



OPEN

Predicting pathogenicity for novel hearing loss mutations based on genetic and protein structure approaches

Paula I. Buonfiglio^{1,9}, Carlos D. Bruque^{1,9}, Vanesa Lotersztein³, Leonela Luce^{1,4}, Florencia Giliberto^{1,4}, Sebastián Menazzi^{1,5}, Liliana Francipane⁵, Bibiana Paoli⁶, Ernesto Goldschmidt⁷, Ana Belén Elgoyhen^{1,8} & Viviana Dalamón¹✉

Hearing loss is a heterogeneous disorder. Identification of causative mutations is demanding due to genetic heterogeneity. In this study, we investigated the genetic cause of sensorineural hearing loss in patients with severe/profound deafness. After the exclusion of *GJB2-GJB6* mutations, we performed whole exome sequencing in 32 unrelated Argentinean families. Mutations were detected in 16 known deafness genes in 20 patients: *ACTG1*, *ADGRV1* (*GPR98*), *CDH23*, *COL4A3*, *COL4A5*, *DFNA5* (*GSDDE*), *EYA4*, *LARS2*, *LOXHD1*, *MITF*, *MYO6*, *MYO7A*, *TECTA*, *TMPRSS3*, *USH2A* and *WSF1*. Notably, 11 variants affecting 9 different non-*GJB2* genes resulted novel: c.12829C>T, p.(Arg4277*) in *ADGRV1*; c.337del, p.(Asp109*) and c.3352del, p.(Gly1118Alafs*7) in *CDH23*; c.3500G>A, p.(Gly1167Glu) in *COL4A3*; c.1183C>T, p.(Pro395Ser) and c.1759C>T, p.(Pro587Ser) in *COL4A5*; c.580+2T>C in *EYA4*; c.1481dup, p.(Leu495Profs*31) in *LARS2*; c.1939T>C, p.(Phe647Leu), in *MYO6*; c.733C>T, p.(Gln245*) in *MYO7A* and c.242C>G, p.(Ser81*) in *TMPRSS3* genes. To predict the effect of these variants, novel protein modeling and protein stability analysis were employed. These results highlight the value of whole exome sequencing to identify candidate variants, as well as bioinformatic strategies to infer their pathogenicity.

Deafness affects approximately 1 out of 500–1000 newborns and is mainly of genetic origin. The genetic causes that lead to hearing loss can be categorized into syndromic and nonsyndromic conditions, which constitute 30% and 70% of the genetic forms, respectively¹. Over 400 distinct syndromes that include hearing impairment are listed in Online Mendelian Inheritance in Man (OMIM), the most common including Alport, Pendred, Usher and Waardenburg^{1–3}. Despite the wide genetic heterogeneity of Nonsyndromic autosomal recessive hearing loss, a few mutations in the *GJB2* and *GJB6* genes (encoding connexin-26 and 30, respectively) account for nearly 50% of the cases in the Mediterranean population^{4–8}. To date, approximately 100 genes and more than 1000 mutations causing nonsyndromic deafness have been reported: more than 72 genes associated with autosomal recessive nonsyndromic deafness (named locus DFNB1 to 108) and nearly 44 with autosomal dominant forms, named locus DFNA1 to 73 (<http://hereditaryhearingloss.org>). Molecular diagnosis of hearing loss (HL) remains a challenge due to the high number of genes involved. Genetic screening consists of analyzing many and sometimes long

¹Laboratory of Physiology and Genetics of Hearing. Instituto de Investigaciones en Ingeniería Genética y Biología Molecular “Dr. Héctor Torres”, Consejo Nacional de Investigaciones Científicas y Técnicas, INGEBI/CONICET, Vuelta de Obligado 2490- (C1428ADN), Ciudad Autónoma de Buenos Aires, Argentina. ²Unidad de Conocimiento Traslacional Hospitalaria Patagónica, Hospital de Alta Complejidad SAMIC - El Calafate, Provincia de Santa Cruz, Argentina. ³Servicio de Genética, Hospital Militar Central “Dr. Cosme Argerich”, C1426 Ciudad Autónoma de Buenos Aires, Argentina. ⁴Laboratorio de Distrofinopatías. Cátedra de Genética, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, C1113AAD Ciudad Autónoma de Buenos Aires, Argentina. ⁵División Genética, Hospital de Clínicas “José de San Martín”, C1120AAR Ciudad Autónoma de Buenos Aires, Argentina. ⁶Sector de Otorrinolaringología Infantil, Hospital de Clínicas “José de San Martín”, C1120AAR Ciudad Autónoma de Buenos Aires, Argentina. ⁷Laboratorio Diagnogen, Ciudad Autónoma de Buenos Aires, Argentina. ⁸Instituto de Farmacología, Facultad de Medicina, Universidad de Buenos Aires, C1121ABG Ciudad Autónoma de Buenos Aires, Argentina. ⁹These authors contributed equally: Paula I. Buonfiglio and Carlos D. Bruque. ✉email: dalamon@dna.uba.ar

genes for mutations, making conventional methods (e.g., Sanger sequencing) expensive and time-consuming⁹. Except for mutations in *GJB2*, most deafness mutations are seen in only a single or few families¹⁰. Moreover, in a limited number of cases, the characteristics of hearing impairment (audiogram shape, hearing impairment onset and progression or family history) pinpoint the gene to test. Therefore, an approach that assays for most or all of the known genes at the same time would ease and accelerate diagnosis. Whole-exome sequencing (WES) has thus become an efficient and cost-effective alternative approach for molecular diagnosis of hearing impairment. After the exclusion of *GJB2* and *GJB6* in patients, new genes and new genetic variants in reported genes have been described using WES in the research and clinical molecular diagnosis of syndromic and nonsyndromic forms of hearing impairment. Moreover, WES is gradually being integrated into the routine genetic diagnostics of Mendelian diseases, including hereditary HL^{11–14}. However, the follow-up of novel variants, in particular missense changes, which can lead to a spectrum of phenotypes and unequivocal genotype-to-phenotype correlations, is not always straightforward. This study presents a custom-designed multistep methodology to evaluate the impact of genetic variants found in patients with HL on protein function. This highlights the importance of developing combined molecular protein structure studies together with database analysis to evaluate and characterize the impact of reported and novel gene variants.

Results

Validation of variants and genetic diagnosis. Thirty-two patients with different forms of hearing loss were studied by WES, followed by data filtering based on 183 genes reported for this pathology. Approximately 90,000 variants were identified for each patient studied. Filtering reduced the list of candidate variants to 1–10 for each patient. After analysis following the American College of Medical Genetics and Genomics (ACMG) and Hearing Loss Variant Curation Expert Panel (HL-EP) parameters, variants were classified from benign to pathogenic in each patient and confirmed by Sanger sequencing and family segregation when available (full data are detailed in Supplementary Figure S1).

Overall, 27 different mutations were identified in the following 16 genes: *ACTG1*, *ADGRV1* (*GPR98*), *CDH23*, *COL4A3*, *COL4A5*, *DFNA5* (*GSDDE*), *EYA4*, *LARS2*, *LOXHD1*, *MITE*, *MYO6*, *MYO7A*, *TECTA*, *TMPRSS3*, *USH2A* and *WSF1* in 20 of the 32 studied patients (62.5%). The other 12 cases that remained undiagnosed had recessive variants in the heterozygous state or variants that failed to segregate with the pathology in the family. Therefore, since they did not fulfill the ACMG criteria, they were not reported as positive. Since neither large insertions/deletions nor repetitions or deep intronic variants were studied, exome data remain available for further analysis.

In positive cases, segregation within the family was performed in 16 out of the 20 cases, and variant inheritance was established. All variants are listed in Table 2, together with the patients' auditory phenotype (pedigrees are detailed in Fig. 1 and Supplementary Figure S1).

Among the 27 different mutations identified in 9 genes, 11 were novel, since they were not reported in deafness and Leiden Open Variation Database (LOVD) databases or in PubMed as related to pathology: c.12829C>T, p.(Arg4277*) in *ADGRV1*; c.337del, p.(Asp109*) and c.3352del, p.(Gly1118Alafs*7) in *CDH23*; c.3500G>A, p.(Gly1167Glu) in *COL4A3*; c.1183C>T, p.(Pro395Ser) and c.1759C>T, p.(Pro587Ser) in *COL4A5*; c.580+2T>C in *EYA4*; c.1481dup, p.(Leu495Profs*31) in *LARS2*; c.1939T>C, p.(Phe647Leu), in *MYO6*; c.733C>T, p.(Gln245*) in *MYO7A* and c.242C>G, p.(Ser81*) in *TMPRSS3* genes. Values for each bioinformatic predictor, as well as its final interpretation using the Varsome platform manually adjusted, are detailed in Supplementary Table S4.

Five of the thirty-two patients under study presented features compatible with syndromic forms of hearing loss: 2 Alport, 2 Usher and 1 Waardenburg type 2 (WS2) (Table 1). Both Alport and Usher cases resulted in causative mutations in genes related to a syndromic phenotype, and the Waardenburg patient remained undiagnosed (nevertheless large deletions and/or insertions cannot be ruled out by this approach). Three additional cases presenting nonsyndromic HL at consultation resulted in pathogenic variants in genes associated both with syndromic and nonsyndromic forms: 2 Usher syndrome and 1 Perrault syndrome. These molecular genetic results might modify the future clinical outcome, genetic counseling and clinical follow-up to the patients and families. Of the 20 patients with identified variants, 13 required no further studies since they had nonsyndromic hearing loss.

Clinical characteristics of relevant cases. Case #1. Two novel variants in *CDH23* (NM_022124.5) were identified, c.337del, p.(Val113*) and c.3353del, p.(Gly1118Alafs*7), in a 23-year-old patient with cochlear implant in childhood (Fig. 1A). Both variants were associated with Usher Type 1D and resulted in truncated proteins, probably with no residual function. According to HL-EP recommendations, they were both interpreted as pathogenic. Both parents were healthy carriers by Sanger sequencing. The patient already showed some initial retinopathy compatible with Usher signs, further indicating that the mutations were causative of the phenotype and that the signs will progress.

Case #2. The patient had postlingual bilateral moderate hearing loss with a U-shaped audiogram, which was caused by a novel heterozygous variant in the *MYO6* gene (NM_004999.4): c.1939T>C in exon 18 p.(Phe647Leu). The patient had four affected siblings (not available for molecular diagnosis). Thus, the variant was linked to a dominant form of inheritance (DFNA22). To further study the effect of the variant in the *MYO6* structure, a model of the affected motor domain bearing the p.(Phe647Leu) mutation was generated (Fig. 3).

Case #3. The patient had postlingual bilateral moderate hearing loss that was caused by a novel heterozygous variant in *EYA4*: c.580+2T>C (splicing site). The variant cosegregated with the affected mother in the family, consistent with a dominant form of inheritance (Fig. 1C).

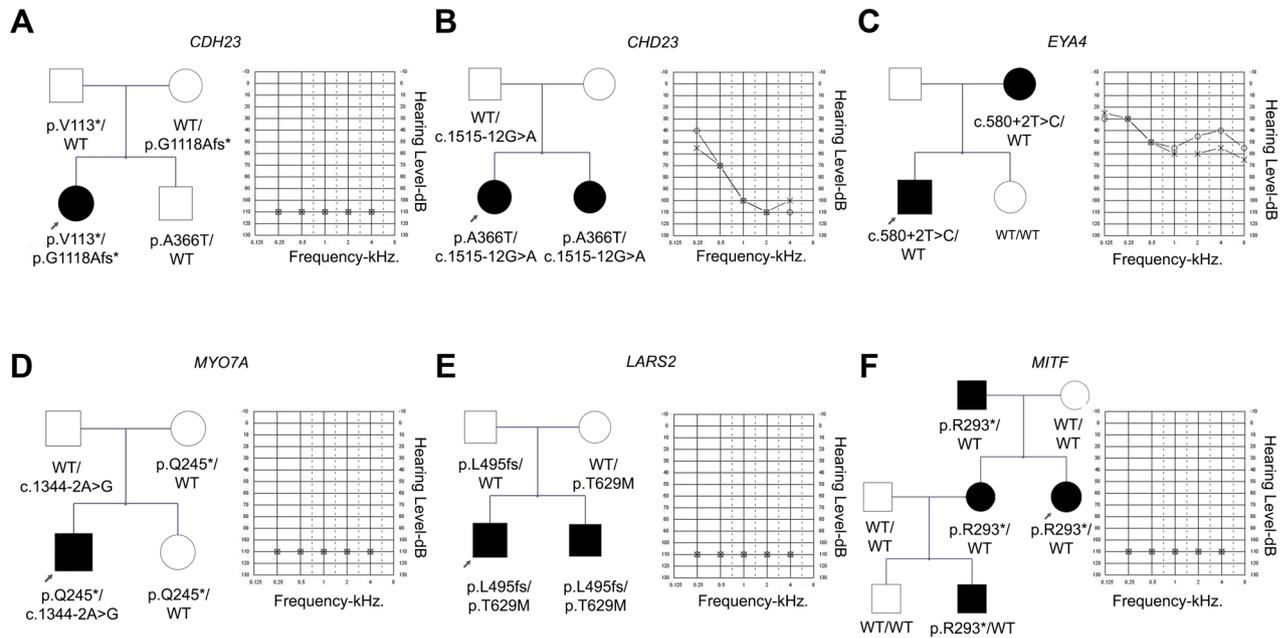


Figure 1. Pedigrees and audiograms of some of the families. All variants were identified by whole-exome sequencing and confirmed by Sanger sequencing. (A) Two novel variants in the *CDH23* gene were identified, c.337del, p.(Val113*) and c.3353del, p.(Gly1118Alafs*7), in a patient with Usher signs. (B) Two previously reported variants were identified in *CDH23*: c.1515-12G > A, reclassified as likely pathogenic after manual curation, and c.1096 G > A, p.(Ala366Thr) classified as benign based on its high population frequency. (C) Postlingual bilateral moderate hearing loss caused by a novel heterozygous variant in *EYA4*: c.580+2T > C (splicing). (D) One-year-old boy with nonsyndromic isolated prelingual hearing loss and no retinal or vestibular pathologies at the time of study. Novel variants c.733C > T, p.(Gln245*) and c.1344-2A > G (splicing site mutation previously reported in ClinVar) in *MYO7A* were detected. (E) Two congenital bilateral profound cochlear implanted sisters with variants in *LARS2*: novel c.1481dup, p.(Leu495Thrfs*31*) and previously reported c.1886C > T, p.(Thr629Met). (F) Previously reported nonsense mutation c.877C > T, p.(Arg293*) in *MITF* cosegregated with pathology in four affected members of the family with nonsyndromic hearing loss.

| Characteristic | Number |
|----------------------------|--------|
| Sex | |
| Male | 14 |
| Female | 18 |
| No family history | |
| Family History | 16 |
| Autosomal recessive | |
| Autosomal dominant | 13 |
| Onset | |
| Congenital | 16 |
| Postlingual | 16 |
| Physical exam | |
| No other signs | 27 |
| Syndromic | |
| Alport | 2 |
| Usher | 2 |
| Waardenburg | 1 |

Table 1. Reported phenotype characteristics of the 32 patients evaluated in this study.

Case #4. Two variants in *MYO7A* (NM_000260.3) were identified: c.1344-2A > G (splicing site mutation reported in ClinVar for Usher Syndrome type I, either in homozygous or compound heterozygous state; rs111033415), in trans with the novel nonsense variant c.733C > T, p.(Gln245*). As this nonsense variant produces a stop codon, it will most likely result in a truncated protein. Thus, both of the variants meet the criteria to be classified as pathogenic for Usher syndrome type 1B or DFNB2 in an autosomal recessive manner. Segrega-

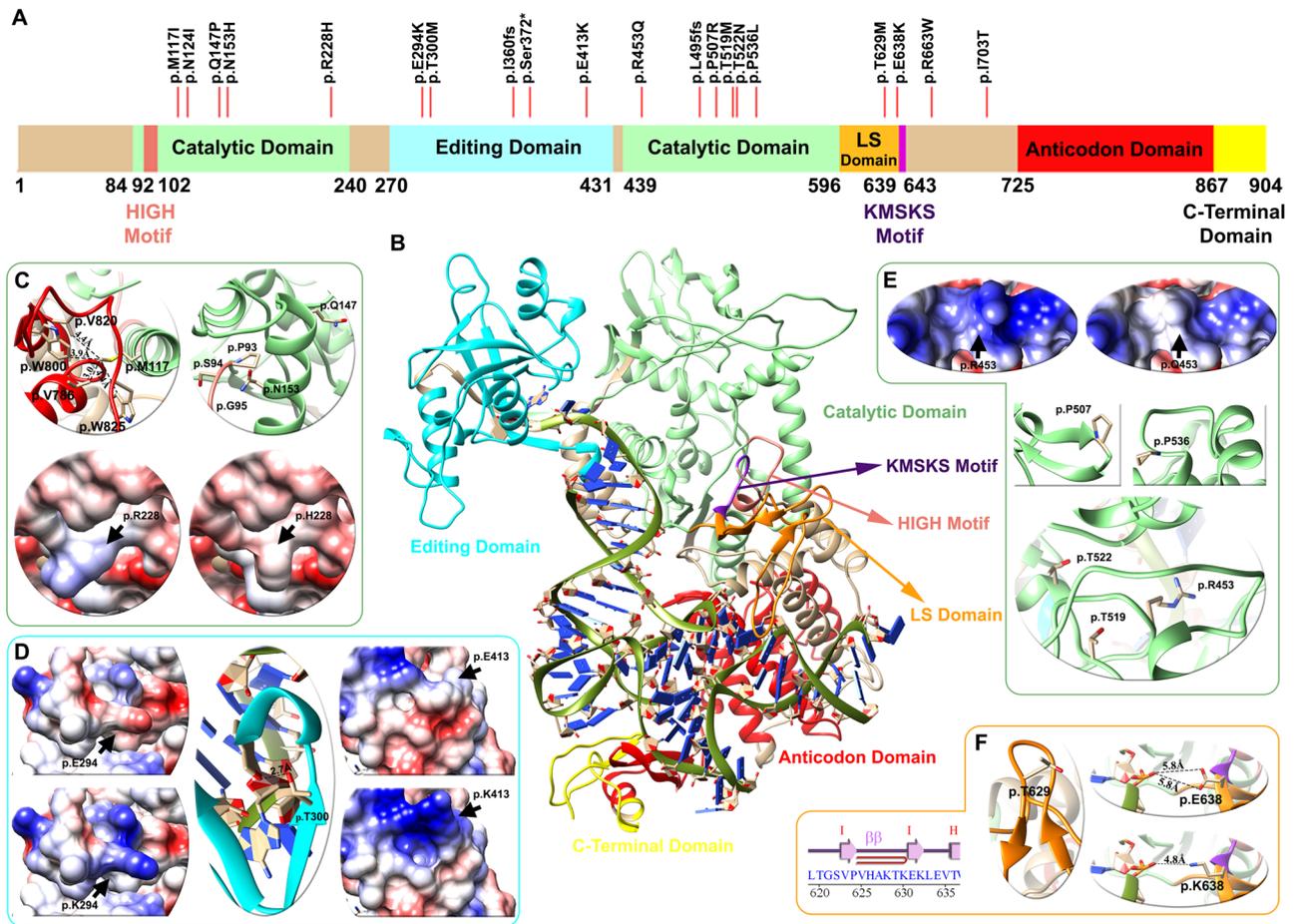


Figure 2. Domain architecture, mapping of variants and structural analysis of the LeuRS protein (*LARS2* gene). (A) Linear representation of the LeuRS protein with its domains and motifs: HIGH motif (pink), catalytic domain (light green), editing domain (cyan), LS domain (orange), KMSKS motif (purple), anticodon domain (red), and C-terminal domain (yellow). Red lines depict the location of the p.Thr629Met variant of case # 5 patient and 19 other pathogenic/likely pathogenic variants found in databases. (B) Human LeuRS molecular homology model, with the representation of domains and motifs. The zoom in shows variant analysis performed in three domains: light green for the catalytic (C,E), cyan for the editing (D) and orange for the leucine-specific (LS) (F) regions. For each domain mutations affect the electrostatic surface of the protein as well as the distance between neighboring residues. The p.Thr629Met variant lies in the LS domain between the hairpin of beta strand I–I, altering the folding of this loop and compromising the stability of the region (Zoom F). Detailed information regarding the genetic variants analyzed can be found in Table 3.

tion in the family was confirmed by Sanger sequencing. The affected proband, a 1-year-old boy, presented isolated prelingual hearing loss at the time of genetic diagnosis, with no retinal or vestibular pathologies (Fig. 1D). Since syndromic and nonsyndromic forms have been reported due to mutations in the *MYO7A* gene, patient clinical follow-up was recommended.

Case #5. A familial case with two affected cochlear implanted siblings (8- and 12-year-old boys). Two variants were identified in *LARS2* (NM_015340.3): c.1886C > T, p.Thr629Met was previously reported only once in ClinVar related to Perrault syndrome, and the novel c.1481dup, p.(Leu495Thrfs*31) (Fig. 1E). This latter novel frameshift mutation is predicted to be pathogenic based on its truncating effect on LeuRS, leading to the loss of its catalytic, leucine-specific and anticodon-binding domains. Pathogenicity was further confirmed by the deleterious effect of the mutation on the LeuRS structure through molecular modeling analysis (Fig. 2). Segregation analysis indicated that the parents were carriers for the mutations.

Case #7. A family case with a proband diagnosed with prelingual bilateral profound sensorineural hearing loss and three other affected members with a similar phenotype (son, sister and mother), who were cochlear implanted. The reported nonsense mutation c.877C > T (NM_000248.3) in *MITF* was found and predicted to be pathogenic, leading to an early truncated and nonfunctional protein p.(Arg293*). The variant cosegregated with the pathology in all of the affected members of the family. The same mutation has been reported in a family with

Waardenburg syndrome type 2 (WS2)¹⁵. It is striking that no Waardenburg signs were observed in any of the members of the family (Fig. 1F).

Case #8. An affected girl (15 years old) with isolated postlingual hearing loss and a sloping audiogram was diagnosed at 11 years old. Heterozygous variants were detected in *TMPRSS3* (NM_024022.2): c.1276 G > A, p.(Ala426Thr), reported several times as likely pathogenic in ClinVar supporting a deleterious effect (rs56264519) and the novel c.733C > T, p.(Ser81*) mutation. Parents were found to be carriers of these mutations, consistent with recessive inheritance.

Case #11. A 27-year-old patient with high frequency hearing loss. Two variants were identified in *CDH23*: c.1515-12G > A, previously reported as variant of uncertain significance (VUS) and now reclassified as likely pathogenic, and c.1096G > A, p.(Ala366Thr) classified as benign based on its high population frequency (BA1 applied) (Fig. 1B). It remains unclear whether these genetic findings are related to the pathology in this family, since the proband had no retinopathies or vestibular abnormalities, as seen in Usher Type D syndromes associated with mutations in *CDH23* (OMIM #601067).

Variant curation. After WES analysis, 28 different variants were found in 20 patients. Sixteen variants were already reported in the ClinVar database: 5 pathogenic, 6 pathogenic/likely pathogenic, 1 VUS, 3 conflicting interpretations and 1 benign (Table 2). We reanalyzed the 16 reported variants according to the Expert Panel specified recommendations for ACMG rules and/or manually adjusted them with evidence of segregation within the family and data from the literature. Eleven of these 16 variants changed their previous category (69%) (Table 2).

For instance, one variant changed from pathogenic to likely pathogenic, [p.(Thr629Met) in *LARS2*], two from likely pathogenic to pathogenic [p.(Arg1890Cys) in *TECTA*] and [p.(Lys118Met) in *ACTG1*], three from likely pathogenic/pathogenic to pathogenic [p.(Ala426Thr) in *TMPRSS3*, p.(Thr3571Met) in *USH2A*, p.(Arg1494*) in *LOXHD1*] and one from likely pathogenic/pathogenic to likely pathogenic [p.(Glu864Lys) in *WFS1*]. Interestingly, two variants changed from conflicting interpretation to benign after reanalysis [p.(Lys41Glu) in *DFNA5* and p.(Glu776Val) in *WFS1*], and another changed to VUS [p.(Val1550Gly) in *COL4A3*]. Variant c.1515-12G > A in *CDH23* reported as VUS in the ClinVar database was reclassified as likely pathogenic, reinforcing its causal relationship with the pathology.

In the case of novel variants, classification was established following the standard protocol as per ACMG/AMP for variant interpretation for genetic hearing loss and the updated recommendations of the ClinGen Hearing Loss Expert Panel, and in some cases (*LARS2* and *MYO6*), their final classification was further established through molecular modeling and in silico strategies.

Combined in silico analysis. Functional assays are essential for the interpretation of missense variants associated with pathology. However, experiments for functional validation are time consuming and not always feasible in the clinical context. Therefore, bioinformatic tools that predict protein malfunction appear to be valid predictable tools of pathogenicity. We implemented full modeling and domain modeling as bioinformatic approaches to determine the in silico implications of missense variants.

Full modeling of *LeuRS* (*LARS2* gene). The *LARS2* gene encodes a mitochondrial leucyl-tRNA synthetase (*LeuRS*) that catalyzes the aminoacylation of a specific tRNA. The protein architecture of *LeuRS* includes motifs that are catalytically important (HIGH and KMSKS) and different domains: catalytic, editing, leucine-specific (LS), anticodon-recognition, and the C-terminal domains (C-ter) (Fig. 2). Sequence variants in *LARS2* have been previously associated with Perrault syndrome, characterized by premature ovarian failure, hearing loss and other severe multisystem metabolic disorders (OMIM #604544).

In Case #9, we identified two variants: c.1481dup, p.(Leu495Thrfs*31) and c.1886C > T (p.Thr629Met). The former is novel and predicted to yield a truncated nonfunctional protein. Although the latter variant has already been reported, we performed a deeper follow-up analysis to understand the impact of the mutation on the translated protein. Thus, we conducted molecular modeling of the entire human *LeuRS* and analyzed its stability, electrostatic surface and tRNA interaction. In addition to the p.Thr629Met variant found in the proband of Case #5, we included in our analysis 17 additional missense variants reported as likely pathogenic or pathogenic for the *LARS2* gene in LOVD, Deafness Variation Database and ClinVar (2 frameshift and 1 nonsense reported variant were not included). These variants are identified above the primary structure of *LeuRS* (Fig. 2A). Most of the variants were located in the catalytic, LS and editing domains, and none were located in the C-terminal or anticodon domain.

The model shows that protein stability was altered in 7 of the variants analyzed (41%), the electrostatic charge in 3 (18%) and the tRNA-protein interaction in 3 (18%) of the previously reported mutations. The analysis was nonconclusive for 4 variants (24%) (Table 3).

The p.Thr629Met variant detected in the patient of Case #5 lies in the LS domain between the hairpin of beta strand I-I altering the folding of this loop, compromising the stability of the region (Fig. 2F and Table 3). An additional previously reported variant, p.(Glu638Lys), also lies in the LS domain and alters the interaction with de t-RNA through the change of a negative to a positive side chain of the residue (Fig. 2F). Ten genetic variants are located in the catalytic domain of the protein. According to the model, p.Met117 (light green helix in Fig. 2C) interacts with residues p.Trp800, p.Val820, p.Trp825 and p.Val786, of the anticodon domain (red helix in Fig. 2C). The methionine to isoleucine change in p.(Met117Ile) results in altered protein stability (Table 3),

| ID | Gene (Transcript ID) | Genotype | Change | dbSNP | Phenotype of patient | Segregation (inheritance) | Reference | ClinVar report | After curation | Criteria applied |
|----|---------------------------------|------------------------|--------------------|-------------|---|---|---------------|---------------------------------------|-------------------|---------------------------------|
| 1 | CDH23 (NM_022124.5) | c.337del | p.(Val113*) | – | Usher syndrome PL, PR, B, PF, CI | Maternal | This work | – | Pathogenic | PM2, PVS1, PM3, PP4 |
| | | c.3353del | p.(Gly1118Alafs*7) | – | | Paternal | This work | – | Pathogenic | PM2, PVS1, PM3 and PP4 |
| 2 | MYO6 (NM_004999.4) | c.1939T>C | p.(Phe647Leu) | rs752585373 | PL, B, M | Non available | This work | – | VUS | PM2, PP3 |
| 3 | EYA4 NG_011596.2 (NM_004100.5) | c.580+2T>C | splicing | – | PL, B, M | Maternal | This work | – | Pathogenic | PM2, PVS1, PP1_Sup |
| 4 | MYO7A NG_009086.2 (NM_000260.4) | c.733C>T | p.(Gln245*) | – | C, PL, B, PF, CI | Maternal | This work | – | Pathogenic | PVS1, PM2, PM3 |
| | | c.1344-2A>G | Splicing | rs111033415 | | Paternal | ⁶⁵ | Pathogenic | Pathogenic | PM2, PVS1, PM3_S, PP4 |
| 5 | LARS2 (NM_015340.3) | c.1481dup | p.(Leu495Thrfs*31) | rs762797278 | C, B, PF, CI | Paternal | This work | – | Pathogenic | PVS1, PM2, PM3, PP1_Sup |
| | | c.1886C>T | p.(Thr629Met) | rs398123036 | | Maternal | ¹⁶ | Pathogenic | Likely Pathogenic | PM2, PM3_S, PP1_Sup, PP4 |
| 6 | ADGRV1 /GPR98 (NM_032119.3) | c.12829C>T | p.(Arg4277*) | – | Usher syndrome B, P, PR | Non available father | This work | – | Pathogenic | PVS1, PM2, PM3, PP1_Sup, PP4 |
| | | c.956dup | p.(Asn319Lysfs*6) | rs752179149 | | Maternal | ⁶⁶ | Pathogenic | Pathogenic | PVS1, PM2, PM3, PP1_Sup, PP4 |
| 7 | MITF (NM_000248.3) | c.877C>T | p.(Arg293*) | – | C, B, PF, CI | Segregation confirmed | ¹⁵ | – | Pathogenic | PM2, PVS1_S, PP1_S |
| 8 | TMPRSS3 (NM_024022.3) | c.1276G>A | p.Ala426Thr | rs56264519 | PL, B, sloping audiometry | Maternal inheritance | ⁶⁷ | Pathogenic/Likely Path | Pathogenic | BS1_Sup, PM3_VS, PP1_S, PS3_Sup |
| | | c.242C>G | p.(Ser81*) | rs757110501 | | Paternal inheritance | This work | – | Pathogenic | PVS1, PM2, PM3, PP1_Sup |
| 9 | WFS1 (NM_006005.3) | c.2590G>A | p.(Glu864Lys) | rs74315205 | B, CI | | ⁶⁸ | Pathogenic/Likely Path | Likely Pathogenic | PM2, PS4_M, PP1_Mod, PP3 |
| 10 | USH2A NG_009497.2 (NM_206933.4) | c.1841-2A>G | Splicing | rs397518003 | PR, B, M. No retinopathies | Maternal inheritance | ⁶⁹ | Pathogenic | Pathogenic | PM2, PM3_VS, PP4 PP1_M, PS3_S |
| | | c.10712C>T | p.(Thr3571Met) | rs202175091 | | Paternal inheritance | ⁷⁰ | Pathogenic/Likely Path | Pathogenic | PM2, PM3_VS, PP4, PP1_M |
| 11 | CDH23 NG_008835.1 (NM_022124.5) | c.1096G>A | p.(Ala366Thr) | rs143282422 | B, High-frequency affected No retinopathies | Maternal | ⁷¹ | Benign | Benign | BA1 |
| | | c.1515-12G>A | splicing | rs369396703 | | Paternal | – | VUS (validated by HL-EP) | Likely Pathogenic | PM2_Sup, PM3, PP1_Sup, PP3, PP4 |
| 12 | COL4A3 (NM_000091.5) | c.3500G>A | p.(Gly1167Glu) | – | Alport Syndrome. Hematuria | De novo (maternal) | This work | – | Pathogenic | PM2, PS2, PM1, PM5, PP3 |
| | | c.4649T>G | p.(Val1550Gly) | rs200655479 | | | ⁷² | Conflicting Interpretation (VUS/LP) | VUS | PM2_Sup, PP3 |
| 13 | DFNA5 (NM_004403.2) | c.119dup | p.(Lys41Glufs*113) | rs758488919 | No familial history | De novo | – | Conflicting Interpretation (VUS/LB) | Benign | BA1, PS2 |
| 14 | COL4A5 (NM_000495.3) | c.1183C>T | p.(Pro395Ser) | – | C, B, S | | This work | – | VUS | PM2 |
| 15 | COL4A5 (NM_000495.3) | c.1759C>T | p.(Pro587Ser) | – | PL, PR, B, M | | This work | – | VUS | PM2, PP3 |
| 16 | COL4A5 (NM_000495.3) | c.3659G>A | p.(Gly1220Asp) | rs104886251 | Alport Syndrome. Hematuria | | ⁷³ | Pathogenic | Pathogenic | PM2, PM1, PP3, PP4, PS4_Sup |
| 17 | WFS1 (NM_006005.3) | c.2327A>T | p.(Glu776Val) | rs56002719 | B, M, PR. High frequencies | Maternal inheritance | ⁷⁴ | Conflicting Interpretation (VUS/B/LB) | Benign | BA1, PP3, PP1_Sup, BS4 |
| 18 | TECTA (NM_005422.4) | c.5668C>T | p.(Arg1890Cys) | rs121909063 | PL, M-S | Segregation confirmed. Paternal Inheritance | ⁷⁵ | Likely pathogenic | Pathogenic | PM2, PP1_VS, PS4_Sup |
| 19 | LOXHD1 (NM_144612.6) | c.4480C>T (homozygous) | p.(Arg1494*) | rs201587138 | C, B, PF | Segregation confirmed | ⁷⁶ | Pathogenic/Likely Path | Pathogenic | PVS1, BS1_Sup, PM3_S, PP1_M |
| 20 | ACTG1 (NM_001614.5) | c.353A>T | p.(Lys118Met) | rs104894544 | PL, B, M-S | Segregation confirmed. Paternal Inheritance | ⁷⁷ | Likely pathogenic | Pathogenic | PS4_Sup, PM2, PP5, PP1_S, PP3 |

Table 2. Relevant Variants identified by WES. All variants were curated following the Hearing Loss Expert Panel recommendations. The phenotype of the patients is indicated as follows: *C* congenital, *PL* postlingual, *PR* progressive, *B* bilateral, *M* moderate, *PF* profound, *S* severe, *CI* cochlear implanted.

| Variant and amino acid change (NM_015340.3) | Effect † | Stability + | ClinVar | Deafness Variation Database/LOVD | Hearing loss expert panel classification + Modeling | Reference |
|---|----------------------------------|--------------|---------|----------------------------------|--|---|
| c.351G>C; p.(Met117Ile) | Stability | 4.11 ± 0.63 | – | P/LP | Likely Pathogenic (PM2, PM3, PP1, PP4, PP3) | PMID: 26,970,254 |
| c.371A>T; p.(Asn124Ile) | Non conclusive | 0.41 ± 0.66 | P | P/– | Likely Pathogenic (PM2, PM3_Strong, PP4) | PMID: 28,708,303 |
| c.440A>C; p.(Gln147Pro) | Stability | 1.49 ± 0.12 | LP | LP/– | Likely Pathogenic (PM2, PM3, PP4, BP4_Supporting, PP3) | SCV000994657.1 |
| c.457A>C; p.(Asn153His) | Stability | 3.36 ± 0.82 | LP | LP/– | Likely Pathogenic (PM2, PM3_PP3, PP4) | PMID: 32,423,379 |
| c.683G>A p.(Arg228His) | Electrostatic Surface | 0.40 ± 0.02 | LP | LP/LP | Likely Pathogenic (PM2, PM3_Supporting, PP3, PP4) | PMID: 28,000,701 |
| c.880G>A; p.(Glu294Lys) | Electrostatic surface | 0.67 ± 0.14 | – | P/LP | Likely Pathogenic (PM2, PM3_Strong, PP4, PP3) | PMID: 28,000,701; 3,276,773; 29,205,794 |
| c.899C>T; p.(Thr300Met) | tRNA interaction | 0.12 ± 0.84 | P | P/P | Likely Pathogenic (PM2, PM3, PP1, PP3, PP4) | PMID: 26,657,938 |
| c.1077del; p.Ile360fs | LoF | – | P | P/P | Pathogenic (PVS1, PS3_Supporting, PM2, PM3, PP4) | PMID: 23,541,342 |
| c.1115C>G; p.(Ser372*) | LoF | – | LP | LP/– | Pathogenic (PVS1, PM2, PP4) | SCV000891207.1 |
| c.1237G>A; p.(Glu413Lys) | Electrostatic surface | 0.14 ± 0.02 | LP | LP/– | VUS (PM2, BP4, PP4) | SCV001244305.1 |
| c.1358G>A; p.(Arg453Gln) | Electrostatic surface/Stability? | 1.42 ± 0.3 | – | P/P | Likely Pathogenic (PM2_Supporting, PM3, PP3, PP4) | PMID: 27,650,058 |
| c.1481dup; p.(Leu495fs) | LoF | – | – | P/– | Pathogenic (PVS1, PM2, PM3, PP1_Supporting) | This study paper |
| c.1520C>G; p.(Pro507Arg) | Stability | 2.20 ± 0.07 | LP | LP/– | Likely Pathogenic (PM2, PM3, PP1_Supporting, PP3) | SCV000731430.1 |
| c.1556C>T; p.(Thr519Met) | & | –0.44 ± 0.25 | – | P/– | Likely Pathogenic (PM2, PM3, PP1_Supporting, PP3, PP4) | PMID: 29,205,794 |
| c.1565C>A; p.(Thr522Asn) | & | –0.59 ± 0.08 | LP | P/LP | Likely Pathogenic (PM2_Supporting, PM3_Strong, PS3_Supporting, PP3, PP4) | PMID: 23,541,342 |
| c.1607C>T; p.(Pro536Leu) | Stability | 6.82 ± 1.87 | LP | LP/– | Likely Pathogenic (PM2, PS3_Supporting, PM3, PP3, PP4) | Accession: SCV000994658.1 |
| c.1886C>T; p.(Thr629Met) | Stability | 2.56 ± 0.19 | P | P/P | Likely Pathogenic (PM2, PM3_Strong, PP1_Sup, PP4) | PMID: 23,541,342 |
| c.1912G>A; p.(Glu638Lys) | tRNA interaction | –0.06 ± 0.50 | P | P/P | Likely Pathogenic (PM2, PM3, PP1_Supporting, PP3, PP4) | PMID: 26,657,938 |
| c.1987C>T; p.(Arg663Trp) | Stability | 2.94 ± 0.32 | P | P/– | Likely Pathogenic (PM2, PS3_Supporting, PM3_Strong, PP3, PP4) | PMID: 28,708,303 |
| c.2108T>C; p.(Ile703Thr) | tRNA interaction | 0.67 ± 0.11 | – | LP/– | Likely Pathogenic (PM2, PM3, PP3, PP4) | PMID: 32,767,731 |

Table 3. Evaluation of genetic variants in LeuRS. †Classification of variants according to structural criteria. + $\Delta\Delta G$ Energy evaluation for pathogenic and likely pathogenic genetic variants, FOLDX: $|X| \pm SD$ ($n = 5$). *LoF* loss of function, *P* pathogenic, *LP*: likely pathogenic. This information was compiled from the LOVD3, ClinVar and deafness variation databases until 26 January 2021. &: both residues are oriented facing the core, causing a probable steric effect. The new model of LeuRS was considered a new parameter (PP3 score applied) to classify the variants reported in databases according to the Hearing Loss Expert Panel classification.

as do p.Gln147Pro and p.Asn153His variants found in the H4 α -Helix. In addition, the p.(Asn153His) change affects protein stability, particularly the interaction between p.Pro93, p.Ser94 and p.Gly95 of the HIGH motif. The variant p.(Arg228His) abolishes a negative charge, altering the protein electrostatic surface (bottom Fig. 2C).

Under the magnifying glass of the present model, variant p.Arg453Gln lies within the catalytic domain of the protein, in contrast with the previously reported model that positioned it in the t-RNA binding domain¹⁶. Thus, the pathogenicity of p.(Arg453Gln) is related to a combination of a change in surface electrostatic charge and a steric effect of this residue in the region (Table 3 and Fig. 2E). Mutations p.(Thr519Met) and p.(Thr522Asn) affect

the structure of the catalytic domain through a steric effect. In addition, p.(Pro507Arg) affects the structure of the loop between the G-G β -strand, and p.(Pro536Leu) affects the structure of the H19-H20 α -helix (Fig. 2E, green).

Three missense variants were found in the editing domain. Variants p.(Glu294Lys) and p.(Glu413Lys) generate a change from glutamic acid to arginine, decreasing the electrostatic surface charge. The p.Thr300 residue, mutated to M in a previously reported patient¹⁷, is crucial for leucine-tRNA edition in the β -strand of this domain, as observed in the center of Fig. 2D.

The new model of LeuRS presented in the present work can be considered a new parameter (PP3 score applied) to classify variants reported in databases according to the Hearing Loss Expert Panel classification. The final classification of reported *LARS2* variants is detailed in Table 3.

Modeling of the N-terminal motor (head) domain of the MYO6 protein. Myosins are actin-based motor molecules with ATPase activity. Myosin VI is a reverse-direction motor protein that moves toward the minus-end of actin filaments and plays a crucial role in the organization of the stereociliary bundle and the maintenance of the cuticular plate anchoring of stereocilia rootlets in hair cells¹⁸. Mutation p.(Phe647Leu) in *MYO6* detected in Case #6 was not previously reported in ClinVar and was not found in the GnomAd (Genome Aggregation Database) or 1000 Genomes databases. To determine the potential effect of the missense variant on *MYO6* function, we further performed in silico studies. The bioinformatic analysis predicted that c.1939T > C, (p.Phe647Leu) is a disease causing variant by Mutation Taster (0,81), damaging by PolyPhen-2 HumDiv (score 0.986), damaging by SIFT (0), deleterious by CADD, damaging by PROVEAN (-5,7) and pathogenic by the UMD-Predictor (score 75), REVEL (0,9).

With the aim of further analyzing the impact of the mutation on the protein, specific protein domain modeling was performed in two conformations. The root mean-square deviation (RMSD) values of the protein backbone distances between the two *MYO6* motor domain conformations (pre-powerstroke and Pi release) were first calculated to determine whether the residue is located in a motile region. In particular, two high RMSD values were obtained for amino acid positions 395 to 405 and 621 to 649 in the two conformations. Therefore, amino acid changes in these regions could affect the interaction with nearby residues, affecting the proper function of the protein (Fig. 3B). Our analysis indicates that residue Phe647 is indeed located in a motile region between both conformations, resulting in a significant influence on the structure of the protein. This change leads to a steric effect in the amino acid side chain, altering interactions with nearby nonpolar residues. In particular, when Phe647 is mutated to Leu, the distance between the side chains of the three nearest residues (less than 5 Å in the wild type), Ile457, Cys476 and Leu651, is significantly increased. For the pre-powerstroke state conformation, the average distance between residue 647 and these amino acids is 4.98 Å (Phe) and 8.42 Å (Leu). For the Pi release state conformation, the average is 3.95 Å (Phe) and 7.64 Å (Leu). Thus, in both the pre-powerstroke and Pi release conformations, the average distance was almost twofold increased in the Leu647 variant (distances are detailed in Fig. 3D). In addition, model analysis shows that the alpha helix containing the Phe647 residue is adjacent to the actin binding region (Fig. 3C, in green β -strand). Hence, it is possible to infer that alterations in the surrounding areas may affect the structure and/or function of the actin binding site.

Discussion

Genotype–phenotype characterization in HL patients is not a straightforward endeavor. Nonequivocal genotypic information is crucial for the clinical care and genetic counseling of HI patients. Gene variants leading to frameshift and nonsense mutations or affecting canonical splice sites most likely lead to a null translation of the mutated allele. However, predicting the effect of DNA substitutions leading to missense mutations is far more complicated. These can lead to a myriad of effects that end in protein malfunction, including altered stability, nonfunctional protein domains and lack of catalytic function. Compiled by field experts gathered in the Hearing Loss Expert Panel, genotype validation is accelerating. In the present work, we report new gene variants in a cohort of Argentinian HI patients. These were curated following the ACMG and Hearing Loss Expert Panel guidelines. Moreover, the pathogenicity of some of the variants was validated through in silico protein analysis. In addition, the pathogenicity of previously reported variants in the ClinVar database was reclassified. The present work adds to the standardization of HI variant interpretation as a crucial step to provide consistent and accurate diagnoses for families and professionals involved, as well as for a better understanding of the mechanisms underlying disease pathogenesis.

Overall, 27 different mutations were identified in 16 hearing loss genes in 20 out of the 32 patients studied. The rate of genetic diagnosis was 63%, significantly higher than the 36% standard of care (*GJB2/6* sequencing only), previously reported in our laboratory^{8,19,20}. There was a significant diversity in the overall diagnostic rate. In patients with a family history, the diagnostic rate was 50% (8/16), while in isolated cases, 69% were diagnosed (11/16). This bias can be attributed to the fact that the isolated patients who were diagnosed had mostly syndromic forms, where the genetic target to be studied is more enclosed and has a higher diagnostic success rate. The curation procedure was effective since eleven of the sixteen (69%) reported variants changed or refined their category previously reported in ClinVar. In this regard, rigorous variant manual curation demonstrates its importance in accurate variant interpretation and hence precise genetic counseling to patients. These outcomes are in accordance with our previous report¹⁹.

The two sisters of Case #11 with nonsyndromic hearing loss exhibited two variants in *CDH23*, a gene that encodes a putative cell adhesion protein with multiple cadherin-like domains. *CDH23* is responsible for both Usher syndrome and DFNB12 nonsyndromic deafness¹. It has been established that nonsense, splice-site, or frameshift mutations in *CDH23* are related to Usher 1D, whereas missense mutations are related to a milder phenotype (nonsyndromic HL, DFNB12). Moreover, concerning the type of hearing loss, previous genotype–phenotype reports showed that the majority of patients have some residual hearing at lower frequencies and a

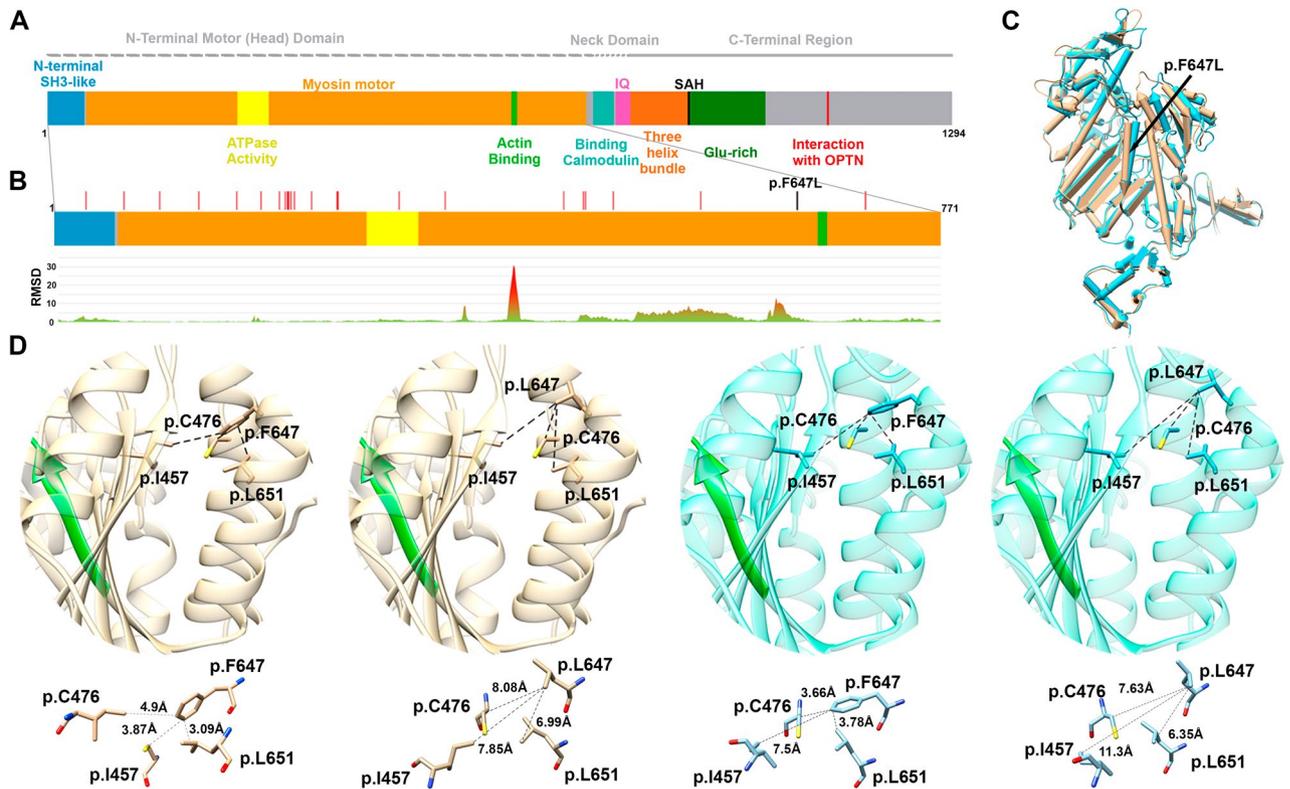


Figure 3. Motor head domain analysis of MYO6 protein. (A) Ideogram of MYO6. (B) RMSD analysis of the motor head domain shows peaks with high RMSD values, revealing that two areas (one including the p.(F647L)) are involved in regions with great motility between pre-powerstroke and Pi release conformations. (C) Motor head domain modeling. Pre-powerstroke conformation in light blue and Pi release in brown. The p.(Phe647Leu) is shown with a black arrow. (D) Zoom in of the alpha-helix at the 647 region. The interaction of the wild-type Phe647 residue or the mutated version Leu647 with the three nearest residues is shown for each protein conformation (Pi release in brown and pre-powerstroke in light blue). The β -strand in green represents the actin binding region. Distances in Å are detailed under each bubble, showing the increase in the distances for the mutated Leu647 residue in both conformations.

characteristic high frequency hearing loss sloping pattern. This is in accordance with the audiogram of our studied siblings^{21,22}. After the curation process, variant c.1515-12G > A in *CDH23*, applying criteria PM2_Supporting, PP3, PM3, PP4 and PP1_Supporting, changed its previous classification in ClinVar from VUS to Likely Pathogenic. Further functional analysis, as well as new reports in Usher patients would provide confident evidence concerning its pathogenicity. Variant c.1096G > A; p.(Ala366Thr) was classified as benign due to its high population frequency in the Ashkenazi population (BA1 criteria applied). Hidden variants in gene regions that we did not explore, such as deep intronic mutations that can disrupt transcription regulatory motifs and/or noncoding RNA regions, might underlie the observed phenotype^{22,23}. Taking into account all the evidence presented we cannot unequivocally conclude that these two variants are the cause of the hearing loss in this family. Nevertheless, due to the reclassification of c.1515-12G > A to likely pathogenic, the case is worth mentioning and discussed.

In Case #10 a nonsense variant was detected in *MITF*. Most mutations in *MITF* have been mostly associated with Waardenburg syndrome type 2, a dominant syndromic form of hearing loss. It is associated with hypopigmentation of the skin, hair and eyes, since *MITF* has a regulatory effect on *TYR* transcription, a key enzyme involved in melanin synthesis²⁴. The variant identified in our patients had already been reported in a Waardenburg type 2 family case¹⁵. Notably, none of the four affected members of this family presented any other signs in addition to profound HL. This finding is in accordance with some reports in which variants in the *MITF* gene cause only hearing loss^{25,26}. The variable phenotype expression could be explained by the presence of modifier genes, as well as interactions with environmental factors^{27,28}.

Several bioinformatic algorithms have been developed to predict the functional consequences of single nucleotide variants in protein coding regions. These in silico approaches are an alternative to tedious and time-consuming experimental approaches to infer pathogenicity. Thus, the new model presented in the present work for the LeuRS human protein aided in defining the plausible pathogenicity of the identified p.Thr629Met variant in Case #9. Moreover, it can be further used to predict the pathogenicity of other reported variants. The p.Thr629Met mutation in the patient lies in the beta hairpin of LeuRS between strands I-I of the LS domains and leads to the disruption of the motif and protein stability. This is in accordance with the observation that beta hairpin motifs are implicated in protein stability^{29,30}. Moreover, the substitution of p.Ala508 in *Escherichia coli* LeuRS (analogous to human LeuRS p.Thr629) with a nonpolar methionine disrupts the structure and/or

position of the leucine-specific domain and thus shifts the location of the KMSKS loop and reduces the catalytic efficiency^{16,31}. It should be noted that the present model of human LeuRS is an improved version of those previously reported^{16,32} since we used a variety of selected crystals. In this regard, the present model shows that most of the variants in the catalytic domain produce a severe effect on protein stability, affecting the proper function of the protein. Our new model shows that variants reported as pathogenic in databases effectively induce significant structural changes of the protein that would cause potential functional changes, playing an important role in genotype/phenotype prediction. Thus, we propose that this model should be used for future *in silico* predictions of variant pathogenicity in patients with Perrault syndrome, becoming a bridge between genomics and structural data to guide the interpretation of human genetic variants.

In Case #6, we identified a novel *MYO6* variant c.1939T > C, p.(Phe647Leu) in a family with late-onset autosomal dominant nonsyndromic hearing loss. Most *MYO6*-known genetic variants present progressive hearing loss^{33–35} and myosin VI is required for structural integrity and proper functioning of inner ear hair cells^{18,36}. More than half of the known pathogenic mutations are located in the motor head domain, indicating that it plays an important role in the function of *MYO6*³⁴. The motor-head domain modeling approach applied in the present work reveals that mutation p.(Phe647Leu) alters the proper function, most likely through the increased distance of this amino acid from the three nearby nonpolar residues. Moreover, the structural change caused by the mutation in the alpha-helix of this highly motile region could potentially affect the contiguous actin binding region. In this regard, mutations in the *MYO6* motor domain alter anchoring of the membrane of stereocilia to actin filaments, leading to disruption of hair bundle organization^{36,37}.

In conclusion, the present work highlights the importance of the curation of genetic variants leading to HL following recommendations of experts for the correct phenotype-genotype correlation. Moreover, we show the importance of the incorporation of integrated workflows for predicting the biomedical impact of the variations identified by exome analysis. Most importantly, we propose a multitarget approach including genomics, protein structure and data analysis to guide the interpretation and standardization of human genetic variants leading to hearing loss.

Materials and methods

Subjects and selection criteria. Thirty-two unrelated Argentinean families were included in this study (Fig. 1 and Supplementary Figure S1). Hearing loss was bilateral and moderate to severe (45–95 dB) or profound (>95 dB), and onset was either congenital or postlingual progressive (average 20–40 years). Audiological evaluation included pure-tone audiometry at four frequencies (0.5, 1, 2, and 4 kHz). The pure-tone average (PTA) was calculated from the audiometric thresholds. The HL patients were divided into three groups based on severity: moderate (41–70 dB HL), severe (71–95 dB HL), and profound (>95 dB HL). The audiometric configurations were classified into low-frequency, middle-frequency (U-shaped), high-frequency and flat types³⁸. Patients underwent auditory brainstem response (ABR), tympanometry, fundus examination, and cardiac and renal ultrasonography to detect undiagnosed syndromic forms. Clinical examinations revealed symptoms suggesting a syndromic form of deafness in five patients: two with visual defects, two others with hematuria, and one with hair pigment abnormalities. All data were reviewed by a clinical geneticist. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of “Administración Nacional de Laboratorios e Institutos de Salud” (ANLIS-19122018) and “Fundación para la Lucha contra las Enfermedades Neurológicas de la Infancia” (FLENI -04,092,020). Written informed consent for testing and publication was obtained from patients or parents in the case of minors.

Genomic DNA was extracted from peripheral blood samples using the standard method. Quality and concentration were measured by agarose gel electrophoresis and absorbance-based nucleic acid quantification (Thermo Scientific–NanoDrop™). A total of 695 patients with different forms of deafness were recruited for *GJB2/GJB6* mutation screening, identifying biallelic pathogenic mutations in 103 of them (15%). Thirty-two patients undiagnosed for *GJB2/GJB6* mutations were selected for WES screening. The age of the patients varied between 6 months and 50 years. Seventeen out of the 32 patients (53%) were females, and fifteen (47%) were males. Among the 32 probands, 50% (16/32) were sporadic and 50% (16/32) had at least two affected relatives with HL (familial cases).

Whole exome sequencing and bioinformatics analysis. Massive parallel sequencing was carried out on an *Illumina NovaSeq6000*. Base calling, read mapping and annotation of variants were performed by Macrogen Genome Sequencing Services (Macrogen, Korea) and data were processed according to the Genome-Analysis-Toolkit (GATK) best practices workflow. Variants for 183 deafness-related genes were filtered from the WES on samples from affected patients and relatives when available (analyzed genes are listed in Supplementary Table S1). The average read length covered 148 bp, with an average exon depth coverage in analyzed genes of 100X; 97% of targeted reads had >tenfold coverage. Aligned reads were compared to the human reference genome (*GRCh37/hg19*). After variant annotation, mutations that arose from known deafness genes were selected by filtering with an *in silico* panel using a homemade Python script pipeline. The missense, nonsense, insertion/deletion and splicing variants were selected from the detected variants. The minor allele frequency threshold (MAF) considered was ≤ 0.01 and 0.005 for recessive and dominant alleles, respectively, when compared with those reported in gnomAD—<http://gnomad.broadinstitute.org/>) and 1000 genomes (<https://www.internationalgenome.org/1000-genomes-browsers/>). The pathogenic potential of selected variants was analyzed using bioinformatic programs and databases: Mutation Taster—<http://www.mutationtaster.org/>³⁹, PolyPhen-2—<http://genetics.bwh.harvard.edu/pph/>⁴⁰, CADD (Combined Annotation Dependent Depletion)—<http://cadd.gs.washington.edu/>, UMD-Predictor—<http://umd-predictor.eu/>⁴¹, ClinVar—<https://www.ncbi.nlm.nih.gov/clinvar/>, LOVD—<https://www.lovd.nl/>, The Human Gene Mutation Database—<http://www.hgmd.cf.ac.uk/>,

dbNSFP—<https://sites.google.com/site/jpopgen/dbNSFP>. All information was compiled and criteria rules were combined to reach a variant classification according to data retrieved from InterVar—<http://wintervar.wglab.org/>, Varsome—<https://varsome.com/> and REVEL—<https://sites.google.com/site/revelgenomics>⁴². Once annotation of variants was performed, the number of significant variations was reduced from more than eighty thousand to a small, manageable number (1 to 10) of putative candidate disease-causative mutations for further validation. This process included various parameters, such as mode of inheritance, mutation localization, mutation type (nonsynonymous variants, splice acceptor or donor site mutations, and coding noninframe In/Dels), frequency, pathogenicity of variants, published reports, and so forth. Pathogenicity prediction and variant classification were assessed taking into account criteria defined by the ACMG and further modifications, as well as recommendations of the HL-EP and data from deafness variation databases (<http://deafnessvariationdatabase.org/>)^{43–45}. Variants were classified as pathogenic, likely pathogenic, VUS, likely benign or benign. Patients were informed of the results, and clinical follow-up was recommended when necessary, since some likely pathogenic variants were identified in syndromic genes, before the onset of additional clinical manifestations. Positive results were always confirmed by Sanger sequencing prior to reporting, and segregation through family was performed when possible.

Variant curation. To further analyze and validate the identified variants in patients, a combined in silico analysis was performed. In the first curation step, the prediction of pathogenicity of the variants previously reported in ClinVar was reanalyzed following the guidelines mentioned before⁴³, collecting new data regarding segregation, new PubMed reports or publications and new updated data of frequency and phenotype/genotype correlations. The putative pathogenic effect of novel variants was manually scored using Varsome⁴⁶.

Bioinformatic analysis. To predict the potential consequences of some of the missense variants on protein function, different bioinformatic analyses were performed. These included a molecular homology modeling (MHM) approach and structural analysis of the mutated proteins. The line protein graphs in figures were generated with the PyGame library in the Python 2.7 programming language (<http://www.python.org>)⁴⁷.

Molecular modeling and structural analysis. MHM was performed for mitochondrial (LeuRS) (UniProt_ID Q15031-1) and the MYO6 motor-head domain (UniProt_ID Q9UM54-3, between amino acid residues 4 and 771). The evaluation criteria to select the template for the MHM were as follows: (1) protein sequence identity as template, (2) X-ray crystal resolution less than 2.4 Å, (3) crystallography conditions, (4) chain length and amount of residues of each template with respect to sequence identity and gaps, (5) structural conformation of each protein template, and (6) species of the template. The alignment between template and target sequences was performed with structural alignments (Expresso extension) in the T-Coffee web server^{48,49}, MODELLER alignment scripts and Kluskal-Wallis in MEGA5 software, taking into account the secondary structures and topology of the regions⁵⁰. Homology modeling was generated using the MODELLER program⁵¹. The models were first optimized with the variable target function method with conjugate gradients and then refined using molecular dynamics with simulated annealing⁵². Model quality evaluation was performed using DOPE⁵³, QMEAN/QMEANDisCo^{54,55} and ProSA⁵⁶ to predict local pairwise residue-residue distances to the assessed model. The UCSF Chimera program⁵⁷ and the backbone-dependent rotamer library were used for structural interpretation and visualization⁵⁸.

For the LeuRS protein model, template structures were selected from the RCSB Protein Data Base⁵⁹ and UniProt (<https://www.uniprot.org/>). Fifteen potential templates were considered using a cutoff in sequence identity larger than 36% and a resolution less than 2.5 Å. For the alignment between templates and the LeuRS sequence, the N-terminal mitochondrial targeting signals (presequences) were deleted using MitoFates⁶⁰. Structural regions that did not have reference templates were not considered for the structural analysis and are not shown in the visualization of the structure. After evaluation, six structures in the editing conformation were selected (Supplementary Table S2). Thirty molecular homology models with the selected templates were generated, and the final structure of the tRNA was determined with the 4AS1 template.

The MYO6 protein structure is divided into three regions: an N-terminal motor-head-domain (which comprises the N-terminal SH3-like, myosin motor, actin binding site and ATP catalytic region), followed by a neck domain (with a calmodulin-binding linker domain and a single IQ motif), and a C-terminal tail region (with a three-helix bundle region, an SAH domain and a unique globular domain required for interaction with other proteins such as cargo-binding)⁶¹ (Fig. 3A). The modeled N-terminal motor (head) contains the N-terminal SH3-like (aa 4–53) and myosin motor (aa 57–771) domains with 98.15% identity between the target sequence and templates and a resolution of less than 2.5 Å. Modeling was performed in two MYO6 conformations according to ATPase cycle states: the pre-powerstroke, which is the key force-generating step (PDBID: 2V26, 4DBR, 4E7Z), and the next state with the release of phosphate, Pi release (PDBID: 4PPF, 4PFO, 4PJN)⁶² (Supplementary Table S3).

Stability analysis. To evaluate the effect of mutations on the stability/folding of LeuRS, we first analyzed the structural and energetic details of the interactions for each mutated residue using FoldX (<http://foldxsuite.org.eu/>)⁶³. Structures were optimized to the FoldX force field command from the molecular homology models. The $\Delta\Delta G$ values were estimated as the difference between the energy of the wild-type protein and the average of five replicas for each point mutation⁶⁴. Values above 1,6 kcal/mol (twice the standard deviation) were considered to significantly destabilize the protein. To favor the wild-type conformation, all residues involved in tRNA interactions (5 Å distance) were fixed when optimizing the structures to the FoldX force field.

Data availability

The genetic variant data that support the findings of this study are openly available in ClinVar at [<http://www.clinvar.com/>], reference number SUB10104745. New LeuRS and MYO6 protein models are openly available in <https://www.modelarchive.org/>, <https://doi.org/10.5452/ma-rc0b9> (LeuRS), <https://doi.org/10.5452/ma-d28xp> (MYO6 pre-powerstroke conformation) and <https://doi.org/10.5452/ma-we6d7> (MYO6 Pi release conformation).

Received: 23 August 2021; Accepted: 14 December 2021

Published online: 07 January 2022

References

- Shearer, A. E., Hildebrand, M. S. & Smith, R. J. H. Hereditary hearing loss and deafness overview. In *GeneReviews* (eds Adam, M. P. *et al.*) (University of Washington, 1999).
- Morton, C. C. & Nance, W. E. Newborn hearing screening—a silent revolution. *N. Engl. J. Med.* **354**, 2151–2164 (2006).
- Wémeau, J.-L. & Kopp, P. Pendred syndrome. *Best Pract. Res. Clin. Endocrinol. Metab.* **31**, 213–224 (2017).
- Hilgert, N., Smith, R. J. H. & Van Camp, G. Forty-six genes causing nonsyndromic hearing impairment: Which ones should be analyzed in DNA diagnostics?. *Mutat. Res.* **681**, 189–196 (2009).
- Feldmann, D. *et al.* Large deletion of the GJB6 gene in deaf patients heterozygous for the GJB2 gene mutation: Genotypic and phenotypic analysis. *Am. J. Med. Genet. A* **127A**, 263–267 (2004).
- Snoeckx, R. L. *et al.* GJB2 mutations and degree of hearing loss: A multicenter study. *Am. J. Hum. Genet.* **77**, 945–957 (2005).
- Del Castillo, I. *et al.* Prevalence and evolutionary origins of the del(GJB6-D13S1830) mutation in the DFNB1 locus in hearing-impaired subjects: A multicenter study. *Am. J. Hum. Genet.* **73**, 1452–1458 (2003).
- Dalamón, V. *et al.* GJB2 and GJB6 genes: Molecular study and identification of novel GJB2 mutations in the hearing-impaired Argentinean population. *Audiol. Neurootol.* **15**, 194–202 (2010).
- Diaz-Horta, O. *et al.* Whole-exome sequencing efficiently detects rare mutations in autosomal recessive nonsyndromic hearing loss. *PLoS ONE* **7**, e50628 (2012).
- Duman, D. & Tekin, M. Autosomal recessive nonsyndromic deafness genes: A review. *Front. Biosci.* **17**, 2213–2236 (2012).
- Atik, T. *et al.* Comprehensive analysis of deafness genes in families with autosomal recessive nonsyndromic hearing loss. *PLoS ONE* **10**, e0142154 (2015).
- Choi, H. J. *et al.* Whole-exome sequencing identified a missense mutation in WFS1 causing low-frequency hearing loss: A case report. *BMC Med. Genet.* **18**, 151 (2017).
- Shearer, A. E., Eliot Shearer, A., Hildebrand, M. S., Sloan, C. M. & Smith, R. J. H. Deafness in the genomics era. *Hear. Res.* **282**, 1–9 (2011).
- Parzefall, T. *et al.* Whole-exome sequencing to identify the cause of congenital sensorineural hearing loss in carriers of a heterozygous GJB2 mutation. *Eur. Arch. Otorhinolaryngol.* **274**, 3619–3625 (2017).
- Jalilian, N. *et al.* A novel pathogenic variant in the *MITF* gene segregating with a unique spectrum of ocular findings in an extended Iranian Waardenburg syndrome kindred. *Mol. Syndromol.* **8**, 195–200 (2017).
- Pierce, S. B. *et al.* Mutations in *LARS2*, encoding mitochondrial leucyl-tRNA synthetase, lead to premature ovarian failure and hearing loss in Perrault syndrome. *Am. J. Hum. Genet.* **92**, 614–620 (2013).
- Soldà, G. *et al.* First independent replication of the involvement of *LARS2* in Perrault syndrome by whole-exome sequencing of an Italian family. *J. Hum. Genet.* **61**, 295–300 (2016).
- Avraham, K. B. *et al.* Characterization of unconventional MYO6, the human homologue of the gene responsible for deafness in Snell's waltzer mice. *Hum. Mol. Genet.* **6**, 1225–1231 (1997).
- Buonfiglio, P. *et al.* GJB2 and GJB6 genetic variant curation in an Argentinean non-syndromic hearing-impaired cohort. *Genes* **11**, 2 (2020).
- Dalamón, V. *et al.* Identification of four novel connexin 26 mutations in non-syndromic deaf patients: Genotype-phenotype analysis in moderate cases. *Mol. Biol. Rep.* **40**, 6945–6955 (2013).
- Miyagawa, M., Nishio, S.-Y. & Usami, S.-I. Prevalence and clinical features of hearing loss patients with *CDH23* mutations: A large cohort study. *PLoS ONE* **7**, e40366 (2012).
- Astuto, L. M. *et al.* *CDH23* mutation and phenotype heterogeneity: A profile of 107 diverse families with Usher syndrome and nonsyndromic deafness. *Am. J. Hum. Genet.* **71**, 262–275 (2002).
- Meng, X., Liu, X., Li, Y., Guo, T. & Yang, L. Correlation between genotype and phenotype in 69 Chinese patients with *USH2A* mutations: A comparative study of the patients with Usher Syndrome and Nonsyndromic Retinitis Pigmentosa. *Acta Ophthalmol.* <https://doi.org/10.1111/aos.14626> (2020).
- Yasumoto, K., Yokoyama, K., Shibata, K., Tomita, Y. & Shibahara, S. Microphthalmia-associated transcription factor as a regulator for melanocyte-specific transcription of the human tyrosinase gene. *Mol. Cell. Biol.* **15**, 1833 (1995).
- Zhang, Z. *et al.* A novel variant in *MITF* in a child from Yunnan-Guizhou Plateau with autosomal dominant inheritance of nonsyndromic hearing loss: A case report. *Mol. Med. Rep.* **17**, 6054–6058 (2018).
- Thongpradit, S. *et al.* *MITF* variants cause nonsyndromic sensorineural hearing loss with autosomal recessive inheritance. *Sci. Rep.* **10**, 12712 (2020).
- Sun, L. *et al.* Molecular etiology and genotype-phenotype correlation of Chinese Han deaf patients with type I and type II Waardenburg Syndrome. *Sci. Rep.* **6**, 35498 (2016).
- Pandya, A. Phenotypic variation in Waardenburg syndrome: mutation heterogeneity, modifier genes or polygenic background?. *Hum. Mol. Genet.* **5**, 497–502 (1996).
- DuPai, C. D., Davies, B. W. & Wilke, C. O. A systematic analysis of the beta hairpin motif in the protein data bank. *bioRxiv* <https://doi.org/10.1101/2020.10.28.359612> (2020).
- DuPai, C. D., Cunningham, A. L., Conrado, A. R., Wilke, C. O. & Davies, B. W. TsrA regulates virulence and intestinal colonization in *mSphere* **5**, 2 (2020).
- Zafar, S. *et al.* Novel mutations in *CLPP*, *LARS2*, *CDH23*, and *COL4A5* identified in familial cases of prelingual hearing loss. *Genes* **11**, 978 (2020).
- Yan, W. *et al.* Modulation of aminoacylation and editing properties of leucyl-tRNA synthetase by a conserved structural module. *J. Biol. Chem.* **290**, 12256–12267 (2015).
- Miyagawa, M., Nishio, S.-Y., Kumakawa, K. & Usami, S.-I. Massively parallel DNA sequencing successfully identified seven families with deafness-associated MYO6 mutations: The mutational spectrum and clinical characteristics. *Ann. Otol. Rhinol. Laryngol.* **124**(Suppl 1), 148S-S157 (2015).
- Tian, T. *et al.* Identification of a novel MYO6 mutation associated with autosomal dominant non-syndromic hearing loss in a Chinese family by whole-exome sequencing. *Genes Genet. Syst.* **93**, 171–179 (2018).
- Cheng, J. *et al.* Exome sequencing identifies a novel frameshift mutation of MYO6 as the cause of autosomal dominant nonsyndromic hearing loss in a Chinese family. *Ann. Hum. Genet.* **78**, 410–423 (2014).

36. Hertzano, R. *et al.* A Myo6 mutation destroys coordination between the myosin heads, revealing new functions of myosin VI in the stereocilia of mammalian inner ear hair cells. *PLoS Genet.* **4**, e1000207 (2008).
37. Oka, S.-I. *et al.* Clinical characteristics and in vitro analysis of MYO6 variants causing late-onset progressive hearing loss. *Genes* **11**, 273 (2020).
38. Mazzoli, M. *et al.* Recommendations for the description of genetic and audiological data for families with nonsyndromic hereditary hearing impairment. *Audiol. Med.* **1**, 148–150 (2003).
39. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361–362 (2014).
40. Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* **7**, 20 (2013).
41. Salgado, D. *et al.* UMD-predictor: A high-throughput sequencing compliant system for pathogenicity prediction of any human cDNA substitution. *Hum. Mutat.* **37**, 439–446 (2016).
42. Ioannidis, N. M. *et al.* REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.* **99**, 877–885 (2016).
43. Oza, A. M. *et al.* Expert specification of the ACMG/AMP variant interpretation guidelines for genetic hearing loss. *Hum. Mutat.* **39**, 1593–1613 (2018).
44. den Dunnen, J. T. *et al.* HGVS recommendations for the description of sequence variants: 2016 update. *Hum. Mutat.* **37**, 564–569 (2016).
45. DiStefano, M. T. *et al.* Correction: ClinGen expert clinical validity curation of 164 hearing loss gene–disease pairs. *Genet. Med.* **21**, 2409–2409 (2019).
46. Kopanos, C. *et al.* VarSome: The human genomic variant search engine. *Bioinformatics* **35**, 1978–1980 (2018).
47. Kolomenski, J. E. *et al.* An update on genetic variants of the NKX2-5. *Hum. Mutat.* **41**, 1187–1208 (2020).
48. Di Tommaso, P. *et al.* T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* **39**, W13–W17 (2011).
49. Armougom, F. *et al.* Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* **34**, W604–W608 (2006).
50. Tamura, K. *et al.* MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
51. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinform.* <https://doi.org/10.1002/cpbi.3> (2016).
52. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
53. Shen, M.-Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).
54. Studer, G. *et al.* QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics* **36**, 1765–1771 (2020).
55. Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343–350 (2011).
56. Wiederstein, M. & Sippl, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410 (2007).
57. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
58. Shapovalov, M. V. & Dunbrack, R. L. Jr. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **19**, 844–858 (2011).
59. Bank, R. P. D. RCSB PDB: Homepage. <https://www.rcsb.org/>.
60. Fukasawa, Y. *et al.* MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol. Cell. Proteomics* **14**, 1113–1126 (2015).
61. Unconventional myosin-VI. <https://www.uniprot.org/uniprot/Q9UM54#Q9UM54-3>.
62. Wulf, S. F. *et al.* Force-producing ADP state of myosin bound to actin. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1844–E1852 (2016).
63. Schymkowitz, J. *et al.* The FoldX web server: An online force field. *Nucleic Acids Res.* **33**, W382–W388 (2005).
64. Bruque, C. D. *et al.* Structure-based activity prediction of CYP21A2 stability variants: A survey of available gene variations. *Sci. Rep.* **6**, 39082 (2016).
65. Maubaret, C., Griffioen, J.-M., Arnaud, B. & Hamel, C. Novel mutations in MYO7A and USH2A in Usher syndrome. *Ophthalmic Genet.* **26**, 25–29 (2005).
66. Sloan-Heggen, C. M. *et al.* Comprehensive genetic testing in the clinical evaluation of 1119 patients with hearing loss. *Hum. Genet.* **135**, 441–450 (2016).
67. Wattenhofer, M. *et al.* Mutations in the TMPRSS3 gene are a rare cause of childhood nonsyndromic deafness in Caucasian patients. *J. Mol. Med.* **80**, 124–131 (2002).
68. Eiberg, H. *et al.* Autosomal dominant optic atrophy associated with hearing impairment and impaired glucose regulation caused by a missense mutation in the WFS1 gene. *J. Med. Genet.* **43**, 435–440 (2006).
69. Bernal, S. *et al.* Mutations in USH2A in Spanish patients with autosomal recessive retinitis pigmentosa: high prevalence and phenotypic variation. *J. Med. Genet.* **40**, e8 (2003).
70. Aller, E. *et al.* Identification of 14 novel mutations in the long isoform of USH2A in Spanish patients with Usher syndrome type II. *J. Med. Genet.* **43**, e55 (2006).
71. Oshima, A. *et al.* Mutation profile of the CDH23 gene in 56 probands with Usher syndrome type I. *Hum. Mutat.* **29**, E37–46 (2008).
72. Domingo-Gallego, A. *et al.* Clinical utility of genetic testing in early-onset kidney disease: seven genes are the main players. *Nephrol. Dial. Transplant* <https://doi.org/10.1093/ndt/gfab019> (2021).
73. Plant, K. E., Green, P. M., Vetrie, D. & Flintner, F. A. Detection of mutations in COL4A5 in patients with Alport syndrome. *Hum. Mutat.* **13**, 124–132 (1999).
74. Smith, C. J. A., Crock, P. A., King, B. R., Meldrum, C. J. & Scott, R. J. Phenotype-genotype correlations in a series of wolfram syndrome families. *Diabetes Care* **27**, 2003–2009 (2004).
75. Plantinga, R. F. *et al.* A novel TECTA mutation in a Dutch DFNA8/12 family confirms genotype-phenotype correlation. *J. Assoc. Res. Otolaryngol.* **7**, 173–181 (2006).
76. Eppsteiner, R. W. *et al.* Prediction of cochlear implant performance by genetic mutation: The spiral ganglion hypothesis. *Hear. Res.* **292**, 51–58 (2012).
77. Zhu, M. *et al.* Mutations in the gamma-actin gene (ACTG1) are associated with dominant progressive deafness (DFNA20/26). *Am. J. Hum. Genet.* **73**, 1082–1091 (2003).

Acknowledgements

We thank the families for their kind contribution to this study.

Author contributions

Conceived and designed the methodology: P.B., C.D.B., A.B.E. and V.D. Analysis of clinical history and medical aspects: L.L., F.G., V.L., S.M., L.F., B.P. and E.G. Compilation of reported genetic variants: P.B., C.D.B. and V.D. Analyzed patients' samples: P.B. and V.D. Analyzed and discussed the data: P.B., C.D.B., A.B.E. and V.D. Contributed reagents/materials/bioinformatics analysis: P.B., C.D.B. and V.D. Wrote the paper: P.B., C.D.B., A.B.E. and V.D. All authors reviewed the manuscript.

Funding

This work was supported by the Grand Prix Foundation Pour L'Audition to A.B.E., a grant for Medical Research from Fundación Alberto J. Roemmers to P.B. and from the "Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT)" PICT-2018-00823 to A.B.E and V.D.; P.B. is supported by a National Scientific and Technical Research Council Research Scholarship (CONICET), and C.D.B. is supported by a Research Fellowship in Health supported by the National Ministry of Health.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-04081-2>.

Correspondence and requests for materials should be addressed to V.D.

Reprints and permissions information is available at www.nature.com/reprints.

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



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/>.

© The Author(s) 2022