



An overload of missense variants in the *OTOG* gene may drive a higher prevalence of familial Meniere disease in the European population

Alberto M. Parra-Perez^{1,2,3} · Alvaro Gallego-Martinez^{1,2,3} · Jose A. Lopez-Escamez^{1,2,3}

Received: 13 September 2023 / Accepted: 14 January 2024 / Published online: 22 March 2024
© The Author(s) 2024

Abstract

Meniere disease is a complex inner ear disorder with significant familial aggregation. A differential prevalence of familial MD (FMD) has been reported, being 9–10% in Europeans compared to 6% in East Asians. A broad genetic heterogeneity in FMD has been described, *OTOG* being the most common mutated gene, with a compound heterozygous recessive inheritance. We hypothesize that an *OTOG*-related founder effect may explain the higher prevalence of FMD in the European population. Therefore, the present study aimed to compare the allele frequency (AF) and distribution of *OTOG* rare variants across different populations. For this purpose, the coding regions with high constraint (low density of rare variants) were retrieved in the *OTOG* coding sequence in Non-Finnish European (NFE).. Missense variants (AF < 0.01) were selected from a 100 FMD patient cohort, and their population AF was annotated using gnomAD v2.1. A linkage analysis was performed, and odds ratios were calculated to compare AF between NFE and other populations. Thirteen rare missense variants were observed in 13 FMD patients, with 2 variants (rs61978648 and rs61736002) shared by 5 individuals and another variant (rs117315845) shared by two individuals. The results confirm the observed enrichment of *OTOG* rare missense variants in FMD. Furthermore, eight variants were enriched in the NFE population, and six of them were in constrained regions. Structural modeling predicts five missense variants that could alter the otogelin stability. We conclude that several variants reported in FMD are in constraint regions, and they may have a founder effect and explain the burden of FMD in the European population.

Keywords Meniere's disease · Hearing loss · Exome sequencing · *OTOG* gene · Population genetics

Introduction

Meniere disease (MD [OMIM 156000]) is a chronic inner ear disorder characterized by recurrent episodes of vertigo, low-to-medium-frequency sensorineural hearing loss (SNHL), and tinnitus or aural fullness (Lopez-Escamez et al. 2015). The condition is associated with an

accumulation of endolymph in the cochlear duct termed “endolymphatic hydrops” (Semaan et al. 2005), and epidemiological studies have reported an increased prevalence of several comorbidities, such as migraine, respiratory allergic disorders (rhinitis, asthma), and autoimmune/autoinflammatory diseases (Gazquez et al. 2011; Kim et al. 2022a, b; Perez-Carpena and Lopez-Escamez 2019; Radtke et al. 2002).

Meniere disease is more commonly observed in the European population, and familial aggregation has been reported in 9–10% of cases in the European and 6% in the East Asian population (Requena et al. 2014; Lee et al. 2015). There are multiple genes associated with MD (Parra-Perez and Lopez-Escamez 2023), including several SNHL genes, showing autosomal dominant or recessive inheritance (Roman-Naranjo et al. 2020) and digenic inheritance (Roman-Naranjo et al. 2021), supporting genetic heterogeneity in the disease.

The most common gene in familial MD (FMD) is *OTOG* (MIM 604,487), with several Spanish families sharing the

✉ Jose A. Lopez-Escamez
jose.lopezescamez@sydney.edu.au

¹ Division of Otolaryngology, Department of Surgery, Instituto de Investigación Biosanitaria, Ibs.GRANADA, Universidad de Granada, Granada, Spain

² Sensorineural Pathology Programme, Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER), Madrid, Spain

³ Faculty of Medicine and Health, School of Medical Sciences, Meniere's Disease Neuroscience Research Program, The Kolling Institute, The University of Sydney, 10 Westbourne St, Sydney, NSW, Australia

same variants with a compound heterozygous recessive inheritance (Roman-Naranjo et al. 2020). Exome sequencing (ES) studies have revealed that a few rare missense variants in the *OTOG* gene may be behind the etiology of a high percentage of FMD cases. Fifteen out of forty-six (33%) Spanish-unrelated MD families with at least one rare missense variant in the *OTOG* gene have been reported. The *OTOG* gene encodes for otogelin, a secreted 2975 amino-acid glycoprotein required for the anchorage of the otolithic and tectorial membranes to the hair cell stereocilia in the sensory epithelia of the vestibule (saccul and utricle) and the organ of Corti. It is also involved in the organization and stabilization of the tectorial membrane structure in the organ of Corti (Avan et al. 2019; Cohen-Salmon et al. 1997).

Previous studies have shown that several non-syndromic SNHL genes have a founder effect in the Asian population (48 variants in 14 genes, 85.7%). However, there are few reported genes showing founder variants in the European population [9 variants in *GJB2* [OMIM 121,011], *TMC1* [OMIM 606,706], and *TMIE* [OMIM 607,237] genes] (Abo-ayge et al. 2023).

We hypothesize that a founder effect may be related to *OTOG* in FMD, and this burden of rare variants will explain the higher prevalence of FMD in the European population. Therefore, this study aims to compare the frequency and distribution of rare variants in the coding region of *OTOG* across different populations to determine whether there are variants that explain the burden of FMD reported in the European population.

Methods

OTOG missense rare variants dataset in gnomAD

Exome data from the gnomAD database v.2.1 (Karczewski et al. 2020) were used to retrieve *OTOG* rare missense variants (allelic frequency [AF] < 0.01) for the non-Finnish European (NFE, $N = 56,885$), African/African American (AFR, $N = 8128$), East Asian (EAS, $N = 9197$), South Asian (SAS, $N = 15,308$), Latino/Admixed American (AMR, $N = 17,296$), and global populations ($N = 125,748$).

OTOG missense rare variants dataset in familial MD

We retrieved *OTOG* rare variants (AF gnomAD NFE < 0.01), from ES data previously reported in FMD (Roman-Naranjo et al. 2020). This dataset generated from 100 unrelated FMD patients includes single-nucleotide variants (SNVs) aligned with the reference genome GRCh38/hg38. FMD patients were diagnosed according to the diagnostic criteria established by the International Classification Committee for

Vestibular Disorders of the Barany Society (Lopez-Escamez et al. 2015).

To determine whether *OTOG* shows an overload of rare variants in different populations, we selected 64 genes with similar coding sequence (CDS) length to *OTOG* (8778 ± 439 bp) (CCDS76390), which were used as controls (Table S1).

Variant annotation and ranking

Variants in *OTOG* were assessed using the combined annotation-dependent depletion (CADD) score and SpliceAI score that predicts the impact on splicing processes (Jagannathan et al. 2019). The standards and guidelines described by the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) were followed (Kim et al. 2022a, b). In addition, selected variants were confirmed in the BAM files to avoid potential false positive calls, using Integrative Genomics Viewer (IGV) (Robinson et al. 2011).

The AF for each variant in the different populations was annotated based on the frequencies described in the gnomAD database v.2.1 (Karczewski et al. 2020). In addition, a GBA was performed using the *OTOG* FMD variants dataset with an AF < 0.01, as previously described (Roman-Naranjo et al. 2020).

Statistical analysis

The aggregated AF for *OTOG* and control genes was calculated for each population and compared with the AF in the FMD cohort using odds ratios (OR) with a 95% confidence interval (CI). Likewise, frequency comparison was performed using OR between the frequency of variants in the NFE population and all other populations. Variants with a p value < 0.05 and $OR \geq 1$ were considered enriched in a specific population.

Linkage analysis

The complete list of *OTOG* common variants (AF > 0.05) was downloaded from the gnomAD database v.2.1 (Karczewski et al. 2020) to calculate the linkage disequilibrium among all known variants in *OTOG*, including variants found in the FMD dataset. The R^2 score was obtained and represented for each pair of variants, using the LDmatrix and LDheatmap function from the LDlinkR (Myers et al. 2020) and LDheatmap (Shin et al. 2006) R packages, respectively. Population genotype data used in LDlinkR were obtained from Phase 3 (Version 5) of the 1000 Genomes Project.

Estimation of missense variant density in the coding sequence

The density of variants in the CDS was calculated to identify regions with an overload of missense variants in the *OTOG* gene. For this, a strategy based on sliding windows was used to calculate the ratio between the window's variants and length. The chosen window length was 100 bp on each side of the evaluated position, and the step between windows was 1 bp, which resulted in a window of 201 bp length. The missense variants used were those described in the gnomAD v.2.1 database (Karczewski et al. 2020) for each population (NFE, AFR, EAS, SAS, and AMR). The high-density threshold was estimated according to the number of missense variants expected in each population according to gnomAD database v.2.1 (Karczewski et al. 2020), calculating it as a direct proportion of the number of expected variants in the overall population. If the computed density in each window was below the predicted density for a given population, that window was considered a low-density region (LDR) or constrained region.

Protein modeling

The canonical otogelin amino acid sequence (NP_001264198.1) was retrieved from RefSeq, and the structural model was predicted using AlphaFold2. The quality of the protein structural models was assessed using several structure validation algorithms, such as Molprobity (Williams et al. 2018), ERRAT (Colovos and Yeates 1993), ProSA-web (Wiederstein and Sippl 2007), QMEANDisCo (Studer et al. 2020), and DeepUMQA (Guo et al. 2022). Molprobity calculates an overall score, a weighted logarithmic combination of geometric scores such as clashscore, percentage of unfavored Ramachandran, and bad sidechain rotamers. Lower values indicate better quality. ERRAT analyzes the interactions between atoms and provides an overall quality factor. ProSA-web uses a *z*-score to measure the energy separation between the native fold and the average of an ensemble of the misfolds in standard deviation units of the protein database. In this case, the length of the structure (2925 aa) is outside the experimental length range used in ProSA-web in its protein database; hence, the *z*-score comparison obtained may be limited. QMEANDisCo evaluates the agreement of pairwise distances between residues and distance constraints from homologous structures. DeepUMQA calculates, per residue, the superposition-free score IDDT that assesses the local distance differences of all atoms in a protein structure. The IDDT result is shown as a mean. Higher scores indicate higher quality models.

The *in silico* model was used to predict the protein stability change ($\Delta\Delta G$) caused by the candidate variants, using the DynaMut2 (Rodrigues et al. 2021), mCSM (Pires et al.

2014), and PremPS (Chen et al. 2020) tools. Variants were classified as neutral when $-0.5 < \Delta\Delta G_{\text{pred}} < 0.5$ (Pancotti et al. 2022). In addition, the pathogenicity prediction of AlphaMissense was calculated for each variant (Cheng et al. 2023).

Results

Overload of *OTOG* variants in familial Meniere disease

Thirteen missense variants with an AF gnomAD NFE < 0.01 were found in 13 Spanish individuals with FMD (Table 1). We confirmed an overload of missense variants in *OTOG* in FMD cases against the NFE population (OR = 3.40 [2.10–5.49], FDR = 5.36E–03) and global population from gnomAD (OR = 3.30 [2.05–5.32], FDR = 9.45E–03).

Comparison of *OTOG* allelic frequencies across different reference populations

The AF of the 13 rare variants observed in the CDS of the FMD cohort was compared between the NFE, AFR, EAS, SAS, and AMR populations (Fig. 1A). Eight rare missense variants were identified to be significantly overrepresented (OR > 1 , *p* value < 0.05) in the NFE population compared with the AFR, EAS, SAS, and AMR populations (Table 2). By contrast, five rare missense variants were significantly overrepresented in the AFR (NC_000011.10:g.17557227G > A; NC_000011.10:g.17638480C > A and NC_000011.10:g.17640936C > T), SAS/EAS (NC_000011.10:g.17641065G > A), and AMR (NC_000011.10:g.17611374C > T) populations, when the allelic frequencies were compared with the NFE population (Table 3). Furthermore, it was observed that most variants found in the *OTOG* gene (10/12) were not present in the East Asian population.

To determine whether *OTOG* showed an overload of missense rare variants, we selected 64 genes with similar lengths in the CDS to *OTOG*. We ranked the genes according to the number of variants for each reference population (NFE, AFR, EAS, SAS, and AMR), and *OTOG* showed a different burden of rare missense variants in the CDS compared to genes with the same CDS in the different populations (Fig. 2). In the NFE population, the number of variants described was in the 22nd percentile, while for the EAS, SAS, and AMR populations, it was above the 55th percentile.

Table 1 Rare missense variants found in the *OTOG* gene for familial Meniere disease (AF < 0.01)

Variant	ID	AA change	CADD	SpliceAI	ACMG	Density region (1)	AF FMD	AF gnomAD NFE	AF gnomAD global
NC_000011.10:g.17553211G>A	rs552304627	NP_001264198.1:p.(Val141Met)	34.0	0.15	Pathogenic (PVS1, PS4, PM2, PP3, BP1)	LDR	0.005 (1/200)	1.34E-03	8.80E-04
NC_000011.10:g.17557227G>A	rs61978648	NP_001264198.1:p.(Val269Ile)	16.59	0	Uncertain significance (PS4, PP1, BP1, BP4)	HDR	0.025 (5/200)	4.43E-03	8.36E-03
NC_000011.10:g.17599671C>T	rs117005078	NP_001264198.1:p.(Pro1240Leu)	33.0	0	Likely pathogenic (PS4, PM2, PP3, BP1)	LDR	0.005 (1/200)	5.55E-03	2.73E-03
NC_000011.10:g.17606001G>A	rs145689709	NP_001264198.1:p.(Arg1353Gln)	23.6	0	Uncertain significance (PS4, PM2, BP1, BP4, BP6)	HDR	0.005 (1/200)	4.08E-03	2.42E-03
NC_000011.10:g.17610645T>C	rs61744602	NP_001264198.1:p.(Leu1794Pro)	7.32	0	Benign (BS1, BS2, BP1, BP6, BP4)	LDR	0.005 (1/200)	3.59E-03	1.63E-03
NC_000011.10:g.17611118C>T	rs748280789	NP_001264198.1:p.(His1952Tyr)	9.44	0	Uncertain significance (PS4, PM2, BP1, BP3, BP4)	LDR	0.005 (1/200)	3.75E-05	1.36E-05
NC_000011.10:g.17611374C>T	rs61736002	NP_001264198.1:p.(Ala2037Val)	8.51	0.07	Uncertain significance (PS4, PM2, BP1, BP4)	LDR	0.025 (5/200)	1.24E-03	2.47E-03
NC_000011.10:g.17612217G>A	rs188527711	NP_001264198.1:p.(Arg2072His)	27.5	0.04	Likely benign (PM2, BP1, BP4)	LDR	0.005 (1/200)	1.60E-03	7.98E-04
NC_000011.10:g.17635125G>A	rs76461792	NP_001264198.1:p.(Arg2556Gln)	23.0	0.02	Benign (PS4, BS1, BS2, BP1, BP4, BP6)	LDR	0.005 (1/200)	4.89E-03	2.76E-03
NC_000011.10:g.17638480C>A	rs61995750	NP_001264198.1:p.(Gln2621Lys)	14.66	0	Benign (BS1, BS2, BP1, BP4, BP6)	HDR	0.005 (1/200)	2.01E-04	9.84E-04
NC_000011.10:g.17640936C>T	rs567966154	NP_001264198.1:p.(Arg2691Cys)	24.7	0	Benign (BS1, BS2, BP1, BP4, BP6)	LDR	0.005 (1/200)	2.40E-04	8.95E-04
NC_000011.10:g.17641065G>A	rs1310923563	NP_001264198.1:p.(Val2734Met)	23.7	0	Uncertain significance (PS4, PM2, BP1, BP4)	HDR	0.005 (1/200)	0	1.39E-05
NC_000011.10:g.17642200G>A	rs117315845	NP_001264198.1:p.(Arg2802His)	16.7	0.03	Uncertain significance (PS4, PM2, BP1, BP4, BP6)	HDR	0.01 (2/200)	2.89E-03	2.61E-03

The notation high-density region (HDR) and low-density region (LDR) indicates the position of the variant in Fig. 4
AA amino acid, ACMG American College of Medical Genetics and Genomics, CADD combined annotation-dependent depletion, FMD Familial Meniere disease, AF allelic frequency, NFE Non-Finnish European

Linkage analysis

The linkage analysis showed that the *OTOG* sequence has a low linkage disequilibrium in the CDS (Fig. 3). Only two variants [NC_000011.10:g.17642200G > A (rs117315845) and NC_000011.10:g.17553211G > A (rs552304627)] were in moderate linkage disequilibrium in the NFE population, with an R^2 of 0.332. These variants have yet to be described in the SAS and EAS populations. Both variants were found in the same MD family.

Two other rare variants were linked in the AFR and AMR populations [NC_000011.10:g.17640936C > T (rs567966154) and NC_000011.10:g.17638480C > A (rs61995750)], with an R^2 of 0.832 and 1, respectively, but not in the NFE, SAS, and EAS populations, since they were not described in the 1000G data. These two variants were found in the same family.

Variant density analysis in the *OTOG* coding sequence

The distribution of LDR in the *OTOG* CDS was similar when compared across different populations (NFE, AFR, SAS, EAS, and AMR) (Fig. 4). The similarity percentage between the regions of the NFE population compared to SAS, AMR, EAS, and AFR was 73.78%, 66.37%, 65.35%, and 60.18%, respectively.

Five variants (NC_000011.10:g.17557227G > A, NC_000011.10:g.17606001G > A, NC_000011.10:g.17638480C > A, NC_000011.10:g.17641065G > A, and NC_000011.10:g.17642200G > A) were found in high-density regions in the *OTOG* CDS. On the other hand, eight rare variants were identified in constrained regions, including NC_000011.10:g.17611374C > T, shared among four FMD patients.

Protein modeling

The otogelin protein model used to evaluate the impact of SNVs on protein stability was obtained by AlphaFold2 modeling (Fig. 1B). According to the geometric validation results (Table S2), a trustworthy model has been obtained compared to experimentally solved structures at the geometric level.

The otogelin model is predicted to have a globular structure and a tail, with secondary structure mainly in β -sheets, which ends with a C-terminal cystine knot (CTCK) domain. In the globular region, there are four Von Willebrand factor type D (VWFD) domains, a trypsin inhibitor-like (TIL) domain, and an epidermal growth factor-like (EGF-like) domain, surrounded by several disordered regions (Fig. 1).

The variants NP_001264198.1:p.(Val141Met), NP_001264198.1:p.(His1952Tyr), NP_001264198.1:p.(Ala2037Val), NP_001264198.1:p.(Arg2072His), and NP_001264198.1:p.(Arg2802His) have been predicted in silico to change the otogelin global stability by at least two different methods (Table S3). In this context, the NP_001264198.1:p.(Val141Met), NP_001264198.1:p.(Ala2037Val), NP_001264198.1:p.(Arg2072His), and NP_001264198.1:p.(Arg2802His) variants have been classified as destabilizing. In contrast, the NP_001264198.1:p.(His1952Tyr) variant is classified as stabilizing (Figure S1).

Nevertheless, the remaining eight variants found in FMD patients are classified as neutral according to the predicted perturbation of protein stability.

Discussion

Our study compares the AF and distribution of *OTOG* rare variants across different populations to assess if the reported FMD variants may have an *OTOG*-related founder effect that would explain the higher prevalence of FMD in NFE population.

MD has a significant familial aggregation in the NFE population, and most reported families with MD have a European ancestry (Requena et al. 2014). Epidemiological studies estimate that the prevalence of FMD is 5–23.5% (Hietikko et al. 2013a, b). In the European population, an overall percentage of 9–15% has been calculated, 9–13% in Spain (Frejo et al. 2017; Requena et al. 2014), 12% in Sweden (Birgersson et al. 1987), and 9.3% in Finland (32.7% for relatives with Meniere-like symptoms) (Hietikko et al. 2013a, b), while in the Asian population, the estimated proportion is approximately 6% in Japan (5.8%) (Mizukoshi et al. 1979) and South Korea (6.3% for relatives with definite FMD and 9.8% FMD-like syndrome) (Lee et al. 2015). The familial aggregation data are likely to be overestimated due to the use of questionnaires and patient interviews and the difficulty of collecting a large cohort of FMD (Morrison et al. 2009), combined with the challenge of MD diagnosis and the different diagnostic criteria in studies published before 1995.

The reported differences in the FMD prevalence may be due to differences in the genetic structure of each population. More frequent and founder variants in hearing impairment genes have already been described in European, African, Asian, and American populations (Aboagye et al. 2023; Adadey et al. 2022); therefore, finding founder variants in specific genes that cause FMD would not be surprising. Founder variants are those variants found with a high frequency within a particular population caused by the presence of the variant in an ancestor or small group of ancestors (Jain et al. 2021).

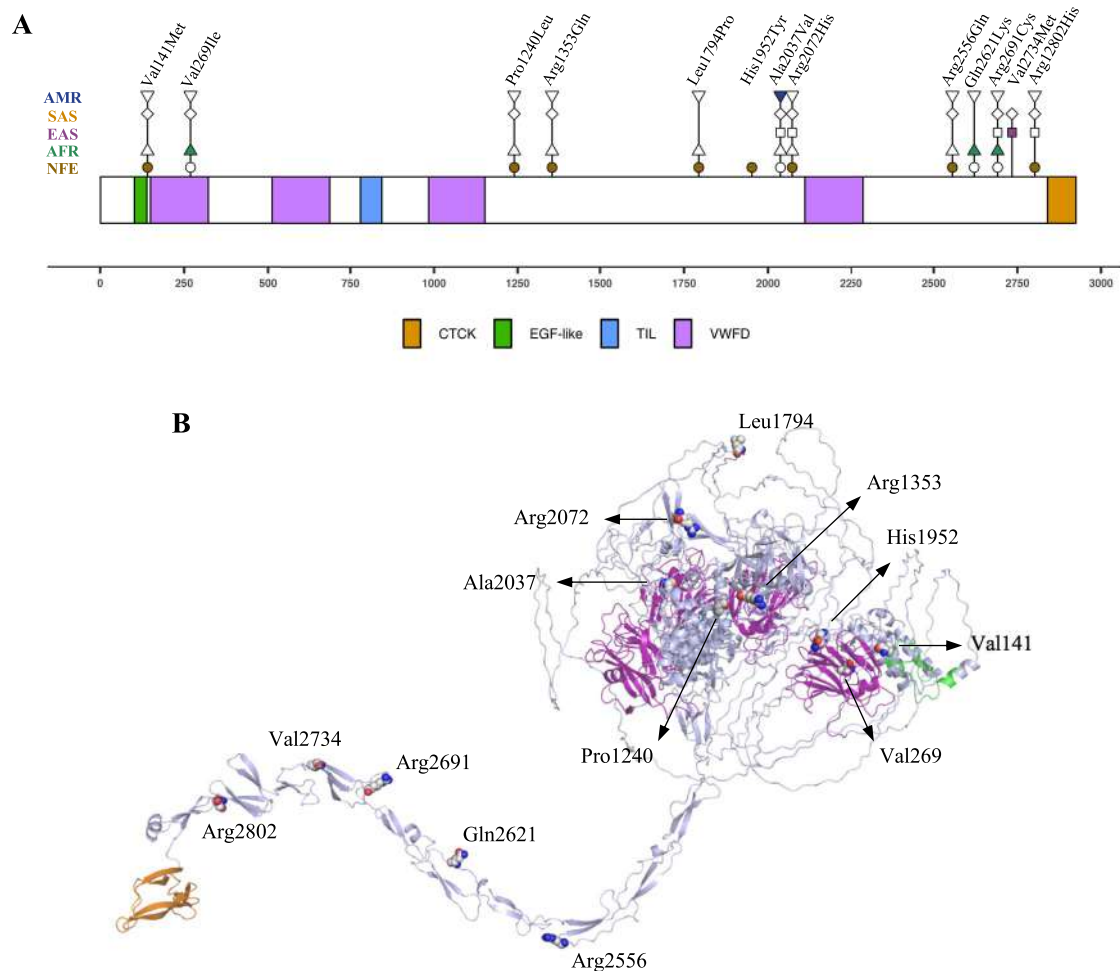


Fig. 1 **A** Representation of the otogelin protein sequence showing the rare variants ($AF < 0.01$) found in the FMD cohort. It is highlighted in which population (NFE, AFR, EAS, SAS, and AMR) they are most frequent. The circle, triangle, square, diamond, and inverted triangle indicate the presence of the variant in the NFE, AFR, EAS, SAS, and AMR populations, respectively. The colored symbol means that that variant is more frequent in that population. **B** Otogelin 3D model outlining the positions of the residues where variants were found. In

addition, the different domains that constitute the protein are shown in color. The domains have been colored using the Uniprot domain annotations (Q6ZRI0). CTCK: C-terminal cystine knot-like domain (orange). EGF-like: epidermal growth factor-like domain (green). TIL: trypsin inhibitor-like cysteine-rich domain (blue). VWFD: von Willebrand factor C-like domain (purple). The NP_001264198.1 sequence has been used as a reference to annotate protein positions

Among the known FMD genes, *OTOG* seems one of the most relevant, with several Spanish families sharing the same variants with a compound heterozygous recessive inheritance pattern (Roman-Naranjo et al. 2020). *OTOG* is expressed in both vestibular and auditory sensory organs, and *Otog* knockout mice exhibit both auditory and vestibular phenotypes (Avan et al. 2019), supporting the hypothesis that variants in *OTOG* may contribute to

MD development. Furthermore, the relative expression of *OTOG* is higher in the apex than at the base of the cochlea, which may be significant given that hearing loss in MD patients is initially observed at low and medium frequencies (El-Amraoui et al. 2001). For this reason, the *OTOG* gene was chosen as the best candidate to study the allelic frequency of its variants across different populations to test whether *OTOG* could explain the higher prevalence

Table 2 Rare missense FMD *OTOG* variants enriched in the NFE population over the AFR, EAS, SAS, or AMR populations

Variant	AF		AFR		EAS		SAS		AMR	
	NFE	AFR	AFR	SAS	EAS	P value	OR (95% CI)	P value	OR (95% CI)	P value
NC_000011.10:g.17553211G>A	1.34E-03	1.48E-04	9.07	8.88E-05	0	4.26E-03	(1.57E-362.482)	-	15.11	(4.03E-127.382)
NC_000011.10:g.17599671C>T	5.55E-03	8.81E-04	6.33	7.10E-04	0	1.64E-09	(2.87E-17.383)	-	7.86	(4.75E-13.923)
NC_000011.10:g.17606001G>A	4.08E-03	4.43E-04	9.24	4.88E-04	0	2.56E-08	(3.11E-45.202)	-	8.39	(4.59E-17.053)
NC_000011.10:g.17610645T>C	3.59E-03	1.47E-04	24.41	0	0	4.08E-09	(4.33E-965.765)	-	-	-
NC_000011.10:g.17611118C>T	3.75E-05	0	0	0	0	-	-	-	-	-
NC_000011.10:g.17612217G>A	1.60E-03	1.47E-04	10.89	2.22E-04	3.71E-04	9.20E-04	(1.90E-434.159)	6.14E-04	7.22	(2.98E-22.799)
NC_000011.10:g.17635125G>A	4.89E-03	1.56E-04	31.53	3.21E-03	0	9.72E-12	(5.60E-1244.4)	-	1.52	(1.16E-2.011)
NC_000011.10:g.17642200G>A	2.89E-03	0	9.28E-05	2.62E-03	0	-	-	2.03E-11	1.10	(0.81E-1.514)

AFR African/African American, AF allelic frequency, AMR Latino/Admixed American, CI confidence interval, EAS East Asian, NFE non-Finnish European, OR odds ratios, SAS South Asian

Table 3 Rare missense FMD *OTOG* variants enriched in the AFR, SAS, EAS, or AMR populations over the NFE population

Variant	AF		AFR		EAS		SAS		AMR	
	NFE	AFR	AFR	SAS	EAS	P value	OR (95% CI)	P value	OR (95% CI)	P value
NC_000011.10:g.17557227G>A	4.43E-03	6.28E-02	0.07	1.42E-03	-	4.10E-237	(0.05E-0.078)	-	3.13	(2.15E-4.678)
NC_000011.10:g.17611374C>T	1.24E-03	4.45E-03	0.28	4.74E-05	13.25	1.06E-07	(0.17E-0.442)	1.13E-04	0.26	(0.18E-0.357)
NC_000011.10:g.17638480C>A	2.01E-04	1.56E-02	0.01	0	-	4.26E-88	(0.00E-0.024)	-	-	-
NC_000011.10:g.17640936C>T	2.40E-04	1.27E-02	0.02	8.89E-05	2.59	5.87E-68	(0.01E-0.034)	4.90E-01	2.70	(0.61E-24.671)
NC_000011.10:g.17641065G>A	0	0	0	9.30E-05	0	1	-	1.65E-01	0	(0-16.115)

AFR African/African American, AF allelic frequency, AMR Latino/Admixed American, CI confidence interval, EAS East Asian, NFE Non-Finnish European, OR odds ratios, SAS South Asian

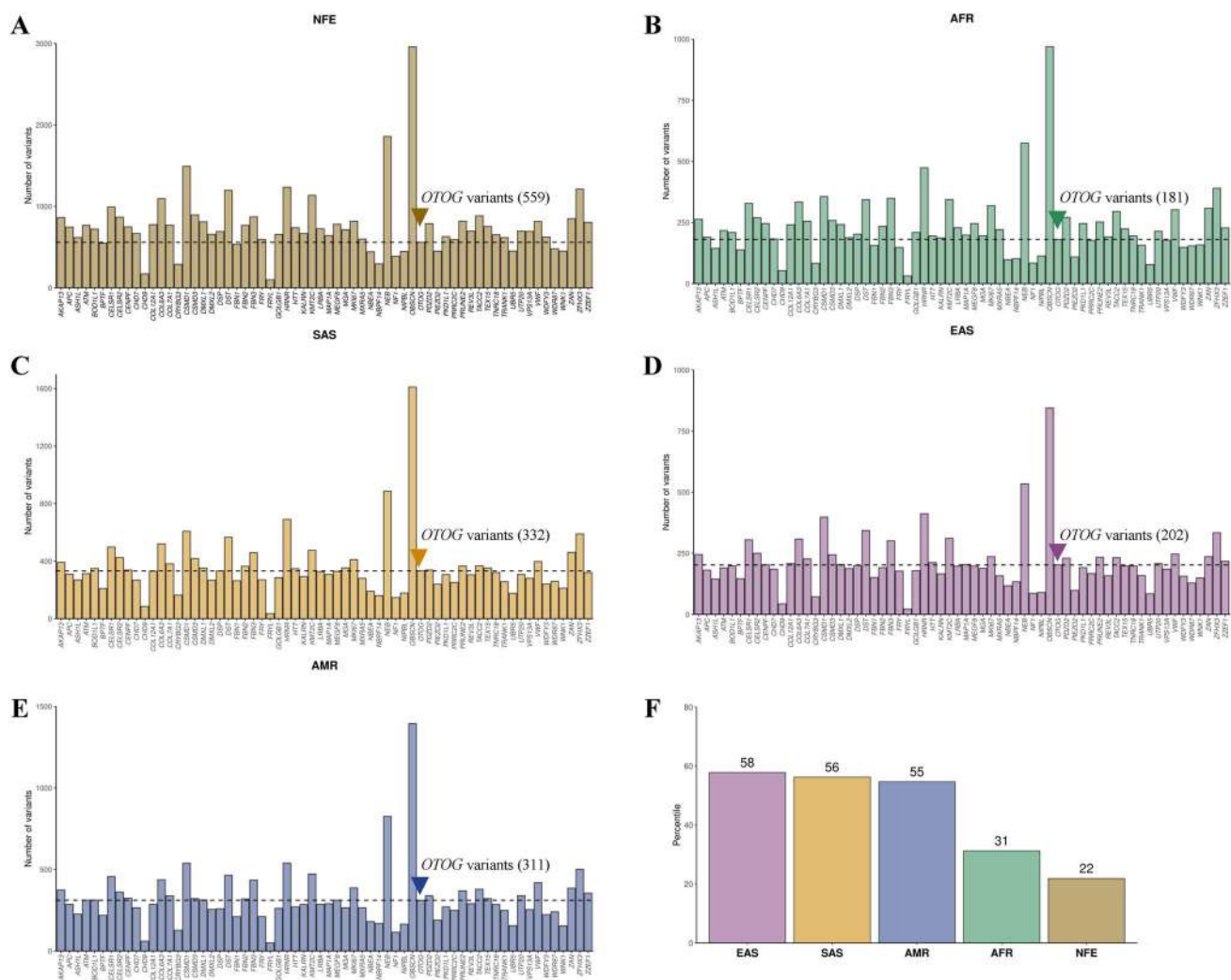


Fig. 2 Bar plot showing the number of rare missense variants found in *OTOG* and control genes for each reference population. **A** NFE, **B** AFR, **C** SAS, **D** EAS, **E** AMR, and **F** percentile of the number of

rare missense variants in the *OTOG* gene in the NFE, AFR, EAS, SAS, and AMR reference populations compared to the number of rare missense variants in control genes (AF < 0.01)

of FMD in the European population and whether there are variants with a founder effect.

Our study shows 13 missense variants in FMD and 8 of them are located in constrained regions in NFE: NC_000011.10:g.17553211G > A (rs552304627), NC_000011.10:g.17599671C > T (rs117005078), NC_000011.10:g.17610645 T > C (rs61744602), NC_000011.10:g.17611118C > T (rs748280789), NC_000011.10:g.17611374C > T (rs61736002), NC_000011.10:g.17612217G > A (rs188527711), NC_000011.10:g.17635125G > A (rs76461792), and NC_000011.10:g.17640936C > T (rs567966154) (Table 1).

Among all variants found, three rare variants are shared in different unrelated FMD patients. The variants NC_000011.10:g.17557227G > A and NC_000011.10:g.17611374C > T are in the same four

FMD patients (three heterozygous and one homozygous). Of note, these two variants are also overrepresented in the AFR and AMR populations, and this could be explained by the mixed ancestry of North African and American with the Spanish population. These variants, which are within the same haplotype and have been previously described in FMD patients, could impact the splicing processing and protein stability, increasing the susceptibility to develop a MD-like phenotype. The variant NC_000011.10:g.17557227G > A, although classified as VUS (variant of uncertain significance) by the ACMG criteria, is in the last position of exon 4. This could disrupt the consensus splicing site, generating possible non-functional alternative forms of the protein. Further functional studies would be required since SpliceAI did not predict a splicing defect. In addition, the variant NC_000011.10:g.17611374C > T has been associated in

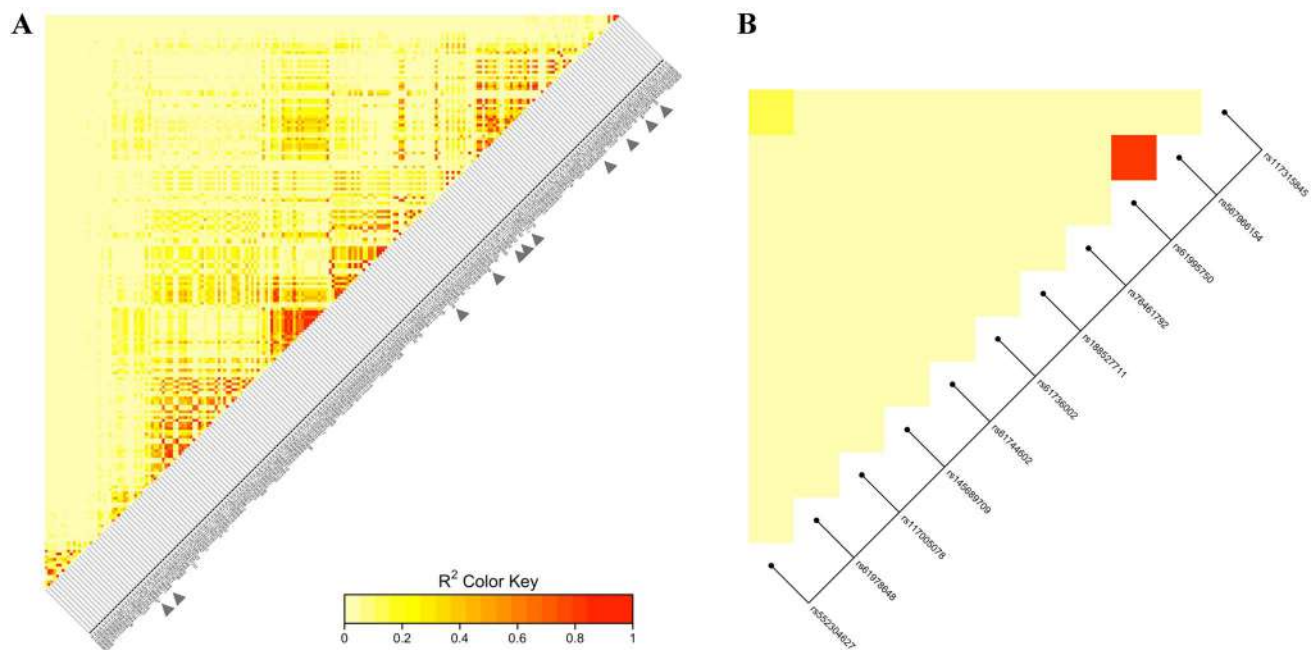


Fig. 3 **A** Heatmap showing the pairwise linkage disequilibrium of variants in the *OTOG* gene (MAF > 0.05 plus *OTOG* variants in the FMD cohort) across all populations. Variants with triangular label are the variants found in the FMD cohort. **B** Heatmap representing the

pairwise linkage disequilibrium of FMD variants in the *OTOG* gene. The two variants not shown in the plots are not annotated in Phase 3 (Version 5) of the 1000 Genomes Project

ClinVar (Landrum et al. 2018) with autosomal recessive non-syndromic hearing loss 18B (DFNB18B), supporting its possible pathogenic effect in FMD. None of the variants found in *OTOG* have been described in published families with DFNB18B (Ganaha et al. 2019).

On the other hand, the NC_000011.10:g.17642200G > A variant has been found in two FMD patients. This variant, classified as VUS and enriched in NFE population, is in the otogelin's tail, close to the CTCK domain, which is involved in the formation of antiparallel homodimers (Avan et al. 2019). Since this variant is predicted to be destabilizing, it could potentially interfere with the correct dimer formation and destabilize the connections between the tectorial and otolithic membranes and the hair bundles of hair cells, as observed in mice (Avan et al. 2019), increasing the susceptibility to MD. However, the CTCK-mediated dimer forms a highly resistant structure, with 11 pairs of cysteine residues forming disulphide bonds (Zhou and Springer 2014), so the effect of the NC_000011.10:g.17642200G > A variant on otogelin functionality is unclear.

The variants NC_000011.10:g.17638480C > A and NC_000011.10:g.17640936C > T, classified as benign and more frequent in the African population, were both found in the same FMD patient and in the same haplotype. Although these variants were found in the otogelin's tail, they were predicted to have no effect on the protein's stability, so it was

not possible to determine whether they will have a functional impact.

The rest of the variants described were each found in different single FMD patients. Since only the exonic variants have been explored, it cannot be excluded that these patients have a second variant in intronic or expression-regulating regions. Further, given the possibility that MD is a multifactorial, polygenic disease in which epigenetics also play a role (Flook et al. 2021), we cannot rule out that other variants and genes may increase the risk of MD, as described in Hui et al. (2023) for hearing loss.

According to the population frequency of the 13 variants, 8 have a higher frequency in the NFE population than in the rest of the populations, which may indicate a founder effect of these variants in the European population. Three arguments supports this hypothesis: (a) the pathogenicity CADD score > 20 (Niroula and Vihinen 2019), (b) the predicted change in protein stability, and (c) their occurrence in constrained regions or regions of low variant density, since constrained regions are enriched in pathogenic variants in ClinVar (Havrilla et al. 2019).

The finding of 8/13 enriched variants in the NFE population, could be because the FMD cohort mainly includes Spanish patients. For this reason, the high proportion of Spanish individuals in this cohort may explain the finding of three variants that were more frequent in the African than in the NFE population, since North African ancestry in the

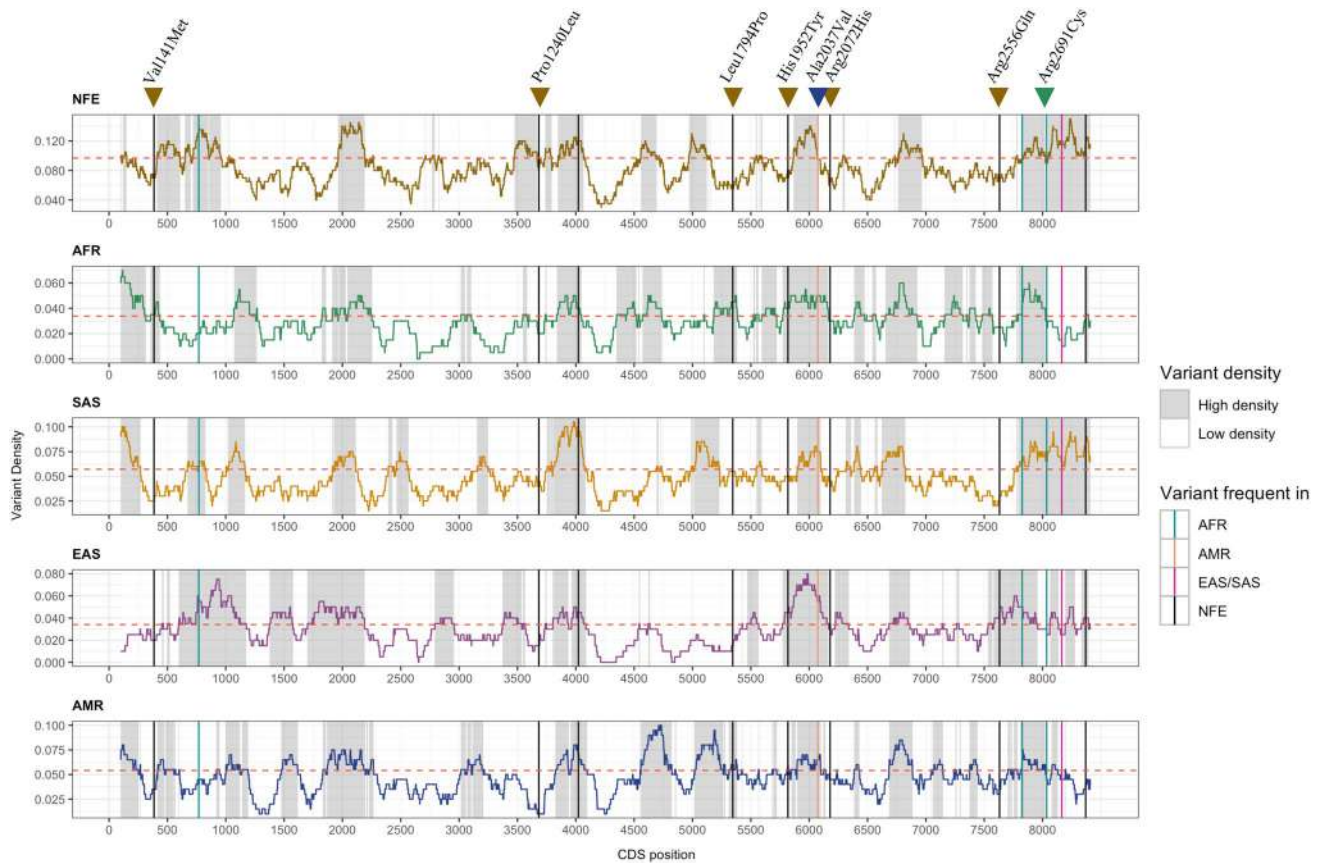


Fig. 4 Variant density profile along the *OTOG* CDS in the NFE, AFR, SAS, EAS, and AMR populations calculated with a 201 bp sliding window. High-density regions (HDR; gray areas) and low-density regions (LDR) are those with a high or low number of variants affecting the value of missense variants expected in each population analyzed according to gnomAD v2.1. The eight variants found in the FMD cohort in low-density regions in the NFE population are indicated with triangles. The triangles' colors (brown, blue, and

green) indicate more frequent variants in NFE, AMR, and AFR, respectively. The Red dashed line indicates the threshold density used for each population according to the number of variants expected in the CDS. Variants in the FMD cohort are represented with vertical lines in each plot. Black lines represent frequent variants in the NFE population, green lines in AFR, pink lines in EAS/SAS, and orange in AMR. The NP_001264198.1 sequence has been used as a reference to annotate protein positions

Spanish population may be up to 11% (Bycroft et al. 2019). Accordingly, it is possible that the prevalence of FMD could be similar in the North African and Spanish populations; thus, current epidemiological studies of MD and familial aggregation (FMD) are needed to support this hypothesis.

We also found that most of the variants found in *OTOG* in our FMD cohort are not found or have a very low frequency in the East Asian population (EAS). This may be due to the greater genetic divergence between both populations (Wang et al. 2018), but it could also anticipate a lower prevalence of *OTOG*-mediated FMD in East Asian population.

Some of these missense variants are in the protein's tail and may interfere with the interaction with other structural proteins or the dimers' formation; however, the physical interactions of otogelin with other tectorial or otolithic membrane proteins are not known.

Limitations

This study is based on exome sequencing datasets that found a burden of missense variants in the CDS in the *OTOG* gene in FMD (Parra-Perez and Lopez-Escamez 2023); however, most families remain undiagnosed and the discovery of variants in intronic or intergenic regulatory regions could change the current picture of the genetic architecture of FMD. Therefore, whole genome sequencing studies are needed. Besides, the lack of genetic studies of FMD in other populations (Escalera-Balsera et al. 2020) prevents estimating the genetic contribution of *OTOG* to FMD.

Conclusions

The *OTOG* gene has an overload of rare missense variants in the NFE population. Several *OTOG* variants observed in FMD are found in constrained regions and could have a founder effect in the NFE population. Further genomic studies on FMD in other populations are needed to know which genes may contribute to its development.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00439-024-02643-8>.

Acknowledgements Alberto M. Parra-Perez is enrolled in the Biomedicine Ph.D. program at Universidad de Granada, Spain and this work is part of his doctoral thesis. The authors want to thank all patients with MD for their enthusiastic participation, and clinicians of the Meniere Disease Consortium for their contribution to improve phenotyping of FMD.

Author contributions AMP-P and JAL-E conceived and designed the study. AMP-P and AG-M performed bioinformatics analyses. AMP-P and JAL-E drafted the manuscript. All authors contributed to the interpretation of the data and approved the final version of the manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Alberto M. Parra-Perez is a Ph.D. Student in the Biomedicine Program at Universidad de Granada, and his salary was supported by Andalusian University, Research and Innovation Department (Grant# PREDOC2021/00343). Alvaro Gallego-Martinez has been funded by the Andalusian University, Research and Innovation Department (Grant DOC_01677) and Andalusian Health Department (Grant PI-0266-2021 GEN4PHEN) and the funding of Cures Within Reach and the Knight family. Jose A Lopez-Escamez has received funds to support research on genetics in MD from The University of Sydney (K7013_B3413 Grant), Instituto de Salud Carlos III (Grant# PI20-1126), CIBERER (Grant# PIT21_GCV21), Andalusian University, Research and Innovation Department (2020 B-CTS-68-UGR20, M3NOMIC; PY20-00303, EPIMEN), Andalusian Health Department (Grant# PI027-2020), Asociación Síndrome de Meniere España (ASMES) and Meniere's Society, UK.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval This study was approved by the local ethics committee (MS/2014/02, Institutional Review Board for Clinical Research, Universidad de Granada, Spain) in accordance with the Helsinki declaration and its amendments. A written informed consent to donate biological samples was obtained from all subjects.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in

the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aboagye ET, Adadey SM, Wonkam-Tingang E, Amenga-Etego L, Awandare GA, Wonkam A (2023) Global distribution of founder variants associated with non-syndromic hearing impairment. *Genes* 14(2):399. <https://doi.org/10.3390/genes14020399>
- Adadey SM, Wonkam-Tingang E, Aboagye ET, Quaye O, Awandare GA, Wonkam A (2022) Hearing loss in Africa: current genetic profile. *Hum Genet* 141(3–4):505–517. <https://doi.org/10.1007/s00439-021-02376-y>
- Avan P, Le Gal S, Michel V, Dupont T, Hardelin J-P, Petit C, Verpy E (2019) Otogelin, otogelin-like, and stereocilin form links connecting outer hair cell stereocilia to each other and the tectorial membrane. *Proc Natl Acad Sci USA* 116(51):25948–25957. <https://doi.org/10.1073/pnas.1902781116>
- Birgerson L, Gustavson KH, Stahle J (1987) Familial Menière's disease: a genetic investigation. *Am J Otol* 8(4):323–326
- Bycroft C, Fernandez-Rozadilla C, Ruiz-Ponte C, Quintela I, Caracado Á, Donnelly P, Myers S (2019) Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula. *Nat Commun* 10(1):551. <https://doi.org/10.1038/s41467-018-08272-w>
- Chen Y, Lu H, Zhang N, Zhu Z, Wang S, Li M (2020) PremPS: predicting the impact of missense mutations on protein stability. *PLoS Comput Biol* 16(12):e1008543. <https://doi.org/10.1371/journal.pcbi.1008543>
- Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, Schneider RG, Senior AW, Jumper J, Hassabis D, Kohli P, Avsec Ž (2023) Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 381(6664):eadg7492. <https://doi.org/10.1126/science.adg7492>
- Cohen-Salmon M, El-Amraoui A, Leibovici M, Petit C (1997) Otogelin: a glycoprotein specific to the acellular membranes of the inner ear. *Proc Natl Acad Sci USA* 94(26):14450–14455
- Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2(9):1511–1519. <https://doi.org/10.1002/pro.5560020916>
- El-Amraoui A, Cohen-Salmon M, Petit C, Simmler M-C (2001) Spatiotemporal expression of otogelin in the developing and adult mouse inner ear. *Hear Res* 158(1):151–159. [https://doi.org/10.1016/S0378-5955\(01\)00312-4](https://doi.org/10.1016/S0378-5955(01)00312-4)
- Escalera-Balsara A, Roman-Naranjo P, Lopez-Escamez JA (2020) Systematic review of sequencing studies and gene expression profiling in familial Meniere disease. *Genes* 11(12):1414. <https://doi.org/10.3390/genes11121414>
- Flook M, Escalera-Balsara A, Gallego-Martinez A, Espinosa-Sanchez JM, Aran I, Soto-Varela A, Lopez-Escamez JA (2021) DNA Methylation signature in mononuclear cells and proinflammatory cytokines may define molecular subtypes in sporadic Meniere disease. *Biomedicines* 9(11):1530. <https://doi.org/10.3390/biomedicines9111530>
- Frejo L, Martin-Sanz E, Teggi R, Trinidad G, Soto-Varela A, Santos-Perez S, Manrique R, Perez N, Aran I, Almeida-Branco MS, Batuecas-Caletrio A, Fraile J, Espinosa-Sanchez JM, Perez-Guillen V, Perez-Garrigues H, Oliva-Dominguez M, Aleman O, Benitez J, Perez P, Lopez-Escamez JA (2017) Extended phenotype and clinical subgroups in unilateral Meniere disease: a cross-sectional

- study with cluster analysis. *Clin Otolaryngol* 42(6):1172–1180. <https://doi.org/10.1111/coa.12844>
- Ganaha A, Kaname T, Yanagi K, Tono T, Higa T, Suzuki M (2019) Clinical characteristics with long-term follow-up of four Okinawan families with moderate hearing loss caused by an OTOG variant. *Hum Genome Var* 6:37. <https://doi.org/10.1038/s41439-019-0068-4>
- Gazquez I, Soto-Varela A, Aran I, Santos S, Batuecas A, Trinidad G, Perez-Garrigues H, Gonzalez-Oller C, Acosta L, Lopez-Escamez JA (2011) High prevalence of systemic autoimmune diseases in patients with Meniere's disease. *PLoS ONE* 6(10):e26759. <https://doi.org/10.1371/journal.pone.0026759>
- Guo S-S, Liu J, Zhou X-G, Zhang G-J (2022) DeepUMQA: ultrafast shape recognition-based protein model quality assessment using deep learning. *Bioinformatics* 38(7):1895–1903. <https://doi.org/10.1093/bioinformatics/btac056>
- Havrilla JM, Pedersen BS, Layer RM, Quinlan AR (2019) A map of constrained coding regions in the human genome. *Nat Genet* 51(1):88–95. <https://doi.org/10.1038/s41588-018-0294-6>
- Hietikko E, Kotimäki J, Männikkö M (2013a) Molecular genetic analysis of Ménière's disease. John Wiley and Sons, Ltd, In Encyclopedia of Life Sciences. <https://doi.org/10.1002/9780470015902.a0025161>
- Hietikko E, Kotimäki J, Sorri M, Männikkö M (2013b) High incidence of Meniere-like symptoms in relatives of Meniere patients in the areas of Oulu University Hospital and Kainuu Central Hospital in Finland. *Eur J Med Genet* 56(6):279–285. <https://doi.org/10.1016/j.ejmg.2013.03.010>
- Hui D, Mehrabi S, Quimby AE, Chen T, Chen S, Park J, Li B, Center RG, Biobank PM, Ruckenstein MJ, Rader DJ, Ritchie MD, Brant JA, Epstein DJ, Mathieson I (2023) Gene burden analysis identifies genes associated with increased risk and severity of adult-onset hearing loss in a diverse hospital-based cohort. *PLoS Genet* 19(1):e1010584. <https://doi.org/10.1371/journal.pgen.1010584>
- Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, Chow ED, Kanterakis E, Gao H, Kia A, Batzoglu S, Sanders SJ, Farh KK-H (2019) Predicting splicing from primary sequence with deep learning. *Cell* 176(3):535–548.e24. <https://doi.org/10.1016/j.cell.2018.12.015>
- Jain A, Sharma D, Bajaj A, Gupta V, Scaria V (2021) Chapter Four—founder variants and population genomes—toward precision medicine. In: Kumar D (ed) *Advances in Genetics*, vol 107. Academic Press, New York, pp 121–152. <https://doi.org/10.1016/bs.adgen.2020.11.004>
- Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, MacArthur DG (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581(7809):434–443. <https://doi.org/10.1038/s41586-020-2308-7>
- Kim SY, Kim BJ, Oh DY, Han JH, Yi N, Kim NJ, Park MK, Keum C, Seo GH, Choi BY (2022a) Improving genetic diagnosis by disease-specific, ACMG/AMP variant interpretation guidelines for hearing loss. *Sci Rep* 12(1):12457. <https://doi.org/10.1038/s41598-022-16661-x>
- Kim SY, Lee CH, Yoo DM, Min C, Choi HG (2022b) Association between asthma and Meniere's disease: a nested case-control study. *Laryngoscope* 132(4):864–872. <https://doi.org/10.1002/lary.29917>
- Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, Karapetyan K, Katz K, Liu C, Maddipatla Z, Malheiro A, McDaniel K, Ovetsky M, Riley G, Zhou G, Maglott DR (2018) ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* 46(D1):D1062–D1067. <https://doi.org/10.1093/nar/gkx1153>
- Lee JM, Kim MJ, Jung J, Kim HJ, Seo YJ, Kim SH (2015) Genetic aspects and clinical characteristics of familial Meniere's disease in a South Korean population. *Laryngoscope* 125(9):2175–2180. <https://doi.org/10.1002/lary.25207>
- Lopez-Escamez JA, Carey J, Chung W-H, Goebel JA, Magnusson M, Mandalà M, Newman-Toker DE, Strupp M, Suzuki M, Trimbini F, Bisdorff A (2015) Diagnostic criteria for Ménière's disease. *J Vestib Res* 25(1):1–7. <https://doi.org/10.3233/VES-150549>
- Mizukoshi K, Ino H, Ishikawa K, Watanabe Y, Yamazaki H, Kato I, Okubo J, Watanabe I (1979) Epidemiological survey of definite cases of Meniere's disease collected by the seventeen members of the Meniere's Disease Research Committee of Japan in 1975–1976. *Adv Otorhinolaryngol* 25:106–111. <https://doi.org/10.1159/000402925>
- Morrison AW, Bailey MES, Morrison GAJ (2009) Familial Ménière's disease: clinical and genetic aspects. *J Laryngol Otol* 123(1):29–37. <https://doi.org/10.1017/S0022215108002788>
- Myers TA, Chanock SJ, Machiela MJ (2020) LDlinkR: an R package for rapidly calculating linkage disequilibrium statistics in diverse populations. *Front Genet* 11:157. <https://doi.org/10.3389/fgene.2020.00157>
- Niroula A, Vihinen M (2019) How good are pathogenicity predictors in detecting benign variants? *PLoS Comput Biol* 15(2):1006481. <https://doi.org/10.1371/journal.pcbi.1006481>
- Pancotti C, Benevenuto S, Birolo G, Alberini V, Repetto V, Sanavia T, Capriotti E, Fariselli P (2022) Predicting protein stability changes upon single-point mutation: a thorough comparison of the available tools on a new dataset. *Brief Bioinform* 23(2):555. <https://doi.org/10.1093/bib/bbab555>
- Parra-Perez AM, Lopez-Escamez JA (2023) Types of inheritance and genes associated with familial Meniere disease. *J Assoc Res Otolaryngol*. <https://doi.org/10.1007/s10162-023-00896-0>
- Perez-Carpena P, Lopez-Escamez JA (2019) Current understanding and clinical management of Meniere's disease: a systematic review. *Semin Neurol* 40(1):138–150. <https://doi.org/10.1055/s-0039-3402065>
- Pires DEV, Ascher DB, Blundell TL (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* 30(3):335–342. <https://doi.org/10.1093/bioinformatics/btt691>
- Radtke A, Lempert T, Gresty MA, Brookes GB, Bronstein AM, Neuhauser H (2002) Migraine and Ménière's disease: is there a link? *Neurology* 59(11):1700–1704. <https://doi.org/10.1212/01.wnl.0000036903.22461.39>
- Requena T, Espinosa-Sanchez JM, Cabrera S, Trinidad G, Soto-Varela A, Santos-Perez S, Teggi R, Perez P, Batuecas-Caletrio A, Fraile J, Aran I, Martin E, Benítez J, Pérez-Fernández N, Lopez-Escamez JA (2014) Familial clustering and genetic heterogeneity in Meniere's disease. *Clin Genet* 85(3):245–252. <https://doi.org/10.1111/cge.12150>
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. *Nat Biotechnol* 29(1):24–26. <https://doi.org/10.1038/nbt.1754>
- Rodrigues CHM, Pires DEV, Ascher DB (2021) DynaMut2: assessing changes in stability and flexibility upon single and multiple point missense mutations. *Protein Science: A Publication of the Protein Society* 30(1):60–69. <https://doi.org/10.1002/pro.3942>
- Roman-Naranjo P, Gallego-Martinez A, Soto-Varela A, Aran I, Moleons Espinosa-Sanchez Amor-Dorado Batuecas-Caletrio Perez-Vazquez Lopez-Escamez MCJMJCAPJA (2020) Burden of rare variants in the OTOG gene in familial Meniere's disease. *Ear Hear* 41(6):1598–1605. <https://doi.org/10.1097/AUD.0000000000000878>

- Roman-Naranjo P, Moleon MDC, Aran I, Escalera-Balsera A, Soto-Varela A, Bächinger D, Gomez-Fiñana M, Eckhard AH, Lopez-Escamez JA (2021) Rare coding variants involving MYO7A and other genes encoding stereocilia link proteins in familial Meniere disease. *Hear Res* 409:108329. <https://doi.org/10.1016/j.heares.2021.108329>
- Semaan MT, Alagramam KN, Megerian CA (2005) The basic science of Meniere's disease and endolymphatic hydrops. *Curr Opin Otolaryngol Head Neck Surg* 13(5):301–307
- Shin J-H, Blay S, McNeney B, Graham J (2006) LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Softw* 16:1–9. <https://doi.org/10.18637/jss.v016.c03>
- Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T (2020) QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics* 36(8):2647–2647. <https://doi.org/10.1093/bioinformatics/btaa058>
- Wang Y, Lu D, Chung Y-J, Xu S (2018) Genetic structure, divergence and admixture of Han Chinese, Japanese and Korean Populations. *Hereditas* 155(1):19. <https://doi.org/10.1186/s41065-018-0057-5>
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35(W407):W410. <https://doi.org/10.1093/nar/gkm290>
- Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci* 27(1):293–315. <https://doi.org/10.1002/pro.3330>
- Zhou Y-F, Springer TA (2014) Highly reinforced structure of a C-terminal dimerization domain in von Willebrand factor. *Blood* 123(12):1785–1793. <https://doi.org/10.1182/blood-2013-11-523639>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.