novel	–	
HL27	Male/7	SLC26A4 (NM_000441): c.1336C>T (p.Gln446Ter);	Het;	Reported;	PMID: 17718863;	Yes
		SLC26A4 (NM_000441): EX5_6 DEL	Het	Reported	PMID: 17443271	
HL28	Female/4	MYO7A (NM_000260): c.562C>G (p.Gln188Glu)	Het	Novel	–	–
HL29	Female/4	GJB2 (NM_004004): c.257C>G (p.Thr86Arg);	Het;	Reported;	PMID: 21329993;	Yes
		GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Het	Reported	PMID: 21329993	
HL30	Female/8	SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	–
		SLC26A4 (NM_000441): c.1149+1G>A	Het	Reported	PMID: 21961810	
HL31	Female/2	GJB2 (NM_004004): c.109G>A (p.Val37Ile);	Het;	Reported;	PMID: 12121355;	Yes
		MYO7A (NM_000260): c.1004C>G (p.Ala335Gly)	Het	Novel	–	
HL32	Male/2	MYO7A (NM_000260): c.586C>G (p.Leu196Val);	Het;	Novel;	–	Yes
		MYO7A (NM_000260): c.1679A>G (p.Tyr560Cys)	Het	Novel	–	
HL33	Male/15	SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	Yes
		OTOF (NM_194248): c.4023+1G>A;	Het;	Reported;	PMID: 21935370;	
		KCNQ4 (NM_172163): c.1905G>A (p.Ser635Ser)	Het	Novel	–	

(*continued*)

Table 1

(continued).

Sample name	Gender/age, yr old	Mutation	Mutation type	Mutation status	Reference	Validation
HL34	Male/7	SLC26A4 (NM_000441): c.919-2A>G	Het	Reported	PMID: 12676893	Yes
HL35	Male/8	SLC26A4 (NM_000441): c.1001+5G>C;	Het;	Reported;	PMID: 27729126;	Yes
		SLC26A4 (NM_000441): c.1226G>A (p.Arg409His);	Het;	Reported;	PMID: 23273637;	
		USH2A (NM_206933): c.2802T>G (p.Cys934Trp)	Het	Reported	PMID: 25649381	
HL36	Male/7	EDN3 (NM_000114): c.142G>A (p.Glu48Lys)	Het	Novel	–	Yes
HL37	Male/3	MYO15A (NM_016239): c.9400C>T (p.Arg3134Ter);	Het;	Reported;	PMID: 23208854;	Yes
		MYO15A (NM_016239): c.10245_10247delCTC	Het	Novel	–	
HL38	Male/5	SLC26A4 (NM_000441): c.281C>T (p.Thr94Ile);	Het;	Reported;	PMID: 17718863;	Yes
		SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	
		TRIOBP (NM_001039141): c.4429_4430insG (p.Trp1477TrpfsX25);	Het;	Novel;	–	
HL39	Female/23	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Hom	Reported	PMID: 21329993	Yes
HL40	Female/–	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3);	Hom;	Reported;	PMID: 21329993;	Yes
		KCNQ1 (NM_000218): c.2016_2017insGATGAGGGGTCCT (p.Pro672ProfsX6);	Het;	Novel;	–	
		TRIOBP (NM_001039141): c.3943_3944insCTCTTCGG (p.Arg1315ProfsX44);			–	
		TRIOBP (NM_001039141): c.3941A>C (p.Glu1314Ala)	Het;	Novel;	–	
			Het	Novel		
HL41	Female/29	MYO15A (NM_016239): c.5964+3G>A;	Het;	Reported;	PMID: 23767834;	Yes
		MYO15A (NM_016239): c.8681_8682insA (p.Ile2894IlefsX32)	Het	Novel	–	
HL42	Male/28	GJB2 (NM_004004): c.235delC (p.Leu79CysfsX3)	Het	Reported	PMID: 21329993	Yes
HL43	Female/27	SLC26A4 (NM_000441): c.919-2A>G;	Het;	Reported;	PMID: 12676893;	Yes
		SLC26A4 (NM_000441): c.2168A>G (p.His723Arg);	Het;	Reported;	PMID: 23705809;	
		TRIOBP (NM_001039141): c.5185-2A>G;	Het;	Novel;	–	
		PTPRQ (NM_001145026): c.3194delT (p.Ile1065IlefsX6)	Het	Novel	–	
HL44	Female/6	TMC1 (NM_138691): c.2050G>C (p.Asp684His)	Hom	Novel	–	Yes
HL45	Male/12	SLC26A4 (NM_000441): c.919-2A>G	Hom	Reported	PMID: 14508505	Yes
HL46	Female/–	–	–	–		–
HL47	Male/11	–	–	–		–
HL48	Male/6	–	–	–		–
HL49	Female/5	–	–	–		–
HL50	Female/3	–	–	–		–
HL51	Female/30	–	–	–		–
HL52	Male/11	–	–	–		–
HL53	Male/30	–	–	–		–
HL54	Male/18	–	–	–		–
HL55	Male/22	–	–	–		–
HL56	Male/–	–	–	–		–
HL57	Female/19	–	–	–		–
HL58	Male/10	–	–	–		–

products of library construction were then subjected to the circularization process. BGISEQ-500 applied DNA NanoBalls (DNBs) technology for sequencing library construction. The following DNBs making, DNBs loading, and sequencing were conducted according to BGISEQ-500 platform protocol. BGISEQ-500 platform was then used to generate the raw images and the results of base calling.

2.3. Bioinformatics analysis and reporting

In this study, we developed a new bioinformatics analysis pipeline, including data filtration, sequence alignment, variants calling, and annotation.

First, raw sequencing reads (FASTQ files) were evaluated using SOAPnuke to generate "clean reads" for further analysis (removing low-quality bases, adapters, and other technical sequences). After reads filtering, we used BWA to align the 50 bp clean reads to hg19 (the human reference genome). Then, Picard was used to sort the BAM files. To improve read mapping, we then used GATK and Picard to perform local realignment and removal of PCR duplicates. Next, we used GAKT to detect sequence variants (SNVs, Indels and copy number variants [CNVs]) as reported before.[22] Variant annotation was performed using our own PERL scripts. In the annotation step, we downloaded 5 well-known databases to facilitate annotation, including HGMD, dbSNP, 1k Genome, HapMap, dbNSFP, and a local database of our own (596 Chinese healthy adults).

We performed the following criteria to detect candidate variants in HL patients:

1. the number of supporting reads for the variant site are more than 10;
2. the percentage for the variant site is more than 20%;
3. a read is kept only when more than 50% sites of the read are tested with a Phred base quality score of more than 5.

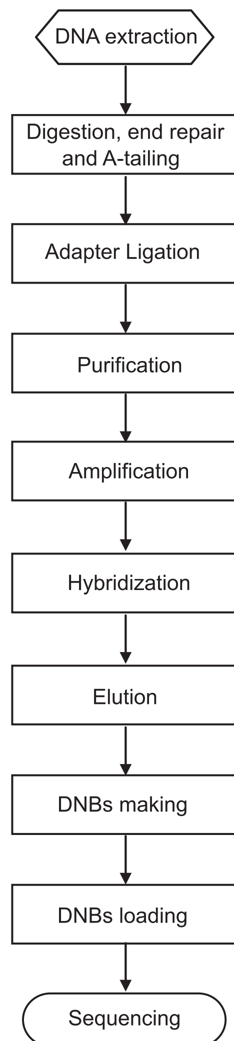


Figure 2. Flow diagram of library construction using BGISEQ-500.

For the nonreported variants, they were recognized as potential variants using the following criteria:

1. the variants were recorded in the following dataset with a frequency of less than 5%: HapMap, NCBI dbSNP, 1K human genome database, and our own polymorphism dataset of 596 Chinese healthy controls;
2. the impact of the potential variants (predicted by dbNSFP (PolyPhen-2,[23] SIFT,[24] MutationTaster,[25] PhyloP[26])) is pathogenic or likely pathogenic.

Finally, validation was carried out using methods other than targeted NGS.

2.4. Variants validation

Clinical disease-causing mutations identified using BGISEQ platform in the patients were confirmed by methods other than targeted NGS (http://links.lww.com/MD/C895). If a pathogenic mutation was detected, segregation analysis was also performed for confirmation when the samples were available. For the validation among the proband family members, we performed Sanger sequencing (ABI 3730) for SNVs and Indels, real-time quantitative PCR (ABI StepOne) for CNVs. Regions of interest were amplified using polymerase chain reaction.

3. Results

3.1. High-throughput NGS

In this study, we developed a targeted NGS assay to detect disease-causing alterations in 58 patients diagnosed with HL using BGISEQ-500 platform.

A total of 58 HL patients and their family members were recruited in this study. We obtained high-quality reads of all the samples by BGISEQ-500 sequencer. After base calling and image analysis using BGISEQ-500 sequencers, primary data was received in the form of FASTQ. Then we used bioinformatics analysis pipeline to analyze the raw sequencing data. The sequencing results of all the patients are as follows: more than 99.12% coverage of the target region was achieved for each sample, and average sequencing depth of 251.37-fold was achieved for each sample. On average, 96.85% of the target region was successfully covered by sequencing data at more than 30-folds, which indicated that this method is reliable in identifying potential disease-causing mutations.

3.2. Diagnostic yield with targeted NGS in HL patients

The development of NGS has accelerated the discovery of genes and mutations that underlie Mendelian diseases. In this study, we used the above mentioned pipeline to detect potential disease-causing mutations. After the filtering process, a total of 62 disease-causing mutations were identified finally, including 31 missense, 17 Indels, 11 splicing, 2 synonymous, and 1 CNV. 58.06% (36/62) of disease-causing mutations has never been reported before (Table 1).

The diagnostic yield is highly related to clinical scientists. In this study, all the suspected HL patients were first evaluated by experienced clinical scientists. Of all the 58 random HL patients, 77.59% (45/58) were found to have deleterious mutations and have molecular confirmation of the clinical diagnosis (Table 1). It was also observed that mutations in gene *SLC26A4* (35.80% of all the disease-causing mutations), *GJB2* (12.35%), and *MYO15A* (11.11%) were the most frequently detected genes in our cohort (Fig. 3).

3.3. Mutation validation

We performed Sanger sequencing (ABI 3730 DNA sequencer) to confirm the reported potential disease-causing SNVs and Indels (Table 1). We performed real-time quantitative PCR (ABI StepOne) to validate all the reported potential CNVs. All the primers used in this study were designed by Primer6.0. Cosegregation analysis was also performed to validate the results when the samples are available. For the 36 novel mutations, 17 of which were successfully validated and cosegregated with the phenotype (HL) in their families.

4. Discussion

In the past few years, large scale genetic screenings of HL have been performed,[12,13] however, most of these studies only focused on several most common variants (such as c.235delC of *GJB2*, c.919-2A>G of *SLC26A4*, and mitochondrial m.1555A>G). Accurate and timely detection of other disease-

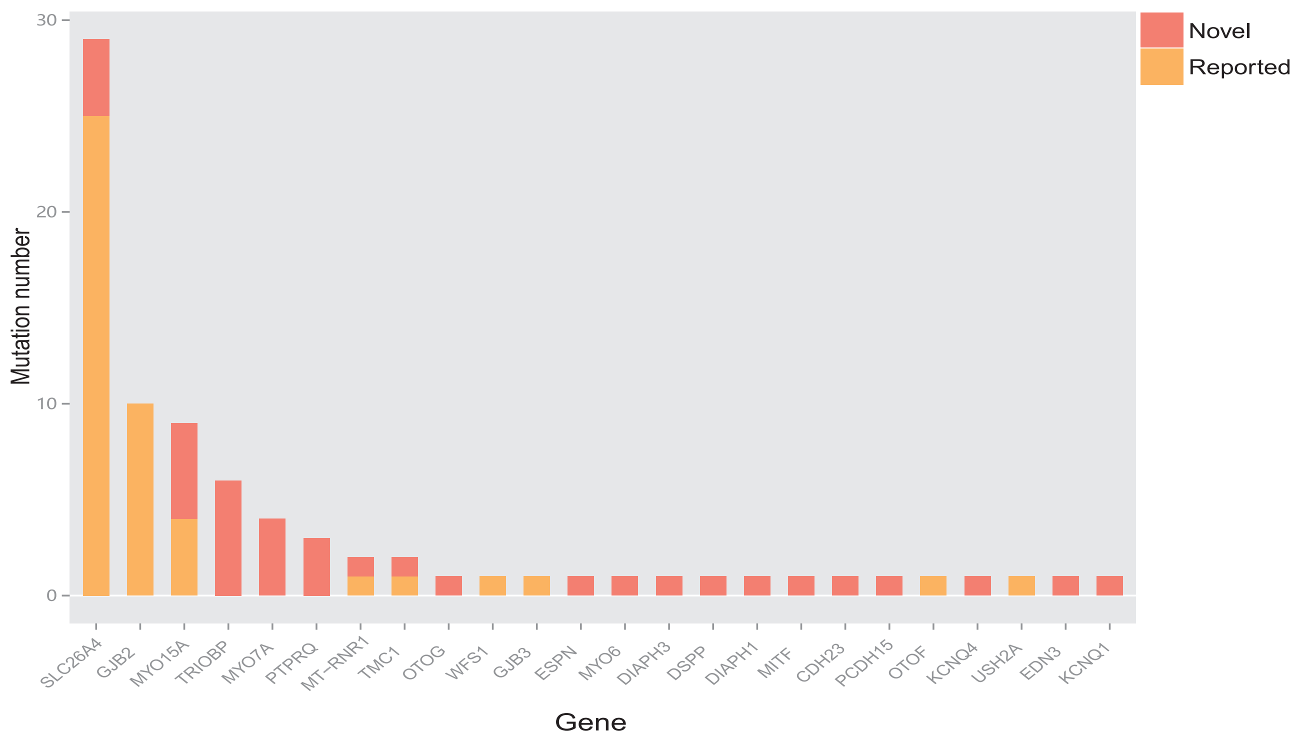


Figure 3. Clinical mutations identified in the 58 clinical cases. Red rectangles indicated the proportion of novel mutations, orange rectangles indicated reported mutations.

causing mutations and related genes play a pivotal role in patients with HL. With the development of NGS, identification of clinical relevant mutations in HL patients has become an option for clinicians. In this study, we established a targeted NGS technology for patients with HL. We performed BGISEQ-500 PE50 sequencing to detect clinically relevant alterations in HL patients in real clinical setting. This newly designed panel targets a total of 127 HL related genes. We applied this panel to 58 clinical HL cases, and identified disease-causing mutations in 45 patients. This new method can greatly improve the diagnostic yield in HL patients and help to make a precise genetic diagnosis. To the best of our knowledge, this is the first report to investigate both syndromic and nonsyndromic HL using BGISEQ-500 platform in the Chinese population.

For the 58 clinical HL cases, most of the disease-causing mutations were found in the *SLC26A4* gene (35.80%), followed by *GJB2* (12.35%), *MYO15A* (11.11%), *TRIOBP* (7.41%), *MYO7A* (4.94%), *PTPRQ* (3.70%), *MT-RNR1* (2.47%), and *TMC1* (2.47%) (Fig. 3). The results showed that the *SLC26A4* (NM_000441) c.919-2A>G mutation is the leading cause in this study. It has been reported that *GJB2* is the leading cause of heredity HL in the Chinese population.[7,8,12,23] Due to the limited number of recruited HL patients, *GBJ2* was found to be the second leading cause in this study (Fig. 3). These results suggested that a minority of HL patients can be defined using a small number of recurrent mutations at predefined positions. NGS technology targeting multiple disease-causing genes is more suitable for the diagnosis of hereditary HL. However, there are still 13 patients with negative results. No disease-causing mutations were identified in the exons of the 127 HL related genes in these patients. The disease-causing mutations of these cases may locate in the intron region, noncoding region, or undiscovered genes. Other factors (multiple genetic and nongenetic causes) may also impact the diagnostic yield.

Interestingly, we detected 2 synonymous mutations in proband HL9 and HL33, which were recognized as the disease-causing mutations in their families. A novel heterozygous variant (c.6024G>A, p.Ser2008Ser) was detected in exon 37 of gene *PTPRQ* in proband HL9. A novel heterozygous variant (c.1905G>A, p.Ser635Ser) was identified in exon 13 of gene *KCNQ4* in proband HL33. These 2 synonymous mutations were not found the 596 healthy controls, and were all validated by Sanger sequencing. Both of these 2 synonymous mutations cosegregated with HL in their families. Although synonymous mutations were predicted to be benign and did not change the sequence of amino acid, we still believe these 2 mutations were the disease-causing mutations of HL. More than 50 diseases afflicting most organ systems have been reported to be associated with synonymous mutations.[27] Several mechanisms have been reported to explain why synonymous mutations are related to the changes in phenotype.[27] In the precursor mRNAs of human genes, exons are separated by noncoding introns. The spliceosome regulates and executes the removal of introns by targeting the correct exon-intron boundary. AG dinucleotides are conserved at the intron ends. The 2 synonymous mutations detected in this study are both resulted in AG sequence, which may provide alternative splicing and affect splicing accuracy of the genes. However, we do know that more experiments need to be performed to illustrate the pathogenic mechanisms.

People with the same variant may represent different phenotypes. In this study, we detected c.538C>T (*GJB3*) and c.1464+2T>A (*ESPN*) in proband HL11. Mutations in gene GJB3 and ESPN both have been described as causative of HL. These variants were also confirmed in the proband father using

Sanger sequencing. However, the proband father shows no syndrome at all. These results showed the high heterogeneity of heredity HL.

In this study, we perform BGISEQ-500 PE50 sequencing in the detection of disease-causing mutations in HL patients in real clinical setting. When we discuss the possibility of using a new method in real clinical setting, cost, turnaround time and throughput should be evaluated first. BGISEQ platform can greatly reduce the cost of sequencing. Turnaround time is the key point for the application of NGS technology in clinical practice. The entire workflow of this method lasts approximately 8 days from the recruitment of sample to clinical reporting. The whole testing workflow includes DNA extraction (0.5 day), library construction (1.5 day), DNBs making and loading (0.5 day), sequencing using BGISEQ-500 (3.5 day), bioinformatics analysis, annotation, and reporting (2 day) for 1 sample.

In summary, we developed a targeted NGS method for profiling of clinically disease-causing mutations in patients with HL. We performed this method in 58 clinical cases and found disease-causing mutations in 45 patients. The results proved that targeted NGS technology using BGISEQ-500 platform could be used for clinical genetic testing in patients with HL.

Acknowledgments

We gratefully thank all the patients.

Author contributions

YS designed the research and wrote the first draft of the article. JY, LMW, WJP, ML, and YY provided patient specimens and collected clinical information. LQD, LWM, and JFM designed and performed the experiments. YS, KK, JY, LMW, WJP, XW, and YY contributed to revising the manuscript. YS, XLC, and CBY developed the bioinformatics pipeline and performed data analysis. All authors reviewed the manuscript.

Data curation: Xiaoli Cui, Chengbin Yan, Lique Du, Liangwei Mao, Jianfen Man.

Formal analysis: Yan Sun.

Funding acquisition: Yan Sun.

Investigation: Yun Yang, Yan Sun, Limin Wu, Min Li, Liangwei Mao, Wei Li.

Methodology: Yan Sun, Wei Li.

Project administration: Yun Yang, Yan Sun, Xuan Wu, Weijun Pan.

Resources: Yan Sun, Jing Yuan, Limin Wu, Min Li, Xiaoli Cui, Chengbin Yan, Lique Du, Jianfen Man, Xuan Wu, Weijun Pan.

Supervision: Yun Yang, Yan Sun, Karsten Kristiansen, Weijun Pan.

Validation: Yan Sun.

Writing – original draft: Yan Sun.

Writing – review and editing: Yun Yang, Yan Sun, Jing Yuan, Karsten Kristiansen, Weijun Pan.

References

[1] Ma D, Zhang J, Luo C, et al. Genetic counseling for patients with nonsyndromic hearing impairment directed by gene analysis. Mol Med Rep 2016;13:1967–74.

[2] Barsky-Firkser L, Sun S. Universal newborn hearing screenings: a three-year experience. Pediatrics 1997;99:E4.

[3] Kennedy CR, McCann DC, Campbell MJ, et al. Language ability after early detection of permanent childhood hearing impairment. New Engl J Med 2006;354:2131–41.

[4] Yoshinaga-Itano C, Sedey AL, Coulter DK, et al. Language of early- and later-identified children with hearing loss. Pediatrics 1998;102:1161–71.

[5] Friedman TB, Griffith AJ. Human nonsyndromic sensorineural deafness. Annu Rev Genomics Hum Genet 2003;4:341–402.

[6] Reiisi S, Sanati MH, Tabatabaiefar MA, et al. The study of SLC26A4 gene causing autosomal recessive hearing loss by linkage analysis in a cohort of Iranian populations. Int J Mol Cell Med 2014;3:176–82.

[7] Dai P, Yu F, Han B, et al. GJB2 mutation spectrum in 2,063 Chinese patients with nonsyndromic hearing impairment. J Transl Med 2009;7:26.

[8] Dai P, Liu X, Han D, et al. Extremely low penetrance of deafness associated with the mitochondrial 12S rRNA mutation in 16 Chinese families: implication for early detection and prevention of deafness. Biochem Biophys Res Commun 2006;340:194–9.

[9] Snoeckx RL, Huygen PL, Feldmann D, et al. GJB2 mutations and degree of hearing loss: a multicenter study. Am J Hum Genet 2005;77:945–57.

[10] Khalifa Alkowari M, Girotto G, Abdulhadi K, et al. GJB2 and GJB6 genes and the A1555G mitochondrial mutation are only minor causes of nonsyndromic hearing loss in the Qatari population. Int J Audiol 2012;51:181–5.

[11] Mahdieh N, Rabbani B, Wiley S, et al. Genetic causes of nonsyndromic hearing loss in Iran in comparison with other populations. J Hum Genet 2010;55:639–48.

[12] Wu CC, Tsai CH, Hung CC, et al. Newborn genetic screening for hearing impairment: a population-based longitudinal study. Genet Med 2017;19:6–12.

[13] Hao Z, Fu D, Ming Y, et al. Large scale newborn deafness genetic screening of 142,417 neonates in Wuhan, China. PloS One 2018;13:e0195740.

[14] Seco CZ, Oonk AM, Dominguez-Ruiz M, et al. Progressive hearing loss and vestibular dysfunction caused by a homozygous nonsense mutation in CLIC5. Eur J Hum Genet 2015;23:189–94.

[15] Schraders M, Ruiz-Palmero L, Kalay E, et al. Mutations of the gene encoding otogelin are a cause of autosomal-recessive nonsyndromic moderate hearing impairment. Am J Hum Genet 2012;91:883–9.

[16] Yariz KO, Duman D, Zazo Seco C, et al. Mutations in OTOGL, encoding the inner ear protein otogelin-like, cause moderate sensorineural hearing loss. Am J Hum Genet 2012;91:872–82.

[17] Xuan J, Yu Y, Qing T, et al. Next-generation sequencing in the clinic: promises and challenges. Cancer Lett 2013;340:284–95.

[18] Shearer AE, DeLuca AP, Hildebrand MS, et al. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. Proc Natl Acad Sci U S A 2010;107:21104–9.

[19] Neveling K, Feenstra I, Gilissen C, et al. A post-hoc comparison of the utility of Sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. Hum Mutat 2013;34:1721–6.

[20] Miyagawa M, Nishio SY, Hattori M, et al. Mutations in the MYO15A gene are a significant cause of nonsyndromic hearing loss: massively parallel DNA sequencing-based analysis. Ann Otol Rhinol Laryngol 2015;124(Suppl 1):158S–68S.

[21] Alford RL, Arnos KS, Fox M, et al. American College of Medical Genetics and Genomics guideline for the clinical evaluation and etiologic diagnosis of hearing loss. Genet Med 2014;16:347–55.

[22] Liu Y, Wei X, Kong X, et al. Targeted next-generation sequencing for clinical diagnosis of 561 Mendelian diseases. PloS One 2015;10:e0133636.

[23] Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;7:

[24] Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res 2001;11:863–74.

[25] Schwarz JM, Rodelsperger C, Schuelke M, et al. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010;7:575–6.

[26] Cooper GM, Stone EA, Asimenos G, et al. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res 2005;15:901–13.

[27] Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nature reviews. Genetics 2011;12:683–91.