Short Report

Whole exome sequencing identifies a novel DFNA9 mutation, C162Y

Gao J, Xue J, Chen L, Ke X, Qi Y, Liu Y. Whole exome sequencing identifies a novel DFNA9 mutation, C162Y [Correction added on 25 October 2012, after first online publication: The title in the citation has been corrected to ‘Whole exome sequencing identifies a novel DFNA9 mutation, C1612Y’.]

We report the genetic analysis of a Chinese family with autosomal dominant non-syndromic progressive sensorineural hearing loss. Taking advantage of next-generation high-throughput DNA sequencing technology, we combined whole exome capture sequencing with Sanger direct sequencing. A novel missense mutation in the coagulation factor C homolog (COCH) gene was identified in a consanguineous Chinese family. This missense mutation in the seventh exon (c.889G>A; p.C162Y) of COCH is most probably a disease-causing mutation and it segregates with the disease. The mutation is not found in the single nucleotide polymorphism (SNP) database, the yhSNP database, the 1000 genomes SNP database or in matching normal controls. It is the first reported autosomal dominant nonsyndromic sensorineural deafness 9 (DFNA9) mutation outside the limulus factor C, cochlin and late gestation lung protein and von Willebrand factor 2 domain, i.e. the first reported DFNA9 mutation in the intervening domain of cochlin, encoded by the COCH gene. In the future, we will focus on functional studies of this mutation.

Conflict of interest
None of the authors have a conflict of interest to declare.

Hearing loss is the most common form of sensory impairment, and hereditary hearing loss (HHL) is an important etiology. HHL can be classified as non-syndromic hearing loss (NSHL) and syndromic hearing loss (SHL) according to the phenotype (1). Screening for mutations in all of these genes with traditional Sanger sequencing is time-consuming and costly because more than 72 genes have been linked to NSHL (http://hereditaryhearingloss.org). In recent years, next-generation sequencing technology, whole exome sequencing, for instance has provided a faster and more cost-effective approach for identifying causative mutations (2–4).

In this study, we identified a missense mutation in the coagulation factor C homolog (COCH) gene (OMIM 603196), a gene causing autosomal dominant nonsyndromic sensorineural deafness 9 (DFNA9) mutation, in a consanguineous family from China. DFNA9 has been clinically characterized by adult-onset, progressive sensorineural hearing loss (SNHL) with vestibular dysfunction. Robertson et al. deduced a very high degree of conservation in the coding region which was then called limulus factor C, cochlin and late gestation lung protein (LCCL), a signal peptide (SP) and two regions of extensive homology to the collagen-binding type A domains of von Willebrand (vWFA1 and vWFA2) (5, 6). To date, 13 missense mutations and 1 deletion mutation have been reported in the COCH gene. Ten are found in the LCCL region while four are located in vWFA2. The mutation we found is a novel missense mutation outside the LCCL and vWFA2 domains.
Materials and methods

Ethical approval for the study was obtained from the health authority ethical committee of Peking University First Hospital. Informed consent was obtained from all the participants and from the parents of minors. We obtained the pedigree of family members after receiving informed consent from all the individuals. Blood samples from all available family members, 49 sporadic patients with hearing loss and all of the 152 normal controls were collected and processed at the Laboratory Center of Peking University First Hospital using the Qiagen blood genomic DNA extraction kit following the protocol provided by the manufacturer.

Whole exome capture and library construction

Human exome capture was performed by accurately following the protocol from Illumina’s TruSeq™ Exome Enrichment Guide (Illumina, San Diego, CA). Illumina’s TruSeq 62 Mb Exome Enrichment kit was used as exome enrichment probe sets; 5 μg of genomic DNA in 80 μl of Elution Buffer (EB) buffer was fragmented in a biorupter to a size of 100–500 bp. DNA concentration was estimated by optical density at 260 nm (OD260) measurement and quantitative real-time polymerase chain reaction. Captured DNA libraries were sequenced with the Illumina HiSeq 2000, yielding 200 (2 × 100) base pairs from the final library fragments using V2 reagent. Base calling was performed with CASAVA 1.8 software (Illumina).

Sanger sequencing

We performed Sanger sequencing on 19 family members, 49 sporadic patients with hearing loss and 152 normal controls. Primers surrounding the suspected variants were designed with the online program PRIMER PREMIER 5.0 (Premier Biosoft). The forward primer is 5′-TGTATGACGATGGCAGTT-3′ and the reverse primer is 5′-ATGTGCTCCTTCTGTCC-3′. Sequencing chemistry was performed using an Applied Biosystems PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (v3.1) (Life Technology, Carlsbad, CA), the PCR amplicon was sequenced on an ABI 3730, and the results analyzed using sequence scanner v1.0.

Bioinformatic analysis

The reads were aligned with the human genome reference sequence [University of California Santa Cruz, human genome assembly 19 (UCSC hg19)] using the Burrows–Wheeler alignment (BWA) 0.5.9rc1 [Correction added on 25 October 2012, after first online publication: The text was corrected to: ‘reads’ were aligned with the human genome reference sequence…]. Variants [single nucleotide polymorphisms (SNPs) and indels] were called with vcftools of SAMTools software version 0.1.16 (7). High VarQuality SNPs were annotated with Perlscript into functional categories such as missense, nonsense, splice sites, coding, non-coding, UTRs. Amino acid substitution (whether an amino acid substitution affects protein function) was annotated with SIFT (Sorting Intolerant From Tolerant).

Results

Clinical features

Family32 is a consanguineous Chinese family. The pedigree of Family32 spans six generations (Fig. 1) and shows autosomal dominant inheritance with no evidence of non-penetrance in the entire pedigree.

![Pedigree of part of Family32](image)
Whole exome sequencing identifies a novel *DFNA9* mutation, C162Y

(a) Scatter diagram of the correlation between age (years) and hearing threshold (dB). We analyzed the data using spss 16.0 on 26 patients (44 ears), grouped by frequencies (kHz). Different frequencies are shown by different colors in the figure. (b) Kaplan–Meier curve to determine the age of onset of hearing loss in Family32. We calculated the data from all family members who were younger than 30 years of age, except for those who were genetically considered as normal. Members older than 30 years were all affected apart from five (individuals 406, 418, 420, 433 and 434). They were not considered in the calculations because we could not be sure of the exact onset time of the disease. Family members 406, 418 and 434 were genetically diagnosed as unaffected and we could not obtain the precise age of the other two, family members 420 and 433, so they were not considered in the process either.

Twenty-eight family members ranging in age from 12 to 67 years were diagnosed as having SNHL by pure tone audiometry (PTA). We performed a statistical analysis on 26 patients (44 ears) between age (years) and hearing level (dB) grouped by frequency (kHz) (Fig. 2a) using spss 16.0. Hearing impairment indicated profound hearing loss at 4 and 8kHz at an early age, and then progressed rapidly to other frequencies at a later age (Fig. 2a). We can also note that hearing loss progressed more slowly at 2kHz than at any other frequencies, especially before the age of 45 (Fig. 2a). In addition, two family members showed sudden sensorineural hearing loss by PTA. We also performed calculations using spss 16.0, shown as a Kaplan–Meier curve (Fig. 2b). We can clearly see that the mean age of onset of hearing loss is around 17 years. In addition to hearing impairment, some affected individuals in their 3rd–4th decade of life developed a precipitous drop in word recognition, and showed a decrease in communication ability. None of the affected individuals complained of any vestibular symptoms. No related clinical tests were carried out because we could not obtain the family members’ approval to perform vestibular evoked myogenic potentials or caloric testing.

**Mutation detection**

We applied whole exome capture sequencing to one affected individual (Family32-414). The reads were aligned with the human genome reference sequence (UCSC hg19) using the BWA tool. Variants (SNPs and indels) were called with vcf tools of SAMTools software (7). Variants found in the SNP database (dbSNP database), the yhSNP database (http://yh.genomics.org.cn/) or the 1000 genomes SNP database were excluded because the disease-causing variants should be rare. Then variants that were located outside the coding sequence (CDS) site were excluded. Finally, only variants that lead to non-synonymous mutations were subjected to downstream analysis. Thus, 340 candidate variants were identified in the entire exome.

We gave preference to those variants that were located in the genes causing the DFNA because Family32 showed autosomal dominant inheritance (http://hereditaryhearingloss.org). Using this method, inspection of the candidate variants revealed only one non-synonymous missense mutation that is related to DFNA. The non-synonymous missense mutation (G>A) is located in the seventh exon of *COCH* (Fig. 3b), a gene known to cause the *DFNA9* mutation. Nineteen biological family members from Family32 were genotyped for this mutation using Sanger sequencing. As expected, this allele segregated precisely with the disease (Fig. 1). Seven affected family members were shown to be heterozygous (Fig. 3a) while three unaffected family members were not. Because seven individuals who were below the age of 17 years had not reached the onset age of hearing loss, they were not considered in the cosegregation analysis. They could still support our conclusion to some degree because they showed no symptoms of hearing loss or abnormal PTA tests and their Sanger sequencing results were homozygous.

We then performed a linkage analysis using a dominant parametric model and the maximum LOD score (logarithm of odds) was 1.48. This is small given the size of the pedigree and the limited number of family members on whom genetic analysis was performed (Fig. 1). The LOD score can support linkage with odds of 1:30. Sanger sequencing was then carried out on 49 sporadic patients with hearing loss and 152 normal matching controls. It turned out that this mutation...
was not found in the population, which indicates that this allele is rare. Furthermore, this mutation was not observed in the dbSNP, yhSNP and 1000 genome SNPs and was not found in the Exome Variant Server (http://evs.gs.washington.edu/EVS/) database either, showing that the mutation is very rare. This G>A mutation changes an amino acid (cysteine to tyrosine) that is conserved in mouse, chimp, rabbit, sheep, cow, horse, dog, microbat, elephant, chicken, lizard and panda. To predict the pathogenicity of this amino acid substitution, the variant (c.889G>A; p.C162Y) was run through the PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) program and this mutation was predicted to be probably damaging with a score of 0.998 (sensitivity: 0.27 and specificity: 0.99). Normal and mutated amino acid sequences were both submitted to SWISS-MODEL to predict the protein structure of cochlin and we found that the mutated protein lacked the a-helix in Gin313.

At the same time, the sequencing result of Family32-511 was heterozygous, while he showed no evidence of hearing loss, according to our clinical tests. Family32-511 was only 13 years old when receiving the clinical tests.

Discussion

In this study, we showed that the combination of whole exome sequencing and Sanger sequencing provides a sensitive, cost-efficient, and time-saving tool for genetic analysis. We found a novel COCH mutation, C162Y, which is most probably disease-causing, in a Chinese pedigree with hereditary sensorineural hearing loss. This mutation is not a common variant and is considered to be most probably disease-causing because: (1) the mutation was found in the COCH gene, a gene related to DFNA9, while according to the results of whole exome sequencing of individual Family32-414, no other variation was found in the genes causing the DFNA mutation; (2) it is located in the CDS, which is likely to result in changes to the amino acid of the translated protein; (3) it was found in all of the eight tested patients of Family32, whereas it was never found in the tested healthy biological family members; (4) the mutation was not found in any of the 49 sporadic patients with hearing loss and the 152 control individuals; (5) the mutated Cys residue is conserved in homologous genes from mouse to human. Further studies will be performed on the function mechanism of C162Y.

Cochlin is the encoded protein of the COCH gene. There are multiple domains in cochlin, including a SP, an LCCL domain, a first intervening domain (Ivd1), vWFA1 domain, and a second intervening domain (Ivd2), vWFA2 domain (6). Cochlin is reported to be one of the most abundant proteins in the inner ear (8), it is a secreted protein and plays an important role in the extracellular matrix and basilar membrane (9). The mutation we identified, c.889G>A, results in a predicted p.C162Y substitution at a conserved cysteine residue of cochlin. The cysteine residues of the cochlin surface can form disulfide bonds, and mutations can alter the intramolecular cochlin disulfide bond formation (10) and lead to mispairing of cysteines during oxidative formation of disulfide bridges (11). On the basis of these observations, it is reasonable to make the hypothesis that the mutation (c.889G>A; p.C162Y) we found in this family, which results in a cysteine to tyrosine substitution, is related to the alteration of disulfide bridges. Moreover, this C162Y variant is considered to be pathogenic using the PolyPhen-2 program (with a score of 0.998), which shows that the mutation is most probably disease-causing.

The mutation we found (c.889G>A; p.C162Y) is located within the Ivd1, and it is the first reported DFNA9 mutation outside the LCCL and vWFA2 domains. Thus, it is interesting to compare the clinical characteristics of patients with the mutation in those domains and our families. Hearing loss in most previously reported DFNA9 families first affected the high frequencies and later involved all frequencies, and the mean age at onset of hearing loss is mostly in adulthood. As for Family32, the mean age at onset of hearing loss is 17 years, and penetrance of hearing loss is complete. Hearing impairment progressed rapidly from high frequency loss to lower frequencies at a later stage while retaining a good hearing level at 2 kHz before age 45; sudden sensorineural hearing loss occurred in some patients in this family. No member of Family32 complained of any vestibular abnormalities.

Among all of the individuals receiving Sanger sequencing, the result for individual Family32-511 was heterozygous (c.889G>A; p.C162Y), while he showed no evidence of hearing loss according to our clinical tests. We have reasons to regard this individual as a potential patient according to the gene diagnosis and disease features of the patients because he was only 13 years old when receiving the clinical tests and the age at onset of hearing impairment in this family is around 17 years. It also gives us the expectation of
implementing prenatal diagnosis for this family in the near future, which could lower the birth rate of potential patients in this family.

Acknowledgements
We wish to thank all the individuals in Family32 who have kindly agreed to participate in this study. The work is supported by the National Natural Science Foundation of China (Project 30872864).

References