# Excess of Rare Missense Variants in Hearing Loss Genes in Sporadic Meniere Disease

Alvaro Gallego-Martinez[1], Teresa Requena[1], Pablo Roman-Naranjo[1] and
Jose A. Lopez-Escamez[1,2]* on behalf of the Meniere Disease Consortium (MeDiC)

[1] Otology and Neurotology Group CTS 495, Department of Genomic Medicine, Centre for Genomics and Oncological Research (GENyO), Pfizer, University of Granada, Andalusian Regional Government, Granada, Spain, [2] Department of Otolaryngology, Instituto de Investigación Biosanitaria (ibs. GRANADA), Hospital Universitario Virgen de las Nieves, Universidad de Granada, Granada, Spain

Meniere's disease (MD) is a clinical spectrum of rare disorders characterized by vertigo attacks, associated with sensorineural hearing loss (SNHL) and tinnitus involving low to medium frequencies. Although it shows familial aggregation with incomplete phenotypic forms and variable expressivity, most cases are considered sporadic. The aim of this study was to investigate the burden for rare variation in SNHL genes in patients with sporadic MD. We conducted a targeted-sequencing study including SNHL and familial MD genes in 890 MD patients to compare the frequency of rare variants in cases using three independent public datasets as controls. Patients with sporadic MD showed a significant enrichment of missense variants in SNHL genes that was not found in the controls. The list of genes includes *GJB2, USH1G, SLC26A4, ESRRB,* and *CLDN14*. A rare synonymous variant with unknown significance was found in the *MARVELD2* gene in several unrelated patients with MD. There is a burden of rare variation in certain SNHL genes in sporadic MD. Furthermore, the interaction of common and rare variants in SNHL genes may have an additive effect on MD phenotype. This study will contribute to design a gene panel for the genetic diagnosis of MD.

**Keywords: SNHL, Meniere's disease, vertigo, variant aggregation, Spanish population**

## INTRODUCTION

Meniere's disease (MD, MIM 156000) is a chronic disorder of the inner ear characterized by episodes of vertigo, associated with low to middle frequency sensorineural hearing loss (SNHL), tinnitus and aural fullness (Lopez-Escamez et al., 2015). The disorder produces an accumulation of endolymph in the membranous labyrinth, and it may affect both ears in 25–40% of patients (termed bilateral MD) and most of cases are considered sporadic (Paparella and Griebie, 1984). However, heterogeneity in the phenotype is observed and some patients may have co-morbid conditions such as migraine or systemic autoimmune disorders (Gazquez et al., 2011; Caulley et al., 2017). This phenotypic spectrum can make the clinical diagnosis challenging considering that some of the symptoms overlap with other vestibular disorders such vestibular migraine (VM) or autoimmune inner ear disease (AIED) (Hietikko et al., 2011; Lempert et al., 2012; Mijovic et al., 2013; Requena et al., 2014b).

Epidemiological evidence showing a genetic contribution in MD is based on familial aggregation studies with a high siblings recurrence risk ratio ($\lambda_s = 16$–48) (Requena et al., 2014a) and the

description of multiple familial cases in European and Asian descendant populations (Arweiler-Harbeck et al., 2011; Hietikko et al., 2013). Exome sequencing has identified private variants in the *FAM136A, DTNA, PRKCB, SEMA3D,* and *DPT* genes in 4 families with autosomal dominant MD, showing incomplete penetrance and variable expressivity (Requena et al., 2015; Martín-Sierra et al., 2016, 2017). Moreover, some relatives in familial MD show partial syndromes, either with SNHL or episodic vertigo, increasing the granularity in the phenotype in a given family. However, the genetic contribution of familial genes in sporadic cases has not been investigated and the occurrence of recessive and novel variants is not known. More than 110 genes and ≈6,000 variants have been related to hereditary non-syndromic hearing loss, making gene sequencing panels an essential tool for genetic diagnosis of hearing loss (The Molecular Otolaryngology and Renal Research Laboratories, The University of Iowa, 0000; Sloan-Heggen et al., 2016). Those genes include 45 genes associated to autosomal dominant SNHL and 71 genes related to recessive hearing loss (Shearer et al., 1999; Van Camp and Smith, 2018).

Targeted gene sequencing panels have been demonstrated to be an excellent tool for molecular diagnosis of rare variants in known genes with allelic heterogeneity (Brownstein et al., 2011; Lionel et al., 2018), as well as in sporadic cases of hearing impairment in specific populations (Gu et al., 2015; Dallol et al., 2016). So, we selected SNHL genes and designed a custom gene panel to search for rare and novel variants in sporadic MD.

In the present study, we describe the genetic variation found in a custom exon-sequencing panel of 69 genes in a large cohort of MD sporadic cases. We report that certain genes such as *GJB2, USH1G, SLC26A4,* and *CLDN14* show an excess of missense variants in sporadic MD cases when compared to controls in the Iberian population, suggesting that several rare variants in these genes may contribute to the SNHL phenotype in sporadic MD.

## MATERIALS AND METHODS

### Editorial Policies and Ethical Considerations

This study protocol was approved by the Institutional Review Board for Clinical Research (MS/2014/02), and a written informed consent to donate biological samples was obtained from all subjects.

### Sample Selection

A total of 890 Spanish and Portuguese patients with MD were recruited. All patients were diagnosed by neurotology experienced clinicians from the Meniere's Disease Consortium (MeDiC), according to the diagnostic criteria for MD formulated by the International Classification Committee for Vestibular Disorders of the Barany Society in 2015 (Lopez-Escamez et al., 2015). Among them, 830 were considered sporadic MD cases while 60 were familial MD cases. Details from the selected cases are described in **Table 1**. As controls, 40 healthy individuals were selected from the same population.

**TABLE 1 |** Participant subjects in this study.

| Diagnosed | *N* | Sex | % | Location | *N* |
|-----------|-----|-----|---|----------|-----|
| SMD | 830 | Male | 40 | North | 500 |
| | | | | South | 320 |
| FMD | 60 | Female | 60 | Center | 90 |
| Control | 40 | | | Portugal | 20 |

*Number of individuals and geographical distribution of the selected cases and controls for targeted-gene sequencing. SMD, sporadic Meniere disease; FMD, familial Meniere disease.*

### Selection of Target Genes

Target genes were selected from a literature search attending to human phenotype (hearing profile, comorbid vestibular symptoms) and phenotype observations in mouse and zebrafish models. Most of them were selected from HearingLoss.org website gene list for monogenic SNHL. Additional genes were added because they have been previously found in familial MD (Requena et al., 2015; Martín-Sierra et al., 2016, 2017), or allelic variations associated with hearing outcome in MD had been described, such as NFKB1 or TLR10 genes (Requena et al., 2013; Cabrera et al., 2014). Mitochondrial genes were added since maternal inheritance is suspected in several families with MD (Requena et al., 2014b). Relevant information about the location, size, bibliography, and other characteristics about each gene included in the panel is presented in **Tables S1, S2**.

The custom panel (Panel ID: 39351-1430751809) were designed by the Suredesign webtool (Agilent) to cover the exons and 50 bp in the flanking regions (5′ and 3′ UTR). This allowed the sequencing around 533.380 kb with more than 98.46% coverage.

### Sample Pooling

Enrichment technology allows the selective amplification of targeted sequences and DNA sample pooling, reducing the costs of reagents and increasing sample size. We decided to pool patient samples according to their geographical origin. Each pool consisted of 10 DNA samples from the same hospital for a total of 93 pools (930 samples).

DNA concentration and quality were measured on each sample using two methods: Qubit dsDNA BR Assay kit (TermoFisher Scientific) and Nanodrop 2000C (ThermoFisher Scientific). All samples had quality ratios ranging 1.8 and 2.0 in 280/260 and 1.6 to 2.0 in 260/230.

### Libraries Preparation Protocol

The HaloPlex Target Enrichment System (Agilent, Santa Clara, CA) was used to prepare the DNA libraries, according to the manufacturer protocol. Validation of the protocol and library performance was analyzed with a 2,100 Bioanalyzer High Sensitivity DNA Assay kit. Expected concentrations were between 1 and 10 ng/ul. Higher concentrations than 10 ng/ul were diluted 1:10 in 10 mM TRIS, 1 mM EDTA. Targeted-sequencing was performed in an Illumina Nextseq500 platform.

## Data Generation Pipelines

Raw data downloaded and sequencing adapters were trimmed following manufacturer indications. The requested depth of coverage for the sequencing panel was 250X. The minimum coverage considered was 30X mean depth for nuclear genes, however mitochondrial sequences reached higher coverages with the enrichment technology. Bioinformatic analyses were performed according to the Good Practices recommended by Genome Analysis ToolKit (https://software.broadinstitute.org/gatk/). Mitochondrial genes were analyzed using the same pipeline.

Two methods were used to find differences in how UnifiedGenotyper and HaplotypeCaller (the old and the most recent tools for variant calling in GATK suite) address sequenced pools. Both custom pipelines use BWA-mem aligner and GATK suite tools following the GATK protocol for Variant Calling against GRCh37/hg19 human reference genome. Left normalization for multi-allelic variants were addressed by separated. Calling was made in the first pipeline with UnifiedGenotyper modifying number of chromosomes per sample (per pool, there are 20 chromosomes). The second pipeline used HaplotypeCaller, which cannot allow the same approach, but can automatically address high number of calls with a different approach. Variants with read depth (RD) <10 and genotype quality (GQ) <20 were excluded in all the calling pipelines following recommended hard filtering steps by GATK suite.

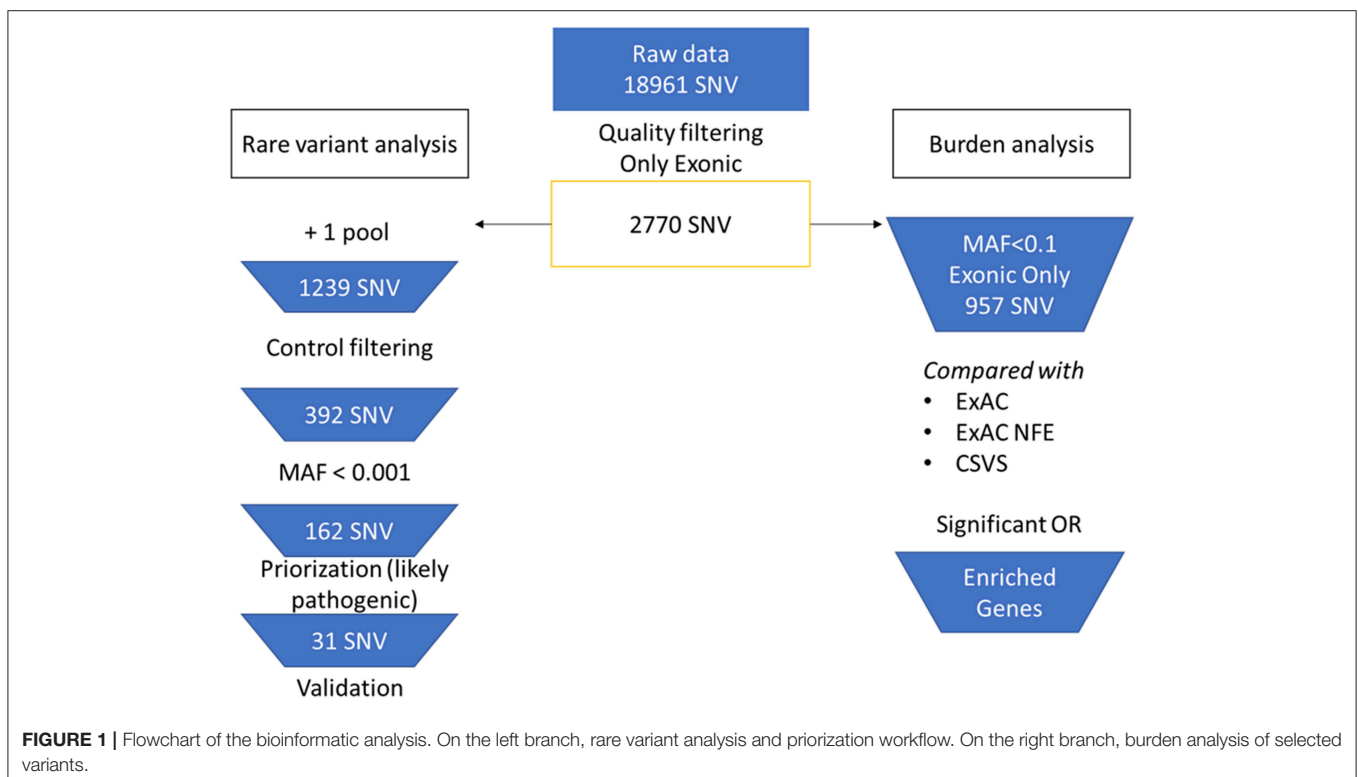A third caller tool, VarScan, was used to filter and annotate quality strand data per variant to compare its output with GATK-based callers. VarScan allows the variant filtering using the information obtained according to each strand polarity. The method retrieves those variants that were only called in one strand, but not in the reverse strand, leading to false positive calls. This step was used as internal quality control to avoid strand bias usually generated in Haloplex data, as it has been reported in other studies (Collet et al., 2015).

## Positive Control SNV Validation

Positive control testing was addressed using samples from patients with familial MD with known variants on certain genes. These individuals come from previous familial studies with independently validated variants by Sanger sequencing. Known variants were also sequenced and validated by Sanger (**Table S3**). Coverage and mapping quality after each pipeline were annotated and measured. Representative chromatographs from validated SNVs are detailed in **Figure S1**.

## Selection and Prioritization of Pathogenic SNV

In order to obtain more information of each SNV, we annotated the merged files using the ANNOVAR tool. Minor allele frequencies (MAF) were retrieved for each candidate variant from gnomAD and ExAC database (total individuals and non-Finnish European (NFE) individuals). Since the estimated prevalence of sporadic MD in Spain is 0.75/1,000 individuals (Morales Angulo et al., 2003), we selected variants with MAF <0.001 for single rare variant analysis and prioritized them according to Combined Annotation Dependent Depletion



**FIGURE 1** | Flowchart of the bioinformatic analysis. On the left branch, rare variant analysis and priorization workflow. On the right branch, burden analysis of selected variants.

(CADD) phred score. For burden analysis of common and rare variants, we chose a higher MAF value <0.1. The Collaborative Spanish Variant Server (CVCS) database including 1,644 unrelated individuals was also used for annotation of exonic rare variants and to retrieve MAF in Spanish population (dataset fully accessible from http://csvs.babelomics.org/) (Dopazo et al., 2016).

KGGseq suite (grass.cgs.hku.hk/limx/kggseq) was used for the selection of rare variants to prioritize the most pathogenic variants according to the integrated model trained algorithm with known pathogenic variants and neutral control variants.

Enrichment analysis for each gene was made with all the exonic variants found with a MAF <0.1. This analysis required to divide the total amount of variants into three groups: those described in global ExAC population, those found in NFE population, and finally those included in CVCS Spanish population. These three reference datasets were used for enrichment analysis comparison.

## Validation of Candidate Pathogenic SNV

Candidate SNV were visually revised in the BAM files using Integrative Genomics Viewer (https://software.broadinstitute.org/software/igv/) and validated in the different pools where they were called using Sanger sequencing.

## Population Statistics

Statistical analysis was performed with IBM SPSS v.20 program, Microsoft Excel suite tools, and diverse python and java encoded public scripts. Due to the overrepresentation of Spanish population in our dataset, most of the selected variants were filtered through exome sequencing data from Spanish controls of CSVS database. The MAF was calculated for each variant in our dataset and rare and previously unreported variants on MD patients were identified in our gene panel. Odd ratios with 95% confidence interval were calculated for each variant using MAF obtained from Spanish population ($N = 1,579$), ExAC ($N = 60,706$), and ExAC NFE ($N = 33,370$) populations as controls.

Gene burden analysis was addressed using $2 \times 2$ contingency tables counting total exonic alternate allele counts per gene in our cases against total and NFE controls in ExAC and CSVS controls. Odds ratios with 95% confidence intervals were calculated using Fisher's exact test and obtaining one-sided p-values. P-values were also corrected for multiple testing by the total amount of variants found for each gene following Bonferroni approach.

## Position of Variants in Significant Enriched Genes

Several models were generated for rare variant-enriched domains in significant enriched genes by using the INSIDER modeling tool (Meyer et al., 2018). The selected variants per gene are detailed in results. Prediction values were annotated with their calculated p values.

# RESULTS

## Single Rare Variant Analysis

We achieved an average capture efficiency rate (percentage of total on-target reads in total sequenced reads) of 69% on the target regions above 30X (minimum depth considered for quality filtering). The mean coverage percentage can be found in **Table S4**. A total of 2,770 SNV in nuclear genes were selected from the raw merged dataset (18,961 SNVs) after filtering by quality controls. The analysis workflow is summarized in **Figure 1**. For rare variants analysis, SNV that were found in more than one pool were selected, remaining 1,239 variants. After that, we filtered by variants observed in the control pools, leaving only 392 exonic SNV in cases (278 missense, 111 synonymous, 2 stopgain, and 1 stoploss).

A final set of 162 SNV with a MAF <0.001 were retrieved (143 missense, 18 synonymous, 1 stoploss, 1 stopgain). All the exonic variants were annotated and scored using different priorization tools. Of them, 136 SNVs were not previously described in any population database and we considered them as potential novel variants.

After prioritizing the exonic variants by CADD phred, 31 rare variants remained (**Table 2**). Six of them were validated by Sanger sequencing in more than two individuals in the following genes: *GJB2, ESRRB, USH1G, SLC26A4* (**Table S3**). The rest of the variants were considered benign or likely benign since they did not reach the pathogenicity threshold predicted for *KGGSeq*. However, a novel synonymous variant in the *MARVELD2* gene was found and validated in three unrelated individuals.

The minor allelic frequencies in SNV of the 24 mitochondrial genes included in the panel were compared with the reference data obtained from MITOmap through its automated mtDNA sequence analysis system Mitomaster (Ruiz-Pesini et al., 2007). However, the candidate variants observed do not belong to the genes targeted in the mitochondrial genome, since they were not validated by Sanger sequencing. We did not found any SNV associated with MD (data not shown).

## Gene Burden Analysis

To analyse the interaction of multiple variants, we considered SNV with a MAF<0.1 for the gene burden analysis. A total of 957 exonic variants were retrieved and their frequencies were compared with the global and NFE frequencies from ExAC, and with the Spanish population frequencies from CSVS.

A gene burden analysis using our gene set was performed using these three reference datasets. After Bonferroni correction, some genes showed a significant enrichment of rare variants in the three comparisons, making them candidate genes to be selected for a diagnosis panel for MD (**Table 3**). Moreover, 6 genes (*FAM136A, ADD1, SLC12A2, POU4F3, RDX, and PRKCB*) presented some novel variants that were validated by Sanger, but they have not been described in global ExAC or CSVS datasets. Although these previously unreported variants could not be sequenced in all the parents of these patients, we considered them as potential *de novo* variants.

A second variant analysis using the missense variants described in CSVS Spanish population database was made

**TABLE 2 |** Prioritized rare SNVs found in the single rare variant analysis for sporadic MD cases.

| Chr. | Position | Ref | Alt | Gene | Exonic function | dbSNP | ExAC MAF | gnomAD MAF | Cohort MAF | CADD phred score |
|------|----------|-----|-----|------|-----------------|-------|----------|------------|------------|------------------|
| chr1 | 6,488,328 | C | T | ESPN | Non-synonymous | rs143577178 | 0.0005 | 0.0005 | 0.004819 | 35 |
| chr4 | 6,303,574 | G | T | WFS1 | Synonymous | | – | – | 0.016867 | – |
| chr4 | 6,303,946 | C | A | WFS1 | Non-synonymous | | – | – | 0.004819 | 22.1 |
| chr4 | 6,304,133 | G | C | WFS1 | Non-synonymous | | | 4.08E-06 | 0.010843 | 21 |
| chr4 | 38,775,706 | G | T | TLR10 | Non-synonymous | | – | – | 0.003614 | 20.7 |
| chr4 | 38,776,070 | C | A | TLR10 | Non-synonymous | | – | – | 0.024096 | 14.27 |
| chr5 | 68,715,821 | G | T | MARVELD2 | Non-synonymous | | – | – | 0.003614 | 23.3 |
| chr7 | 107,336,408 | A | C | SLC26A4 | Non-synonymous | rs200511789 | 0.0004 | 0.0004 | 0.003614 | 24.3 |
| chr10 | 14,563,299 | G | T | FAM107B | Non-synonymous | | – | – | 0.003614 | 24.9 |
| chr11 | 76,885,923 | G | A | MYO7A | Non-synonymous | rs781991817 | 0.0002 | 0.0003 | 0.004819 | 34 |
| chr11 | 76,892,617 | G | C | MYO7A | Non-synonymous | rs200641606 | 0.0007 | 0.0007 | 0.003614 | 25.5 |
| chr12 | 65,672,602 | C | T | MSRB3 | Synonymous | rs149757878 | 0.0002 | 0.0005 | 0.004819 | - |
| chr13 | 20,763,612 | C | A | GJB2 | Non-synonymous | rs72474224 | – | – | 0.007229 | 23.8 |
| chr14 | 76,957,891 | G | A | ESRRB | Non-synonymous | rs201344770 | 0.0003 | 0.0002 | 0.008434 | 23.8 |
| chr14 | 76,966,336 | G | A | ESRRB | Non-synonymous | rs200237229 | 0.0007 | 0.0005 | 0.003614 | 1.198 |
| chr14 | 76,966,347 | C | T | ESRRB | Non-synonymous | rs201448899 | 0.001 | 0.0007 | 0.008434 | 15.41 |
| chr16 | 24,046,852 | C | T | PRKCB | Synonymous | rs115645964 | 0.0003 | 0.0003 | 0.003614 | - |
| chr17 | 72,915,919 | C | T | USH1G | Non-synonymous | rs151242039 | 0.0006 | 0.0008 | 0.003614 | 8.91 |
| chr17 | 72,916,543 | T | G | USH1G | Non-synonymous | | – | – | 0.004819 | 14.31 |
| chr17 | 79,478,028 | G | A | ACTG1 | Synonymous | rs187127467 | 0.0002 | 0.0003 | 0.008434 | - |
| chr18 | 32,462,094 | G | T | DTNA | Non-synonymous | rs533568822 | 2.47E-05 | 3.66E-05 | 0.003614 | 25 |
| chr19 | 50,784,974 | A | C | MYH14 | Non-synonymous | | – | – | 0.003614 | 20.3 |
| chr22 | 38,119,197 | G | T | TRIOBP | Non-synonymous | | – | – | 0.003614 | 17.9 |
| chr22 | 38,119,969 | C | A | TRIOBP | Non-synonymous | | – | - | 0.010843 | 15 |
| chr22 | 38,119,977 | G | T | TRIOBP | Non-synonymous | | – | – | 0.003614 | 0.022 |
| chr22 | 38,120,106 | A | C | TRIOBP | Non-synonymous | | – | – | 0.003614 | 2.968 |
| chr22 | 38,120,116 | C | A | TRIOBP | Non-synonymous | | – | – | 0.024096 | 16.55 |
| chr22 | 38,120,263 | C | A | TRIOBP | Non-synonymous | | - | - | 0.020482 | 22.3 |
| chr22 | 38,120,302 | C | A | TRIOBP | Non-synonymous | | – | - | 0.003614 | 23.8 |
| chr22 | 38,168,691 | G | C | TRIOBP | Non-synonymous | rs373236761 | 0.0001 | 0.0002 | 0.003614 | 26.6 |

*Minor allele frequency for each SNV is detailed as annotated by ExAC and gnomAD (exomes). Pathogenicity prediction is detailed according to CADD phred score.*

(**Table 4**). Eighteen genes showed an excess of missense variants (a total of 46 variants, detailed in **Table S5**). Of note, five genes causing autosomal recessive SNHL showed the highest accumulation of missense variants when they were compared with NFE and Spanish population datasets: *SLC26A4, GJB2, CLDN14, ESRRB,* and *USH1G*. The variants in these five genes were validated through Sanger sequencing and considered Spanish population-specific variants.

## Excess of Rare Variants in Hearing Loss Genes in Familial Cases

We used exome sequencing datasets from familial MD cases previously reported to search for rare variants identified in our panel in the sporadic cases. Although no single missense variant was found segregated in all the cases in the same family, we found several rare missense variants in at least one case per family in genes such as *GJB2, GRHL2, TRIOBP, RDX, KCNQ4, WFS1,* and *ADD1*. These MD families show phenotypic differences in terms of age of onset, hearing profile and disease progression

and the presence of rare variants can be addressed as potential modulators of the phenotype in each familial case (**Table 5**).

## Effect of Rare Variant Interaction

We selected exonic variants from the gene burden analysis to analyze their potential additive effect at the protein-protein interaction interfaces by the tool INSIDER for our selected five genes. However, protein interfaces for *ESRRB, CLDN14* and *SLC26A4* genes could not be loaded and processed on the database (lacking predicted interfaces on ÉCLAIR database or crystalized protein structures on Protein Data Bank (PDB) database). Of note, most relevant affected interaction is observed in the self-interaction *GJB2-GJB2* by the known variants observed in the burden analysis (significant spatial clustering with 4 SNV, $p = 0.0009$) rs111033218:G>C (p.Phe83Leu), rs80338945:A>G(p.Leu90Pro), rs374625633:T>C(p.Ile30Val), and rs2274084:C>T(p.Val27Ile) (**Figure 2A**).

Other interactions of interest were founded between the *USH1G—USH1C* genes, but the involved variants were not located in the known interaction surface of *USH1G* (**Figure 2B**).

**TABLE 3** | Gene burden analysis 1.

| Gene | # variants | Odds ratio ExAC[a] | P value | P corrected[b] | Odds ratio ExAC NFE[a] | P value | P corrected[b] | Odds ratio CSVS[a] | P value | P corrected[b] |
|---|---|---|---|---|---|---|---|---|---|---|
| MYH14 | 50 | 13.54 (5.85–31.37) | 1.20E-09 | 6.01E-08 | 56.94 (25.13–128.98) | <1.00E-15 | <1.00E-15 | 38.52 (16.95–87.56) | <1.00E-15 | <1.00E-15 |
| MYO7A | 43 | 3.11 (1.73–5.57) | 1.41E-04 | 6.08E-03 | 4.64 (2.65–8.13) | 7.72E-08 | 3.32E-06 | 3.85 (2.18–6.81) | 3.60E-06 | 1.55E-04 |
| WFS1 | 36 | 1.24 (0 6–2.57) | 5.67E-01 | 1 | 1.84 (0.94–3.62) | 7.70E-02 | 1 | 3.28 (1.76–6.1) | 1.83E-04 | 6.57E-03 |
| ADD1 | 29 | 249.89 (96.17–649.29) | <1.00E-15 | <1.00E-15 | 318.03 (122.45–826) | <1.00E-15 | <1.00E-15 | 747.23 (287.95–1939.08) | <1.00E-15 | <1.00E-15 |
| WHRN | 28 | 24.63 (7.96–76.17) | 2.68E-08 | 7.51E-07 | 39.97 (13.03–122.58) | 1.12E-10 | 3.13E-09 | 62.91 (20.62–191.98) | 3.44E-13 | 9.64E-12 |
| TPRN | 26 | 36.63 (10.61–126.49) | 1.24E-08 | 3.22E-07 | 46.93 (13.64–161.49) | 1.03E-09 | 2.68E-08 | 81.9 (23.93–280.26) | 2.24E-12 | 5.81E-11 |
| USH1G | 24 | 4.04 (1.91–8.51) | 2.46E-04 | 5.90E-03 | 18.38 (9.26–36.48) | <1.00E-15 | <1.00E-15 | 8.12 (4–16.49) | 6.57E-09 | 1.58E-07 |
| USH1C | 22 | 5.12 (2.58–10.17) | 3.13E-06 | 6.88E-05 | 10.68 (5.54–20.59) | 1.54E-12 | 3.38E-11 | 5.12 (2.58–10.18) | 3.06E-06 | 6.73E-05 |
| P2RX2 | 22 | 20.34 (9.75–42.44) | <1.00E-15 | <1.00E-15 | 23.74 (11.41–49.41) | <1.00E-15 | <1.00E-15 | 23.51 (11.29–48.94) | <1.00E-15 | <1.00E-15 |
| ESPN | 19 | 27.88 (17.06–45.55) | <1.00E-15 | <1.00E-15 | 24.35 (14.88–39.84) | <1.00E-15 | <1.00E-15 | 17.39 (10.59–28.56) | <1.00E-15 | <1.00E-15 |
| RDX | 17 | 94.37 (36.06–247) | <1.00E-15 | <1.00E-15 | 57.38 (21.85–150.66) | <1.00E-15 | <1.00E-15 | 75.26 (28.72–197.23) | <1.00E-15 | <1.00E-15 |
| TJP2 | 17 | 2.78 (1.59–4.88) | 3.58E-04 | 6.09E-03 | 7.99 (4.79–13.32) | <1.00E-15 | <1.00E-15 | 2.92 (1.67–5.11) | 1.67E-04 | 2.85E-03 |
| SLC26A4 | 15 | 1.79 (1.02–3.15) | 4.31E-02 | 6.46E-01 | 4.14 (2.5–6.86) | 3.37E-08 | 5.06E-07 | 3.65 (2.19–6.09) | 6.62E-07 | 9.93E-06 |
| ESRRB | 14 | 12.54 (7.26–21.63) | <1.00E-15 | <1.00E-15 | 10.67 (6.16–18.48) | <1.00E-15 | <1.00E-15 | 7.54 (4.31–13.18) | 1.41E-12 | 1.98E-11 |
| DTNA | 13 | 6.07 (4.68–7.87) | <1.00E-15 | <1.00E-15 | 4.87 (3.74–6.35) | <1.00E-15 | <1.00E-15 | 4.78 (3.67–6.23) | <1.00E-15 | <1.00E-15 |
| PRKCB | 13 | 27.62 (18.65–40.91) | <1.00E-15 | <1.00E-15 | 17.6 (11.84–26.16) | <1.00E-15 | <1.00E-15 | 40.79 (27.61–60.28) | <1.00E-15 | <1.00E-15 |
| SLC12A2 | 12 | 21.91 (15.95–30.1) | <1.00E-15 | <1.00E-15 | 55.14 (40.31–75.43) | <1.00E-15 | <1.00E-15 | 79.32 (58.04–108.41) | <1.00E-15 | <1.00E-15 |
| KCNQ1 | 11 | 7.93 (5.92–10.63) | <1.00E-15 | <1.00E-15 | 13.75 (10.33–18.29) | <1.00E-15 | <1.00E-15 | 6.96 (5.18–9.34) | <1.00E-15 | <1.00E-15 |
| NFKB1 | 10 | 1.87 (1.3–2.71) | 7.81E-04 | 7.81E-03 | 1.92 (1.33–2.77) | 4.69E-04 | 4.69E-03 | 3.79 (2.72–5.29) | 4.44E-15 | 4.44E-14 |
| DPT | 10 | 1.88 (1.2–2.92) | 5.36E-03 | 5.36E-02 | 8.03 (5.5–11.74) | <1.00E-15 | <1.00E-15 | 4.38 (2.94–6.51) | 3.01E-13 | 3.01E-12 |
| GJB2 | 9 | 0.59 (0.31–1.12) | 1.06E-01 | 9.51E-01 | 2.81 (1.77–4.44) | 1.04E-05 | 9.32E-05 | 2.5 (1.57–3.98) | 1.21E-04 | 1.09E-03 |
| CEACAM16 | 9 | 1.25 (1.04–1.49) | 1.56E-02 | 1.40E-01 | 1.33 (1.12–1.59) | 1.48E-03 | 1.33E-02 | 2.45 (2.09–2.87) | <1.00E-15 | <1.00E-15 |
| FAM136A | 8 | 31.18 (27.22–35.71) | <1.00E-15 | <1.00E-15 | 38.36 (33.5–43.92) | <1.00E-15 | <1.00E-15 | 14.71 (12.81–16.89) | <1.00E-15 | <1.00E-15 |
| GRHL2 | 8 | 11.72 (9.35–14.7) | <1.00E-15 | <1.00E-15 | 9.28 (7.39–11.66) | <1.00E-15 | <1.00E-15 | 6.45 (5.11–8.15) | <1.00E-15 | <1.00E-15 |
| EYA4 | 7 | 36.65 (28.83–46.59) | <1.00E-15 | <1.00E-15 | 25.31 (19.88–32.22) | <1.00E-15 | <1.00E-15 | 7.68 (5.97–9.88) | <1.00E-15 | <1.00E-15 |
| COCH | 7 | 3.41 (2.54–4.58) | <1.00E-15 | <1.00E-15 | 3.39 (2.52–4.56) | <1.00E-15 | <1.00E-15 | 3.95 (2.95–5.28) | <1.00E-15 | <1.00E-15 |
| CCDC50 | 5 | 1.25 (0.99–1.59) | 6.22E-02 | 3.11E-01 | 1.62 (1.3–2.03) | 2.51E-05 | 1.25E-04 | 3.22 (2.63–3.94) | <1.00E-15 | <1.00E-15 |
| KCNJ10 | 5 | 4.38 (3.9–4.93) | <1.00E-15 | <1.00E-15 | 3.14 (2.78–3.55) | <1.00E-15 | <1.00E-15 | 2.3 (2.02–2.61) | <1.00E-15 | <1.00E-15 |
| SEMA3D | 4 | 1.56 (1.03–2.38) | 3.60E-02 | 1.44E-01 | 1.13 (0.72–1.77) | 5.84E-01 | 1 | 3.79 (2.63–5.47) | 1.09E-12 | 4.35E-12 |
| CLDN14 | 4 | 4.51 (2.56–7.94) | 1.74E-07 | 6.97E-07 | 22.86 (13.56–38.55) | <1.00E-15 | <1.00E-15 | 4.72 (2.69–8.29) | 6.61E-08 | 2.65E-07 |
| POU4F3 | 4 | 7.31 (6.53–8.18) | <1.00E-15 | <1.00E-15 | 4.92 (4.38–5.52) | <1.00E-15 | <1.00E-15 | 9.2 (8.23–10.28) | <1.00E-15 | <1.00E-15 |
| MSRB3 | 4 | 13.35 (11.34–15.71) | <1.00E-15 | <1.00E-15 | 10.85 (9.21–12.79) | <1.00E-15 | <1.00E-15 | 4.06 (3.4–4.83) | <1.00E-15 | <1.00E-15 |

[a]Odds ratios were calculated in the 95% confidence interval.
[b]P-values were corrected with Bonferroni method.
List of 29 genes showing a significant excess of missense exonic variants in patients with sporadic MD, according to the MAF observed in global ExAC population (N = 60,706), non-Finnish European ExAC population (NFE) (N = 33,370) and Spanish population from CSVS (N = 1579).

**TABLE 4** | Gene burden analysis 2.

| Gene | Number variants | Percentage variants retained | Odds ratio ExAC[a] | P value | P corrected[b] | Odds ratio ExAC NFE[a] | P value | P corrected[b] | Odds ratio CSVS[a] | P value | P corrected[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GJB2 | 6 | 80 | 0.5 (0.27–0.93) | 2.91E-02 | 1.75E-01 | 3.2 (2.12–4.83) | 2.75E-08 | 1.65E-07 | 2.06 (1.33–3.19) | 1.14E-03 | 6.85E-03 |
| SEMA3D | 2 | 50 | 1.1 (0.76–1.61) | 6.13E-01 | 1 | 0.8 (0.53–1.21) | 2.85E-01 | 5.70E-01 | 2.67 (1.94–3.68) | 2.03E-09 | 4.06E-09 |
| CLDN14 | 2 | 50 | 4.47 (2.55–7.83) | 1.67E-07 | 3.35E-07 | 23.18 (13.81–38.9) | <1.00E-15 | <1.00E-15 | 4.64 (2.65–8.11) | 7.47E-08 | 1.49E-07 |
| SLC26A4 | 6 | 40 | 1.18 (0.72–1.93) | 5.04E-01 | 1 | 2.88 (1.89–4.38) | 8.13E-07 | 4.88E-06 | 2.33 (1.51–3.59) | 1.23E-04 | 7.37E-04 |
| NFKB1 | 3 | 30 | 1.37 (0.99–1.91) | 5.92E-02 | 1.78E-01 | 1.43 (1.03–1.98) | 3.37E-02 | 1.01E-01 | 2.73 (2.03–3.66) | 2.21E-11 | 6.62E-11 |
| POU4F3 | 1 | 25 | 1.56 (1.47–1.66) | <1.00E-15 | <1.00E-15 | 1.06 (1–1.14) | 6.16E-02 | 6.16E-02 | 1.84 (1.73–1.95) | <1.00E-15 | <1.00E-15 |
| ESRRB | 3 | 21 | 4.41 (3.31–5.89) | <1.00E-15 | <1.00E-15 | 3.39 (2.52–4.55) | <1.00E-15 | <1.00E-15 | 1.84 (1.33–2.54) | 2.04E-04 | 6.12E-04 |
| USH1G | 5 | 21 | 2.51 (1.38–4.56) | 2.66E-03 | 1.33E-02 | 20.27 (12.06–34.06) | <1.00E-15 | <1.00E-15 | 4.67 (2.68–8.17) | 6.11E-08 | 3.05E-07 |
| CCDC50 | 1 | 20 | 0.54 (0.44–0.66) | 7.23E-10 | 7.23E-10 | 0.7 (0.59–0.84) | 1.41E-04 | 1.41E-04 | 1.38 (1.19–1.61) | 3.38E-05 | 3.38E-05 |
| P2RX2 | 3 | 14 | 2.57 (1.93–3.42) | 8.71E-11 | 2.61E-10 | 3.14 (2.38–4.14) | <1.00E-15 | <1.00E-15 | 2.67 (2.01–3.55) | 1.12E-11 | 3.36E-11 |
| FAM136A | 1 | 13 | 140.36 (131.25–150.11) | <1.00E-15 | <1.00E-15 | 76.75 (71.75–82.09) | <1.00E-15 | <1.00E-15 | 3.68 (3.41–3.97) | <1.00E-15 | <1.00E-15 |
| RDX | 2 | 12 | 4.49 (3.59–5.61) | <1.00E-15 | <1.00E-15 | 2.69 (2.12–3.4) | <1.00E-15 | <1.00E-15 | 3.33 (2.65–4.2) | <1.00E-15 | <1.00E-15 |
| TPRN | 3 | 12 | 2.24 (1.7–2.96) | 1.46E-08 | 4.37E-08 | 6.55 (5.1–8.4) | <1.00E-15 | <1.00E-15 | 2.95 (2.26–3.86) | <1.00E-15 | <1.00E-15 |
| ESPN | 2 | 11 | 10.69 (9–12.69) | 0.00E+00 | <1.00E-15 | 10.26 (8.64–12.19) | <1.00E-15 | <1.00E-15 | 2.01 (1.64–2.46) | 9.24E-12 | 1.85E-11 |
| SLC12A2 | 1 | 8 | 281.71 (246.32–322.17) | <1.00E-15 | <1.00E-15 | 308.08 (269.39–352.33) | <1.00E-15 | <1.00E-15 | 14.77 (12.86–16.96) | <1.00E-15 | <1.00E-15 |
| PRKCB | 1 | 8 | 15.1 (13.03–17.48) | <1.00E-15 | <1.00E-15 | 10.5 (9.05–12.19) | <1.00E-15 | <1.00E-15 | 5.54 (4.75–6.46) | <1.00E-15 | <1.00E-15 |
| MYH14 | 3 | 6 | 14.44 (10.42–20.03) | <1.00E-15 | <1.00E-15 | 28.12 (20.38–38.79) | <1.00E-15 | <1.00E-15 | 5.86 (4.16–8.25) | <1.00E-15 | <1.00E-15 |
| ADD1 | 1 | 3 | 563.42 (492.71–644.27) | <1.00E-15 | <1.00E-15 | 308.08 (269.39–352.33) | <1.00E-15 | <1.00E-15 | 14.77 (12.86–16.96) | <1.00E-15 | <1.00E-15 |

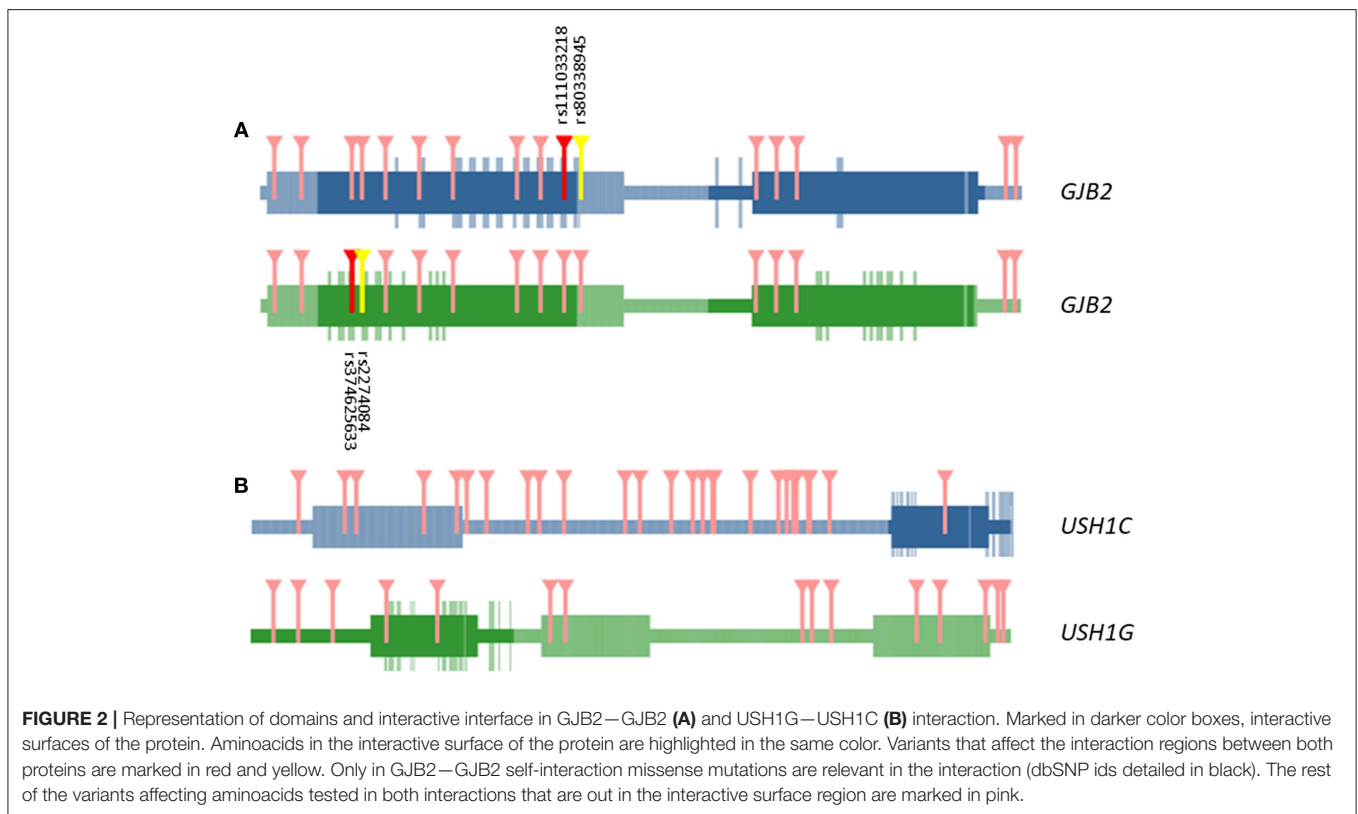[a]Odds ratios were calculated in the 95% confidence interval.

[b]P-values were corrected with Bonferroni method.

List of 18 genes showing a significant excess of previously reported missense exonic variants in patients with sporadic MD, according to the MAF observed in CSVS Spanish database (N = 1579), compared with global ExAC population (N = 60,706) and non-Finnish European ExAC population (N = 33,370). In bold, selected genes with higher percentage of variants retained (>20%) and significant OR on Spanish and NFE populations. Selected SNV can be consulted in **Table S5**.

**TABLE 5 |** Missense variants found in familial MD cases.

| Cases | Family | SNV | Gene | RefSeq | Described in | Pathogenicity | MAF (gnomAD) | MAF (ExAC) |
|---|---|---|---|---|---|---|---|---|
| All | 1 | chr2:70527974 G>A | FAM136A | NM_032822.2 | (Requena et al., 2015) | Pathogenic | | |
| 1 | 1 | chr4:6303197 G>A | WFS1 | NM_006005.3 | No | Conflicting interpretations of pathogenicity | 0.0041 | 0.0045 |
| 2 | 1 | chr13:20763264 C>T | GJB2 | NM_004004.5 | No | Benign/likely benign | 0.0094 | 0.0106 |
| All | 1 | chr18:32462094 G>T | DTNA | NM_001198938.1 | (Requena et al., 2015) | Likely pathogenic | 3.66E-05 | 2.47E-05 |
| 2 | 2 | chr1:41296865 G>A | KCNQ4 | NM_004700.3 | No | Unknown significance | 2.13E-05 | 3.49E-05 |
| All | 2 | chr1:168665849 G>A | DPT | NM_001937.4 | (Martín-Sierra et al., 2017) | Likely pathogenic | 2.03E-05 | 2.5E-05 |
| 2 | 2 | chr4:2900221 A>G | ADD1 | NM_014189.3 | No | Unknown significance | 8.12E-06 | 8.24E-06 |
| All | 3 | chr7:84642128 C>T | SEMA3D | NM_152754.2 | (Martín-Sierra et al., 2017) | Pathogenic | | |
| All | 3 | chr8:102555482 G>T | GRHL2 | NM_024915.3 | No | Unknown significance | | |
| All | 4 | chr16:23999898 G>T | PRKCB | NM_002738.6 | Martín-Sierra et al., 2016 | Pathogenic | | |
| 1 | 4 | chr22:38119405 C>T | TRIOBP | NM_001039141.2 | No | Likely benign | 4.06E-05 | 2.49E-05 |
| 1 | 4 | chr3:191098660 A>G | CCDC50 | NM_178335.2 | No | Benign | 0.0061 | 0.0065 |
| 1 | 4 | chr11:110134833 T>C | RDX | NM_002906.3 | No | Unknown significance | | |

*Variants were retrieved from familial cases segregating a partial phenotype in different families.*



**FIGURE 2 |** Representation of domains and interactive interface in GJB2—GJB2 **(A)** and USH1G—USH1C **(B)** interaction. Marked in darker color boxes, interactive surfaces of the protein. Aminoacids in the interactive surface of the protein are highlighted in the same color. Variants that affect the interaction regions between both proteins are marked in red and yellow. Only in GJB2—GJB2 self-interaction missense mutations are relevant in the interaction (dbSNP ids detailed in black). The rest of the variants affecting aminoacids tested in both interactions that are out in the interactive surface region are marked in pink.

To assess if the SNHL genes showing enrichment of missense variants were located in genomic regions with a higher recombination rates, we retrieved recombination rates from deCODE genetics maps for the *ESRRB, GJB2, USH1G, CLDN14,* and *SLC26A4* genes and calculated linkage disequilibrium correlations for candidate missense variants in these five genes.

*USH1G* and *ESRRB* genes have the highest recombination rates and they seem to be in genomic regions considered as hotspots (**Table S6**). However, most of the rare missense variants found were not clustered and showed a scattered distribution along the different exons with a low recombination rate (**Table S7**).

# DISCUSSION

This study shows that patients with sporadic MD have an enrichment of few rare variants in certain hearing loss genes such as *GJB2*, *SLC26A*, or *USH1G*. This excess of missense variants in some genes may increase the risk to develop hearing loss in MD and may contribute to explain the heterogeneity observed in the phenotype (Francioli et al., 2015). To understand the relevance of population frequencies in our cohort, we performed the association analysis between variants observed in MD cases against their respective frequencies on a healthy population for each gene of the panel (Lek et al., 2016). From the total amount of variants, we selected rare coding variants for all the targeted genes. We applied a stronger filter for the selection of missense variants by choosing previously described variants for each gene significantly overrepresented in MD cases.

Since many missense variants were not found in the Spanish population from CSVS (Dopazo et al., 2016), a third comparison limited to previously reported variants in CSVS database was carried out. We followed this conservative approach to reduce false positive findings in the readings. This third restrictive gene analysis was limited to 132 variants observed at least once in the Spanish reference population.

From the final analysis, we found that some genes such as *SLC26A4*, *ESRRB*, *CLDN14*, *GJB2*, and *USH1G* retained the higher number of missense variants among Spanish MD patients. We also found one novel synonymous variant in the *MARVELD2* gene in 3 unrelated patients. Besides from its functional implications, it may also generate a cryptic splice site. However, more testing is needed to confirm this finding.

## Multiallelic Model for MD

The excess of missense variants in SNHL genes may point to core gene for hearing loss in MD. Our hypothesis is that common cis-regulatory variants and rare variants in one or more genes will contribute to the phenotype in MD. The model will need the additive effect of at least a common and a rare variant in the same gene in a given individual (Castel et al., 2018). In the simplest bi-allelic hypothesis, we will have:

$$Ind\ 1 = cv\ a + rv\ z\ (geneA)$$
$$Ind\ 2 = cv\ b + rv\ y\ (geneB)$$
$$Ind\ 3 = cv\ a + rv\ x\ (geneA)$$
$$Ind\ 4 = cv\ b + rv\ w\ (geneB)$$

Where *cv* is a common variant and *rv* represents a rare variant; however, this model could be more complex for a single gene:

$$Ind\ 1 = cv\ a + cv\ c + rv\ z\ (geneA)$$
$$Ind\ 2 = cv\ b + cv\ d + rv\ z\ (geneB)$$

So, several rare variants will be targeting the core genes (*rv z*, *rv x* for gene A; *rv y*, *rv w* for gene B) and common variants in the same genes will explain variable expressivity of the MD phenotype. Finally, in a more complex scenario, it could involve several genes (oligogenic multiallelic hypothesis):

$$Ind\ n = cv\ a + rv\ z\ (gene\ A) + cv\ b + rv\ y\ (gene\ B)$$
$$+ \ldots + cv\ n + rv\ m\ (gene\ N)$$

## Gene Panel for Familial MD

The Genomics England project (https://www.genomicsengland.co.uk/) has designed gene panels for the diagnosis of many genetic disorders including familial MD (https://panelapp.genomicsengland.co.uk/panels/394/). This panel is in an early stage of development because it only considers 130 genes with limited evidence to few families. The results of this study can be used to improve the design of panels for the diagnosis of MD.

For the design of our panel, we chose a total of 69 genes. Most of the genes were selected according to the hearing loss profile (low frequency or pantonal hearing loss). However, more than 90 genes have been related to hearing loss, so more hearing loss genes could be involved in the phenotype (fully accessible from Hereditary Hearing Loss Homepage: http://hereditaryhearingloss.org/). Genetic evidence of hearing loss has been obtained from linkage analyses until the emergence of NGS techniques (Shearer et al., 1999), that have facilitated the clinical development of genetic diagnosis in hearing loss. Custom panels and microarrays have been the flags of a new age of discovery of novel and rare variants for genetic diagnostic of hearing loss (Brownstein et al., 2011; Shearer and Smith, 2015).

Our panel was designed considering hearing loss as the main symptom shared by all patients with MD, since the vestibular phenotype and other associated co-morbidities such as migraine or autoimmune disorders are more variable. To improve the diagnostic yield of MD and to decrease this granularity in the phenotype, it will be recommendable to select sporadic patients with an early age of onset for future studies.

## Rare Missense Variants in Hearing Loss Genes in Sporadic MD

The frequency of hearing loss related genes is population-specific (Sloan-Heggen et al., 2016). Herein, we present a study for MD patients in the Spanish population. As a part of the study, we consider a panel of genes related to hearing loss and other symptoms. Besides from the validated variants in singletons, only a few rare variants such as *ESRRB* rs201448899:C>T, *MARVELD2* rs369265136:G>A, *SLC26A4* rs200511789:A>C, and *USH1G* rs151242039:C>T have been validated in more than one sporadic case in the entire cohort. All these genes had been previously considered as pathogenic for hearing loss, but they have never been involved with MD.

*ESRRB* encodes the estrogen-related receptor beta, also known as nuclear receptor subfamily 3, group B, member 2 or NR3B2. This gene encodes for a protein like the estrogen receptor but with a different and unknown role. Mutations in the mouse ortholog have been involved in the placental development and autosomal recessive SNHL (Collin et al., 2008; Weber et al., 2014).

*MARVELD2* encodes a protein found in the tight junctions, between epithelial cells. The encoded protein seems to forge barriers between epithelial cells such the ones in the organ of Corti, Defects in this gene are associated with DFNB49 (Mašindová et al., 2015).

*SLC26A4* gene encodes pendrin, a protein extensively studied in hearing loss. Its alteration is one of the most common causes of syndromic deafness and autosomal recessive SNHL. It is also associated with enlarged vestibular aqueduct syndrome (EVAS) [(Yang et al., 2007), 36].

*USH1G* is a gene translating to a protein that contains three ankyrin domains, a class I PDZ-binding motif and a sterile alpha motif. This protein is well-known to interact with harmonin (*USH1C*) in the stereocilia of hair cells, a protein associated with Usher syndrome type 1C (Weil et al., 2003). This protein plays a role in the development and maintenance of the auditory and visual systems and functions in the cohesion of hair bundles formed by inner ear sensory cells. Alterations in the integrity of the protein seem to be the cause of Usher syndrome type 1G (Weil et al., 2003; Miyasaka et al., 2016).

However, *ESRRB* rs201448899:C>T has been observed in more Spanish controls than in global or NFE in ExAC. This increased frequency on the Iberian population when compared with other known largest frequencies as NFE, suggests that this is a population specific variant rather than a MD disease variant. Only the *MARVELD2* rs369265136:G>A variant remains as a proper novel related to MD cases. However, the functional effect of a synonymous variant is unknown and functional studies will be required to decipher the relevancy of this variant in MD cases in the future.

## Burden Analysis of Rare Missense Variants in Sporadic MD

Our results demonstrate a burden of rare missense variants in few SNHL genes, including *GJB2, ESRRB, CLDN14, SLC26A4,* and *USH1G*. We speculate that the additive effect of several missense variants in the same gene could interact with the same or other genes at the protein level resulting in the hearing loss phenotype.

Population analysis was addressed in order to obtain a better image of our cohort. Despite the limitation that represents the small number of genes considered in our panel, we have found a significant increase of missense variants on several hearing loss genes in the Iberian population (**Table 4**). These findings suggest the involvement of multiple missense variants in the same gene and may explain several clinical findings in MD. So, incomplete phenotype found in relatives of patients with familial MD or even the variable expressivity observed could be explained on the differences found in multiple rare variants with additive effect among individuals of the same family (Requena et al., 2015; Martín-Sierra et al., 2016, 2017). In addition, some sporadic cases where a single rare variant with unknown significance cannot explain the phenotype could be singletons individuals with low frequency variants probably following a compound heterozygous recessive pattern of inheritance. Our results start to decipher the complex interaction between rare and ultrarare variations (MAF <0.0001) with common variants in the same or different genes in sporadic MD, adding more evidence to understand the genetic architecture of MD. However, one of the limitations of this study is the lack of availability of a replication cohort with different ethnicity in which to validate these findings.

Another limitation of our dataset is that the method used for resequencing mitochondrial genes may not be able to distinguish mitochondrial from nuclear sequences, as capture panels such as those based in the *Haloplex* technology may sequence all mitochondrial genome fragment replicas that are dispersed throughout the nuclear genome. Hence, variants observed may not belong to the genes targeted in the mitochondrial genome, but to their pseudogenes in the nuclear genome.

Several hypotheses could explain the excess of missense variants in SNHL genes in MD. First, the variable expressivity of SNHL in MD phenotype, could be the result of additive effect of low frequency or rare variants in the same gene. The combination of low frequency variants in the same gene can be a rare situation, as rare as the disease. As much changes are added to the protein, its integrity could be affected, showing a suboptimal functioning and finally, a loss of function. In our case, GJB2, that forms a hexamer with a transmembrane channel function, has been determined as possible affected by these changes in their interactions. Previous studies have determined how certain changes in the monomer can affect to the develop of the hexamer hemichannel (Bicego et al., 2006; Jara et al., 2012). Here, bioinformatics models show how the interaction of low frequency variants found in MD patients can impact the interaction between two connexins monomers, but this effect could be amplified in a model including the 6 connexins that form the connexon. However, this hypothesis is difficult to reconcile with the fact that for some small genes such as GJB2, complex alleles with several point mutations are exceedingly rare.

A second hypothesis points to the interaction of common and rare variants in one or several genes in the disease phenotype, following its complex disease definition (Becker, 2004; Mitchell, 2012). So, the excess of rare variants will be targeting core genes for hearing loss in MD. In this case, high significant genes in our study could be added to the panel of candidate targets of the disease, although a single variant could not be enough to explain the disease phenotype. So, the interaction between cis-regulatory variants with rare variant in some of our candidate genes and other, a priori, not related SNHL genes could be relevant in the expressivity. *USH1G* interacts with *USH1C*, a known gene involved in Usher syndrome. *USH1G* has been observed to have a minor role in Usher syndrome in Spanish population (Aller et al., 2007), but not in MD, even though they share similar hearing loss profile. Although no one of the missense variants in *USH1G* were in an interaction domain, this could be of interest when considering interaction between different proteins as a main factor to develop a mild phenotype. This hypothesis was reinforced through the data found in familial cases. For instance, the variant rs748718975 in *DPT* gene was only associated with the SNHL phenotype in the family where it was described, but these cases showed different characteristics in the age of onset or hearing loss outcome. These differences between the cases can be explained with other variants found in *KCNQ4* (rs574794136:G>A) and *ADD1* (rs372777117:A>G) genes, although these variants were previously described as variants of unknown significant. So, the variant rs574794136:G>A was found in two sisters with MD, but not in the third one, that was carrier of rs372777117:A>G. This excess of rare variants in certain genes observed in familial cases could explain the differences in expressivity in a given family.

This panel was made as an early screening diagnostic panel. Here we have found that certain SNHL gene variants can be related to MD in the Iberian population and the results show that multiple rare allelic variants in the same gene should be consider

as likely pathogenic. Although there are large differences in the coverage for some genes between the MD panel and the WES databases, these are not the ones with excess of missense variants. Our results will contribute to design a novel gene panel for the genetic diagnosis of MD.

## DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

TR and JL-E as project managers conceived the main idea of the work. Sample preparation and protocol were carry out by TR and AG-M. AG-M performed the bioinformatics and statistical analysis. Validations of tested SNV were done by AG-M and PR-N. AG-M, and JL-E took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

## FUNDING

## ACKNOWLEDGMENTS

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00076/full#supplementary-material

## REFERENCES

Aller, E., Jaijo, T., Beneyto, M., Nájera, C., Morera, C., Pérez-Garrigues, H., et al. (2007). Screening of the USH1G gene among Spanish Patients with Usher Syndrome. Lack of mutations and evidence of a minor role in the pathogenesis of the syndrome. *Ophthalm. Genet.* 28, 151–155. doi: 10.1080/13816810701537374

Arweiler-Harbeck, D., Horsthemke, B., Jahnke, K., and Hennies, H. C. (2011). Genetic aspects of Familial Ménière's disease. *Otol. Neurotol.* 32, 695–700. doi: 10.1097/MAO.0b013e318216074a

Becker, K. G. (2004). The common variants/multiple disease hypothesis of common complex genetic disorders. *Med. Hypotheses* 62, 309–317. doi: 10.1016/S0306-9877(03)00332-3

Bicego, M., Beltramello, M., Melchionda, S., Carella, M., Piazza, V., Zelante, L., et al. (2006). Pathogenetic role of the deafness-related M34T mutation of Cx26. *Hum. Mol. Genet.* 15, 2569–2587. doi: 10.1093/hmg/ddl184

Brownstein, Z., Friedman, L. M., Shahin, H., Oron-Karni, V., Kol, N., Abu Rayyan, A., et al. (2011). Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in middle Eastern Families. *Genome Biol.* 12:R89. doi: 10.1186/gb-2011-12-9-r89

Cabrera, S., Sanchez, E., Requena, T., Martinez-Bueno, M., Benitez, J., Perez, N., et al. (2014). Intronic variants in the NFKB1 gene may influence hearing forecast in patients with unilateral sensorineural hearing loss in Meniere's disease. *PLoS ONE* 9:e112171. doi: 10.1371/journal.pone.0112171

Castel, S. E., Cervera, A., Mohammadi, P., Aguet, F., Reverter, F., Wolman, A., et al. (2018). Modified penetrance of coding variants by cis-regulatory variation contributes to disease risk. *Nat. Genet.* 50, 1327–1334. doi: 10.1038/s41588-018-0192-y

Caulley, L., Quimby, A., Karsh, J., Ahrari, A., Tse, D., and Kontorinis, G. (2017). *Autoimmune Arthritis in Ménière's Disease: A Systematic Review of the Literature. Semin. Arthritis Rheum.* 48, 141–147. doi: 10.1016/j.semarthrit.2017.11.008

Collet, A., Tarabeux, J., Girard, E., Enghien, C. D. D., Golmard, L., and Deshaies, V., Lermine, A., et al. (2015). Pros and cons of haploplex enrichment in cancer predisposition genetic diagnosis. *Genetics* 2, 263–280. doi: 10.3934/genet.2015.4.263

Collin, R. W., Kalay, E., Tariq, M., Peters, T., van der Zwaag, B., Venselaar, H., et al. (2008). Mutations of ESRRB encoding estrogen-related receptor beta cause autosomal-recessive nonsyndromic hearing impairment DFNB35. *Am. J. Hum. Genet.* 82, 125–138. doi: 10.1016/j.ajhg.2007.09.008

Dallol, A., Daghistani, K., Elaimi, A., Al-Wazani, W. A., Bamanie, A., Safiah, M., et al. (2016). Utilization of amplicon-based targeted sequencing panel for the massively parallel sequencing of sporadic hearing impairment patients from Saudi Arabia. *BMC Med. Genet.* 17(Suppl. 1):67. doi: 10.1186/s12881-016-0329-8

Dopazo, J., Amadoz, A., Bleda, M., Garcia-Alonso, L., Alemán, A., García-García, F., et al. (2016). 267 Spanish exomes reveal population-specific differences in disease-related genetic variation. *Mol. Biol. Evol.* 33, 1205–1218. doi: 10.1093/molbev/msw005

Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., et al. (2015). Genome-wide patterns and properties of de novo mutations in humans. *Nat. Genet.* 47, 822–826. doi: 10.1038/ng.3292

Gallego-Martinez, A., Requena, T., Roman-Naranjo, P., and J. A., Lopez-Escamez (2018). Excess of rare missense variants in hearing loss genes in sporadic meniere disease. *BioRxiv [preprint]* 393322. doi: 10.1101/393322

Gazquez, I., Soto-Varela, A., Aran, I., Santos, S., Batuecas, A., Trinidad, G., et al. (2011). High prevalence of systemic autoimmune diseases in patients with Menière's disease. *PLoS ONE* 6:e26759. doi: 10.1371/journal.pone.0026759

Gu, X., Guo, L., Ji, H., Sun, S., Chai, R., Wang, L., et al. (2015). Genetic testing for sporadic hearing loss using targeted massively parallel sequencing identifies 10 novel mutations. *Clin. Genet.* 87, 588–593. doi: 10.1111/cge.12431

Hietikko, E., Kotimäki, J., Kentala, E., Klockars, T., Sorri, M., and Männikkö, M. (2011). Finnish Familial Meniere disease is not linked to chromosome 12p12.3, and anticipation and cosegregation with migraine are not common findings. *Genet. Med.* 13, 415–420. doi: 10.1097/GIM.0b013e3182091a41

Hietikko, E., Kotimäki, J., Sorri, M., and Männikkö, M. (2013). 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, 279–285. doi: 10.1016/j.ejmg.2013.03.010

Jara, O., Acuña, R., García, I. E., Maripillán, J., Figueroa, V., Sáez, J. C., et al. (2012). Critical role of the first transmembrane domain of Cx26 in regulating oligomerization and function. *Mol. Biol. Cell* 23, 3299–3311. doi: 10.1091/mbc.e11-12-1058

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536 (7616): 285–291. doi: 10.1038/nature19057

Lempert, T., Olesen, J., Furman, J., Waterston, J., Seemungal, B., Carey, J., et al. (2012). Vestibular migraine: diagnostic criteria. *J. Vestib. Res.* 22, 167–172. doi: 10.3233/VES-2012-0453

Lionel, A. C., Costain, G., Monfared, N., Walker, S., Reuter, M. S., Hosseini, S. M., et al. (2018). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet. Med.* 20, 435–443. doi: 10.1038/gim.2017.119

Lopez-Escamez, J. A., Carey, J., Chung, W. H., Goebel, J. A., Magnusson, M., Mandalà, M., et al. (2015). Diagnostic criteria for Menière's disease. *J. Vestib. Res.* 25, 1–7. doi: 10.3233/VES-150549

Martín-Sierra, C., Gallego-Martinez, A., Requena, T., Frejo, L., Batuecas-Caletrío, A., and Lopez-Escamez, J. A. (2017). Variable expressivity and genetic heterogeneity involving DPT and SEMA3D genes in autosomal dominant familial Meniere's disease. *Eur. J. Hum. Genet.* 25, 200–207. doi: 10.1038/ejhg.2016.154

Martín-Sierra, C., Requena, T., Frejo, L., Price, S. D., Gallego-Martinez, A., Batuecas-Caletrio, A., et al. (2016). A novel missense variant in PRKCB segregates low-frequency hearing loss in an autosomal dominant family with Meniere's disease. *Hum. Mol. Genet.* 25, 3407–3415. doi: 10.1093/hmg/ddw183

Mašindová, I., Šoltýsová, A., Varga, L., Mátyás, P., and Ficek, A., Hučková, M., et al. (2015). MARVELD2 (DFNB49) mutations in the hearing impaired central european roma population - prevalence, clinical impact and the common origin. *PLoS ONE* 10:e0124232. doi: 10.1371/journal.pone.0124232

Meyer, M. J., Beltrán, J. F., Liang, S., Fragoza, R., Rumack, A., Liang, J., et al. (2018). Interactome INSIDER: a structural interactome browser for genomic studies. *Nat. Methods* 15, 107–114. doi: 10.1038/nmeth.4540

Mijovic, T., Zeitouni, A., and Colmegna, I. (2013). Autoimmune sensorineural hearing loss: the otology–rheumatology interface. *Rheumatology* 52, 780–789. doi: 10.1093/rheumatology/ket009

Mitchell, K. J. (2012). What is complex about complex disorders? *Genome Biol.* 13:237. doi: 10.1186/gb-2012-13-1-237

Miyasaka, Y., Shitara, H., Suzuki, S., Yoshimoto, S., Seki, Y., Ohshiba, Y., et al. (2016). Heterozygous mutation of Ush1g/Sans in mice causes early-onset progressive hearing loss, which is recovered by reconstituting the strain-specific mutation in Cdh23. *Hum. Mol. Genet.* 25, 2045–2059. doi: 10.1093/hmg/ddw078

Morales Angulo, C., Gómez Castellanos, R., García Mantilla, J., Bezos Capelastegui, J. T., and Carrera, F. (2003). Epidemiology of Menière's disease in Cantabria. *Acta Otorrinolaringol. Esp.* 54, 601–605. doi: 10.1016/S0001-6519(03)78456-8

Paparella, M. M., and Griebie, M. S. (1984). Bilaterality of Meniere's disease. *Acta Oto Laryngol.* 97, 233–37.

Requena, T., Cabrera, S., Martín-Sierra, C., Price, S. D., Lysakowski, A., and Lopez-Escamez, J. A. (2015). Identification of two novel mutations in FAM136A and DTNA genes in autosomal-dominant familial Meniere's disease. *Hum. Mol. Genet.* 24, 1119–1126. doi: 10.1093/hmg/ddu524

Requena, T., Espinosa-Sanchez, J. M., Cabrera, S., Trinidad, G., Soto-Varela, A., Santos-Perez, S., et al. (2014a). Familial clustering and genetic heterogeneity in Meniere's disease. *Clin. Genet.* 85, 245–252. doi: 10.1111/cge.12150

Requena, T., Espinosa-Sanchez, J. M., and Lopez-Escamez, J. A. (2014b). Genetics of dizziness: cerebellar and vestibular disorders. *Curr. Opin. Neurol.* 27, 98–104. doi: 10.1097/WCO.0000000000000053

Requena, T., Gazquez, I., Moreno, A., Batuecas, A., Aran, I., Soto-Varela, A., et al. (2013). Allelic variants in TLR10 Gene may influence bilateral affectation and clinical course of Meniere's disease. *Immunogenetics* 65, 345–355. doi: 10.1007/s00251-013-0683-z

Ruiz-Pesini, E., Lott, M. T., Procaccio, V., Poole, J. C., Brandon, M. C., Mishmar, D., et al. (2007). An enhanced MITOMAP with a global MtDNA mutational phylogeny. *Nucleic Acids Res.* 35, D823–D828. doi: 10.1093/nar/gkl927

Shearer, A. E., Hildebrand, M. S., and Smith, R. J. H. (1999). "Hereditary hearing loss and deafness overview," in *GeneReviews®*, eds M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, and A. Amemiya (Seattle, WA: University of Washington). Available online at: http://www.ncbi.nlm.nih.gov/books/NBK1434/ (Accessed December 20, 2018).

Shearer, A. E., and Smith, R. J. (2015). Massively parallel sequencing for genetic diagnosis of hearing loss: the new standard of care. *Otolaryngol. Head Neck Surg.* 153, 175–182. doi: 10.1177/0194599815591156

Sloan-Heggen, C. M., Bierer, A. O., Shearer, A. E., Kolbe, D. L., Nishimura, C. J., Frees, K. L., et al. (2016). Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss. *Hum. Genet.* 135, 441–450. doi: 10.1007/s00439-016-1648-8

The Molecular Otolaryngology and Renal Research Laboratories, The University of Iowa. *Deafness Variation Database*. Available online at: http://deafnessvariationdatabase.org/ (Accessed June 30, 2018).

Van Camp, G., and Smith, R. J. H. (2018). *Hereditary Hearing Loss Homepage*. Available online at: https://hereditaryhearingloss.org (Accessed December 20, 2018).

Weber, M. L., Hsin, H. Y., Kalay, E., Brožková, D. S., Shimizu, T., Bayram, M., et al. (2014). Role of Estrogen Related Receptor Beta (ESRRB) in DFN35B hearing impairment and dental decay. *BMC Med. Genet.* 15:81. doi: 10.1186/1471-2350-15-81

Weil, D., El-Amraoui, A., Masmoudi, S., Mustapha, M., Kikkawa, Y., Lainé, S., et al. (2003). Usher syndrome type I G (USH1G) is caused by mutations in the gene encoding SANS, a protein that associates with the USH1C protein, harmonin. *Hum. Mol. Genet.* 12, 463–471. doi: 10.1093/hmg/ddg051

Yang, T., Vidarsson, H., Rodrigo-Blomqvist, S., Rosengren, S. S., Enerback, S., and Smith, R. J. (2007). Transcriptional control of SLC26A4 is involved in pendred syndrome and nonsyndromic enlargement of vestibular aqueduct (DFNB4). *Am. J. Hum. Genet.* 80, 1055–1063. doi: 10.1086/518314