SHORT COMMUNICATION

Exome sequencing identifies *POU4F3* as the causative gene for a large Chinese family with non-syndromic hearing loss

Xin Zhang Cai[1,2,5], Ying Li[2,5], Lu Xia[2], Yu Peng[2], Chu Feng He[1], Lu Jiang[1], Yong Feng[1,2], Kun Xia[2,3], Xue Zhong Liu[4], Ling Yun Mei[1] and Zheng Mao Hu[2,3]

Hearing impairment, or deafness (in its most severe form), is one of the most common human sensory disorders. There have been several reports of autosomal dominant mutations in the *POU4F3* gene, which is associated with non-syndromic hearing loss. In this study, we identified a novel heterozygous mutation (c.602delT, p.L201fs) in the gene *POU4F3* by taking advantage of whole-exome sequencing, which was validated by Sanger sequencing and completely co-segregated within a large hearing impaired Chinese family. We have focused on this pedigree since 2002, and we have mapped a deafness locus named DFNA42 (which has been renamed DFNA52, OMIM entry 607683) via a genome-wide scan. Furthermore, we analyzed this mutational variant and found that it was located at the beginning of the first functional domain of *POU4F3*, which could theoretically impair the function of POU4F3. We have identified a novel frameshift mutation in the *POU4F3* gene. Further functional studies of variants of this specific gene are needed to illustrate the pathogenic mechanism(s) that underlie hearing impairment.
Journal of Human Genetics (2017) **62**, 317–320; doi:10.1038/jhg.2016.102; published online 18 August 2016

INTRODUCTION

Hearing impairment, or deafness (in its most severe form), is one of the most common human sensory disorders. In the USA, ~ 1 in 1000 children is born with a congenital or prelingual profound hearing loss, and another 1 in 1000 children develops deafness before maturity.[1,2] According to its clinical characteristics, hearing loss can be categorized as syndromic hearing loss or non-syndromic hearing loss (NSHL).[3] NSHL is a common sensory deafness and can be divided into the following two types: prelingual and postlingual deafness. In the prelingual NSHL category, ~ 75–80% are autosomal recessive, 20–25% are dominant and 1–1.5% are X-linked. Furthermore, inheritance patterns are autosomal dominant in nearly all of the hereditary postlingual NSHL families.[3,4] Currently, 57 loci (DFNA1-57) for autosomal dominant, non-syndromic, sensorineural hearing loss have been mapped to chromosomal regions; however, only 32 causative genes coding for a wide variety of proteins have been cloned since 1997 (http://hereditaryhearingloss.org/). This is because the identification of the predisposing gene for many hearing loss cases/ families is fundamentally restricted by a small pedigree size, locus

heterogeneity, limited power of traditional positional cloning strategies and so on.[5]

We have focused on the hearing impaired Chinese pedigree since 2002, and we have mapped a deafness locus named DFNA42 (which has been named DFNA52, OMIM entry 607683) via a genome-wide scan, in addition to previous screenings of *POU4F3* by traditional methods. Unfortunately, we did not identify any disease-causing mutations at that time.[3]

POU4F3 is a transcription factor, which contains two functional domains, a POU-specific (POUS) domain and a POU-homeo domain (POUHD), and POU4F3 is necessary for the maturation and survival of hair cells in the inner ear.[6] There have been several reports about autosomal dominant mutations in the *POU4F3* gene that are associated with NSHL.[7–10] The *POU4F3* gene encodes for a protein, which exclusively distributes in the nucleus; however, the mutant protein is localized both to the nucleus and the cytoplasm. Furthermore, the mutant protein's half-life is longer than that of the wild type.[11–13]

[1]Department of Otolaryngology-Head and Neck Surgery, Xiangya Hospital, Central South University, Changsha, China; [2]State Key Laboratory of Medical Genetics, Central South University, Changsha, China; [3]School of Biological Science and Technology, Central South University, Changsha, China and [4]Department of Otolaryngology-Head and Neck Surgery, Leonard M. Miller School of Medicine, University of Miami, Miami, FL, USA
[5]These two authors contributed equally to this work.
Correspondence: Professor XZ Liu, Department of Otolaryngology-Head and Neck Surgery, Leonard M Miller School of Medicine, University of Miami, Miami, FL 33136, USA.
E-mail: xliu@med.miami.edu
or Professor LY Mei, Department of Otolaryngology-Head and Neck Surgery, Xiangya Hospital, Central South University, Changsha, China.
E-mail: entmly@163.com
or Professor ZM Hu, State Key Laboratory of Medical Genetics, Central South University, 110 Xiangya Road, Changsha 410078, China.
E-mail: huzhengmao@sklmg.edu.cn
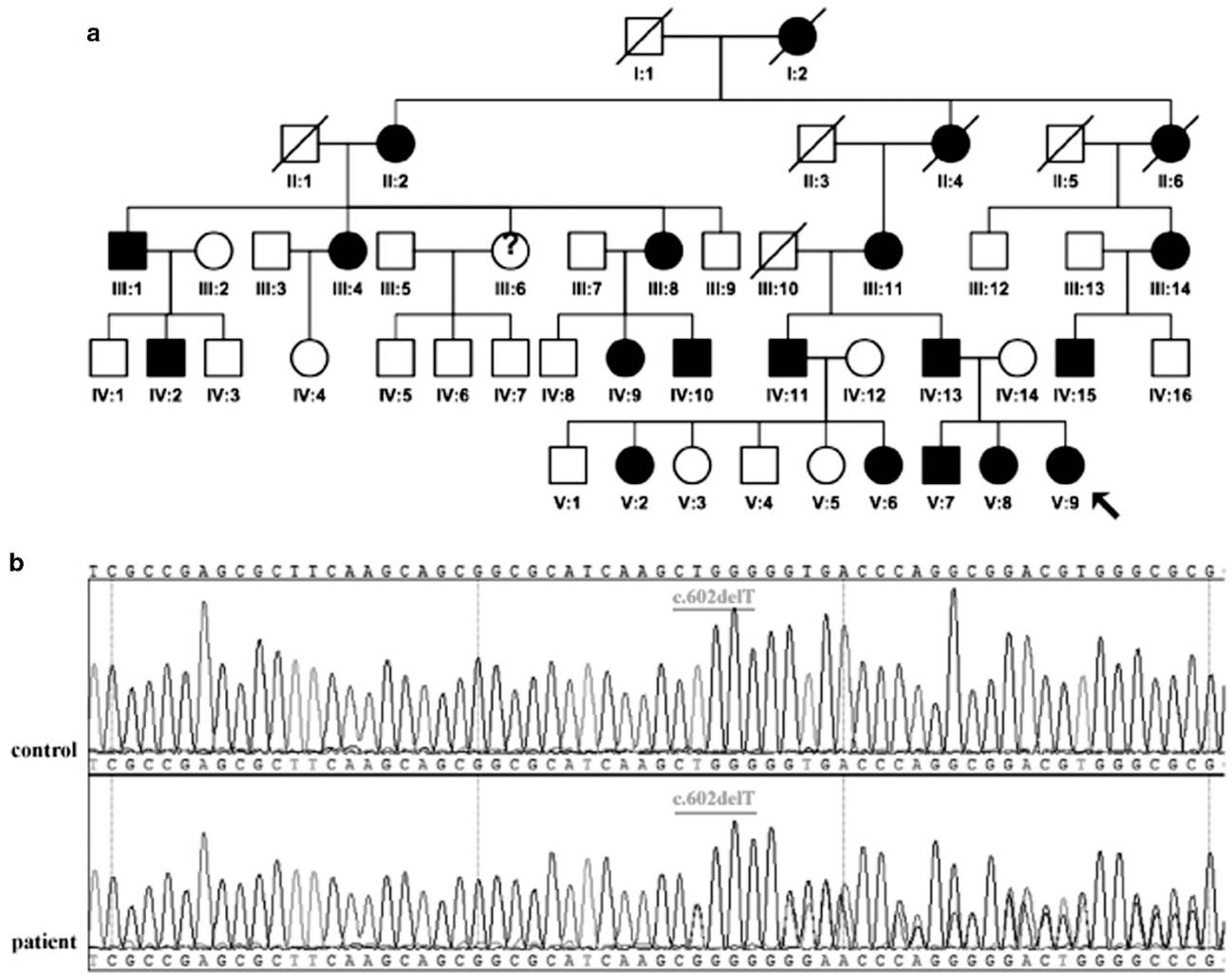Received 2 February 2016; revised 1 June 2016; accepted 30 June 2016; published online 18 August 2016

Figure 1 Identification of *POU4F3* mutation in a hearing loss Chinese family. (**a**) The autosomal dominant, non-syndromic hearing impairment pedigree. Roman numerals refer to generations, and individuals within a generation are numbered from left to right, as is the standard convention. Proband was noted with an arrow. Filled symbol refers to the patient. Open symbols with a question mark refer to the individuals whose data are unknown. Open symbols refer to unaffected individuals. (**b**) The DNA sequence of the mutation p.L201fs in *POU4F3* in the Chinese hearing loss pedigree. The first line represents the control sequence, and the second line represents the patient sequence. The mutation information is indicated by red lettering. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

Table 1 Direct identification of the causal genes by exome sequencing

	IV:13 (whole/locus)	V:6 (whole/locus)	V:3 (whole/locus)	Two affected IV:13+V:6 (whole/locus)	IV:13+V:6-V:3 (whole/locus)
Non-synonymous SNV/Indel	10 431/111	10 545/100	10 692/100	5965/76	1476/33
Not in dbSNP137/1000 Genomes Project/ ESP-6500/ our 500 in-house control samples	326/6	360/5	419/5	118/4	47/4
Predicted to be damaging	206/5	222/4	232/4	78/3	32/3

Abbreviations: dbSNP, database single-nucleotide polymorphisms; ESP-6500, NHLBI GO Exome Sequencing Project; Indel, short insertions or deletions; SNV, single-nucleotide variants. Rows show the effect of excluding from consideration variants found in dbSNP137, 1000 Genomes Project, ESP-6500 and our 500 in-house control samples with exomeSeq data. Columns show the effect of requiring that non-synonymous SNV/Indel variants be observed in each affected or unaffected individual (columns 2–4) or two affected individuals (column 5). Column 6 provides non-synonymous SNV/Indel variants found in affected individuals but not in unaffected individual. (whole/locus): indicate non-synonymous SNV/Indel variants be observed in whole-exome or in locus region by linkage analysis that has been reported before.

MATERIAL AND METHODS

Subject recruitment and clinical examination

A family with five generations of hearing impairment, which came from the Hunan Province of China, was identified by the Department of Otorhinolaryngology at the Xiangya Hospital of Central South University, Changsha, China. In this family, there were 17 affected and 21 unaffected individuals (Figure 1a). The onset age of hearing loss in this family ranged from 16 to 30 years. The characteristics of the

hearing loss in this pedigree were consistent with bilateral, postlingual, progressive, sensorineural and non-syndromic hearing impairment. Clinical and audio logical examinations were performed on all of the family members, and the clinical characteristics of this family have been described previously.[3] Written informed consent was obtained from all of the study participants. The research study was approved by the Human Ethics Committee of the State Key Laboratory of Medical Genetics, Central South University.

Whole-exome sequencing and bioinformatics analysis

Genomic DNA was extracted from peripheral blood lymphocytes obtained from all of the recruited family members using a standard phenol–chloroform method. For each individual selected for whole-exome sequencing, 1 μg of genomic DNA was fragmented, selected and amplified with a PCR. We used the Agilent SureSelect Human All Exon V4+UTRs enrichment platform (Agilent Technologies, Santa Clara, CA, USA) to collect the protein-coding regions of the human genomic DNA, according to the manufacturer's instructions. An exon-enriched library was sequenced on an Illumina Hiseq 2000 platform following the manufacturer's instructions (Illumina, San Diego, CA, USA).

Generated reads were aligned to the human reference genome (University of California Santa Cruz (UCSC, http://genome.ucsc.edu) hg19 and National Center for Biotechnology Information (NCBI) b37.1). Sequencing quality was evaluated based on the Illumina GERALD report and the fast Quality Control. Reads that passed were then aligned to the human reference genome (UCSC hg19) using the Burrows-Wheeler Aligner (University of California, Santa Cruz, CA, USA).[14]

We used Samtools and/or the Genome Analysis Toolkit (The Broad Institute of Harvard and MIT, Cambridge, MA, USA) to identify single-nucleotide polymorphisms (SNPs) and insertions and deletions.[15] The SNPs with a read depth $\geqslant 8 \times$ and a quality $\geqslant 30 \times$ were reserved for subsequent analysis. If SNPs and insertions and deletions affected coding sequence, SeattleSeq annotation was annotated to predict the effect of the variants on protein function. Considering the basis of the SNP database (dbSNP), 1000 genome annotations and the supplied PolyPhen prediction, every non-synonymous variant not assigned a 'benign' prediction was considered to be damaging.

Sanger sequencing of the implicated gene

Taking advantage of the PCR amplification and Sanger sequencing, we validated the presence of variants in the candidate genes identified via exome sequencing. Primers were designed using the online tool Primer3. PCR amplification was performed using a GeneAmp PCR System 2720 (Applied Biosystems, Foster City, CA, USA) in a 10 μl mixture containing 5 μl of Premix Ex Taq (Takara Bio, Dalian, China), 30 ng of primers and 30 ng of genomic DNA. PCR amplification was performed as follows: an initial denaturing step at 95 °C for 5 min, followed by 32 cycles of denaturating at 95 °C for 30 s, annealing at 55–60 °C for 30 s, extension at 72 °C for 30 s and then finally extension at 72 °C for 10 min. PCR products were verified by 6% polyacrylamide gel electrophoresis and silver staining.

RESULTS

Whole-exome sequencing identified three candidate genes

We performed whole-exome sequencing on two affected individuals (IV: 13 and V: 6) and one unaffected individual (V: 3) from the hearing impaired family (Figure 1a). An average of 9.9 gigabases (Gb) was generated per individual, and ~ 99% passed the quality assessment and were aligned to the human reference sequence. Approximately 2.4 Gb were mapped to the targeted exome (NCBI consensus coding sequence) with a mean depth of ~ 60-fold. By comparing our non-synonymous single-nucleotide variants and insertions and deletions variants against dbSNP137, we first excluded high-frequency (minor allele frequency > 0.01) polymorphisms present in the 1000 Genomes Project and NHLBI GO Exome Sequencing Project (ESP-6500) and then excluded variants in our 500 in-house control samples using exomeSeq data. Second, we extracted the segregated variants using whole-exome sequencing and removed the shared SNPs in the three individuals. We were able to subsequently select three candidate variants in three genes (PCDHGB3, SH3RF2 and POU4F3) that were carried by both of the patients for further analysis (Table 1). One is a frameshift mutation (c.602delT, p.L201fs) in the gene POU4F3 (NM_002700.2), another is a missense mutation (c.655A > G, p.I219V) in the gene SH3RF2 (NM_152550.3) and the last one is a missense mutation (c.1739G > A, p.R580H) in the gene PCDHGB3 (NM_018924.4).

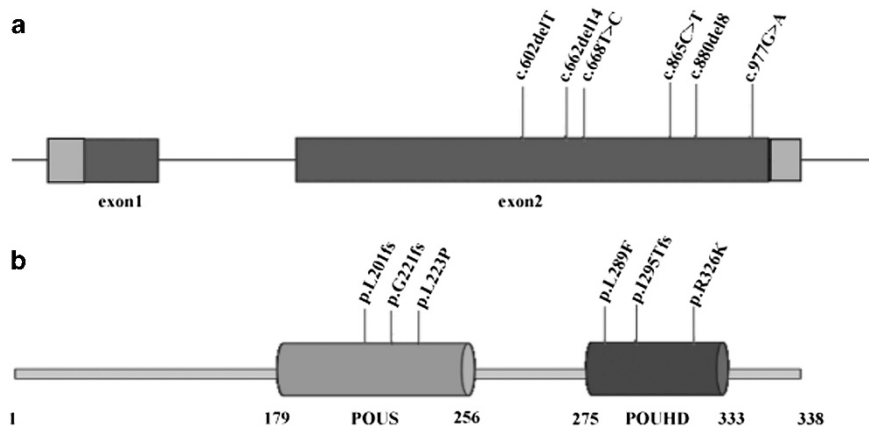

Figure 2 POU4F3 mutations. (a) Schematic of the POU4F3 gene with the position of the mutations shown based on the human genomic sequence. (b) POU4F3 protein with the positions of alterations. POU4F3 is a multi-domain protein carrying a POU-specific (POUS) domain and a POU-homeo domain (POUHD). The p.L201fs, p.G221fs and p.L223P alterations were clustered in the POUS domain. The p.L289F, p.I295Tfs and p.R326K alterations were clustered in the POUHD domain. A full color version of this figure is available at the Journal of Human Genetics journal online.

320

Co-segregating of the three candidate genes

We used Sanger sequencing to validate the variations of *PCDHGB3*, *SH3RF2*, and *POU4F3* and analyze the co-segregating status of the variant in these family members. After screening the available DNA of the pedigree, we found that all of the patients carried heterozygous p.L201fs mutations in *POU4F3* (Figure 1b). Furthermore, the mutation was not detected in other unaffected members of this family and 500 unrelated controls from the same ethnic background. Unfortunately, neither *SH3RF2* nor *PCDHGB3* was co-segregated within this Chinese family. These results suggested that this novel mutation in *POU4F3* is not a rare polymorphism, but rather is a causative mutation for autosomal dominant, non-syndromic hearing impairment in this Chinese family.

Results of clustalx1.83 software

Taking advantage of the clustalx1.83 software (Conway Institute UCD, Dublin, Germany) , we analyzed variant c.602delT and found that it was located at the beginning of the POU$_S$ domain of POU4F3 (Figure 2b). We found that a termination codon occurred early because of the mutation, resulting in a truncated protein. Compared with previous studies, the mutation found in this study affects both the POU$_S$ and POU$_{HD}$ domains.

DISCUSSION

We demonstrated that whole-exome sequencing of just a few samples from one family can be applied to identify the causative gene. Furthermore, identification of disease-causing genes may be achieved using the combined strategy of exome sequencing and traditional analysis with even fewer samples required. This approach, together with continuing improvement of sequencing technology, will make medical genomics of hearing impairment a reality.

To date, approximately five mutations in *POU4F3* (c.977G>A, c.880del8, c.865C>T, c.668T>C and c.662del14) have been found (Figure 2a). We have made a brief genotype–phenotype analysis according to limited clinical characteristics that included four reported families. We classified them into two groups in terms of the location of their variants. Accordingly, the Chinese family (c.602delT) and the Korean family (c.662del14) belonged to the first group where the mutations were located in the first domain of *POU4F3*, whereas the Dutch family (c.865C>T) and family H (c.880del8) belonged to the second group where the mutations were located in the second domain of *POU4F3*.[16] Interestingly, all of the four families were characterized with a large variability in terms of onset age and progressive autosomal dominant non-syndromic hearing loss.[11,12,17] However, the pedigrees in the second group displayed abnormal function of the vestibular apparatus, and only half of the patients had tinnitus, whereas families in the first group showed no malformations of the inner ear by computed tomography examination or abnormal function by vestibular testing. However, all of the patients in the Chinese family had tinnitus.[3] Moreover, examination of the history and clinical records of the affected members in our family suggested that high-frequency loss in this pedigree begins in the second or third decade and progresses to profound deafness involving all frequencies. This feature of the other three families was unknown. We know that the variants of the first group are located in the first domain, which would impair both of the domains in *POU4F3*, whereas the mutations in the second group only impaired the second domain. Therefore, we inferred that the first domain could be related to tinnitus, and the second domain may be sufficient for the occurrence of hearing loss.

Therefore, further functional studies of variants of this specific gene are needed to illustrate the pathogenic mechanism(s) that underlie hearing impairment.

1 Morton, N. E. Genetic epidemiology of hearing impairment. *Ann. NY Acad. Sci.* **630**, 16–31 (1991).
2 Kim, S. Y., Kim, A. R., Kim, N. K., Kim, M. Y., Jeon, E. H., Kim, B. J. *et al.* Strong founder effect of p.P240L in CDH23 in Koreans and its significant contribution to severe-to-profound nonsyndromic hearing loss in a Korean pediatric population. *J. Transl. Med.* **13**, 263 (2015).
3 Xia, J., Deng, H., Feng, Y., Zhang, H., Pan, Q., Dai, H. *et al.* A novel locus for autosomal dominant nonsyndromic hearing loss identified at 5q31.1-32 in a Chinese pedigree. *J. Hum. Genet.* **47**, 635–640 (2002).
4 Venkatesh, M. D., Moorchung, N. & Puri, B. Genetics of non syndromic hearing loss. *Med. J. Armed Forces India* **71**, 363–368 (2015).
5 Ku, C. S., Naidoo, N. & Pawitan, Y. Revisiting Mendelian disorders through exome sequencing. *Hum. Genet.* **129**, 351–370 (2011).
6 Fukui, H., Wong, H. T., Beyer, L. A., Case, B. G., Swiderski, D. L., Di Polo, A. *et al.* BDNF gene therapy induces auditory nerve survival and fiber sprouting in deaf Pou4f3 mutant mice. *Sci. Rep.* **2**, 838 (2012).
7 Baek, J. I., Oh, S. K., Kim, D. B., Choi, S. Y., Kim, U. K., Lee, K. Y. *et al.* Targeted massive parallel sequencing: the effective detection of novel causative mutations associated with hearing loss in small families. *Orphanet J. Rare. Dis.* **7**, 60 (2012).
8 Frydman, M., Vreugde, S., Nageris, B. I., Weiss, S., Vahava, O. & Avraham, K. B. Clinical characterization of genetic hearing loss caused by a mutation in the POU4F3 transcription factor. *Arch. Otolaryngol. Head Neck Surg.* **126**, 633–637 (2000).
9 Nolan, L. S., Jagutpal, S. S., Cadge, B. A., Woo, P. & Dawson, S. J. Identification and functional analysis of common sequence variants in the DFNA15 gene, Brn-3c. *Gene* **400**, 89–97 (2007).
10 Xiang, M., Gan, L., Li, D., Chen, Z. Y., Zhou, L., O'Malley, B. W. Jr. *et al.* Essential role of POU-domain factor Brn-3c in auditory and vestibular hair cell development. *Proc. Natl Acad. Sci. USA* **94**, 9445–9450 (1997).
11 Collin, R. W., Chellappa, R., Pauw, R. J., Vriend, G., Oostrik, J., van Drunen, W. *et al.* Missense mutations in POU4F3 cause autosomal dominant hearing impairment DFNA15 and affect subcellular localization and DNA binding. *Hum. Mutat.* **29**, 545–554 (2008).
12 Lee, H. K., Park, H. J., Lee, K. Y., Park, R. & Kim, U. K. A novel frameshift mutation of POU4F3 gene associated with autosomal dominant non-syndromic hearing loss. *Biochem. Biophys. Res. Commun.* **396**, 626–630 (2010).
13 Weiss, S., Gottfried, I., Mayrose, I., Khare, S. L., Xiang, M., Dawson, S. J. *et al.* The DFNA15 deafness mutation affects POU4F3 protein stability, localization, and transcriptional activity. *Mol. Cell. Biol.* **23**, 7957–7964 (2003).
14 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
15 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
16 Vahava, O., Morell, R., Lynch, E. D., Weiss, S., Kagan, M. E., Ahituv, N. *et al.* Mutation in transcription factor POU4F3 associated with inherited progressive hearing loss in humans. *Science* **279**, 1950–1954 (1998).
17 van Drunen, F. J., Pauw, R. J., Collin, R. W., Kremer, H., Huygen, P. L. & Cremers, C. W. Vestibular impairment in a Dutch DFNA15 family with an L289F mutation in POU4F3. *Audiol. Neurootol.* **14**, 303–307 (2009).