dCollection

전사인자 유전자 돌연변이에 의한 유전성 난청 임상 표현형 및 기능 분석

Clinical and Functional Characterization of Genetic Hearing Loss Caused by Variants in Transcription Factors

2023년 8월

서울대학교 대학원

의학과 이비인후과학 전공

이 상 연

전사인자 유전자 돌연변이에 의한 유전성 난청 임상 표현형 및 기능 분석

지도 교수 이 준 호

이 논문을 의학박사 학위논문으로 제출함
2023년 4월

서울대학교 대학원
의학과 이비인후과학 전공
이 상 연

이상연의 의학박사 학위논문을 인준함
2023년 7월

위 원 장   오승하   (인)

부위원장   이준호   (인)

위   원   성문우   (인)

위   원   박무균   (인)

위   원   이규엽   (인)

Abstract

Clinical and Functional Characterization of Genetic Hearing Loss Caused by Variants in Transcription Factors

<div align="right">

Sang-Yeon Lee

Medicine

Department of Otorhinolaryngology

The Graduate School

Seoul National University

</div>

Sensorineural hearing loss is one of the most common hereditary sensory disorders. With recent developments in genomics, over 150 deafness genes have been identified and functional classifications of genetic hearing loss, based on the molecular mechanisms and the spatiotemporal expression in the inner ear, are currently being developed. The functional assignments of genetic hearing loss have elucidated the natural course of residual hearing, revealed genotype-phenotype correlations, and facilitated the development of target drug and gene therapy. Transcription factors (TFs) recognize specific DNA sequences to control transcription by

forming a complex that guides genome expression. More than 1,600 human TFs have been documented, and their TF variants have been implicated in diverse human diseases. Transcriptional regulation, one of the functional classifications of genetic hearing loss, also serves a critical role in developing and maintaining hearing function. However, only a handful of TF variants known to cause hearing loss are currently understood, with their clinical phenotypes, genotypes, and molecular mechanisms in the context of hearing loss remaining poorly defined. Here, this study aimed to elucidate the clinical phenotypes and genotypes of non－syndromic hearing loss resulting from pathogenic variants in TF genes. The DNA of 1280 probands was subjected to molecular genetic testing, and 720 probands with disease－causing variants were identified. Ultimately, 33 probands (2.6%) had non－syndromic hearing loss caused by pathogenic variants in TF genes. The genetic landscape was exclusively clustered in only four TF genes (*POU3F4*, *POU4F3*, *LMX1A*, and *EYA4*), indicating a narrow molecular etiologic spectrum. Specifically, the phenotype－genotype analysis of the four TF genes showed that *LMX1A*－associated deafness is characterized by asymmetric hearing loss. This study also presented diverse functional aspects of novel *POU4F3* variants and identified 14 downstream target genes associated with inner ear development

using patient-derived lymphoblastoid cell lines. The transcriptome profiles of patient-derived lymphoblastoid cells showed a significant correlation with cochlear hair cells, providing a breakthrough for cases where human cochlear sample collection was unfeasible. Collectively, the results of this study present the phenotype-genotype map of TF variants underlying non-syndromic hearing loss, including the asymmetric hearing loss phenotype underlying *LMX1A*-associated DFNA7, and refine previously proposed molecular mechanisms underlying *POU4F3*-associated DFNA15.

Keywords: Genetic hearing loss, Transcriptional factor, *LMX1A*, *POU4F3*

Student Number: 2019-34627

Table of Contents

List of Tables

List of Figure

Chapter 1. Introduction

1.1. Study Background

Sensorineural hearing loss (SNHL) is the most common sensory disorder in humans. Genetic analysis enhances the understanding of the pathogenic mechanisms of hearing loss, of which over 50% is congenital or prelingual deafness, [1,2] and a significant proportion postlingual deafness.[3,4] Given the developments in genomics, over 200 genes that cause hearing loss have been identified. The delineation of specific audiological phenotypes based on the genetic etiology aids the understanding of some types of inherited hearing loss in terms of the prediction of clinical course of residual hearing, revelation of genotype—phenotype correlations and application of appropriate audiological rehabilitation.[5,6] To derive this information and enhance such understanding, many variants of the corresponding deafness gene and audiologic data must be available. Technically, establishing a correlation between genotype and audiologic phenotype can be challenging, particularly in cases involving recently discovered or novel deafness genes. Thus, a thorough analysis of the audiological data and genotypes of these

rare cases is mandatory.

Recently, functional classifications of genetic hearing loss based on the pathogenic mechanisms and the tonotopic expressions in the inner ear are being developed, including (1) hair bundle development and functioning; (2) synaptic transmission; (3) cellular adhesion and maintenance; (4) cochlear ion homeostasis; (5) extracellular matrix; (6) oxidative stress, metabolism, and mitochondrial defects; and (7) transcriptional regulation.[7] The application of these functional assignments provides a better understanding of genetic hearing loss, including the development of target drug and gene therapy.

The "central dogma" refers to the transfer of sequence information between RNA, DNA, and proteins within a biological system. It describes how the information embedded in DNA is transferred to mRNA (transcription) and how amino−acid chains are synthesized from mRNA (translation).[8] Transcription factors (TFs) recognize specific DNA sequences to control transcription by forming a complex that guides genome expression.[9] TFs generally contain several domains (effector, DNA−binding, and regulatory domains) that regulate their localization, chromatin accessibility, and transcriptional activity. More than 1600 human TFs have been documented in the literature,[9] and their variants have been

implicated in diverse diseases and syndromes, including cardiovascular diseases, cancer, neurological disorders, autoimmune diseases, and diabetes.[10] Transcriptional regulation serves as a critical role in the development and maintenance of hearing function. However, only a handful of TF variants are known to cause hearing loss, and their clinical phenotypes, genotypes, and molecular mechanisms in the context of hearing loss remain poorly understood.

1.2. Purpose of Research

First, this study aimed to elucidate the genetic landscape of disease-causing TF variants, as one of the functional classifications of genetic hearing loss (transcriptional regulation) and assess their clinical phenotype in the large-scale in-house database.

Second, this study aimed to suggest the novel auditory phenotype of variants in *LMX1A*, a poorly understood TF gene related to hearing loss, with the aim of revealing genotype-phenotype correlations.

Third, this study aimed to functionally characterize novel variants of *POU4F3*, one of the representative TF genes related to autosomal dominant hearing loss, via computational structural modeling and diverse aspects of the molecular studies, uncovering pathogenic mechanisms underlying POU4F3-associated DFNA15.

Chapter 2. Materials and Methods

2.1. Ethical Approval

All procedures in this study were approved by the Institutional Review Boards of Seoul National University Hospital (IRB-H-0905-041-281) and Seoul National University Bundang Hospital (IRB-B-1007-105-402). Written informed consent was obtained from both affected and unaffected members of the families. In the case of pediatric participants, written informed consent was obtained from their parents or guardians.

2.2. Molecular Genetic Testing

Genomic DNA was extracted from peripheral blood using the standard procedure and was subjected to exome sequencing using a Sure Select 50 Mb Hybridization and Capture Kit and a HiSeq2000 platform in four proband samples. The paired-end read length was 100 bp, and the reads were aligned using the University of California Santa Cruz (UCSC) hg19 reference genome browser

(https://genome.ucsc.edu/). As described previously,[11,12] bioinformatics analysis and strict filtering were then performed to retrieve candidate variants of deafness genes via the following filtering process: (i) Non-synonymous single nucleotide polymorphisms (SNPs) with quality scores > 30 and read depths > 20 were selected (ii) Each variants with minor allele frequencies (MAFs) ≤0.001 were chosen based on several database, including the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/); (iii) filtering was performed based on known deafness genes; (iv)) Each variants with MAFs ≤ 0.001 were included using ethnically-matched controls (Korean Reference Genome Database (KRGDB), http://152.99.75.168:9090/KRGDB/welcome.jsp) consisting of 1722 Korean individuals (3444 alleles); (v) The pathogenic potential of each variant was determined using in-silico tools (Combined Annotation Dependent Depletion (CADD), https://cadd.gs.washington.edu/ and Rare Exome Variant Ensemble Learner (REVEL), https://sites.google.com/site/revelgenomics/). In addition, Genomic Evolutionary Rate Profiling (GERP++) score from the UCSC Genome Browser (http://genome.ucsc.edu/) was utilized to estimate the evolutionary conservation of the amino acid sequences. Further, compatibility with inheritance patterns and

1 3

audiological phenotypes was evaluated; (vi) The candidate variants were confirmed through Sanger sequencing, and a segregation study was performed using paternal DNA samples. The pathogenicity of the variants was classified using the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) guidelines for genetic hearing loss.[13,14]

2.3. Clinical Phenotyping

The affected individuals underwent comprehensive evaluations, including medical history reviews, physical examinations, imaging, auditory phenotyping, and molecular genetic testing. Also, vestibular function tests were performed when possible. Audiological assessments included pure-tone audiometry or electrophysiological tests [of the auditory brainstem response threshold (ABRT) and auditory steady state response (ASSR)], depending on patient age. Auditory phenotyping focused on hearing loss onset, severity, progression, and asymmetry. The average hearing threshold was calculated by averaging the air-conduction thresholds at 0.5, 1, 2, and 4 kHz; hearing loss severity was classified as mild (20–40 dB), moderate (41–55 dB), moderately severe

(56–70 dB), severe (71–90 dB), or profound (>90 dB).[15] The audiogram configurations were subclassified into four categories: down–sloping, rising, U–shaped, or flat across the frequencies.[16] Asymmetric hearing loss was defined as a between–ear difference in the average hearing threshold > 15 dB when the hearing thresholds at all frequencies in both ears were > 25 dB, as previously described.[17]

2.4. Structure Modeling and Analysis

The AlphaFold Protein Structure Database was used to generate protein structures.[18,19] POU homeodomain and POU–specific domain were assembled with the DNA binding cleft in between. The homeobox protein HOX–B1/DNA ternary complex (PDB ID: 1B72) was aligned to the *POU4F3* structure to allow DNA–binding analysis.[20] The stabilities of truncated LMX1A and POU4F3 were evaluated based on the predicted aligned error (PAE) score, which reflects inter–domain accuracy. All the figures were generated using PyMOL (ver. 2.5.2) software (PyMOL Molecular Graphics System ver. 2.0, Schrödinger Inc., New York, NY, USA).

2.5. Cloning and Plasmid Construction

The human *POU4F3* cDNA clone (ORIGENE, CAT# RC211206) served as templates for *in vitro* mutagenesis. The *POU4F3* cDNA-loaded plasmids were subcloned into the pCMV expression vector using oligonucleotides that introduced Mlul and Asisl restriction sites into the 5′ and 3′ of the cDNA. The site-directed mutagenesis was performed to create plasmid construct encoding mutant *LMX1A* and *POU4F3*. Specifically, in the *POU4F3* p.Ala189Serfs*26 plasmid, 376 bp of the cDNA sequence from the in-frame stop codon to the Myc-DDK codon was deleted, and ligation was subsequently performed.

2.6. Cell Culture and Transfection

Human embryonic kidney 293 T cell (HET293T) line was obtained from American Type Culture Collection (ATCC; Manassas, VA, USA). HEK293T cells were transfected with constructs encoding wild-type and mutant proteins fused to C-terminal Myc-DDK tags using Lipofectamine 3000 transfection reagent (Invitrogen). After transfection for 48 hours, cells were fixed in 4% paraformaldehyde

for 15 minutes, permeabilized in phosphate-buffered saline (PBS) containing 0.3% Triton X-100 for 10 minutes, and then blocked in PBS containing 10% donkey serum for 1 hour at 37° C in a humid atmosphere.

2.7. Immunocytochemistry

The transfected cells were fixed in 4% paraformaldehyde for 15 min, followed by PBS washing, and the process was repeated three times. The HEK293T cells were incubated at 24° C with primary antibodies (anti-Myc [mouse, cell signaling, #2276,1:4000] or anti-DDK [goat, cell signaling, #14793,1:800, 2 h]) (Sigma Aldrich Corp., St. Louis, MO, USA)] for 2 h and Rhodamine phalloidin (Invitrogen, R415, 1:100) for 1 h. They were then washed three times with PBS (4℃), followed by consecutive incubation with secondary antibodies (anti-Myc; 488 goat anti-mouse, Invitrogen, A-11017, 1:400 for 2 h or anti-DDK; goat-anti-mouse Alexa680, Invitrogen, Seoul, Korea; 1:400 for 2 h) and Rhodamine phalloidin (Invitrogen, R415) 1:100 staining for 1h. After washing three times with PBS, the samples were mounted with DAPI (Vector Laboratories Inc., Burlingame, CA) at room temperature for 90 min.

The samples were examined with a laser scanning confocal microscope (Zeiss LSM 510, Carl Zeiss, Germany).

2.8. Western Blot and Cycloheximide Chase Assay

Whole proteins were separated using 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to 0.45 μm polyvinylidene difluoride (PVDF) membranes (Millipore; Billerica, MA, USA). The membranes were incubated with 5% skim milk to block nonspecific binding at room temperature for 1 h. Membrane blots were incubated against Myc-tag from Cell signaling Technology (Danvers, MA, USA) and β-actin from Santa Cruz biotechnology (Santacruz, CA, USA). The membranes with bound primary antibodies were incubated with anti-mouse secondary antibodies that were conjugated horseradish peroxidase (HRP) (Santacruz) for 1 h at room temperature. For cycloheximide chase assay, 80 ug/ml of cycloheximide was treated for indicated times. After incubation with cycloheximide, the cells were washed with PBS and lysed with RIPA lysis buffer. Total cell lysates were centrifuged at 13,000 rpm for 15 min and the supernatants were collected for the protein analysis. The protein band was detected using chemiluminescence (ATTA, Tokyo, Japan).

X-ray films (Agfa, Mortsel, Belgium) were used for detection. β-actin antibodies were used as loading controls. The intensity of bands was measured using the Image J software.

2.9. Luciferase Reporter Assay

HEK293T cells were transfected with the -711 bp upstream SNAP-25 promoter region, including KpnI (GGTAC/C) and XhoI (C/TCGAG) sites, cloned into the pGL4.12[luc2CP] vector (Promega) with pCMV6 vector (Myc-DDK), pCMV6-wt POU4F3 cDNA, pCMV6-p.Ala189Serfs*26 POU4F3 cDNA, pCMV6-p.Leu248Pro POU4F3 cDNA, pCMV6-p.Phe293Leu POU4F3 cDNA, and pCMV6-p.Val318Met POU4F3 cDNA (Fig. 1). Forty-eight hours later, the cells were lysed with luciferase cell lysis buffer (200 μL) and luciferase activities measured under conditions that minimize the ceiling effect (i.e., empty-luc 2 μg and enhance-luc 4 μg). The transcriptional activities were normalized to that of the internal control (Myc-DDK) in terms of fold changes. Experiments were performed in duplicate, and measurements were performed three times to identify Luciferase activity. Statistical significance was assessed by one-way ANOVA or the t-test.

2.10. RNA Sequencing and Bioinformatic Analysis

The RNA sequencing library was prepared and sequenced by Macrogen (Seoul, Korea). Further analyses were done by the Genomics Core Facility in Seoul National University Hospital (Seoul, Korea). RNA sequencing libraries were generated in accordance with the manufacturer's protocols (TruSeq Stranded TotalRNA LT sample prep kit; Illumina). After confirming library size, approximately 400-500 bp, paired-end sequencing was performed. Average total reads were 121,382,372 with over 95% of Q30 reads. For mapping and alignment, raw data were trimmed using the Trimmomatic program to remove adaptor sequences.[21] In addition, a window size of 4, mean quality of 15, and minimum length < 36 bp were set for trimming. Trimmed reads were mapped against the reference genome (GRCh37) using the HISAT2 program.[22] Over 95% of the processed reads were mapped and aligned using the Stringtie program to acquire transcript quantification.[23] In some cases, raw reads were analyzed using the Kallisto program as indicated.[24] Individual samples were further analyzed to achieve the differentially expressed gene (DEG) list using the DESeq2 program[25] with RLF normalization and nbinomWaldtest to list fold change over 2, and a p-value < 0.05. For the correlation analyses,

normalized reads were used to calculate the Spearman's coefficient. DEGs were further analyzed using Hierarchical and K-mean clustering with Euclidean distance and average linkage, and visualized as a heatmap using either Morpheus or the Multiple Experiment Viewer software (MeV, v4.9.0).[26] The Protein-Protein Interaction (PPI) analysis was performed using the Search Tool Retrieval of Interacting Genes/Proteins (STRING) v11.5 database.[27] Using the DEG list, g: profiler[28] or DAVID[29] ontology analyses were performed to examine the GO terms of biological process, cellular component, molecular function, and the KEGG pathway. Enriched GO terms were further analyzed using ReviGo[30] or QuickGo[31]. Plots were prepared using Excel, Grapad Prism, and R programs.

2.11. Real-time quantitative reverse transcription-Polymerized Chain Reaction

For validation of the RNA sequencing, expression levels of selected genes were evaluated using RT-qPCR. Total RNA was extracted from the lymphoblastoid cell line using RNeasy Mini Kit (Qiagen Inc., Hilden, Germany). cDNA was synthesized using the PrimeScript Reverse Transcriptase-reagent Kit (TaKaRa, RR037A) according

to the manufacturer's protocols. RT-qPCR was performed using a Light-Cycler 480 Instrument Ⅱ, using the Light Cycler 480 probes master kit (Roche; Indianapolis, IN, USA) and SYBR Premix Ex Taq II (TaKaRa, RR420A). The following primers were used for MYO6 genotyping: forward (5′-CCTGACCACTTAGCAGAGTTGG-3′) and reverse (5′-TTTAATGCAGGCTTCAGCTCGATA-3′). The following primers were used for BMP2 genotyping: forward (5′-TGTATCGCAGGCACTCAGGTCA-3′) and reverse (5′-CCATCCGTTTCTGGTACTTCTTC-3′). The following primers were used for AHI1 genotyping: forward (5′-GCTCAGTAGACACAGAACCTGG-3′) and reverse (5′-CTCCTGCATTTAGTGAGAAGAGG-3′). The relative gene expression was calculated using the 2-ddCt analysis method with GAPDH as the endogenous control.

Chapter 3. Results

3.1. Genetic Load of Alternations of Transcriptional Factor Genes in Hearing Loss

The DNA of 1280 probands was subjected to molecular genetic testing regardless of any specific audiologic phenotypes or modes of inheritance, and 720 probands in whom causative deafness variants were identified. Among them, 55 probands with genetically confirmed disease-causing TF variants were included (4.3%). Twenty-two Families harboring pathogenic variants implicated in syndromic deafness (1.7%), primarily Waardenburg syndrome and branchio-oto-renal syndrome, were excluded. Ultimately, 33 families (2.6%) with TF-associated non-syndromic deafness were included. The causative TF genes of these families were *POU3F4* (*n* = 16, 48.5%), *POU4F3* (*n* = 6, 18.2%), *LMX1A* (*n* = 6, 18.2%), and *EYA4* (*n* = 5, 15.2%). The disease-causing TF variants, which were exclusively clustered in only four TF genes (*POU3F4*, *POU4F3*, *LMX1A*, and *EYA4*), indicating a narrow molecular

etiologic spectrum and highlighting their role in non−syndromic deafness in Korea.

3.1.1 *LMX1A*: Genotypes and Associated Clinical Phenotypes

Nine patients from six *LMX1A*−associated families were identified. In most cases, the pedigrees indicated an autosomal dominant inheritance pattern. In one family, a de novo heterozygous missense variant (c.595A>G:p.Arg199Gly) was previously reported.[32] Four of the six variants were in the homeodomain, and the remaining two were truncated variants in LIM2 and the C−terminus, respectively. Remarkably, we identified four novel *LMX1A* heterozygous variants related to DFNA7 (Fig. 2). Of these novel variants, three (p.Arg208*, p.Gln240Arg, and p.Val241Met) were located in the homeodomain and one (p.Gln297Thrfs*41) in the C−terminus. One novel nonsense variant (p.Arg208*), creating a premature termination codon in the homeodomain, is extremely rare. This residue was highly conserved among the *LMX1A* orthologs of several species and in the *LMX1B* paralog. Moreover, this variant p.Arg208* was predicted to be disease−causing via *in silico*

analyses. Accordingly, p.Arg208* is "pathogenic" based on the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) guidelines. The remaining two novel missense variants (p.Gln240Arg and p.Val241Met) lie in the homeodomain and were also absent from population databases. Specifically, the missense variant p.Val241Leu has been previously reported in the same residue as the variant detailed herein (p.Val241Met).[33] These residues, Gln240 and Val241, in proteins encoded by the *LMX1A* orthologs of several species and the *LMX1B* paralog, are highly conserved, suggesting that they are functionally important. Additionally, they scored consistently highly on CADD and REVEL analyses, and were thus predicted to be "disease-causing". Thus, p.Gln240Arg and p.Val241Met were classified as "variant of uncertain significance" and "likely pathogenic" respectively, based on the ACMG/AMP guidelines. Finally, the novel frameshift variant (p.Gln297Thrfs*41) lay in the last exon of *LMX1A*, which is predicted to escape nonsense-mediated mRNA decay. The truncated variant is associated with premature termination of translation at codon 338; associated with loss of > 10% of the LMX1A protein. The variant is absent from population databases, and the affected residue (p.Gln297) is highly conserved among *LMX1A* orthologs and the *LMX1B* paralog. Accordingly,

p.Gln297Thrfs*41 is "pathogenic" based on the ACMG/AMP guidelines. In summary, the novel *LMX1A* variants are classified as "pathogenic" (p.Arg208* and p.Gln297Thrfs*41), "likely pathogenic" (p.Val241Met), and "variant of uncertain significance" (p.Gln240Arg) (Table 2).

In the four families segregated with *LMX1A* novel variants, asymmetric hearing loss (interaural difference > 15dB) was identified in most affected individuals for whom audiological evaluations were possible. Overall, based on the systematic review of the auditory phenotype of patients with *LMX1A*-related DFNA7 (Table 3), asymmetric hearing loss was previously identified in 7 of 10 affected individuals with *LMX1A*-related conditions using the same criteria for asymmetric hearing loss that we employ here. The radiological evaluations did not reveal any cochleovestibular malformation that might explain the loss. Three of the four *LMX1A* patients who were eligible for follow-up audiometry reported progressive hearing loss. In one patient, hearing deteriorated to profound hearing loss in her left ear and the asymmetric hearing loss remained. The patient eventually underwent unilateral cochlear implantation (CI), with significant improvement in her speech perception scores 3 and 6 months postoperatively.

3.1.2 *POU4F3*: Genotypes and Associated Clinical Phenotypes

Ten patients from six *POU4F3*-associated families were identified, and in all of them an autosomal dominant inheritance pattern was determined. All familial variants were missense or frameshift variants within the two functional domains, including the POU-specific and POU homeodomain. Exome sequencing revealed four novel *POU4F3* variants, segregating as a dominant trait in the four unrelated Korean families: one frameshift variant produced a premature termination codon in the POU-specific domain (p.Ala189Serfs*26) that lacked both mono- and bi-partite nuclear localization signals (NLSs); two missense variants (p.Leu248Pro and p.Phe293Leu) in which the alterations were located within the POU-specific domain and POU homeodomain, respectively, but outside the NLSs; and one missense variant (p.Val318Met) in which the alteration was located within the bipartite NLS (Fig. 4). Co-segregation of the variants with the phenotypes of the family members, including both parents, was confirmed. The novel frameshift variant (p.Ala189Serfs*26) produced a truncated protein, which was predicted to escape the nonsense-mediated mRNA decay. The truncated variant was associated with premature

termination of translation at codon 215, and a loss of > 10% of the *POU4F3* protein. This variant was not only absent from the KRGDB but also demonstrated a low global MAF. Conservation of residue Ala189 in proteins encoded by *POU4F3* orthologs between several species. In addition, this variant was consistently predicted to be disease-causing by *in silico* analyses based on both CADD and REVEL scores. Next, a novel missense variant (p.Leu248Pro), also located in the C-terminus of the POU-specific domain, demonstrated an extremely low MAF. Conservation of residue Leu248 in proteins encoded by *POU4F3* orthologs in several species and evidenced by a high GERP++ score of 5.44, suggests its functional importance. Also, this variant was consistently predicted to be disease-causing by *in silico* analyses (CADD and REVEL scores). The remaining two novel missense variants (p.Phe293Leu and p.Val318Met), located in the POU-specific domain and POU homeodomain, respectively (the latter variant within the bipartite NLS), were also absent from public databases. These residues, Phe293 and Val318, in proteins encoded by the *POU4F3* orthologs of several species, are highly conserved. Indeed, these variants were consistently predicted to be disease-causing by *in silico* analyses (CADD and REVEL scores). Accordingly, based on the ACMG/AMP guidelines for hearing loss, the novel

POU4F3 variants were classified as " pathogenic " (p.Ala189Serfs*26) and " variant of uncertain significance " (p.Leu248Pro, p.Phe293Leu, and p.Val318Met) (Table 4).

Nine of the ten patients (90%) in the POU4F3 group had SNHL, except one patient who had mixed hearing loss (Table 1). The audiograms had a U−shaped configuration, characterized by a mid− frequency notch at 1–2 kHz in five patients (50.0%). Down−sloping (n = 3, 30.0%), mixed hearing loss (n = 1, 10.0%), and flat (n = 1, 10.0%) configurations characterized the audiograms of the remaining patients. The severity of hearing loss tended to be moderate to moderately severe initially but progressed thereafter. In three patients, their hearing loss eventually deteriorated to severe−to−profound, and they underwent CI at a mean age of 41.3 years (SD: 13.1). One patient was implanted bilaterally in a single procedure, and the other two patients were implanted unilaterally. The CI outcomes were favorable, with K−CID, PB, and spondee scores above 96%, 70%, and 70% at the 1−year post−operative exam, respectively. One patient displayed bilateral moderate SNHL in her early 30s and opted for bilateral middle ear implantation (MEI) surgery rather than a hearing aid due to unsatisfactory experience with conventional hearing aids. She has been a satisfied user of middle ear implantation for 6 years.

3.2 Functional Study of Novel *POU4F3* variants associated with autosomal dominant hearing loss

In this cohort study, it has been ascertained that the prominent TF gene associated with genetic hearing loss, *POU4F3*, is recognized as a primary causative gene for autosomal dominant hearing loss. Exome sequencing was used to identify four novel *POU4F3* variants (c.564dupA: p.Ala189Serfs∗26, c.743T>C:p.Leu248Pro, c.879C>G:p.Phe293Leu, and c.952G>A:p.Val318Met), and diverse aspects of the molecular consequences of their protein expression, stability, subcellular localization, and transcriptional activity were investigated. Furthermore, this study investigates whether *POU4F3* variants can impact the expression of downstream target genes, potentially affecting inner ear development, using patient−derived cell lines.

3.2.1 3D Protein Modeling and Structure Analysis

The DNA binding interface of Alpha−fold generated POU4F3 model structure is depicted (Fig. 5). Alpha−fold generated model

structure of POU4F3 was used to examine the effects of the identified variants on POU4F3 protein structure, compared to wild-type POU4F3 protein (Fig. 5a). Leu248 amino acid residue is present in the helix-d of the POU-specific domain. Intra-helical proline substitution at Leu248 causes helical kinks, resulting in dramatic conformational changes in the POU-specific domain (Fig. 5b). Next, Phe293 amino acid residue is present in the helix-b of the POU-homeodomain. The missense variant p.Phe293Leu collapses the interhelical interface (aromatic ring stacking) between Phe293, Trp321, and Phe322 in helix-a (left) by disrupting biochemical interactions between helix-a and helix-b, which in turn destabilize the helical assembly of POU-homeodomain (Fig. 5c). Val318 amino acid residue is present in the helix-a of the POU-homeodomain. The long side chain of the missense variant Val318Met collapses the hydrophobic interactions with Ile307, Leu289, and Leu311, and induces hydrogen-bonding, destabilizing the POU-homeodomain helical assembly by causing molecular clashes with the adjacent Ile307 and Leu289 (Fig. 5d). PAE analysis also indicated that p.Ala189Serfs*26-induced premature termination of translation destroys the POU4F3 protein structure, including the DNA-binding functional domains, in turn destabilizing the protein (Fig. 6). Collectively, the *POU4F3* variants

compromised protein stability, and probably impaired the DNA-binding ability.

3.2.2 Protein Expression and Stability

The western blot analysis demonstrated that the wild-type and the three mutant proteins carrying missense variants (p.Leu248Pro, p.Phe293Leu, and p.Val318Met) were expressed as a single band corresponding to the correct molecular weight (36 kDa), indicating that the staining was derived explicitly from the tagged POU4F3 proteins (Fig. 7a). Compared to the wild-type protein, the three mutant proteins carrying missense variants had weaker intensities, probably due to protein instability (Fig. 7b). Additionally, the p.Ala189Serfs*26 POU4F3 protein was stably expressed with a smaller molecular weight (21 kDa) due to premature termination of the POU4F3 protein. Interestingly, the expression of a truncated protein (p.Ala189serfs*26) was stronger than the wild-type protein (Fig. 7b).

To determine whether *POU4F3* variants destabilize POU4F3 protein, we performed cycloheximide (CHX) chase assays to block protein synthesis. HEK293T cells were transfected with wild type and four mutant POU4F3 vectors for 24 h, followed by treatment

with CHX (80 μg/ml) for 1, 2, and 3 h, respectively. Three missense variants decreased the stability of POU4F3 protein compared with wild-type protein. Conversely, the half-life of the truncated mutant protein (p.Ala189serfs*26) lacking both mono- and bi-partite NLSs showed a longer trend compared with the wild-type protein, suggesting the mutant protein (p.Ala189serfs*26) was more stable than the wild-type protein (Fig. 7c).

3.2.3 Subcellular localization

Subcellular localization of transcriptional factors in the nucleus is necessary for its transcriptional activity that regulates target gene expression. All the mutant POU4F3 proteins showed significantly reduced reporter gene expression compared to the wild-type protein. The HEK293T cells transfected with the empty vector (negative control) demonstrated no cytoplasmic or nuclear fluorescence, confirming that the small tags attached to empty vectors did not induce any additional trafficking of the cloned protein to the target cells and organelles. Further, HEK293T cells were transfected with constructs encoding wild-type and mutant proteins fused to C-terminal Myc-DDK tags (Fig 8a,b). Notably,

the mutant POU4F3 (p.Ala189Serfs*26), which lacks both mono-
and bi-partite NLSs, localized to both the cytoplasm and the
nucleus, but the nuclear proportion (approximately 2%) was
significantly lower than that of the cytoplasm (approximately 98%).
In contrast, all other mutant POU4F3 proteins (p.Leu248Pro,
p.Phe293Leu, and p.Val318Met) localized exclusively to the nuclei
(approximately 98%), consistent with the localization of the wild-
type protein (Fig. 8c).

3.2.4 Transcriptional activity

To investigate the transcriptional activities of the mutant POU4F3
proteins, an *in vitro* luciferase reporter assay incorporating the
SNAP-25-Luc reporter construct was performed. HEK293T cells
were transfected with six pCMV6 plasmid constructs encoding
Myc-DDK only (negative control), wild-type POU4F3, mutant
POU4F3 (p.Ala189Serfs*26), mutant POU4F3 (p.Leu248Pro),
mutant POU4F3 (p.Phe293Leu), and mutant POU4F3
(p.Val318Met). The fold changes in the luciferase activities of
mutant POU4F3 proteins compared to those of the wild-type
protein were analyed. The experimental condition yielding optimal
wild-type POU4F3-induced transcription efficiency was

determined to be $2 \mu g$ *SNAP-25-Luc*, to minimize the influence of the ceiling effect. While the wild-type POU4F3 increased the luciferase activity approximately four-fold, the mutant POU4F3 showed only two-fold increase in luciferase activity, demonstrating a significantly poorer transcriptional activity of mutant POU4F3 (p < 0.001) to elicit transcription of the downstream target genes of *POU4F3* (Fig. 9).

3.2.5 RNA sequencing and bioinformatic analyses

RNA sequencing analysis was performed to investigate comprehensively the molecular pathways affected by *POU4F3* variants identified from the hearing-loss families. Specifically, the patient-derived lymphoblastoid cell lines (LCLs) to mimic the original molecular regulatory programs were utilized. These cell lines were, at least in part, altered by these variants, thereby replicating the original pathogenic circumstances. Figure 10a depicts the experiment and analysis flow. For subsequent sequencing analysis, we read over 100,000,000 reads from eight RNA-sequencing libraries. Around 95% of the reads passed quality inspection (Q > 30), and an average of 95% reads were successfully mapped to the human reference genome (GRCh37),

covering approximately 20,000 human genes.

Prior to the transcriptome analyses, a correlation analysis was conducted to determine whether these cells exhibited the molecular features of the cochlea. Mouse transcriptome data, including adult cochlear inner and outer hair cells, and early postnatal cochlea, was also used because public RNA—sequencing data for human cochlear tissues were unavailable. The adult mouse testis transcriptome data was used as negative controls. Table 5 summarizes the public data used in this study. The Spearman's correlation coefficients ranged from 0.4 to 1, indicating a positive correlation (Fig. 11a). Statistically significant p—values were found in two reference sets: 0.0399 for postnatal day 4 cochlea transcriptomes with patient samples and 0.0404 for postnatal day 7 cochlea transcriptomes with patient samples. There is a moderate correlation between the reference transcriptome and the patient cell transcriptome in each case, as shown by the Spearman's correlation coefficients of 0.53 and 0.521. Of note that both postnatal day 4, and postnatal day 7 cochlear transcriptome have a slightly stronger correlation with adult cochlear transcriptome with an average coefficient of roughly 0.75, suggesting that these transcriptomes favorably share similar molecular pathways. However, there was no correlation between the sample transcriptome, adult testis transcriptome, and the other

reference transcriptome (Fig. 11b, c). Overall, the RNA-seq data, as detailed in this study, would reflect more of the molecular signature of the early postnatal cochlea (postnatal day 4 and day 7) rather than other response pathways reported in adult tissues.

After observing that RNA-seq data contain more early postnatal cochlear regulatory molecular pathways, differential gene expression between wild-type and hearing-loss groups were analyzed. 630 genes had statistically significant expression variations, which were represented using a volcano plot and a heatmap split into upregulated (n = 203) and downregulated (n = 427) groups (Fig. 10b, c). We used GO (Gene Ontology) analysis to elucidate the biological processes underpinning the observed dysregulation. Approximately 360 GO terms associated with biological processes were substantially enhanced. Revigo was used to further visualize the top 30 GO terms, displaying them as representative. The most distinguishable groups consisted of cellular differentiation, cellular developmental processes, and other GO keywords pertaining to development. This trend was also seen in the top ten GO terms. Two of the most enriched GO-terms (p = 1.18E-11 and 1.8E-11, respectively) were cell differentiation (GO:0030154) and cellular developmental processes (GO:0048869) (Fig. 10d). Developmental process is an ancestor term, and inner

ear developmental GO-term is its child term (GO:0048839). Noteworthy is the fact that 14 genes are differentially expressed in GO:0048839 in the hearing loss group, with a p-value of 0.01. Its other ancestor GO-term is either ear development (GO: 0043583) or anatomical structure development (GO:0048856), for which 15 or 236 genes are enriched in each term, respectively. Using this ancestor chart, we determined that inner ear development is the most enriched child GO-term in the hearing-loss population. This study next examined the differential expression of the 14 genes in the inner ear developmental GO term (Fig. 10e).

To determine if these genes and *POU4F3* are physically and functionally associated, we performed STRING analyses under the assumption that mRNA expression level is linearly correlated with translation. The results observed possible protein associations in three groups (Fig. 12). There was the association of 14 dysregulated genes with the Notch pathway (*DLL1*), BMP pathway (*BMP2*), and Wnt (*SDC4*, *PTK7*, and *CELSR1*) pathway, which are connected to *Sox9*. These pathways have been reported as critical regulators to produce hair cells and dysregulations of them could lead to the hearing loss.[34] Initially, misregulation of BMP pathway was validated by measuring *BMP2* expression level quantitatively (Fig. 13).

Interestingly, *POU4F3* clustered together with *Myo6* and *DFNA5*, mutations or dysregulations of which have been found in hearing loss studies.[35-38] Since we observed *Myo6* expression significantly dysregulated in the hearing loss group, we further validated by qRT-PCR. Even though marginal, quantitative analyses reached statistical significance, suggesting the *POU4F3* variant regulates *Myo6* expression in the hearing loss pathology (Fig. 13). In the case of *AHI1*, even though we did not see the association in the STRING analyses, we further confirmed its expression level by qRT-PCR (Fig. 13), since it has been associated with non-syndromic deafness.[39] Overall, we confirmed dysregulation of *BMP2*, *Myo6* and *AHI1* by qRT-PCR (Fig. 13). These data concluded that the *POU4F3* variations might regulate Wnt, Notch, and/or BMP pathways, specifically leading to the misregulation of *BMP2*, *Myo6*, and *AHI1* in the pathogenesis of hearing loss.

In addition to the 14 enriched genes involved in the development of the inner ear, we also analyzed the expressoin level of known *POU4F3* target genes, including *Lhx3, Gfi1, Bdnf, Ntf3, Myo6, Caprin1, and Nr2f2* (Fig. 14). The majority of these genes, with the exception of *Lhx3* and *Ntf3*, were expressed based on transcriptome analysis, with average read-counts of 6224.024 and an average normalized value of 677.85. Statistically significant

downregulation was seen in *Bdnf* and *Myo6* in the hearing loss group, compared to the hearing loss group. The expression level of *Bdnf* was repressed by 93 percent, representing the most significant reduction. A reduced pattern was observed except for *Gfi1* and *Caprin1*, but the p-value exceeded 0.05. Overall, the dysregulation of *POU4F3* downstream targets was reaffirmed.

After that, we explored whether the four variants had distinct transcriptome signatures. Spearman's correlation analysis revealed that the truncating variant (p.Ala189Serfs*26) was less correlated to the other three variants (Fig. 15a). Notably, the truncating variant (p.Ala189Serfs*26) altered the subcellular location. DEseq2 was used to pinpoint the dysregulation between the truncating variant (p.Ala189Serfs*26) and the others. We used fold-changes and expression levels following the visualization of the MA-plot (Fig. 15b). These patterns of expression clearly demonstrated upregulation or downregulation (Fig. 15c). Then, we studied the enriched biological processes to observe numerous categories, including cellular processes, synthesis, stimuli responses, cellular process regulation, metabolic process regulation, and cellular localization (Fig. 15d). Nuclear import (GO: 0051170) was one of the significantly enriched GO-terms (Fig. 15d). We narrowed them down further to 51 DEGs. The *POU4F3* truncating

variant (p.Ala189Serfs*26) exhibited a diminished pattern (Fig. 15e), dysregulating the nuclear import process at a molecular level.

Chapter 4. Discussion

4.1. Genetic Load of Alternations of Transcription Factor Genes in Non－Syndromic Deafness

This study is the first to provide detailed genotype and audiological phenotypes associated with TF variants inducing non－syndromic deafness. In the clinical exome sequencing era, many questions regarding the pathogenic mechanisms of hearing loss have been answered, allowing a functional classification of the etiology of genetic hearing loss. Based on the in－house databases of genetic hearing loss, TF genes were implicated in ~3% of the study patients. Notably, 33 potentially pathogenic variants were observed, including nine novel variants, accounting for non－syndromic deafness clustered in only four TF genes (*POU3F4*, *POU4F3*, *LMX1A*, and *EYA4*), indicating a narrow molecular etiologic spectrum within the enormous number of TF genes reported thus far in humans (up to 1600 genes). The limited genetic spectrum of TF genes accounting for non－syndromic deafness suggests the functional redundancy of many other TF genes in inner ear

development or the maintenance of function. Alternatively, fetuses with variants in developmentally lethal, deafness−related TF genes may be spontaneously aborted. The results provide further insights into the genetic landscape of TF−related non−syndromic deafness and thus a basis for the implementation of a personalized, genetically tailored approach for audiological treatment and rehabilitation in these patients.

4.2. Novel Molecular Genetic Etiology of Asymmetric Hearing Loss: Autosomal−Dominant *LMX1A* Variants

Previous studies demonstrated that Lmx1a is predominantly expressed in the developing hindbrain and inner ear.[40] The hindbrain provides various extrinsic signals, including Lmx1a, for inner ear development, segregation, and patterning.[40,41] Specifically, the reciprocal negative interaction between Lmx1a and Lmo4 (LIM−domain only protein within the inner ear) is a key mechanism in patterning various components of the inner ear.[40] Furthermore, Lmx1a independently forms the endolymphatic duct and the hair cells in the basal cochlea.[40] In this study, the novel *LMX1A* variants identified herein compromised the DNA−binding ability and

significantly reduced the transcriptional activity,[42] proposing this possible causal relationship between loss of function of *LMX1A* and hearing loss .

Several risk factors were associated with asymmetric hearing loss, including congenital cytomegalovirus infection, inner ear malformations, brainstem lesions, and meningitis.[43-45] In particular, Waardenburg syndrome due to *MTIF* or *PAX3* variants has also been reported to manifest variable type of hearing loss, including single-sided deafness or asymmetric hearing loss.[46,47] Similarly, the variability in the severity of the hearing loss between both ears, reminiscent of the phenotype of asymmetric hearing loss, was reported in Waardenburg syndrome type II families with *KITLG* variants.[48] Additionally, some cases with *GJB2*, *SLC26A4* and *CLDN9* biallelic variants manifest non-syndromic asymmetric hearing loss.[49] More specifically, the *GJB2* c.235delC homozygous variant has been reported to account for a significant proportion of asymmetric hearing loss.[50] Despite previous evidence, the discovery of a genetic etiology that consistently exhibits asymmetric hearing loss has remained elusive, particularly in cases of non-syndromic hearing loss. The identification of such a gene would significantly advance the understanding and management of asymmetric hearing loss.

For the first time, the present study suggests that *LMX1A* is one of the candidate genes which, if altered, could be associated with dominantly inherited asymmetric hearing loss. More specifically, an auditory phenotype of between−ear asymmetry, with a characteristic severity and configuration, was identified in all probands, although the extent of asymmetry varied, suggesting that genetic testing for *LMX1A* should be prioritized during etiological examination of patients with asymmetric hearing loss and a familial history thereof. A subset of subjects with the *LMX1A*−dominant variants who first noticed hearing loss at various ages evidenced progressive asymmetric hearing loss in later life. Although the incidence of asymmetric hearing loss differs among studies,[51] any discrepancy may be explained by the different criteria used to define such loss and the age at ascertainment. Given the progressive nature of *LMX1A* variants,[51] the extent of asymmetry between the ears may be diluted over time, because the hearing of the better ear is more likely to deteriorate. The insight regarding the auditory phenotype of *LMX1A* variants could potentially lead to timely and appropriate audiological rehabilitation as a good example of precision medicine.

Although the exact mechanism of how *LMX1A* variants may lead to asymmetric hearing loss remains enigmatic, this phenotype

of asymmetric hearing loss may be attributable to incomplete penetrance in some ears with *LMX1A*-dominant variants. Additionally, various transcription factors have been reported to control left-right axis determination through transcriptional effectors and downstream target regulators to rescue the laterality defect.[52-54] In this sense, regulation of transcriptional effectors or downstream target of *LMX1A* may control left-right asymmetry in the inner ear properties, which might have contributed to the phenotype of asymmetric hearing loss in individuals with *LMX1A*-dominant variants. Given the significance of the inner ear source of *Lmx1a* on the formation of the hair cells in the cochlea,[40] there is a possibility of differential degree of degeneration of hair cells between two ears in the prenatal stage, leading to severe phenotype in the more penetrant ear.

4.3. Ramifications of *POU4F3* variants associated with autosomal dominant hearing loss in various molecular aspects.

POU transcription factor domains (POU-specific and POU-homeodomain) were associated with high-affinity DNA binding.[55,56]

Given that novel *POU4F3* variants were located in key regions encoding DNA−binding sites, it was hypothesized that the ability of the resultant mutant proteins to bind with their DNA targets would be compromised and would fail to induce sufficient target gene expression. Consistent with previous studies,[57,58] the 3D protein modeling and structure analysis showed that novel *POU4F3* missense variants disrupted the interhelical interface of the DNA− binding functional domains, in turn reducing protein expression and stability. Furthermore, western blot analysis demonstrated significantly weaker bands of the mutant proteins (p.Leu248Pro, p.Phe293Leu, and p.Val318Met), further revealing a greater instability of these mutant proteins compared to the wild−type protein based on the CHX chase assay. Moreover, the α −helix of the POU−specific domain and the third helix of the POU− homeodomain were essential for high−affinity DNA binding.[57] In this study, changes in the tertiary protein structure of the interhelical interface of the DNA−binding functional domains, caused by *POU4F3* variants, may affect the DNA−binding ability; thus, downstream target gene expression related to inner ear hair cell function could not be induced.[38,59] Together, the *POU4F3* variants identified in this study produced aberrant proteins that possibly disturb binding of target genes with the predicted POU4F3

recognition sequences, resulting in a significant reduction of the transcriptional activity necessary to induce regulation of downstream target gene expression.

Several downstream targets of *POU4F3*, including *Lhx3*[60], *Gfi1*[61], *Bdnf*[62], *Ntf3*[62], *Myo6*[63], *Caprin1*[64], and *Nr2f2*[65], are relevant to certain inner ear hair cell functions, which play important roles in inner ear development and maintenance. This study further identified altered expression of 14 downstream target genes associated with inner ear development using patient-derived lymphoblastoid cell lines, which was verified using RT-qPCR. Consistent with previous reports, *POU4F3* variants downregulated the expression of myosin 6, essential for maintenance of stereocilia of the hair cells, which is responsible for auditory mechanoelectrical transduction.[63] Altered expression of downstream *POU4F3* targets may provide a mechanistic basis for *POU4F3* variant-induced hearing loss (DFNA15). First, PCR and Sanger sequencing confirmed that *POU4F3* was obviously expressed in both human embryonic kidney cell line (HEK293T) and patient-derived lymphoblastoid cell lines, suggesting that these cell lines could at least provide the transcriptional environment to assess *POU4F3* function. Next, the downstream targets (*Gfi1*, *Bdnf*, *Myo6*, *Caprin1*, and *Nr2f2*) of *POU4F3* was identified to be expressed in the

lymphoblastoid cell lines, and some of them were misregulated in the patient-derived cell lines. Further, *POU4F3* was connected functionally and physically with 14 mis-regulated genes identified in the lymphoblastoid cell lines as evidenced by an observation that the Notch, Wnt, and BMP pathways that were association with 14 dysregulated genes were functionally linked with *Myo6* and *DFNA5* which clustered together with *POU4F3*. Finally, but not least, there was a significant moderate correlation of the expression profile between patient-derived cells and the cochlear hair cells. Collectively, the results provide evidence that patient-derived lymphoblastoid cell lines can be implicated in the transcriptome study of genetic hearing loss for studying *POU4F3* transcriptional function. Nonetheless, given their restricted expression pattern in the hair cells and the developing and adult sensory neurons, transfection into neuron-derived cell lines, such as PC12 and ND7, or inner ear sensory cells would be more relevant to assessing the transcriptional function of disease-causing *POU4F3* variants. As stated previously, this was a significant breakthrough for further transcriptome studies of genetic hearing loss in cases where cochlear tissue harvest is not clinically feasible.

  POU4F3 variants were associated with distinct subcellular localization patterns that merited further investigation. The POU

homeodomain contains two putative NLSs required for proper POU4F3 trafficking into the nucleus: N-terminal monopartite (amino acids 274 to 278; RKRKR) and C-terminal bipartite NLS (amino acids 314 to 331; KKNVVRVWFCNLQRQKQKR)[57]. Weiss *et al.* (2003) demonstrated that variants affecting monopartite NLS, bipartite NLS, or both, exerted differential effects on abnormal subcellular localization during the nuclear import process.[57] As such, it is conceivable that the frameshift variant producing a truncated protein (p.Ala189serfs*26), lacking both the mono- and bi-partite NLSs, localized exclusively in the cytoplasm. Weiss *et al.* (2003) previously reported that a truncated *POU4F3* protein lacking both the mono- and bi-partite NLSs exhibited significantly aberrant localization in the cytoplasm (> 80%), compared to other truncated *POU4F3* proteins lacking only the monopartite (approximately 23%) or the bipartite NLSs (approximately 47%). This suggests that both NLSs contribute significantly to nuclear trafficking of *POU4F3*[57]. In addition, the results observed differentially expressed transcripts associated with cellular localization between the truncating variant (p.Ala189serfs*26) and three other missense variants, suggesting that aberrant nuclear import (i.e., cytoplasmic localization of transcriptional factor) affects the ability to activate downstream targets. Among 51 DEGs related to nuclear import, expression of

Frizzled-7 receptor (FZD7) was the most downregulated. Specifically, *FZD7* interacts with the Wnt/β-catenin pathway,[66,67] which is required for cochlear hair cell differentiation.[68] Thus, a truncated *POU4F3* protein lacking both the mono- and bi-partite NLSs led to the downregulation of *FZD7* expression, which might reduce nuclear β-catenin accumulation and, in turn, affect cochlear air cell differentiation. Of note that these genes are expressed in the RNA-seq data, further supporting that experimental design and conclusion reflect the hearing loss pathology.

Interestingly, western blot analysis showed that the expression of the mutant protein (p.Ala189serfs*26) was stronger than the wild-type protein. Moreover, the mutant (p.Ala189serfs*26) was more stable than the wild-type protein upon protein stability assays. This was in agreement with a previous study, which demonstrated that the mutant (p.Ile295Thrfs*5) was more stable than the wild-type protein.[57]. It has also been demonstrated that the mutant protein had a significantly greater half-life than the wild-type protein.[57] Although the exact mechanism remains poorly understood, recent studies have demonstrated that the ubiquitin-proteasome system is interrelated with the bipartate NLS function with regard to regulation of protein stability.[69,70] The cellular protein quality control system (i.e., ubiquitin-proteasome system)

regulates t h e half-lives of various regulatory proteins and removes misfolded proteins.[71] NLS defects, including those of lysine ubiquitination sites, could decrease protein degradation while upregulating t h e i r h a l f - l i v e s .[69] In this sense, the mutant (p.Ala189serfs*26) without the bipartite NLS (including lysine ubiquitination sites) would be likely to be more stable than the wild-type protein, similar to the extended half-life of *POU4F3* with the mutant (p.Ala189serfs*26). However, this variant (p.Ala189serfs*26) might undergo nonsense-mediated mRNA decay *in vivo*, in contrast with the *in vitro* situation.

In this study, the predominantly nuclear localization of the p.Val318Met-POU4F3, residing within a bipartite NLS, did not perfectly align with the classical hypothesis. Lin *et al*. (2017) demonstrated that the missense variant (p.Lys328Glu), in which bipartite NLS amino acid residues were affected, was associated with aberrant POU4F3 subcellular localization,[70] which was in contrast with the present study. It has been suggested that the substitution of a basic lysine with an acidic glutamate (p.Lys328Glu) may alter bipartite NLS molecular properties, highlighting the importance of correct basic amino acid cluster alignment of bipartite NLS in maintaining the POU4F3 protein localization.[70] Indeed, bipartite NLS basic amino acids, the first two

(KK) and the last two (KR), are considered essential for nuclear localization of POU4F3.[57] The missense variant in this study (p.Val318Met), with conserved bipartite NLS basic amino acids and molecular properties, may therefore show normal nuclear localization, albeit with disrupted hydrophobic interactions. Further studies using naturally *POU4F3*-expressing inner ear hair cells are likely to provide additional information.

Chapter 5. Conclusions

In summary, the present study elucidates the clinical phenotypes and genotypes of non-syndromic deafness resulting from variants in TF genes. The DNA of 1280 probands was subjected to molecular genetic testing, and 720 probands in whom causative deafness variants were identified. Ultimately, 33 probands (2.6%) had non-syndromic deafness due to defective TF genes, which were exclusively clustered in only four TF genes (*POU3F4*, *POU4F3*, *LMX1A*, and *EYA4*), indicating a narrow molecular etiologic spectrum. Through the genotype-phenotype map of TF variants underlying non-syndromic hearing loss, this study defines the audiological phenotype of *LMX1A*-associated deafness characterized by asymmetric hearing loss. Furthermore, the results of this study present diverse functional aspects of novel *POU4F3* variants and identify 14 downstream target genes associated with inner ear development using patient-derived lymphoblastoid cell lines, which showed a significant correlation with cochlear hair cells, providing a breakthrough for cases where human cochlear sample collection was unfeasible. The insights into the auditory phenotype,

genotypes, and molecular mechanisms of hearing loss caused by TF variants could potentially pave the way for timely and appropriate audiological rehabilitation. This severs as a good example of precision medicine.

Bibliography

1   Morton, C. C. & Nance, W. E. Newborn hearing screening—a silent revolution. *New England Journal of Medicine* **354**, 2151-2164 (2006).

2   Lieu, J. E., Kenna, M., Anne, S. & Davidson, L. Hearing loss in children: a review. *JAMA* **324**, 2195-2205 (2020).

3   Lee, S.-Y. *et al.* The molecular etiology of deafness and auditory performance in the postlingually deafened cochlear implantees. *Scientific reports* **10**, 1-12 (2020).

4   Rim, J. H. *et al.* Differential genetic diagnoses of adult post-lingual hearing loss according to the audiogram pattern and novel candidate gene evaluation. *Human genetics*, 1-13 (2021).

5   Lee, S.-Y. *et al.* Flexible real-time polymerase chain reaction-based platforms for detecting deafness mutations in koreans: a proposed guideline for the etiologic diagnosis of auditory neuropathy spectrum disorder. *Diagnostics* **10**, 672 (2020).

6   Kim, B. J. *et al.* Rising of LOXHD1 as a signature causative gene of down-sloping hearing loss in people in their teens

and 20s. *Journal of Medical Genetics* (2021).

7   Delmaghani, S. & El-Amraoui, A. Inner ear gene therapies take off: current promises and future challenges. *Journal of clinical medicine* **9**, 2309 (2020).

8   Crick, F. H. On protein synthesis. *Symp Soc Exp Biol* **12**, 138-163 (1958).

9   Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **172**,   650-665   (2018). https://doi.org:10.1016/j.cell.2018.01.029

10   Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237-1251 (2013). https://doi.org:10.1016/j.cell.2013.02.014

11   Lee, S.-Y. *et al.* Severe or profound sensorineural hearing loss caused by novel USH2A variants in Korea: potential genotype-phenotype correlation. *Clinical and experimental otorhinolaryngology* **13**, 113 (2020).

12   Lee, S.-Y. *et al.* Novel KCNQ4 variants in different functional domains confer genotype-and mechanism-based therapeutics in patients with nonsyndromic hearing loss. *Experimental & Molecular Medicine* **53**, 1192-1204 (2021).

13   Oza, A. M. *et al.* Expert specification of the ACMG/AMP variant interpretation guidelines for genetic hearing loss.

Human mutation **39**, 1593-1613 (2018).

14   Abou Tayoun, A. N. *et al.* Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Human mutation* **39**, 1517-1524 (2018).

15   Lee, S.-Y. *et al.* Natural Course of Residual Hearing with Reference to GJB2 and SLC26A4 Genotypes: Clinical Implications for Hearing Rehabilitation. *Ear and Hearing* **42**, 644-653 (2021).

16   Lee, S.-Y. *et al.* No auditory experience, no tinnitus: lessons from subjects with congenital-and acquired single-sided deafness. *Hearing research* **354**, 9-15 (2017).

17   Lin, P.-H. *et al.* Etiologic and audiologic characteristics of patients with pediatric-onset unilateral and asymmetric sensorineural hearing loss. *JAMA otolaryngology-head & neck surgery* **143**, 912-919 (2017).

18   Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

19   Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research* (2021).

20   Piper, D. E., Batchelor, A. H., Chang, C.-P., Cleary, M. L. &

Wolberger, C. Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* **96**, 587-597 (1999).

21   Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

22   Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nature methods* **12**, 357-360 (2015).

23   Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature biotechnology* **33**, 290-295 (2015).

24   Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology* **34**, 525-527 (2016).

25   Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 1-21 (2014).

26   Saeed, A. *et al.* TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374-378 (2003).

27   Szklarczyk, D. *et al.* STRING v11: protein-protein association

networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research* **47**, D607-D613 (2019).

28   Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g: Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic acids research* **35**, W193-W200 (2007).

29   Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57 (2009).

30   Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS one* **6**, e21800 (2011).

31   Binns, D. *et al.* QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics* **25**, 3045-3046 (2009).

32   Lee, S. Y. *et al.* Novel genotype-phenotype correlation of functionally characterized LMX1A variants linked to sensorineural hearing loss. *Hum Mutat* **41**, 1877-1883 (2020). https://doi.org:10.1002/humu.24095

33   Wesdorp, M. *et al.* Heterozygous missense variants of LMX1A lead to nonsyndromic hearing impairment and vestibular dysfunction. *Human genetics* **137**, 389-400

(2018).

34   Müller, U. & Barr-Gillespie, P. G. New treatment options for hearing loss. *Nature reviews Drug discovery* **14**, 346-365 (2015).

35   Zhu, M. *et al.* Mutations in the γ-actin gene (ACTG1) are associated with dominant progressive deafness (DFNA20/26). *The American Journal of Human Genetics* **73**, 1082-1091 (2003).

36   Van Wijk, E. *et al.* A mutation in the gamma actin 1 (ACTG1) gene causes autosomal dominant hearing loss (DFNA20/26). *Journal of medical genetics* **40**, 879-884 (2003).

37   Ahmed, Z. M. *et al.* Mutations of MYO6 are associated with recessive deafness, DFNB37. *The American Journal of Human Genetics* **72**, 1315-1322 (2003).

38   Vahava, O. *et al.* Mutation in transcription factor POU4F3 associated with inherited progressive hearing loss in humans. *Science* **279**, 1950-1954 (1998).

39   Elsayed, S. M. *et al.* Non-manifesting AHI1 truncations indicate localized loss-of-function tolerance in a severe Mendelian disease gene. *Human molecular genetics* **24**, 2594-2603 (2015).

40   Huang, Y. *et al.* Reciprocal negative regulation between

Lmx1a and Lmo4 is required for inner ear formation. *Journal of Neuroscience* **38**, 5429-5440 (2018).

41   Wu, D. K. & Kelley, M. W. Molecular mechanisms of inner ear development. *Cold Spring Harbor perspectives in biology* **4**, a008409 (2012).

42   Lee, S.-Y. *et al.* Novel Molecular Genetic Etiology of Asymmetric Hearing Loss: Autosomal-Dominant LMX1A Variants. *Ear and Hearing* **43**, 1698-1707 (2022).

43   Vila, P. M. & Lieu, J. E. Asymmetric and unilateral hearing loss in children. *Cell and tissue research* **361**, 271-278 (2015).

44   Clemmens, C. S. *et al.* Unilateral cochlear nerve deficiency in children. *Otolaryngology--Head and Neck Surgery* **149**, 318-325 (2013).

45   Friedman, A. B. *et al.* Risk analysis of unilateral severe-to-profound sensorineural hearing loss in children. *International journal of pediatric otorhinolaryngology* **77**, 1128-1131 (2013).

46   Wang, G. *et al.* Analysis of genotype-phenotype relationships in 90 Chinese probands with Waardenburg syndrome. *Human Genetics*, 1-14 (2021).

47   Kim, S. H. *et al.* Molecular etiology of hereditary single-side

deafness: Its association with pigmentary disorders and Waardenburg syndrome. *Medicine* **94** (2015).

48   Seco, C. Z. *et al.* Allelic mutations of KITLG, encoding KIT ligand, cause asymmetric and unilateral hearing loss and Waardenburg syndrome type 2. *The American Journal of Human Genetics* **97**, 647-660 (2015).

49   Ramzan, M. *et al.* Variants of human CLDN9 cause mild to profound hearing loss. *Human mutation* **42**, 1321-1335 (2021).

50   Guo, C. *et al.* Hearing Phenotypes of Patients with Hearing Loss Homozygous for the GJB2 c. 235delc Mutation. *Neural Plasticity* **2020** (2020).

51   Wesdorp, M. *et al.* Heterozygous missense variants of LMX1A lead to nonsyndromic hearing impairment and vestibular dysfunction. *Human genetics* **137**, 389-400 (2018).

52   Wu, C.-S., Lu, Y.-F., Liu, Y.-H., Huang, C.-J. & Hwang, S.-P. L. Zebrafish Cdx1b modulates epithalamic asymmetry by regulating ndr2 and lft1 expression. *Developmental Biology* **470**, 21-36 (2021).

53   Concepcion, D., Hamada, H. & Papaioannou, V. E. Tbx6 controls left-right asymmetry through regulation of Gdf1.

Biology open **7**, bio032565 (2018).

54   Fillatre, J. *et al.* TEADs, Yap, Taz, Vgll4s transcription factors control the establishment of Left—Right asymmetry in zebrafish. *Elife* **8**, e45241 (2019).

55   Ruvkun, G. & Finney, M. Regulation of transcription and cell identity by POU domain proteins. *Cell* **64**, 475—478 (1991).

56   Verrijzer, C. P. *et al.* The DNA binding specificity of the bipartite POU domain and its subdomains. *The EMBO Journal* **11**, 4993—5003 (1992).

57   Weiss, S. *et al.* The DFNA15 deafness mutation affects POU4F3 protein stability, localization, and transcriptional activity. *Molecular and cellular biology* **23**, 7957—7964 (2003).

58   Collin, R. W. *et al.* Missense mutations in POU4F3 cause autosomal dominant hearing impairment DFNA15 and affect subcellular localization and DNA binding. *Human mutation* **29**, 545—554 (2008).

59   Cui, T.—Y. *et al.* Four novel variants in POU4F3 cause autosomal dominant nonsyndromic hearing loss. *Neural plasticity* **2020** (2020).

60   Hertzano, R. *et al.* Lhx3, a LIM domain transcription factor, is regulated by Pou4f3 in the auditory but not in the vestibular

system. *European Journal of Neuroscience* **25**, 999-1005 (2007).

61   Hertzano, R. *et al.* Transcription profiling of inner ears from Pou4f3 ddl/ddl identifies Gfi1 as a target of the Pou4f3 deafness gene. *Human molecular genetics* **13**, 2143-2153 (2004).

62   Clough, R. L. *et al.* Brn-3c (POU4F3) regulates BDNF and NT-3 promoter activity. *Biochemical and biophysical research communications* **324**, 372-381 (2004).

63   Ma, D. B. *et al.* Inhibition of Myo6 gene expression by co-expression of a mutant of transcription factor POU4F3 (BRN-3C) in hair cells. *Molecular medicine reports* **9**, 1185-1190 (2014).

64   Towers, E. R., Kelly, J. J., Sud, R., Gale, J. E. & Dawson, S. J. Caprin-1 is a target of the deafness gene Pou4f3 and is recruited to stress granules in cochlear hair cells in response to ototoxic damage. *Journal of cell science* **124**, 1145-1155 (2011).

65   Tornari, C., Towers, E. R., Gale, J. E. & Dawson, S. J. Regulation of the orphan nuclear receptor Nr2f2 by the DFNA15 deafness gene Pou4f3. *PloS one* **9**, e112247 (2014).

66   Gurney, A. *et al.* Wnt pathway inhibition via the targeting of

Frizzled receptors results in decreased growth and tumorigenicity of human tumors. *Proceedings of the National Academy of Sciences* **109**, 11717-11722 (2012).

67   Kim, M. *et al.* Functional interaction between Wnt3 and Frizzled-7 leads to activation of the Wnt/β-catenin signaling pathway in hepatocellular carcinoma cells. *Journal of hepatology* **48**, 780-791 (2008).

68   Jacques, B. E. *et al.* A dual function for canonical Wnt/β-catenin signaling in the developing mammalian cochlea. *Development* **139**, 4395-4404 (2012).

69   An, L. *et al.* Dual-utility NLS drives RNF169-dependent DNA damage responses. *Proceedings of the National Academy of Sciences* **114**, E2872-E2881 (2017).

70   Lin, Y.-H. *et al.* A novel missense variant in the nuclear localization signal of POU4F3 causes autosomal dominant non-syndromic hearing loss. *Scientific reports* **7**, 1-6 (2017).

71   Yun, Y. *et al.* Proteasome activity in the plasma as a novel biomarker in mild cognitive impairment with chronic tinnitus. *Journal of Alzheimer's Disease* **78**, 195-205 (2020).

Table 1. Phenotypes and genotypes associated with non-syndromic deafness caused by transcription factor variants.

Patient	Sex	Timing of HL	Genotype	Age at HL Detection	Type of HL	Audiogram Configuration	Degree of HL (Most Recent)	Asymmetry	HL Progression	Age at CI
SB2-1	M	prelingual	POU3F4[NM_000307.4]c.626A>G:p.Gln229Arg	1 month	SNHL	Flat	profound	No	No	R) 2 yr, L) 12 mo
SB2-2	M	prelingual	POU3F4[NM_000307.4]c.626A>G:p.Gln229Arg	6 months	SNHL	Flat	profound	No	No	R) 6 yr, L) 7 yr
SB7	M	prelingual	POU3F4[NM_000307.4]c.1060delA:p.Thr354Glnfs*115	12 months	SNHL	Flat	profound	No	No	2 yr
SB8	M	postlingual	POU3F4[NM_000307.4]c.950dupT:p.Leu317Phefs*12	35 months	MHL	Mixed HL	severe	No	Yes[d] (2.25 dB HL/yr)	(−)
SB9	M	postlingual	POU3F4[NM_000307.4]c.632C>T:p.Thr211Met	3 years	MHL	Mixed HL	severe	Yes	Yes (0.7 dB HL/yr)	(−)
SB11	M	prelingual	POU3F4[NM_000307.4]c.1084T>C:p.X362Argext*113	3 years	SNHL	Flat	profound	No	No	12 yr
SB13	M	prelingual	POU3F4[NM_000307.4]c.623T>A:p.Leu208*	15 months	SNHL	Flat	profound	No	Yes (0.8 dB HL/yr)	R) 6 yr, L) 2yr
SH17	M	prelingual	Xq21.2, 80851535-82597832 bp	1 month	SNHL	Flat	profound	No	Yes (4.7 dB HL/yr)	R) 13 mo, L) 25 mo
SB19	M	prelingual	Xq21.2, 81810457-82810060 bp	14 months	SNHL	Flat	profound	No	No	29 yr
SH54	M	postlingual	POU3F4[NM_000307.4]c.540C>A:p.Cys180*	1 month	MHL	Mixed HL	severe	No	Yes (1.4 dB HL/yr)	(−)
SH65	M	prelingual	POU3F4[NM_000307.4]c.910C>A:p.Pro303His	1 month	SNHL	Downsloping	severe	Yes	Yes (0.5 dB HL/yr)	3 yr
SH149	M	prelingual	POU3F4[NM_000307.4]c.458delC:p.Pro153Leufs*88	3 months	MHL	Mixed HL	profound	Yes	Yes (2.5 dB HL/yr)	3 yr
SH228	M	prelingual	POU3F4[NM_000307.4]c.989G>A:p.Arg330Lys	unknown [a]	MHL	Mixed HL	severe	No	No	(−)
SB332	M	prelingual	Xq21.2, deletion	1 month	MHL	N/A [b]	severe	No	No	(−)
SB430	M	prelingual	Xq21.2, deletion	1 month	MHL	N/A [b]	severe	No	No	21 mo
SH565-1	M	prelingual	POU3F4[NM_000307.4]c.958G>T:p.Glu320*	1 month	MHL	Mixed HL	moderate	No	No	(−)
SH565-2	M	prelingual	POU3F4[NM_000307.4]c.958G>T:p.Glu320*	2 month	MHL	Mixed HL	moderately severe	No	No	(−)
SB736	M	prelingual	POU3F4[NM_000307.4]c.626A>G:p.Gln229Arg	12 months	MHL	Mixed HL	profound	No	No	10 yr
SB218	F	postlingual	POU4F3[NM_002700.2] c.564dupA:p.Ala189Serfs*26	30 years	SNHL	U-shaped	moderate	No	Yes (1.6 dB HL/yr)	(−)
SB307	F	postlingual	POU4F3[NM_002700.2] c.743T>C:p.Leu248Pro	26 years	SNHL	U-shaped	moderately severe	No	Yes	(−)
SB347-1	F	postlingual	POU4F3[NM_002700.2] c.952G>A:p.Val318Met	16 years	MHL	Mixed HL	profound	No	Yes (16.7 dB HL/yr)	36 yr
SB347-2	F	postlingual	POU4F3[NM_002700.2] c.952G>A:p.Val318Met	20 years	SNHL	Flat	profound	No	Yes	52 yr
SB438-1	F	postlingual	POU4F3[NM_002700.2] c.879C>A:p.Phe293Leu	unknown [a]	SNHL	Downsloping	mild	No	Yes	(−)
SB438-2	F	postlingual	POU4F3[NM_002700.2] c.879C>A:p.Phe293Leu	37 years	SNHL	Downsloping	moderately severe	No	Yes (2.3 dB HL/yr)	(−)
SB618-1	M	prelingual	POU4F3[NM_002700.2] c.952G>A:p.Val318Met	unknown [a]	SNHL	U-shaped	moderate	No	Yes (5 dB HL/yr)	(−)
SB618-2	M	unknown [a]	POU4F3[NM_002700.2] c.952G>A:p.Val318Met	unknown [a]	SNHL	U-shaped	severe	Yes	unknown [a]	(−)
SB618-3	F	unknown [a]	POU4F3[NM_002700.2] c.952G>A:p.Val318Met	unknown [a]	SNHL	Downsloping	R) severe, L) profound	Yes	unknown [a]	(−)
SB709	F	postlingual	POU4F3[NM_002700.2] c.662_675del:p.Gly221Glufs*77	39 years	SNHL	U-shaped	profound	No	Yes (10.6 dB HL/yr)	36 yr
SB481	M	prelingual	LMX1A[NM_177398.4] c.595A>G:p.Arg199Gly	1 months	SNHL	N/A [b]	R) profound, L) severe	Yes	unknown [a]	(−)
SB727	F	postlingual	LMX1A[NM_177398.4] c.622C>T:p.Arg208*	13 years	SNHL	Downsloping	R) moderate, L) profound	Yes	Yes	32 yr
SH407	F	postlingual	LMX1A[NM_177398.4] c.887dup:p.Gln297Thrfs*41	20 years	SNHL	Downsloping	R) moderately severe, L) moderate	Yes	fluctuation	(−)
SB742-1	F	prelingual	LMX1A[NM_177398.4] c.719A>G:p.Gln240Arg	2 months	SNHL	N/A [b]	R) moderate, L) profound	Yes	No	(−)
SB742-2	F	postlingual	LMX1A[NM_177398.4] c.719A>G:p.Gln240Arg	20 years	SNHL	Downsloping	moderately severe	No	Yes	(−)
SH421-1	F	prelingual	LMX1A[NM_177398.4] c.721G>A:p.Val241Met	4 months	SNHL	N/A [b]	moderate	No	unknown [a]	(−)
SH421-2	M	postlingual	LMX1A[NM_177398.4] c.721G>A:p.Val241Met	17 years	SNHL	Downsloping	R) profound, L)	Yes	No	(−)

SH512-1	F	prelingual	*LMX1A*[NM_177398.4] c.331del:p.Gln111Argfs*7	3 months	SNHL	N/A [b]	moderate	No	unknown [a]	(−)
SH512-2	F	prelingual	*LMX1A*[NM_177398.4] c.331del:p.Gln111Argfs*7	1 year	SNHL	Downsloping	severe	No	unknown [a]	(−)
SB302-1	F	postlingual	*EYA4*[NM_004100.5] c.697C>T:p.Gln233*	35 years	SNHL	U−shaped	moderate	No	Yes	(−)
SB302-2	F	postlingual	*EYA4*[NM_004100.5] c.697C>T:p.Gln233*	40 years	SNHL	Flat	severe	No	Yes	(−)
SB545	F	postlingual	*EYA4*[NM_004100.5] c.208+1del	50 years	MHL	Downsloping	severe	No	Yes	80 yr
SB865	F	postlingual	*EYA4*[NM_004100.5] c.578dup:p.Tyr193*	10 years	MHL	Downsloping	severe	Yes (54 dB)	No	(−)
SH537	F	postlingual	*EYA4*[NM_004100.5] c.1468G>T:p.Glu490*	45 years	SNHL	Flat	moderately severe	No	Yes	(−)
SH117-1	F	postlingual	*EYA4*[NM_004100.5] c.1194del:p.Met401Trpfs*3	15 years	SNHL	Downsloping	moderate	No	Yes	(−)
SH117-2	M	postlingual	*EYA4*[NM_004100.5] c.1194del:p.Met401Trpfs*3	unknown [a]	SNHL	Downsloping	severe	No	unknown [a]	(−)
SH117-3	F	postlingual	*EYA4*[NM_004100.5] c.1194del:p.Met401Trpfs*3	unknown [a]	SNHL	Downsloping	moderate	No	unknown [a]	(−)

Abbreviations: M, male; F, female; B, both; R, right; L, left; HL, hearing loss; dB, decibel; SNHL, sensorineural hearing loss; MHL, mixed hearing loss; HA, hearing aid; CI, cochlear implant; MEI, middle ear implant; yr, year; mo, month.

[a] Unknown, due to lack of record.

[b] Not available, because the patient was too young to undergo pure tone audiometry.

[c] Cochlear implant was done on both ears simultaneously.

[d] Hearing loss progression was observed in bone conduction only.

6 8

Table 2. *LMX1A* novel variants in the current study and its pathogenicity prediction analysis

| Family | Genomic position (GRCh37/hg19) | HGVS | | Domain | Zygosity | In-silico Prediction | | | | MAF | ACMG/AMP guideline |
		Coding DNA change	Protein change			CADD	REVEL	GERP	KRGDB (1722 individuals)	gnomAD	Classification
SB727	chr1:165182925	c.622C>T	p.Arg208*	Homeodomain	Het	47.00	0.581	5.75	Absent	Absent	Pathogenic
SB742	chr1:165179964	c.719A>G	p.Gln240Arg	Homeodomain	Het	27.0	0.892	5.61	Absent	Absent	VUS
SH421	Chr1:165179962	c.721G>A	p.Val241Met	Homeodomain	Het	25.7	0.928	5.61	Absent	Absent	Likely Pathogenic
SH407	Chr1:165175201	c.887dup	p.Gln297Thrfs*41	NA (C-terminus)	Het	NA	NA	NA	Absent	Absent	Pathogenic

Abbreviations: MAF, Minor allele frequency; Het, heterozygote; VUS, variant uncertain significance; NA, not available
Refseq transcript accession number NM_177398.4; Refseq protein accession number NP_796372.1
HGVS: Human Genome Variation Society (https://www.hgvs.org/)
Sequence Variant Nomenclature (http://varnomen.hgvs.org/)
CADD: Combined Annotation Dependent Depletion (https://cadd.gs.washington.edu/)
REVEL: Rare Exome Variant Ensemble Learner (https://sites.google.com/site/revelgenomics/)
KRGDB: Korean Reference Genome Database (http://coda.nih.go.kr/coda/KRGDB/index.jsp)
ExAC: Exome Aggregation Consortium databases
gnomAD: The Genome Aggregation Database (https://gnomad.broadinstitute.org/)
ACMG/AMP 2018 guideline (http://wintervar.wglab.org/)

Table 3. Audiological phenotype of previously discovered and newly identified heterozygous variants in *LMX1A* associated with DFNA7

Origin	Family number	Sex/ Age	LMX1A (NM_177398.3; NP_796372.1)			Audiological phenotype						Reference
			HGVS nucleotide change	HGVS protein change	Exon (rank/total) Domain	Audiologic evaluation	Hearing threshold (0.5−1−2−4kHz) and Average: Rt	Hearing threshold (0.5−1−2−4kHz) and Average: Lt	Asymmetry[a]	Progression		Reference
Dutch	II:7	F/56	c.721G>C	p.Val241Leu	6/9 Homeodomain	PTA	40−80−80−90 (72.5)	85−90−90−85 (87.5)	Yes	Yes		Wesdorp et al. 2018
		F/73				PTA	60−80−75−80 (73.75)	115−115−120−120 (117.5)				
	III:8	F/24	c.721G>C	p.Val241Leu	6/9 Homeodomain	PTA	20−40−45−50 (38.75)	30−25−20−55 (32.5)	No	Yes		Wesdorp et al. 2018
		F/44				PTA	40−60−70−70 (60)	40−60−60−55 (53.75)				
	IV:2	F/6	c.721G>C	p.Val241Leu	6/9 Homeodomain	PTA	45−55−60−55 (53.75)	10−20−35−40 (26.25)	Yes	Yes		Wesdorp et al. 2018
		F/9				PTA	40−45−50−50 (46.25)	20−15−35−70 (35)				
Dutch	I:2	F/85	c.290G>C	p.Cys97Ser	4/9, LIM2	PTA	80−75−85−85 (81.25)	80−90−80−80 (82.5)	No	N/A		Wesdorp et al. 2018
	II:2	F/26	c.290G>C	p.Cys97Ser	4/9, LIM2	PTA	35−15−20−45 (28.75)	10−0−5−15 (7.5)	Yes	Yes		Wesdorp et al. 2018
		F/54				PTA	50−60−65−65 (60)	20−30−30−45 (31.25)				
	II:3	M/15	c.290G>C	p.Cys97Ser	4/9, LIM2	PTA	80−80−80−80 (80)	80−90−90−80 (85)	No	Yes		Wesdorp et al. 2018
		M/52				PTA	100−115−120−120 (113.75)	100−120−120−120 (115)				
	II:4	M/30	c.290G>C	p.Cys97Ser	4/9, LIM2	PTA	20−20−40−40 (30)	30−25−25−30 (27.5)	No	Yes		Wesdorp et al. 2018
		M/40				PTA	40−45−50−60 (48.75)	55−55−60−45 (53.75)				
Korea	SB481− 927	M/3m	c.595A>G	p.Arg199Gly	5/9 Homeodomain	ASSR	100−100−100−110 (102.5)	90−90−70−70 (80)	Yes	N/A		Lee et al. 2020
Korea	SB727− 1294	F/31	c.622C>T	p.Arg208*	5/9 Homeodomain	PTA	35−50−50−60 (48.75)	75−80−80−100 (83.75)	Yes	Yes		Lee et al. 2021
Korea	SB742− 1317	F/2m	c.719A>G	p.Gln240Arg	6/9 Homeodomain	ASSR	50−50−50−50 (50)	100−100−100−100 (100)	Yes	N/A		Lee et al. 2021
	SB742− 1320	F/60	c.719A>G	p.Gln240Arg	6/9 Homeodomain	PTA	50−60−70−65 (61.25)	40−70−60−65 (58.75)	No	N/A		Lee et al. 2021
Korea	SH421− 906	F/2m	c.721G>A	p.Val241Met	6/9 Homeodomain	ASSR	50−40−40−40 (42.5)	50−70−60−50 (57.5)	Yes	N/A		Lee et al. 2021
	SH421− 907	M/24	c.721G>A	p.Val241Met	6/9 Homeodomain	PTA	110−110−110−110 (110)	40−40−45−50 (43.75)	Yes	Yes		Lee et al. 2021
		M/31				PTA	95−115−115−120 (106.25)	25−40−50−75 (47.5)				

| Korea | SH407-878 | M/21 | c.887dup | p.Gln297Thr fs*41 | 8/9 C-terminus | PTA | 55-70-60-55 (60) | 50-55-35-30 (40) | Yes | N/A | Lee et al. 2021 |

Abbreviation: M, male; F, female; m, months; Rt, right; Lt, left; PTA, pure tone audiometry; ASSR, auditory steady state response; N/A, not available; HGVS, Human Genome Variation Society (https://www.hgvs.org/)

[a] Note that asymmetric hearing loss was defined as a between-ear difference in the average hearing threshold.

Table 4. *POU4F3* novel variants in the current study and in-silico prediction analysis

Proband	Genomic Position: Change (GRCh37/hg19)	HGVS		Location (Exon /Domain)	Zygosity/ Inheritance	Insilico Predictions		Alternative Allele Frequency		ACMG/AMP guideline
		Nucleotide change	Amino Acid change			CADD Phred	REVEL	KRGDB (1722 individuals)	GMAF (gnomAD)	Classification
SB218-423	Chr5:145719554A>AA	c.564dupA	p.Ala189Ser*fs**26	Exon2 / POU	Het / Autosomal dominant	NA	NA	Absent	Absent	Pathogenic
SB307-610	Chr5:145719733T>C	c.743T>C	p.Leu248Pro	Exon2 / POU	Het / Autosomal dominant	29.3	0.950	Absent	Absent	VUS
SB438-852	Chr5:145719869C>G	c.879C>G	p.Phe293Leu	Exon2 / Homeobox	Het / Autosomal dominant	24.8	0.913	Absent	Absent	VUS
SB347-679	Chr5:145719942G>A	c.952G>A	p.Val318Met	Exon2 / Homeobox	Het / Autosomal dominant	29.6	0.936	Absent	Absent	VUS

Abbreviations: MAF, Minor allele frequency; Het, heterozygote; VUS, variant uncertain significance; NA, not available
Refseq transcript accession number NM_002700.2; Refseq protein accession number NP_002691
HGVS: Human Genome Variation Society (https://www.hgvs.org/)
Sequence Variant Nomenclature (https://mutalyzer.nl/)
CADD: Combined Annotation Dependent Depletion (https://cadd.gs.washington.edu/)
REVEL: Rare Exome Variant Ensemble Learner (https://sites.google.com/site/revelgenomics/)
KRGDB: Korean Reference Genome Database (http://152.99.75.168:9090/KRGDB/welcome.jsp)
gnomAD: The Genome Aggregation Database (https://gnomad.broadinstitute.org/)

Table 5. A summary of the public data used in the current study

| SRR ID	Source Tissue	Organism	Age	Total Reads	Mapped Reads	Reference
SRR6798475	Cochlear inner hair cell	mouse	Adult	7,332,016	6,657,566	Yi Li et al, Scientific Data volume 5, Article number: 180199 (2018)
SRR6798476	Cochlear inner hair cell	mouse	Adult	229,368	2,085,153	
SRR6798477	Cochlear inner hair cell	mouse	Adult	12,585,587	11,397,869	
SRR6798479	Cochlear outer hair cell	mouse	Adult	14,468,752	12,296,204	
SRR6798481	Cochlear outer hair cell	mouse	Adult	12,215,142	10,377,025	
SRR6798482	Cochlear outer hair cell	mouse	Adult	7,161,499	6,087,056	
SRR1534779	Cochlea	mouse	Postnatal day 0	117,562,038	23,321,842	Scheffer D et al, The Journal of Neuroscience, April 22, 2015, 35(16):6366–6380
SRR1534787	Cochlea	mouse	Postnatal day 4	29,941,458	20,392,791	
SRR1534792	Cochlea	mouse	Postnatal day 7	36,343,531	23,813,133	
SRR15597783	Testis	mouse	Adult	33,528,710	27,553,478	Han G, Cho C Series GSE175633 (2021)
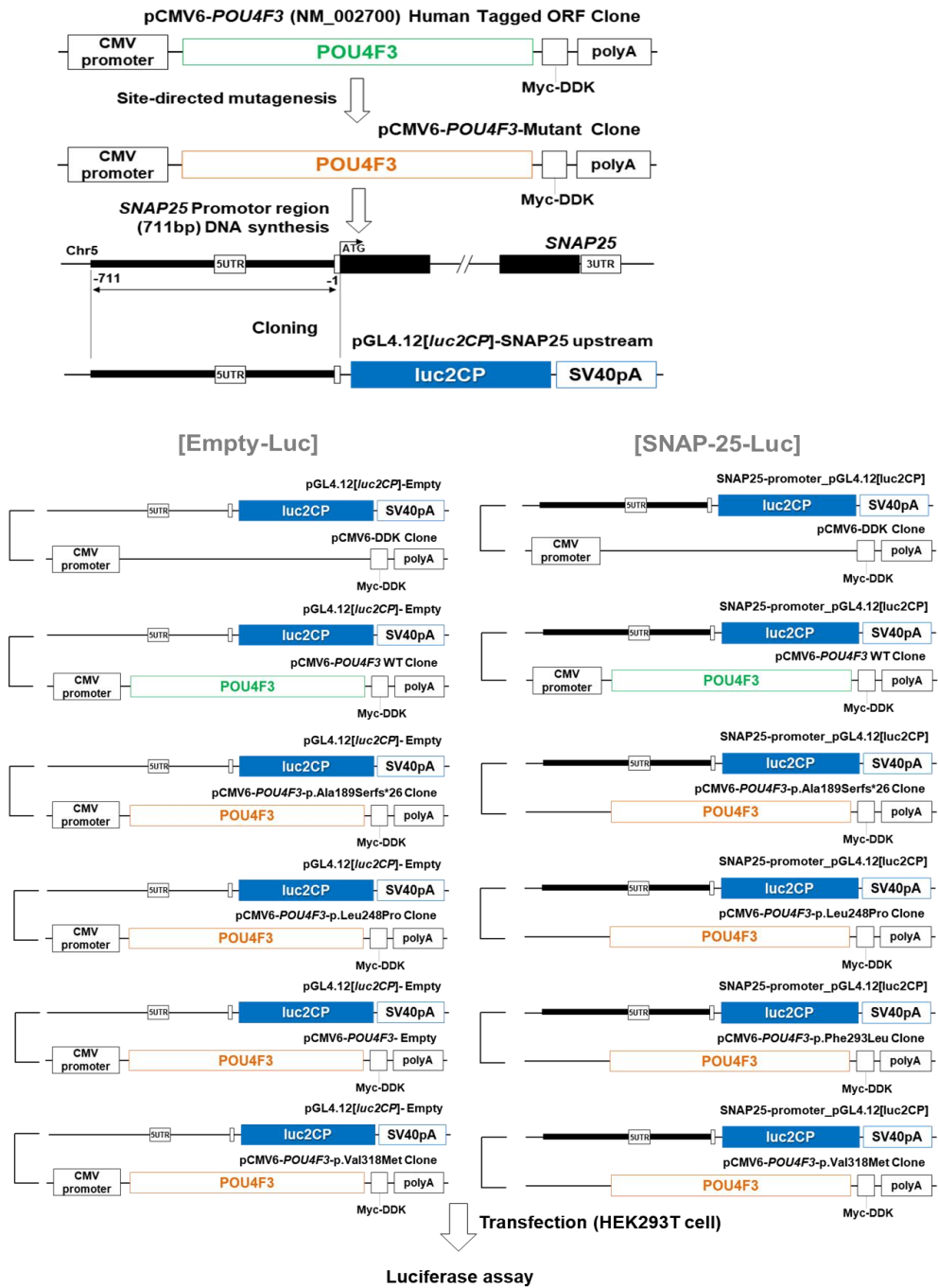
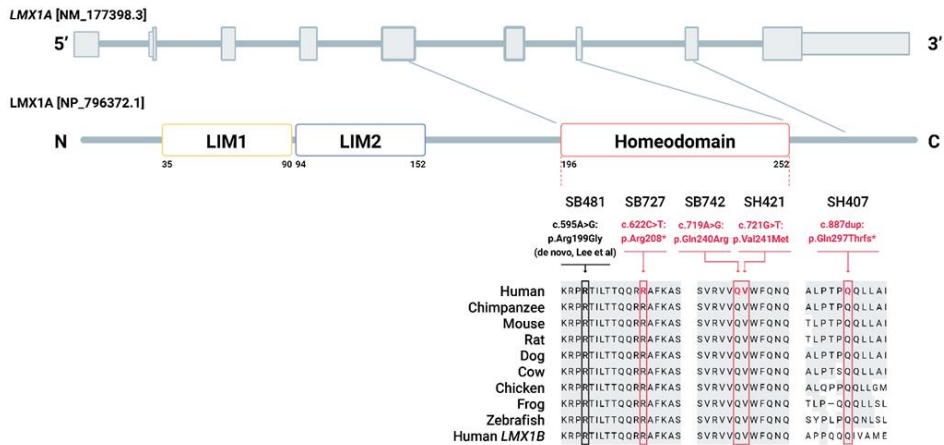Fig 1. A schematic figure of plasmid constructs for luciferase reporter assay.

Fig 2. The LMX1A protein contains two LIM domains (LIM1 and LIM2) and one homeodomain. The domains are depicted as in the Universal Protein Resource (UniProt) database. The four novel variants are depicted in the schematic of the LMX1A protein; three (p.Arg208*, p.Gln240Arg, and p.Val241Met) are located in the homeodomain and one (p.Gln297Thrfs*41) in the C-terminus. A de novo, heterozygous, missense *LMX1A* variant (c.595A>G:p.Arg199Gly) is located in the N-terminal arm encoding the homeodomain (as previously reported by Lee et al.), and is also depicted in the schematic of the LMX1A protein. Conservation of the affected residues among species and the human paralog *LMX1B* was documented for all *LMX1A* variants observed in the present study. The exons corresponding to homeodomain region and Gln297 residue are connected by lines, respectively.
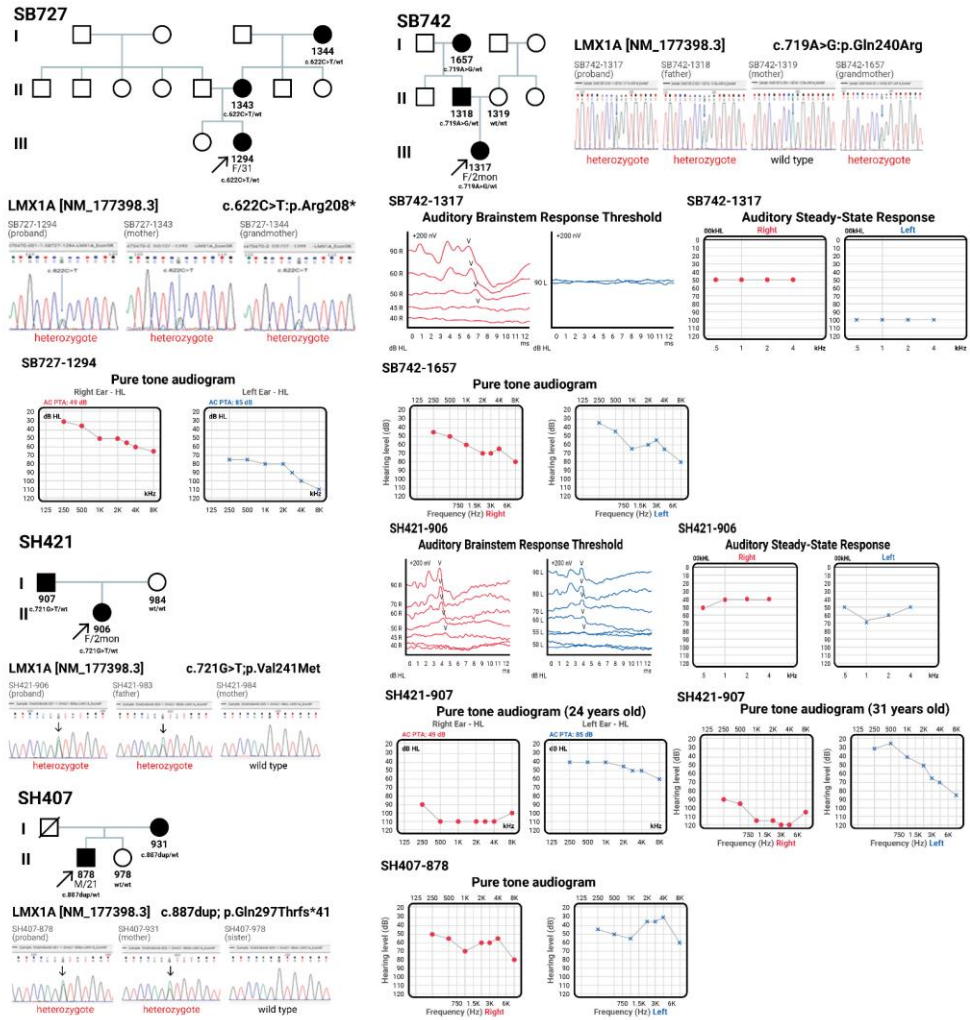
Fig 3. The pedigrees of the four families and Sanger sequence chromatograms of the respective *LMX1A* variants exhibiting segregation of c.622C>T:p.Arg208*, c.719A>G:p.Gln240Arg, c.721G>T:p.Val241Met, and c.887dup:p.Gln297Thrfs*41. Asymmetric hearing loss was identified in most affected individuals for whom audiological evaluations were possible.
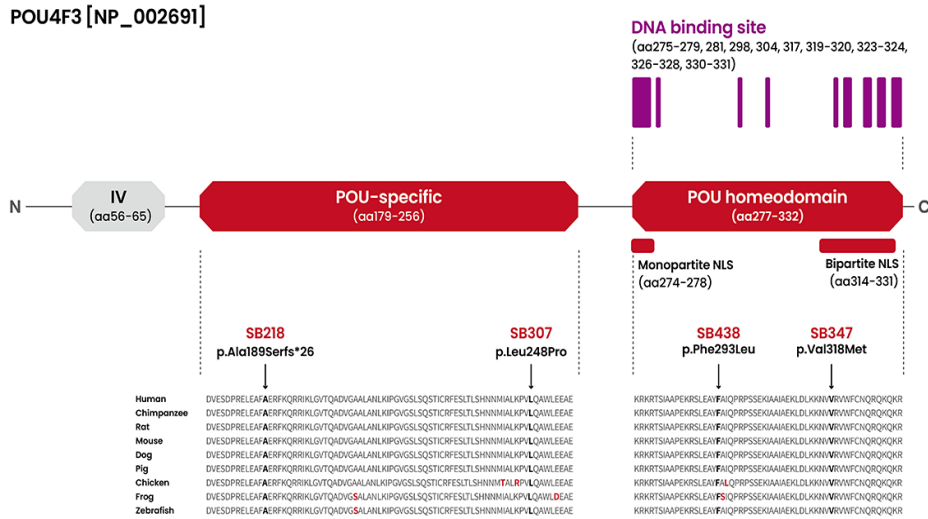
Fig 4. Four *POU4F3* novel variants within the functional DNA−binding domains. Two (p.Ala189Serfs*26 and p.Leu248Pro) were located in the POU−specific domain, while the remaining two (p.Phe293Leu and p.Val318Met) were in the POU−homeodomain. Conservation of the affected residues among species was documented for all *POU4F3* variants identified in the study.
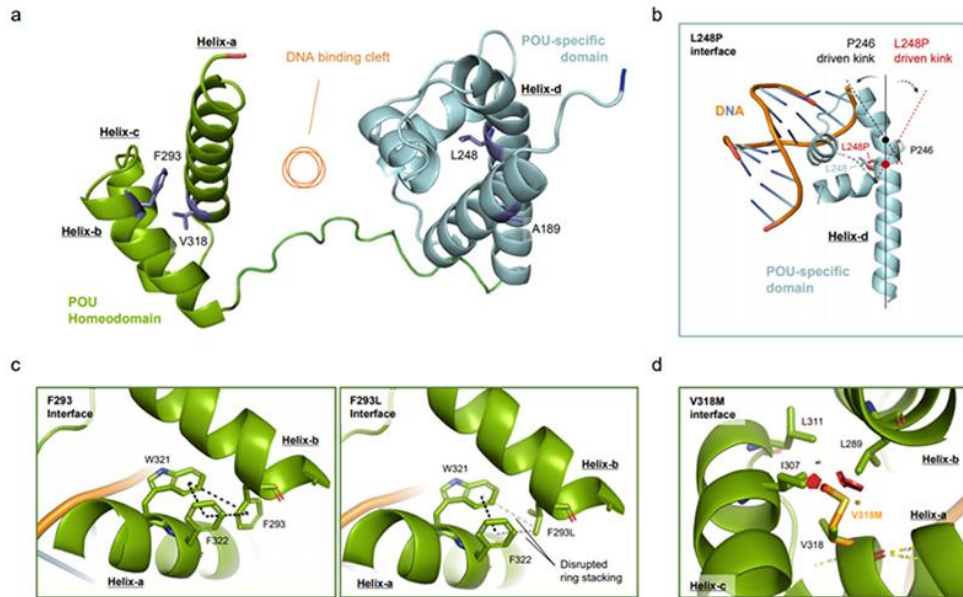
Fig 5. Novel *POU4F3* variants destabilize the inter-helical interactions, impairing the transcriptional activity of POU4F3. (a) Sideview of Alphafold generated model structure of POU4F3. (Jumper et al., Highly accurate protein structure prediction with AlphaFold., Nature (2021)). POU homeodomain (green) and POU-specific domain (cyan) assembled with DNA binding cleft (orange circle) in between. Val318 and Phe293 are present in the Helix-a and Helix-b of the homeodomain (green), respectively, while Leu248 is in the Helix-d of the POU-specific domain (cyan). All the mutant residues are facing intra-helical spaces, not directly interacting with DNA. (b) Intra-helical proline substitution at Leu248P causes helical kinks. A 27-amino acid long helix-d has a natural kink (black dotted line) driven by Pro246 in the middle. Additional proline substitution induces the formation of an additional kink (red dotted line) starting from Leu248Pro (red dot), causing dramatic conformational changes in the POU-specific domain. (c) Phe293 forms aromatic ring stacking (black dotted line) with

Trp321 and Phe322 in helix-a (left), stabilizing the interhelical interface. The Phe293Leu variant largely disrupts biochemical interactions between helix-a and helix-b, destabilizing helical assembly of POU-homeodomain. (d) Key amino acid residues of intramolecular hydrophobic cavity of POU homeodomain. Val318 forms hydrophobic interactions with Ile307, Leu289, and Leu311. The Val318Met mutant with long side chain clashes (red polygons), with the adjacent Ile307 and Leu289, changing the distance between helices. L, Leu; V, Val; P, Pro; F, Phe; W, Trp; I, Ile; M, Met.
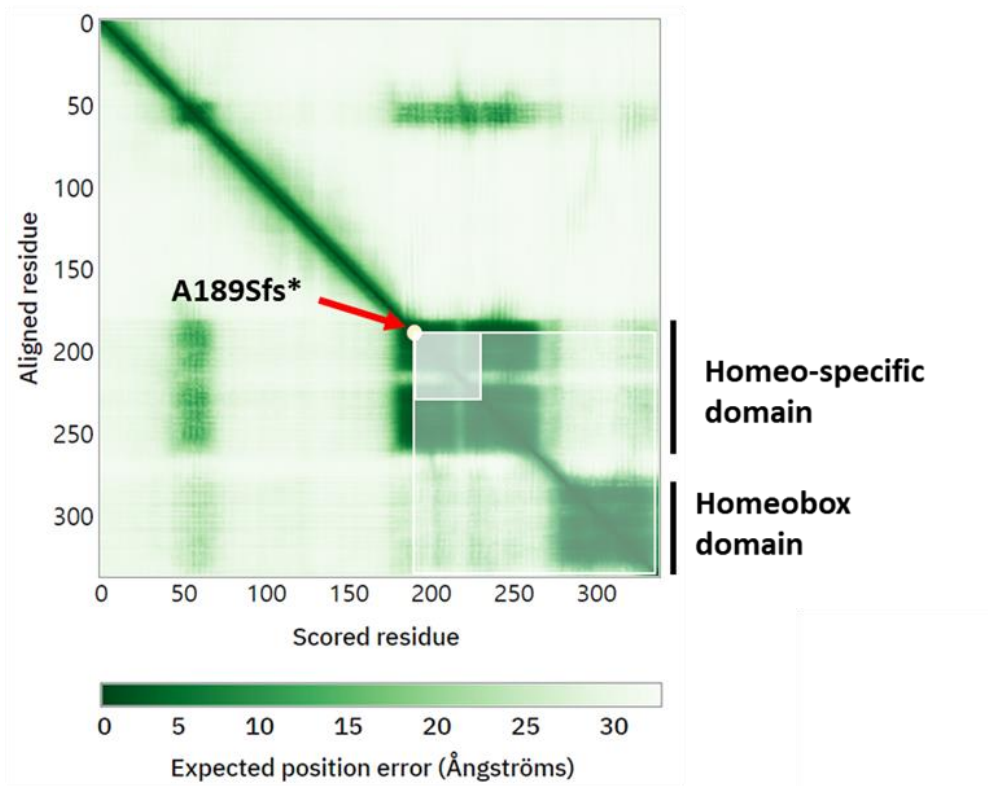
Fig 6.　The *POU4F3* Ala189Serfs*26 variant destabilized POU4F3 protein stability, as demonstrated by the predicted aligned error (PAE) score.
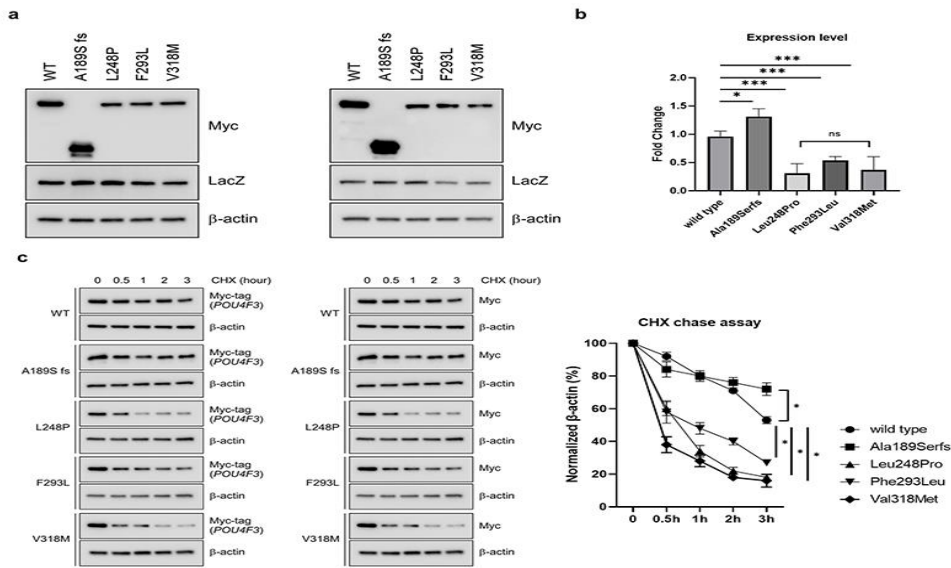
Fig 7. Western blot analysis for POU4F3 wild-type, frameshift, missense mutations by transient transfection at HEK293T cell. (a) Expressions of POU4F3 wild-type and mutants were detected by western blotting in HEK 293T cells. Molecular weight of wild-type and mutant proteins (p.Leu248Pro, p.Phe293Leu, and p.Val318Met) are 36kDa, whereas molecular weight of truncated mutant protein (p.Ala189Serfs*26) is 21kDa. The immunoblots are representative of independent repetitive experiments. LacZ is used as a transfection control. (b) The bands intensity was quantified by Image J. The band intensity was normalized to β-actin. Intensity data was presented as means ± standard deviations from two independent plots in a triplicate manner. (c) Comparison of the stability of wild-type and mutant POU4F3 using protein stability assays in the transient overexpression system. HEK 293 cells, overexpressing POU4F3, were treated with cycloheximide (80 μg/ml) for up to 3h to block the general translation. LacZ, transfection control; β-actin, loading control; ns, no statistical significance; *p <0.05, ***p < 0.001, ANOVA with Bonferroni comparisons.
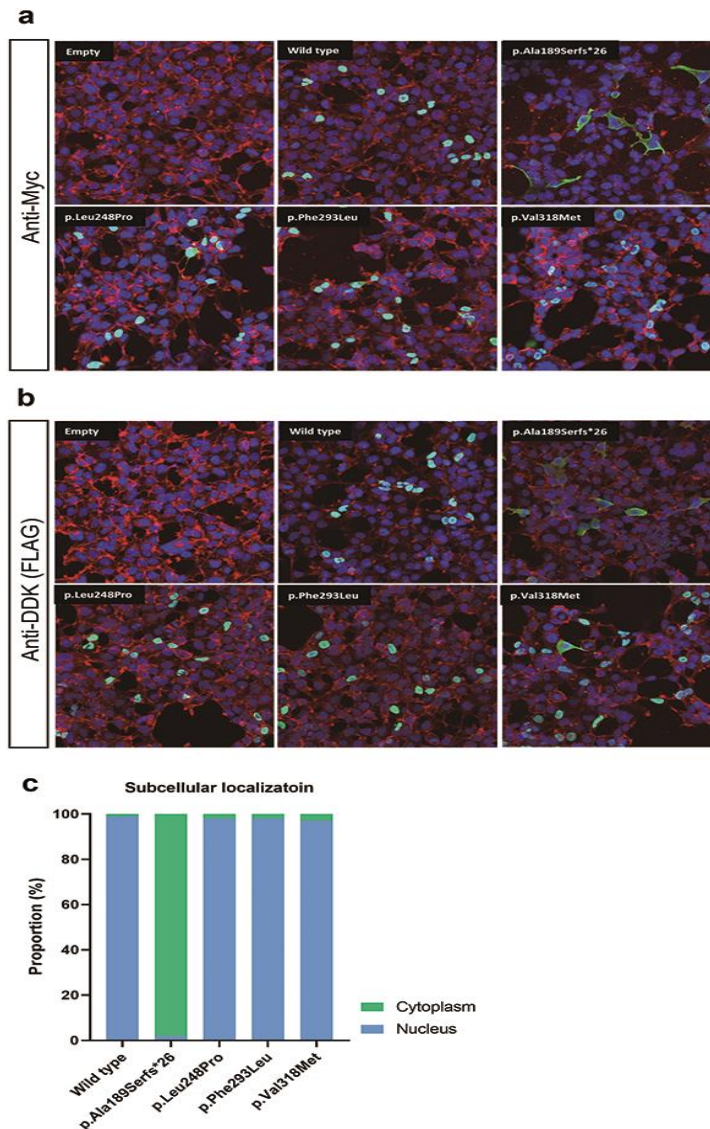
Fig 8. Immunofluorescence of the wild-type and mutant POU4F3 proteins. (a) Cells were immuno-stained with anti-Myc (green) and phalloidin (red). The nuclei were stained with DAPI (blue). (b) Cells were immuno-stained with anti-DDK (green) and Rhodamine-phalloidin (red). Rhodamine-phalloidin (red) staining was used to label F-actin and stabilize actin filaments in vitro. The nuclei were stained with DAPI (blue). (c) Quantitation of cytoplasmic and nuclear localization of POU4F3, depending on the variants.
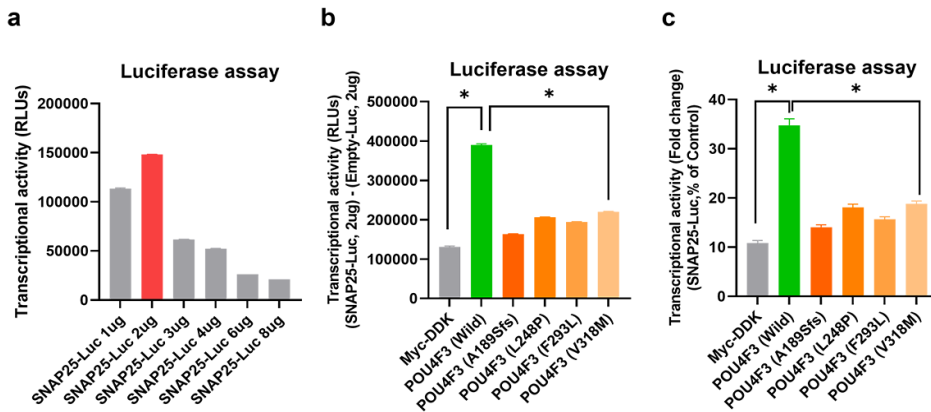
Fig 9. Transcriptional activity of novel *POU4F3* variants. (a) Luciferase activities measured under specific conditions. To minimize the ceiling effect, the condition (i.e., empty-Luc 2 μg and SNAP25-Luc 2 μg) was determined as the luciferase vector system. (b, c) The transcriptional activities of the wild-type and mutant POU4F3 proteins in the SNAP25-Luc vector were normalized to that of the internal control (Myc-DDK). Using the luciferase vector system, the transcriptional activity in the wild-type and the four *POU4F3* variants were analyzed. All variants exhibited significantly reduced transcriptional activities compared to the wild type. ∗ a statistical significance
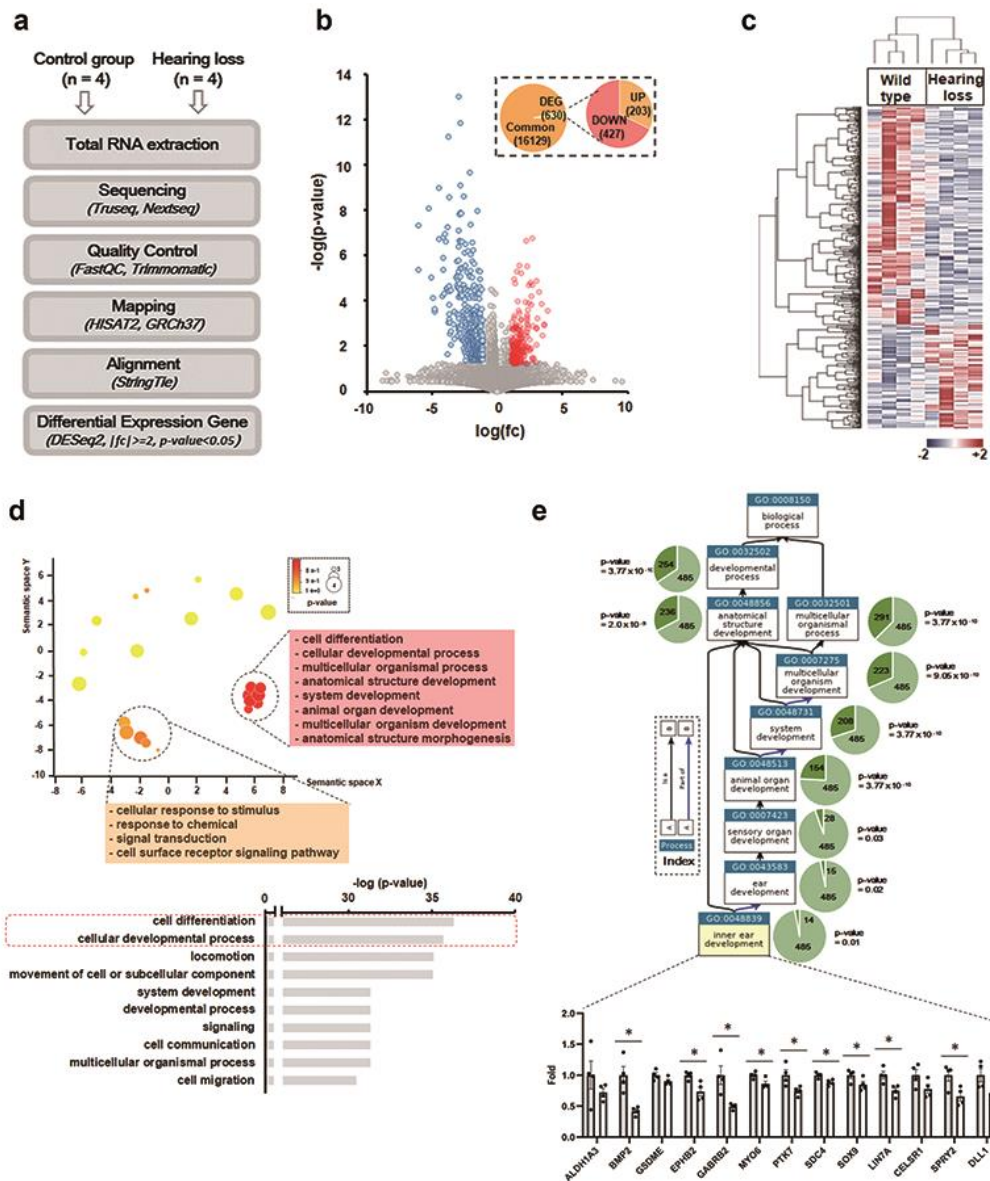
Fig 10. RNA sequencing analysis. (a) Schematic diagram of the analysis flow. (b) Volcano plot of significantly different genes (n = 630). Upregulation (red dot) and downregulation (blue dot) gene numbers were summarized as a pie graph (inlet). (c) Heatmap analyses of differential gene expression. The higher expression level was shown as red color while the lower expression was shown as blue. (d) (Upper) Revigo visualization of the top 30 gene

ontology (GO) data. Clustered terms were listed in each box. (Bottom) Top 10 GO terms in biological process. The red dot box showed top2 GO terms, including cell differentiation and cellular developmental process. (e) Ancestor chart view of the QuickGO. In each GO term, enriched gene number was shown in the dark green pie while term size was shown as green pie with the p-value. The colored arrow showed the relationship between the Ancestor term and the Child term. Fourteen genes which belong to the GO-term (GO: 0048839) were significantly enriched (p-value = 0.01) to confirm their dysregulation (bottom).

a
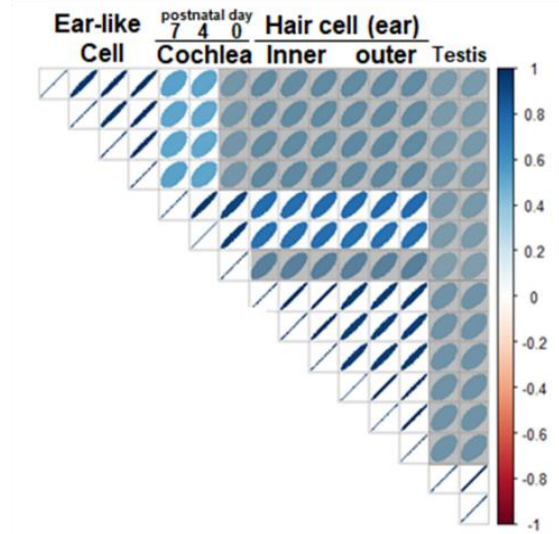
b

coefficient	h-earlike				Cochlea_p_7day	Cochlea_p_4day	Cochlea_p_0day	inner hair cell (adult)			outer hair cell (adult)			testis (adult)	
h-earlike	1.000	0.957	0.970	0.970	0.531	0.521	0.498	0.581	0.578	0.581	0.581	0.581	0.579	0.469	0.468
	0.957	1.000	0.969	0.971	0.529	0.519	0.496	0.586	0.582	0.587	0.585	0.586	0.584	0.466	0.466
	0.970	0.969	1.000	0.977	0.529	0.517	0.491	0.589	0.585	0.590	0.590	0.591	0.588	0.473	0.472
	0.970	0.971	0.977	1.000	0.535	0.524	0.498	0.595	0.592	0.596	0.595	0.595	0.593	0.477	0.476
Cochlea_p_7day	0.531	0.529	0.529	0.535	1.000	0.959	0.912	0.762	0.757	0.766	0.766	0.766	0.764	0.478	0.476
Cochlea_p_4day	0.521	0.519	0.517	0.524	0.959	1.000	0.945	0.747	0.742	0.751	0.747	0.748	0.745	0.469	0.467
Cochlea_p0_day	0.498	0.496	0.491	0.498	0.912	0.945	1.000	0.679	0.673	0.682	0.676	0.676	0.674	0.440	0.437
inner hair cell (adult)	0.581	0.586	0.589	0.595	0.762	0.747	0.679	1.000	0.984	0.992	0.946	0.945	0.944	0.519	0.519
	0.578	0.582	0.585	0.592	0.757	0.742	0.673	0.984	1.000	0.985	0.939	0.938	0.937	0.516	0.516
	0.581	0.587	0.590	0.596	0.766	0.751	0.682	0.992	0.985	1.000	0.947	0.947	0.945	0.519	0.519
outer hair cell (adult)	0.581	0.585	0.590	0.595	0.766	0.747	0.676	0.946	0.939	0.947	1.000	0.993	0.991	0.517	0.517
	0.581	0.586	0.591	0.595	0.766	0.748	0.676	0.945	0.938	0.947	0.993	1.000	0.992	0.517	0.516
	0.579	0.584	0.588	0.593	0.764	0.745	0.674	0.944	0.937	0.945	0.991	0.992	1.000	0.516	0.515
testis (adult)	0.469	0.466	0.473	0.477	0.478	0.469	0.440	0.519	0.516	0.519	0.517	0.517	0.516	1.000	0.997
	0.468	0.466	0.472	0.476	0.476	0.467	0.437	0.519	0.516	0.519	0.517	0.516	0.515	0.997	1.000

c

	h-earlike				Cochlea_p_7day	Cochlea_p_4day	Cochlea_p_0day	inner hair cell (adult)			outer hair cell (adult)			testis (adult)	
h-earlike	0	1.47E-15	4.85E-17	5.40E-17	0.041	0.041	0.059	0.150	0.151	0.148	0.153	0.154	0.152	0.115	0.115
	1.47E-15	0	3.41E-17	1.44E-17	0.042	0.042	0.059	0.167	0.169	0.164	0.170	0.171	0.169	0.110	0.110
	4.85E-17	3.41E-17	0	3.63E-19	0.038	0.037	0.052	0.167	0.169	0.164	0.170	0.172	0.170	0.119	0.119
	5.40E-17	1.44E-17	3.63E-19	0	0.040	0.039	0.054	0.171	0.172	0.168	0.174	0.175	0.173	0.115	0.115
Cochlea_p_7day	0.041	0.042	0.038	0.040	0	6.04E-14	5.03E-09	0.018	0.019	0.017	0.017	0.017	0.018	0.056	0.055
Cochlea_p_4day	0.041	0.042	0.037	0.039	6.04E-14	0	6.07E-11	0.031	0.032	0.029	0.031	0.031	0.031	0.059	0.058
Cochlea_p_0day	0.059	0.059	0.052	0.054	5.03E-09	6.07E-11	0	0.107	0.110	0.102	0.108	0.108	0.109	0.061	0.060
inner hair cell (adult)	0.150	0.167	0.167	0.171	0.018	0.031	0.107	0	1.39E-20	3.15E-24	1.81E-11	2.04E-11	2.23E-11	0.138	0.141
	0.151	0.169	0.169	0.172	0.019	0.032	0.110	1.39E-20	0	1.48E-20	3.75E-11	4.17E-11	4.50E-11	0.139	0.142
	0.148	0.164	0.164	0.168	0.017	0.029	0.102	3.15E-24	1.48E-20	0	1.58E-11	1.77E-11	1.95E-11	0.135	0.138
outer hair cell (adult)	0.153	0.170	0.170	0.174	0.017	0.031	0.108	1.81E-11	3.75E-11	1.58E-11	0	3.06E-25	7.34E-24	0.135	0.138
	0.154	0.171	0.172	0.175	0.017	0.031	0.108	2.04E-11	4.17E-11	1.77E-11	3.06E-25	0	2.61E-24	0.134	0.136
	0.152	0.169	0.170	0.173	0.018	0.031	0.109	2.23E-11	4.50E-11	1.95E-11	7.34E-24	2.61E-24	0	0.136	0.139
testis (adult)	0.115	0.110	0.119	0.115	0.056	0.059	0.061	0.138	0.139	0.135	0.135	0.134	0.136	0	1.03E-28
	0.115	0.110	0.119	0.115	0.055	0.058	0.060	0.141	0.142	0.138	0.138	0.136	0.139	1.03E-28	0

Fig 11. Correlation analyses of transcriptome between patient–derived cell lines and mouse models' RNA–sequencing data. (a) Spearman's correlation coefficient with statistical test results was visualized. A stronger positive correlation coefficient showed a linear shape with blue color. Shading indicates p–values greater than 0.05. (b) Spearman's correlation coefficients were shown as a table with higher values as red to lower as white. H–earlike sample is patient–derived transcriptome. Cochlea_p_7day = postnatal 7–day Cochlea; Cochlea_p_4day = postnatal 4–day Cochlea; Cochlea_p_0day = postnatal 0–day Cochlea. (c) p–value table from Spearman's correlation measurement. The red color showed a p–value lower than 0.05.
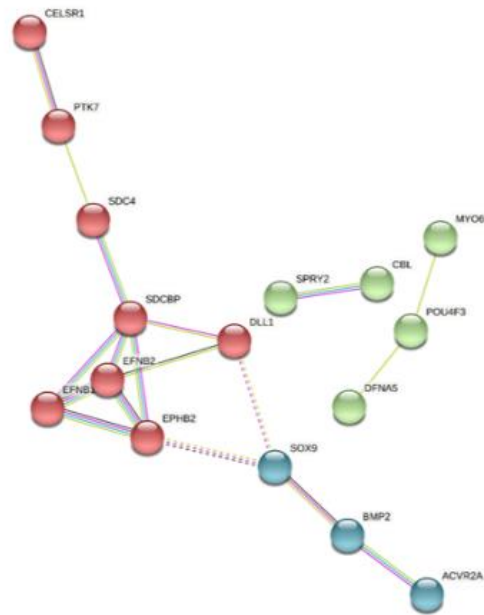
Fig 12. Functional and physical association of 14 enriched genes with *POU4F3*. Fourteen enriched genes and *POU4F3* were clustered in 3 groups. The dotted line indicates the edge of the cluster. Three parallel lines showed interaction evidence. The thick line means strength of data support. Each colored line showed interactions. Sky blue line is known interaction with curated databases. The purple one is experimentally determined known interactions. Predicted interactions showed as green (gene neighborhood), red (gene fusion), and blue (gene co-occurrence). Another interaction is shown as either textmining (yellow-green) or co-expression (black).
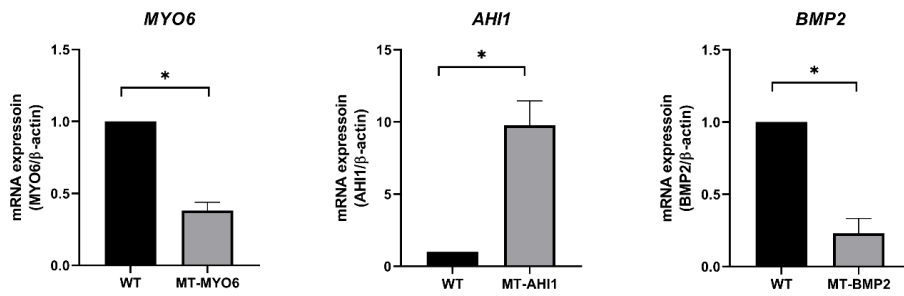
Fig 13. Validation of transcriptome analysis for three target genes (*MYO6*, *AHI1*, *BMP2*) associated with inner ear development using RT−qPCR.
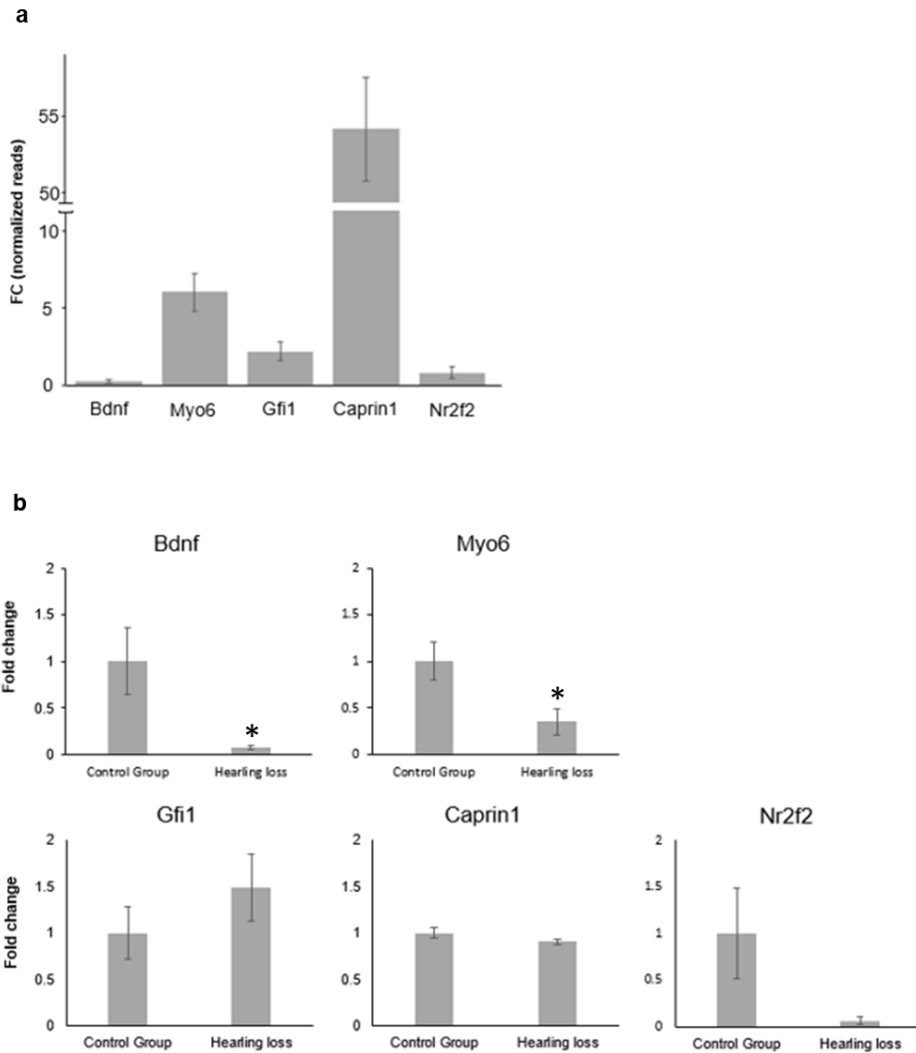
Fig 14. The expression level of known target genes. (a) Normalized reads showed expressional fold changes ranging from 1 to 55. (b) *Bdnf* and *Myo6* showed statistically significant repression in the hearing-loss group (∗; p-value <0.05). *Gfi1*, *Caprin 1*, and *Nr2f2* showed their expression, but dysregulation was not significant.
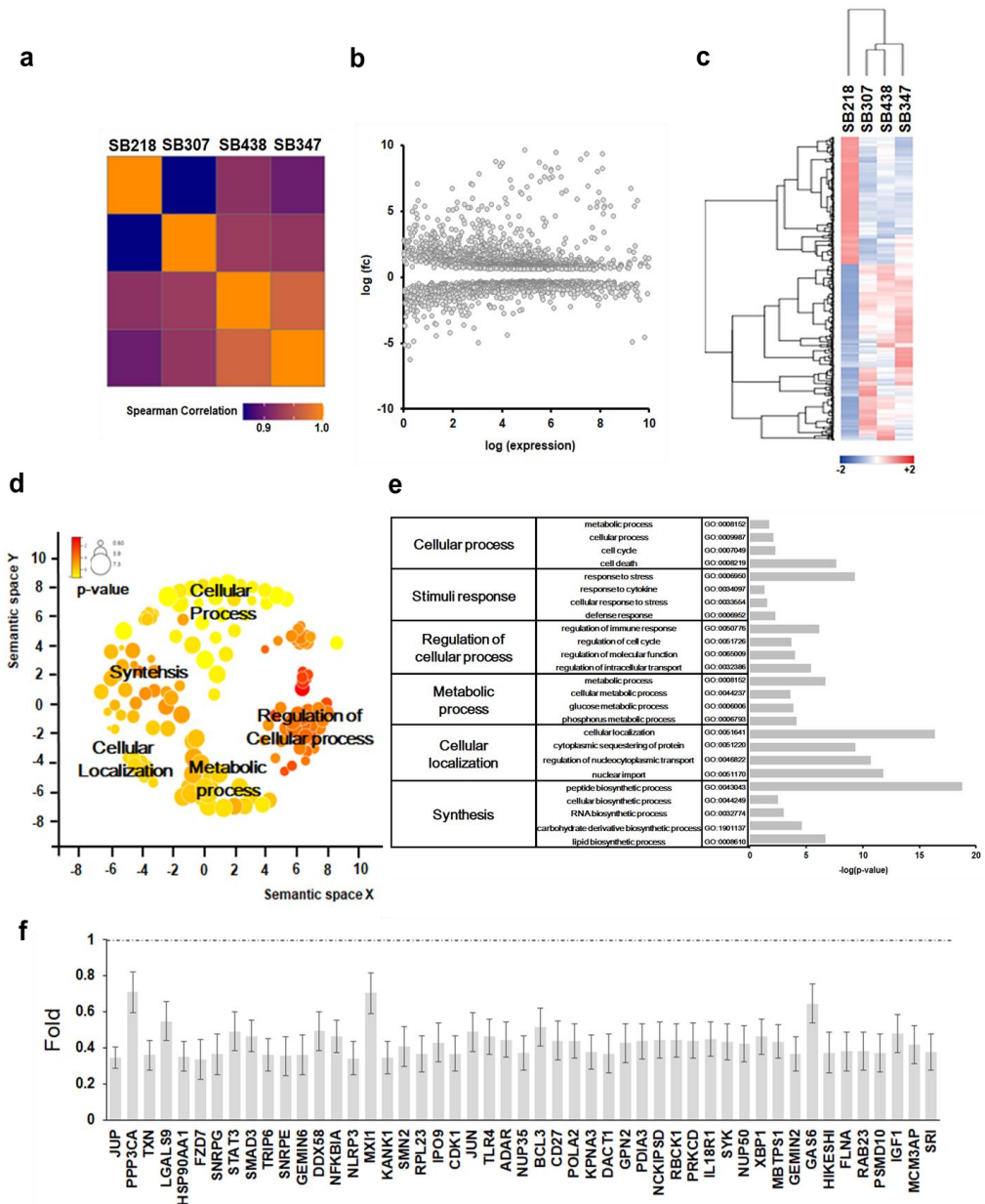
Fig 15. Transcriptome profiles between four *POU4F3* variants (a) Spearman's correlation coefficient was shown as a heatmap. The yellowest is the highest expression. (b) MA−plot of the transcriptome of four patients. The x−axis indicates an individual's expression level, while the y−axis shows fold change. (c) Heatmap of the four patients' transcriptomes. The red color indicates higher expression. (d) Gene ontology (GO) analyses the listed genes

9 1

under the Revigo visualization to show representative categories, including Cellular process, Synthesis, Cellular localization, Metabolic process, and regulation of the cellular process. (e) Representative lists of the significantly enriched GO terms in the biological process. (f) Gene expression fold−change among patients visualized as a bar graph in the nuclear import cluster.

요약 (국문초록)

감각신경성 난청은 가장 흔한 감각질환으로, 소아 감각신경성 난청 환자의 절반 이상에서 유전성 난청으로 진단된다. 난청을 일으키는 150여 개 이상의 유전자가 발견되었고, 최근에는 난청 유전자의 내이 내 병인 메커니즘과 시간적 공간적 발현정보를 기반으로 한 유전성 난청의 기능적 분류가 제시되었다. 난청 유전자의 기능적 분류 중 한 부분인 전사인자는 일반적으로 특정 DNA 서열을 인식하며 특정 유전자의 전사와 발현을 조절한다. 인간에서 1600개 이상의 전사인자가 알려져 있고, 전사인자 유전자의 돌연변이 발생은 다양한 질병을 야기한다. 하지만, 난청과 관련된 전사인자 유전자의 임상 표현형, 유전형 및 분자유전학적 메커니즘에 대해서는 잘 알려져 있지 않다. 본 연구에서는 1,280 유전성 난청 가계의 데이터베이스를 이용하여 유전형 및 임상 표현형을 분석을 시행하였다. 약 2.6%에서 전사인자 유전자의 돌연변이에 의한 비증후군성 난청으로 진단되었다. 특이적으로, 네 개의 전사인자 유전자 (*POU3F4, POU4F3, LMX1A, EYA4*)의 돌연변이가 비증후군성 난청을 유발하였고 이는 비증후군성 난청과 관련된 전사인자 유전자의 스펙트럼은 매우 국한되어 있음을 제시한다. 전사인자에 의한 기능적 분류에 따른 유전성 난청의 유전형과 표현형 지도를 제작한 결과, 그동안 잘 알려져 있지 않은 전사인자 유전자 *LMX1A*의 돌연변이는 비대칭성 난청 표현형과 유의한 연관성이 있음을 확인하였다. 마지막으로, 성인의 유전성 난청의 대표적 전사인자 유전자인 *POU4F3*의 신규 돌연변이를 확인하였고 기능 분석에 따른 다양한 분자유전학적 메커니즘을 제시하였다. 또한, 환자 세포주를 이용한 전사체 분석을 통해 *POU4F3*의 돌연변이는 내이 발달과 관련된 14개의 표적 유전자의 발현을 조절함을 확인하였다. 종합하면, 본 연구를 통해 비증후군성 난청을 유발하는 전사인자 유전자 지도를 제시하였고 최초로 *LMX1A* 전사인자 유전자와 비대칭성 난청의 연관성에 대해 보고하였으며 기능분석을 통해 전사활성에 영향을 미치는 *POU4F3* 돌연변이의 다양한 병인 메커니즘 및 내이 발달과 관련된 표

적 유전자를 확인하였다. 난청 유전자의 기능적 분류를 통해 향후 기존 연구에서 정의되지 않은 새로운 유전형−표현형 상관성을 규명하고 유전성 난청을 위한 표적 약물 및 유전자 치료 개발에 도움을 줄 수 있을 것으로 기대된다.

주요어 : 유전성 난청, 전사인자 유전자, *LMX1A*, *POU4F3*
학   번 : 2019−34627