Genome-wide association study meta-analysis brings monogenic hearing loss genes into the polygenic realm

Clifford Royce[1,2], Johnson Jacquelyn[1,3], Mackey Caroline E[3], Mikita Elizabeth A[1,3], Ryan Allen F[1,2], VA Million Veteran Program, Maihofer Adam X[1,3], Nievergelt Caroline M[1,3]

1. Veterans Affairs San Diego Healthcare System, Research Service, San Diego, CA, USA
2. University of California San Diego, Division of Otolaryngology – Head and Neck Surgery, La Jolla, CA, USA
3. University of California San Diego, Department of Psychiatry, La Jolla, CA, US

Corresponding authors: Nievergelt and Clifford

Acknowledgements/Conflicts/Funding Sources

This research is based on data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration, and was supported by MVP_000, MVP_060 and MVP_039 as well as award #IO1RX004293. This publication does not represent the views of the Department of Veteran Affairs or the United States Government.

Abstract

For 5% of the global population, disabling age-related and noise-induced hearing loss (SNHL) is associated with isolation, depression, cognitive decline, and dementia. On an economic level, SNHL is correlated with decreased employment, particularly for manual labor. Here we present one of the largest genome-wide association studies (GWAS) on SNHL to date, using clinical diagnosis from the Million Veteran Program (MVP) (210,240 cases and 265,275 controls). GWAS findings are contrasted and then meta-analyzed with the United Kingdom Biobank (UKB) self-reported hearing loss GWAS (87,056 cases and 163,333 controls). We identify substantial genetic overlap despite differences in demographics, experiences, and source of phenotypes. Meta-analysis of the two cohorts identifies 108 loci, including novel and known familial hearing loss genes. Significant gene ontology pathways included “sensory perception of mechanical stimulus,” “ear development,” “actin-binding,” and “cytoskeletal protein binding,” indicative of mechanosensory structure and function in the cochlea. Eighteen congenital hearing loss genes are suggested, expressing proteins in subcellular regions of inner and outer hair cells, including stereocilia, rootlets, and mechanoelectrical transduction tip-links, as well as supporting cells, Type 1 neurons and stria vascularis. SNP-based heritability estimates over familial hearing loss genes showed a 3.26 fold enrichment relative to genes not listed as familial hearing loss genes. However, due to the small number of familial hearing loss genes, the majority of heritability was accounted for by genes not yet designated as hearing loss genes.

Introduction

The effects of permanent hearing loss range from annoyance at a partner’s “mumbling” to a disabling disorder that currently affects 1.6 billion people globally, a number that is predicted to increase by greater than 50% over the next three decades[1]. Its health burden is currently estimated at 43.4 million years lived with disability (YLD), or 28% of all YLDs in the Global Health Data Exchange database[1]. On a personal level, hearing loss is associated with isolation, depression, cognitive decline, frailty and dementia[2-4], and on an economic level, with decreased employment, particularly for manual labor[5,6]. Hearing aids and cochlear implants are at best palliative; however, their use is constrained to approximately one-third of those eligible and heavily skewed towards those with upper incomes[7].

Twin-studies have calculated a heritability for age-related sensorineural hearing loss (SNHL) of 0.36-0.67[8-10], indicating that a large portion is genetically based. It has become evident that

common SNHL is polygenic[11], consisting of multiple genes with small effect sizes. In recent years, genome-wide association studies (GWAS) have focused on large cohorts (ref 13, 16, 17, 26) and meta-analysis of multiple cohorts (ref 12, 14, 15, 19) to achieve the statistical power required to identify significant loci of these smaller effect sizes[12-20]. Although the technology to cure hearing impairment caused by monogenic disorders by gene therapy is rapidly advancing with some syndromes currently undergoing clinical trials[21], ultimate treatment of the polygenic forms of SNHL will require identification and characterization of multiple genes, variants, and pathways using technology yet to be determined. Given that statistical power to identify risk loci through GWAS mandates large studies, otologic research has thus far has included “minimal phenotyping,” defined as the use of one or two self-reported items for disease identification for large numbers, then added to smaller cohorts with audiogram or ICD data.[12-15,17] Questionnaire responses are a reasonable proxy for hearing loss[23], identifying substantial heritability in both audiogram and self-report data[23]. However, epidemiologic studies comparing the audiogram to self-report have indicated a wider sensitivity range of 30%-78% and specificity range of 48%-88%, dependent on hearing loss measure and prevalence of hearing loss in the cohort.[24,25]

Here we introduce a more objective phenotype from the Million Veteran Program (MVP), an aging population currently with over one million participants and diverse ancestral make-up. The phenotype consists of International Classification of Diagnoses (ICDs) of the most common forms of SNHL, i.e., age-related and noise-induced, typically derived from audiograms.

To address the question of 'minimal phenotyping', we then contrast MVP ICD results with a GWAS on self-reported hearing loss in the United Kingdom Biobank (UKB) Data[26]. Finally, we perform the largest meta-analysis to date on SNHL and describe the genetic architecture of SNHL.

Results

Ancestry-specific and multi-ancestry hearing loss GWAS in MVP

Hearing loss was defined based on an ICD diagnosis of SNHL from the electronic health record. After excluding participants for mixed, conductive, and uncommon forms of SNHL (e.g. Meniere's disease and ototoxicity), 210,240 cases and 265,275 controls were available for GWAS. GWAS was stratified by ancestry group (Supplementary Table 1). GWAS of European ancestry participants (MVP EUA; 176,393 cases and 172,117 controls) identified 43 genome-wide significant (GWS) loci (Supplementary Table 2, Fig. 1b). The inflation of test statistics present (Genomic control λ = 1.33) was almost entirely accounted for by polygenic effects (LDSC attenuation ratio: 94.1%; LDSC intercept = 1.06; SE = 0.009). GWAS of African ancestry participants (MVP AFA; 22,757 cases and 74,842 controls) identified one significant locus (Chr11:8,078,850-8,079,171; leading SNP rs59174224 p-value = 4.12E-08) within the gene *TUB*. *Tub* is a known murine gene leading to profound hearing loss. A modifier *Moth1* protects hearing, and *Mtap1* then mediates *Moth1* at the synaptic terminal[27]. GWAS of MVP indigenous American ancestry participants (MVP IAA; 11,090 cases, 18,316 controls) did not identify any significant loci.

Multi-ancestry meta-analysis of MVP EUA, AFA, and IAA ancestries (MVP ICD multi-ancestry meta, N = 475,515 subjects) identified 52 loci; 14 of which were not significant in individual ancestry analysis. Functional mapping and annotation of risk loci from the MVP multi-ancestry meta-analysis using the FUMA pipeline are summarized in Supplementary Table 3. Of these 52 loci, 17 contained a GWS SNP with a CADD score predicted to be deleterious to function (>12.37) that was within the exon region of a gene. We identified 58 independent loci across all MVP analyses. Regional association plots comparing ancestry-specific analyses for the 58 loci are shown in Supplementary Table 4.

Gene-based, gene-set, and gene-tissue analyses of MVP meta-analysis

To aggregate genetic markers to the level of genes and allow integration with gene expression data, we performed gene-based analyses using Multi-marker Analysis of GenoMic Annotation (MAGMA)[28]. Gene-based analyses identified 63 genes reaching gene-wide significance (Supplementary Table 5), several of which had not already been implicated by the positional mapping of significant SNPs to genes. Notably, for the first time, the African American GWAS

manifests a significant hit, TUB (p-value 4.97E-09). To ascertain specific biological pathways, a competitive set-based analysis of 17,012 pre-defined curated gene sets and gene ontology (GO) terms obtained from Molecular Signatures Database (MsigDB) was performed (Supplementary Table 6). Significant curated gene sets identified were in the gene ontology biological processes area, i.e., GOBP_EAR_DEVELOPMENT (p-value 1.52E-07), GOBP_SENSORY_PERCEPTION_OF_MECHANICAL_STIMULUS (p-value 3.19E-07), and NIKOLSKY_BREAST_CANCER_14Q22_AMPLICON (p-value 2.50E-06) (Supplementary Table 7). We note that the Nikolsky breast cancer amplicon pathway significance level was primarily driven by one locus with two mapped genes with SNPs in partial LD, *NID2* (p-value 2.61E-14) and *C14orf116* (*FOXN3,* p-value 7.25E-10) (further details in supplementary text).

Genetic comparisons of MVP and UKB hearing loss GWAS

To obtain additional discovery power, we performed a meta-analysis of the MVP GWAS and the previously published UKB EUA GWAS[26], which had identified 41 loci. Given potential study differences later discussed, we made between-study comparisons of genetic architecture before performing the meta-analysis.

h^2_{SNP} estimated from the MVP was 0.0761 (SE = 0.0039, on the liability scale at 30% prevalence) (Supplementary Table 9). In contrast, h^2_{SNP} estimated from the UKB was higher at 0.1162 (SE = 0.0059, liability scale at 30% population prevalence). Differences in h^2_{SNP} were explored with univariate MiXeR, noting that h^2_{SNP} is proportional to the number of causal variants and the effects sizes of these variants. Univariate MiXeR analyses suggest that the difference in h^2_{SNP} was due to significantly more causal variants identified within UKB (5,240 causal variants, SE = 250) than in MVP (3,790 causal variants, SE = 198) (z = 4.5, p-value 5.4 x 10^{-6}), rather than study differences in variant effect sizes (z-test comparing causal effect size variance estimates: z = 1.3, p-value 0.18) (Fig. 1a and Supplementary Table 10a).

Shared and unshared genetics between studies were explored with bivariate MiXeR analyses. MiXeR estimated that MVP and UKB shared 3,624 (SE = 204) causal variants, comprising 95.6% of the total MVP causal variants and 69.2% of the total UKB causal variants (Fig. 1a and Supplementary Table 10b). Genetic correlation among shared variants was 0.73 (SE = 0.03), which was higher than the standard genome-wide estimate of genetic correlation 0.58 (SE = 0.01). For unshared genetics, MiXeR estimated that 1,617 (SE = 235) variants were unique to the UKB, and 167 (SE = 126) variants were unique to MVP. Given that there were relatively few

MVP-specific variants, we note that modeling MVP specific variation did not produce better model fit than a reduced model which assumed no MVP-specific variation (best vs. max AIC = -0.44).

To make variant-level comparisons, we examined the effect size correlation of the leading variants of loci that were GWS in either study. There was an r = 0.77 correlation between studies of leading variant effect sizes, establishing that the existing identified GWS loci have similar effects across studies. Following this, we performed cross-study meta-analysis. We compared leading SNPs from individual analyses to meta-analysis: Of the 43 MVP EUA GWS loci, 70% became more significant in meta-analysis. Of 40 UKB EUA GWS loci (one SNP not found, with no proxy), 65% became more significant the meta-analysis, indicating a common genetic signal between the two studies.

Multi-study, multi-ancestry meta-analysis of MVP and UKB

A meta-analysis of the MVP multi-ancestry study and the UKB GWAS identified 108 GWS loci (Supplementary Table 11 and Fig. 2b). Regional association plots comparing ancestry-specific analyses for these loci are shown in Supplementary Table 4. Functional mapping and annotation of risk loci using the FUMA pipeline are summarized in Supplementary Table 12. Of the 108 loci, 28 (26%) contained an exonic GWS SNP with a predicted deleterious CADD score (>12.37).

Gene-based analyses (MAGMA) identified 192 genome-wide significant genes (Supplementary Table 13). Competitive set-based analysis of pre-defined curated gene sets and gene ontology (GO) terms identified three significant gene sets (Supplementary Table 14), GOBP_SENSORY_PERCEPTION_OF_MECHANICAL_STIMULUS (p-value 6.42E-09), GOMF_ACTIN_BINDING (p-value 2.55E-06) and GOMF_CYTOSKELETAL_PROTEIN_BINDING (p-value 1.63E-06).

To identify tissues and pathways underlying these gene-based associations in the cross-study meta-analysis, we performed a MAGMA tissue expression analysis with 30 general human tissue types and found significant enrichment in the nerve ($p = 2\text{x}10^{-4}$) and pituitary ($p = 1.3\text{x}10^{-3}$). Analysis of 54 different tissue types identified significant enrichment in the cerebellum ($p = 3.85\text{x}10^{-6}$) and cerebellar hemisphere ($p = 1.38\text{x}10^{-5}$) (Supplementary Table 16 and Supplementary Fig. 5). However, these gene-tissue analyses were based on GTEx v8 data sets, which do not include the human cochlea, a tissue relevant to hearing loss. We therefore

performed enrichment analyses based on available mouse data from two recent, large studies on cochlear cells[29] and the organ of Corti[30] (Supplementary Table 17). The cochlear data included 34 cell types, and the organ of Corti study included outer hair cells, inner hair cells, pillar, Deiter's, and melanocytes. In the cochlear cell type data, no cell type was significantly enriched when conditioned on the average expression of all cell types as the baseline. The organ of Corti analysis showed a significant enrichment of all cell types except for Deiter's cells.

Examination of implicated genes in the context of congenital hearing loss

Eighteen genes that were implicated by our GWAS from either gene-based analyses or positional mapping of significant loci, were also present in the Hereditary Hearing Loss Homepage database (Table 2 and Supplementary Table18). Many of these genes (12/18) were members of significant GO terms from our gene-set analysis. Within "sensory perception of mechanical stimulus", the significant congenital hearing loss genes fall into functional groups including structural, maintenance, and ion channel elements in different subcellular areas of hair cells. *COL11A1* (min. p-value 2.32E-17, is a collagen Type 11 pro α-chain providing structural support, controlling lateral growth of collagen II fibers, while *ESPN* (min. p-value 6.97E-08) regulates the organization of actin within the stereocilia and *TRIOBP* (min. p-value 1.95E-13) binds to F-actin in stereocilia, maintaining its shape and rigidity. *CLIC5* (min. p-value 1.45E-07) is a chloride channel protein that stabilizes the attachment of the plasma membrane to the base of the stereocilia. *CLRN2* (min. p-value 1.71E-09) maintains the structure and function of stereocilia. *LOXHD1* (min. p-value 3.97E-07) is required for functional mechanicoelectrical transduction channels (MET) in mature cochleae. *CDH23* (min. p-value 2.40E-10), is part of the stereocilia tiplinks and its mutations lead to Usher's Syndrome, *MYO7A* (min. p-value1.19E-06) is an unconventional myosin, an actin-based motor molecule which helps maintain tension on this MET. *TMPRSS* (min. p-value 5.02E-09) is a transmembrane protease required for the maintenance of hair cells. *ATP2B2* (min. p-value 5.48E-10) expresses a membrane protein which transports Ca^{2+} from the cell. "Actin-binding" and its parent GO term "cytoskeletal binding" contain four known congenital hearing loss genes. Besides *ESPN, MYO7A,* and *TRIOBP, GAS2* (min. p-value 8.24E-07) encodes a controller of actin shape rearrangements.

Discussion

Million Veteran Program multi-ancestry analysis

In this large-scale GWAS and meta-analysis of hearing loss, we provide new targets for investigation, with additional evidence for several previously implicated genes, including known congenital hearing loss genes. In tying together these monogenic hearing loss genes with age-related hearing loss, polygenic risk scores from the UK Biobank's self-report study have been shown to correlate with higher odds of children's hearing loss[31]. Childhood hearing loss is known to be at least 50% monogenic in origin[32]. Taken together, this suggests common genetic signals for both congenital hearing loss and adult age-related and noise-induced loss.

Biological pathways identified in this study include many known familial hearing loss genes. Significant gene ontology (GO) pathways in the MVP multi-ancestry GWAS encompassed "ear development" and "sensory perception of mechanical stimulus". The ear development pathway contains well-described congenital hearing loss genes identified as significant in our gene-based analyses, including *COL11A1*, *EYA4* (p-value 1.35E-09), and *TRIOBP* (Supplementary Table 7)[33-35]. The "sensory perception of mechanical stimulus" pathway contains C*OL11A1*, *TRIOBP, TMPRSS3* and *ATP2B2*[36-38], among others (Supplementary Table 7). Both non-syndromal and syndromal congenital hearing loss genes are identified as significant within this pathway, located in multiple subcellular compartments of hair cells, including stereociliae, rootlets, tip links attaching adjacent stereocilia, and mechanoelectrical transduction channels (MET)(Fig. 4).

Sound waves travel through the external ear, vibrating through the tympanic membrane and ossicles before entering the cochlea or inner ear (Fig. 4). The Organ of Corti is suspended within the center of the cochlea where the sound/fluid wave travels from base to apex. In response to sound waves, the stiff stereocilia of hair cells bend at their rootlets. F-actin filaments with multiple cross-links maintain the structural integrity of stereocilia of both inner hair cells (IHCs) and outer hair cells (OHCs). These parallel actin fibers are critical for stiffness in response to the fluid wave. Within the stereocilia, *ESPN* expresses an actin-binding protein required for assembly and stabilization[39]. *TMPRSS3* and *CLRN2* express a stereociliary protein critical for hair cell survival[40].

The attachment of the stereocilia to the apical surface of the hair cell is at its rootlet, and its flexibility allows for bending in response to the fluid wave. *TRIOBP, CLIC5*, and *ATP2B2*, all non-syndromal congenital hearing loss genes, are expressed within or near the rootlet (Fig. 4c)[35,36,41,42]. The rootlet is within the area of vibration most sensitive to sound as measured by high-resolution optics[43]. *TRIOBP* wraps around actin in the rootlet providing stiffness and support, and *CLIC5*, within the stereocilia plasma membrane, may stabilize the connection between stereocilia actin and the membrane[43-45]. *ATP2B2* expresses a transmembrane Ca^{++} ATPase at the lower half of the stereocilia and cuticular plate which regulates Ca^{++} concentration within the cytosol. Although not on the hereditary gene list, *SPTBN1* expression (p-value 1.14E-13) is a part of the support structure at the area of the rootlet of stereocilia and is associated with an autosomal dominant neurodevelopmental disorder which includes SNHL (Fig. 4c)[46]. In further support of the importance of the “sensory mechanical stimulus pathway”, our analysis of the organ of Corti indicates outer and inner hair cell enrichment (p-value 3.49E-05 and 9.74E-04, respectively).

Tip links are chained protein threads connecting rows of stereocilia, from the tip of the shorter to the side of the next taller structure (Fig. 4c). *CDH23* constitutes the upper half of its length, and *MYO7A* tethers the link to an actin bundle within the stereocilium. These two genes are responsible for different forms of Usher’s syndrome, one of the most common genetic cause of congenital deafness. Both *CLRN2* and *LOXHD1* are engaged in MET maintenance at the tip of the stereocilia[47].

Bending at the rootlet allows for tension of the tip links, which in turn controls the opening of the mechanotransduction ion channel (MET). Adjacent to the lower end of each tip link, the channel opens in response to the fluid sound wave, allowing Ca^{2+} and K^{+} influx to depolarize both OHCs and IHCs to different effects. Tension on the tip links of the OHCs opens MET channels for ion influx, driving a central feedback mechanism adjusting the length of OHC bodies. This dampening effect focuses the wave on a particular tonotopic focus of inner hair cells (IHCs). In reaction to this focused fluid wave, tip links joining IHC stereociliae tighten, opening their own MET channel, leading to depolarization and subsequent neurotransmitter release at their basal ribbon synapses, firing a signal through the spiral ganglion cochlear nerve fibers within the auditory pathway to the auditory cortex for sound perception.

Movement of OHCs is restricted at both the apical stereocilia side and in the supporting Deiter’s cell basal lamina area. OHC stereocilia tips are embedded in the tectorial membrane (TM) (Fig. 4b)[47]. This gelatinous structure provides a static force to the shearing motion of the sound wave

endolymph motion against the longest of the outer hair cell stereocilia. *COL11A1* and *COL9A3* (p-value 7.03E-05, within a locus with lead SNP rs61734651, p-value 6.33E-11), causes of variants of Stickler's Syndrome, code for two α-chain collagens providing support within the tectorial membrane[48]. Other significant support genes include *GAS2*, identified in pillar and Deiters' cells as a microtubule regulatory gene, inducing rearrangement of the actin skeleton[49].

Our cochlear cell-type analysis noted enrichment within melanocytes, a part of the stria vascularis maintaining the endocochlear potential (p-value 2.04E-06). *ATP6V0A4*, also identified within "sensory perception of mechanical transduction", is a known familial hearing loss gene associated with distal renal tubular acidosis.[38] *ATP6V0A4* is expressed in basal stria cells, acting as a proton pump to maintain the pH of endolymph bathing the apical region of hair cells. *SOX10* (p-value 0.28, in a locus with lead SNP rs5756795, p-value 1.18E-22), is expressed within the stria vascularis as well, constituting an aquaporin repressor[50].

Although not related to our significant GO terms, significant congenital hearing loss genes were identified related to neurons and the ribbon synapse. Studies of both age-related and noise-induced SNHL have found that after injury and prior to any changes on audiogram, up to 50% of ribbon synapses can be lost [51,52]. These synapses connect the IHC to Type I spiral ganglion nerve cells[53]. Located within this synapse area are *LMX1A* (p-value 7.66E-07), whose expression is identified in Type I spiral ganglion nerves, and *REST* (p-value 3.32E-03, in locus with lead SNP rs13152711, p-value 1.13E-08), a neuro-restrictive silencer[54]. It is posited that early loss of synapses prior to loss of hair cells, while not immediately leading to a measurable hearing loss, may be associated with loss of comprehension in the face of normal hearing[55]. Nevertheless, eliciting a significant GO term related to neural synapses may require a different phenotype consisting of the measurement of speech comprehension in noisy environments.

A unifying theme between these GO terms, "sensory perception of a mechanical stimulus" and "actin binding" (Supplementary Table 15), may lie in the structure and function of stereocilia. Significant actin-related genes contribute to the stiffness of stereocilia in line with the stereocilia's function in mechanical transduction of sound. Actin-binding proteins are abundant in other supporting structures in the cochlea as well, i.e., pillar cells, etc. Actin and collagen-rich structures make up a structural portion of the cochlea absorbing the stress and vibration of the sound wave, including the tectorial membrane and basement membrane. In addition to enrichment of IHCs and OHCs, cell-type enrichment analysis shows significant expression in supporting pillar cells (p-value 3.62E-03). The diversity of the cell enrichment findings correlates with the polygenicity of this phenotype.

Differences between the Million Veterans Program and the United Kingdom Biobank

This GWAS adds the MVP cohort to the hearing loss GWAS field and with its connection to the electronic health record, has documentation of ICD's to add to self-report studies. This GWAS elicited a total of 58 distinct loci, thus our phenotyping approach was broadly successful for variant identification. A benefit that our approach has over self-report is the exclusion of many other diseases that lead to hearing loss, e.g. Meniere's or ototoxicity, which may have distinct genetic architectures that can lead to inflation in the results. Despite differences in phenotyping approach and populations under investigation, there is nevertheless a broad similarity of the genetics of hearing loss between the MVP and UKB datasets. Importantly, this similarity indicates that the data are suitable for aggregation in meta-analysis, results of which more than doubled the number of loci identified.

While broadly similar, we were able to detect some genetic differences between MVP and UKB, i.e., MiXeR analysis indicated 1,617 or roughly one-third of its variants, unique to UKB with 4.4% unique to MVP. Several factors may be responsible for this difference. Self-report may elicit variants related to psychological effects of hearing loss. Hearing loss is known to be associated with depression, isolation, and cognition[2] and participants may have responded with reference to the impact on their lives. The MVP phenotype was designed as a narrower definition, focusing on the most common age-related and noise-induced hearing loss.

A second difference between cohorts lies in the demographics and environmental exposures of MVP versus UKB. The Million Veteran Program is a group of US Veterans consisting of greater than 90% males, who have been exposed to loud noises, including blast, artillery, gunfire, ships, and aircraft during their military career[56]. In addition, they have sustained a 20% prevalence of head injury, another cause for hearing loss[57]. Thus, this is a cohort with noise, blast, and head injury exposure in addition to age-related hearing loss. In contrast, the UK Biobank reflects an out-patient civilian population, with a 3.5% participation in military experience, 10.4% exposed to loud music for > 1 year, and 19.6% exposed to industrial noise for > 1 year[58]. Industrial noise exposure in the UK is restricted to a maximum of 85 dB for an eight-hour day, while in the US, the Occupational Safety and Health Administration actionable level is 90 dB time-weighted average (TWA). Since the decibel is a logarithmic scale, this equates to a discernable increase of 25% in exposure. The disparity between exposures over a 45-year work lifetime is anticipated to incur a median threshold shift of 7 dB at 85 dB noise level versus 15 dB at 90 dB noise level

at 4kHz [59], a perceptible difference. Thus, although the genetic architecture of the two populations is similar, the MVP cohort may have elicited the genetic expression of variants related to these exposures, whereas in a cohort with the same genetic background, these identical variants may not have been “tested” environmentally. Nevertheless, the strength of the study lies in its meta-analysis – the number of loci, variants, and genes more than doubled any previous study, and many of the suggested genes are monogenic, suggesting that different variants of the same monogenic genes may lead to age-related and noise-induced hearing loss.

This study has limitations. The MVP cohort is roughly 90% male. There may be some sex differentials to be elicited with a cohort that has a larger percentage of females. Although an ICD diagnosis is a step towards an objective phenotype, it does not describe the state of the individual participants’ frequencies, and heritability studies as well as animal research has indicated that genomics of SNHL may be frequency-specific[10]. We cannot yet extricate age-related hearing loss from noise-induced loss, noise being the potentially preventable portion of SNHL. This is particularly important in the MVP cohort, an aging population who has been exposed to more noise than civilian populations. The fact that combining these two disparate cohorts more than doubles the significant loci identified may indicate that this research is still in its infancy and is just at the “bottom of the curve,” as far as numbers required for identifying genes with sufficient power to be relevant.

Having said that, here we have identified over 100 significant loci for hearing loss by combining two non-overlapping populations. Our study indicates a variety of cochlear-relevant gene ontology terms, pointing to the function of both inner and outer hair cells with multiple known hearing loss genes that have contributed to this genotype. Our next tasks will be two-fold. First, we will interpret findings of audiograms to ascertain the impact of variants in different tonotopic areas of the cochlea. Second, our efforts will include global acquisition of data sources with more precise phenotypes to increase the percentage of known heritability, identify relevant loci, variants, and genes for this pervasive problem.

Methods

MVP Phenotype

MVP provided data from 963,753 participants in version 23_1 (released Apr 29, 2024), recruited 2011 through Sep 30, 2023, consisting of ICD records from the US Veterans Administration Corporate Data Warehouse[60]. Information was linked to individual, de-identified health records. Hearing loss was assessed using ICD codes; after exclusion for 216 hearing-related ICD codes with other types of hearing loss (Supplementary Table X), 475,515 participants remained. Cases included ICD diagnosis of sensorineural hearing loss H90.3, H90.5, H91.1, H91.10, H91.11, H91.12, H91.13, 388.01, 389.1, 389.10, 389.11, 389.12, 389.14, 389.16, and 389.18., while controls consisted of those who had no evidence of the above ICDs nor H91.9, H91.90. H91.91, H91.92, H91.93, Z45.3, Z46.1, Z96.2, Z97.4, 389.9, and V53.2. All participants provided written informed consent, and the study was approved by the University of California San Diego and VA Central Institutional Review Boards.

Genotyping, quality control, and imputation

Analyses were based on release 4 of the MVP imputed genotype data containing 658,301 participants. Details of the genotyping, quality control, and imputation procedures used have been reported in detail[60]. In brief, MVP samples were genotyped on a customized version of the Affymetrix Axiom biobank array and standard genotype quality control procedures were followed. Genotype data was phased using Eagle version 2.4[61] and imputed using minimac version 4[62] with the Haplotype Reference Consortium (https://doi.org/10.1038/ng.3643) reference panel.

Assessment of ancestry

Ancestry was determined using SNPweights[63] with a global reference panel[64], including 93,172 SNPs that overlap between the reference data and MVP genotype data. Reference SNPs were selected to have MAF > 1% in reference populations and are in low LD (based on independent pairwise LD pruning, using a 1000 kb window, 50 SNP step size, and r-squared of 0.2, as calculated in PLINK[65]. Participants were placed into the three following large groupings: 1. European ancestry (EUA; individuals with ≥ 90% European ancestry); 2. African ancestry (AFA; individuals with ≥ 5% African ancestry, < 90% European ancestry, < 5% East Asian, Native American, Oceanian and Central-South Asian ancestry, or individuals with ≥ 50% African ancestry, < 5% Native American, Oceanian and < 1% Asian ancestry); and Indigenous American

Ancestry, including individuals from Latin America (IAA; individuals with ≥ 5% Native American ancestry, < 90% European, and < 5% African, East Asian, Oceanian and Central-South Asian ancestry). Native Americans (individuals with ≥ 60% Native American ancestry, < 20% East Asian, < 15% Central-South Asian and < 5% African and Oceanian ancestry) were included in IAA.

Calculation of principal components (PCs)

Sample relatedness was estimated using KING[66]. For each pair of related (defined as kinship coefficient > 0.0884) participants, one was removed, with the preference to retain cases. If case status was identical between the pair, one of the pair was removed at random. PCs were calculated with unrelated subjects of the same ancestry using FlashPCA2[67]. SNPs were excluded for MAF < 1%. Remaining SNPs were pruned for LD over 1 MB windows stepped over 50 variants at a time with an r^2 threshold of 0.05.

ICD Hearing Loss GWAS

GWAS was performed for each of the 3 ancestry groups separately in PLINK 2[65], using logistic regression including sex, age, age2, and 10 PCs as covariates. Only SNPs with MAF > 1% and imputation info score > 0.6 (calculated within analytic sample) were analyzed. Genome-wide significance was assessed at a *p* value of 5.0×10^{-8}.

UK Biobank Self-Reported Hearing GWAS

Summary GWAS data for self-reported hearing difficulty[26] in the UKB were downloaded from the UKB Returned Dataets Catalog. Phenotyping and data analysis methods have been previously described in detail[26]. In brief, the cohort consisted of those of white British ancestry recruited from 2007 through 2013 who volunteered to participated in the UKB study. Participants completed a questionnaire including field code 2247 and 2257 regarding hearing difficulties and difficulties hearing in ambient noise. Cases were those who answered “yes” to both questions, controls were those who answered “no” to both. In total, there were 87,056 cases and 163,333 controls included in the UKB GWAS. For compatibility with MVP, summary statistics were converted to the log odds ratio scale[68].

Meta-analysis

Inverse variance weighted fixed effects meta-analysis was performed in METAL[69]. Only SNPs present in two or more input Dataets were included in meta-analysis. Three meta-analyses were performed in total: 1. Multi-ancestry meta-analysis of MVP EUA, AFA, and IAA; 2. MVP EUA + UKB; 3. MVP multi-ancestry + UKB EUA.

Functional mapping and annotation

Functional annotation of the GWAS results were performed with the FUMA pipeline version v1.5.2[70]. Annotations are based on human genome assembly GRCh37 (hg19). FUMA was used with default settings unless stated otherwise. The SNP2Gene module was used to define independent genomic risk loci and variants in LD with lead SNPs ($r^2 > 0.6$, calculated using ancestry appropriate 1KGPp3 reference genotypes: EUA were used as reference genotypes for EUA, AMR used for IAA, and AFR used for AFA samples). Genomic risk loci were combined to have non-overlapping base pair positions. SNPs were annotated to the nearest gene. Functional consequences of SNPs were obtained by mapping the SNPs on their chromosomal position and reference alleles to databases containing known functional annotations, including ANNOVAR[71], Combined Annotation Dependent Depletion[72] and RegulomeDB[73].

Regional association plots

Regional visualizations of genome-wide significant loci were produced using LocusZoom 1.4[74]. Linkage Disequilibrium (LD) was calculated using 1KGPp3 data, where EUR samples were used as reference genotypes for EUA data, AMR used for IAA, and AFR used for AFA data.

Gene-based and gene set, and tissue-enrichment analyses with MAGMA

The MAGMA v1.08[28] tool implemented in FUMA was used to perform gene-based, gene-pathway, and human-based tissue enrichment analyses. For gene-based analysis, SNPs were mapped to 18,809 protein coding genes. For each gene, its association with SNHL was determined as the weighted mean squared test statistic of SNPs mapped to the gene, where LD

patterns were calculated using ancestry appropriate 1KGPp3 reference genotypes. The significance of genes was set at a Bonferroni-corrected threshold of p-value 0.05/18,809 = 2.658e-06. To identify specific biological pathways implicated, gene-based test statistics were used to perform a competitive set-based analysis of 17,012 pre-defined curated gene sets and GO terms obtained from MsigDB[75]. The significance of pathways was set at a Bonferroni-corrected threshold of p-value 0.05/17,012 = 2.94e-06. To test if tissue-specific gene expression was associated with this hearing loss phenotype, gene set-based analysis was also used with expression data from GTEx v8 RNA-seq, where the expression of genes within specific tissues were used to define the gene properties used in the gene-set analysis model.

Cochlear cell type enrichment analyses

Cochlear cell type enrichment analyses were performed via MAGMA gene-tissue analysis of the EUA gene-based GWAS results paired with mice expression data. Mouse gene expression data was derived from cochlear cells and nuclei from Jean *et al.* (20-day old mice only)[29] and Hoa et al[30]. Additional pre-processing steps were applied to data from Jean *et al.*; cells were retained if the RNA count was between 1,000 and 30,000, gene/non-coding RNA between 500 and 5,000, and mitochondrial genes less than 2% using the R package Seurat v4.9.9.9058. Pseudobulk gene expression was calculated by log2 normalizing the mean raw expression across each cell type. Data from Hoa et al. for the 5 tissues from the organ of Corti had been pre-processed into reads per kilobase million by study authors.

Mouse gene Ensembl IDs were converted to human Entrez IDs using a homology map from the Mouse Genome Database[76]. Only genes with one-to-one mappings were kept.

MAGMA v1.10 was used to annotate the SNPs based on GRCh37 and perform gene-set analysis. Gene-set analyses were conditioned on the log2 mean expression across the cochlear cell types for each gene.

SNP-based heritability and genetic correlation

SNP-based heritability (h2SNP) was evaluated in EUA data using LD score regression. Input LD scores were computed from 1KGPp3 EUA samples. The LDSC intercept was used to test for artifactual inflation of test statistics, and the attenuation factor was used to estimate the

proportion of inflation coming from polygenic signal. h2SNP was converted to the liability scale[77]. Cross-trait LDSC[78] was used to estimate the genetic correlation between Dataets.

Univariate and bi-variate Gaussian mixer model (MiXeR) analysis

We used univariate MiXeR v1.3[79] to estimate the genetic architecture of MVP EUA and UKB EUA phenotypes. MiXeR estimates SNP-based heritability and two subcomponents whose product is proportional to heritability: the proportion of non-null SNPs (polygenicity) and variance of effect sizes of non-null SNPs (discoverability). MiXeR was applied to GWAS summary statistics under the default settings with the supplied EUA LD reference panel. The results for the number of influential variants reflect the number of SNPs necessary to explain 90% of SNP based heritability. Bivariate MiXeR[80] was used to estimate phenotype specific polygenicity and the shared polygenicity between phenotypes. Goodness-of-fit of the MiXeR model relative to simpler models of polygenic overlap was assessed using AIC values. Heritability, polygenicity and discoverability estimates were contrasted between datasets using the *z*-test.

Hearing genes from HereditaryHearingLoss.org

We extracted 196 protein-coding genes associated with non-syndromic and syndromic hearing loss from the Hereditary Hearing Loss database (https://hereditaryhearingloss.org, accessed 8/9/2024). This excludes Mitochondrial Non-syndromic Hearing Loss genes and *MIR96,* which encodes microRNA. We then aligned these genes with our genomic risk loci (mapped genes) and gene-based analysis (MAGMA). As genes may have multiple symbol aliases, these were checked using the NIH gene database (https://www.ncbi.nlm.nih.gov/gene/ ; accessed 8/28/2024).

Fig. Legends

Fig. 1: Genetic architecture of ICD coded hearing loss in MVP EUA and self-reported hearing loss in UKB EUA.

a, Polygenic overlap between hearing loss ICD in MVP EUA and self-reported hearing loss in the previously published UKB EUA GWAS[26], indicating an estimated 3,624 causal variants are shared, accounting for 95.6% of the variants influencing MVP EUA and 69.2% of the variants influencing UKB EUA. The numbers (in thousands, with standard errors in parenthesis) in the Venn diagram indicate the estimated quantity of causal variants per component. The size of the circles reflects the degree of polygenicity. Genetic correlation (r_g) estimated between the two phenotypes is shown below the Venn diagram. **b,** Manhattan plot of the MVP EUA analysis, showing 43 genome-wide significant (GWS) loci. The y axis represents -log10 p-values from two-sided z-test for meta-analyses effect estimates. The red line represents genome-wide significance at $p < 5 \times 10{-}8$. **c,** Manhattan plot of the UKB EUA analysis of hearing loss GWAS [26], showing 41 genome-wide significant (GWS) loci. The X chromosome was not included in this analysis.

Fig. 2: MVP multi-ancestry and UKB EUA meta-analysis of hearing loss.

a, Manhattan plot of MVP multi-ancestry and UKB EUA meta-analysis of hearing loss GWAS, showing 108 genome-wide significant (GWS) loci. The y axis represents -log10 p-values from two-sided z-test for meta-analyses effect estimates. The red line represents genome-wide significance at $p < 5.0E-08$. Note that the X chromosome was not included in this analysis. **b,** Manhattan plot of the gene-based hearing loss meta-analysis across 725,904 subjects in the MVP multi-ancestry and UKB EUA meta-analysis. A selection of genome-wide significant hits, including genes identified in HereditaryHearingLoss.org (in bold). The y axis represents -log10 p values from the two-sided F-test for gene-based analysis. The red line indicates the gene-wide significance threshold at $p < 2.53E-06$ (Bonferroni correction for 19,728 genes tested).

Fig. 3. Cochlear function may be hampered by variants within significant congenital hearing loss genes. Genes that are not color-coded are ubiquitous, involved in development or in maintenance.

A, Cross-section of outer, middle, and inner ear. Sound waves travel through the external auditory canal, vibrating the tympanic membrane and ossicular chain, where it continues as a fluid wave into the cochlea or inner ear. The Organ of Corti (OC), arranged tonotopically, is suspended within the center of the cochlea (green) and the sound/fluid wave travels from base to apex. **b,** Cross-section of the Organ of Corti (OC). The OC is supported by the basilar membrane below (BM) and movement is restricted by gelatinous tectorial membrane above (TM), where tips of longest OHC stereocilia are embedded. The shearing motion of the fluid endolymph wave motion against the OHC stereocilia induces a centrally based dampening reflex movement, redirecting and focusing the wave towards the locally juxtaposed inner hair cells (IHCs). In response, IHCs transduce this mechanical movement into an electric signal, releasing neurotransmitter at the ribbon synapse. Post-synaptic neurons of the cochlear nerve transmit signals through the auditory pathway to the brain's auditory cortex. The stria vascularis maintains the electric potential of endolymph that bathes the apices of hair cells. **c,** Expanded area of apical hair cells from b., showing stereociliary movement that applies to both OHCs and IHCs. Fluid/sound wave movement pushes stereocilia laterally, stretching tip-links attaching a taller stereocilium to its adjacent shorter one. This movement opens a mechanoelectrical transduction (MET) channel at its tip (black), allowing influx of Ca++ and K+ and triggering depolarization at the base of the cell. The stereocilium's parallel f-actin fibers are critical for stiffness and rigidity of the stereocilia, while allowing for bending at the reinforced rootlet area. Tight junctions between hair cells and supporting cells maintain the specificity of endolymph that bathes the stereocilia of IHCs and OHCs. **Abbr:** OC – Organ of Corti, TM – tectorial membrane, OHC – outer hair cell, IHC – inner hair cells, MET – mechanicoelectrical transduction channel.

1. Nocini, R., Henry, B.M., Lippi, G. & Mattiuzzi, C. Estimating the worldwide burden of health loss due to hearing loss. *Eur J Public Health* **33**, 146-148 (2023). PMID 36377968
2. Dawes, P. *et al.* Hearing loss and cognition: the role of hearing AIDS, social isolation and depression. *PLoS One* **10**, e0119616 (2015). PMID 25760329
3. Yévenes-Briones, H. *et al.* Association Between Hearing Loss and Impaired Physical Function, Frailty, and Disability in Older Adults: A Cross-sectional Study. *JAMA Otolaryngol Head Neck Surg* **147**, 951-958 (2021). PMID 34554203
4. Jafari, Z., Kolb, B.E. & Mohajerani, M.H. Age-related hearing loss and tinnitus, dementia risk, and auditory amplification outcomes. *Ageing Res Rev* **56**, 100963 (2019). PMID 31557539
5. Shan, A. *et al.* Hearing loss and employment: a systematic review of the association between hearing loss and employment among adults. *J Laryngol Otol* **134**, 387-397 (2020). PMID 32468973
6. Garcia Morales, E.E. *et al.* Labor Force Participation and Hearing Loss Among Adults in the United States: Evidence From the National Health and Nutrition Examination Survey. *Am J Audiol* **31**, 604-612 (2022). PMID 35623104
7. Bainbridge, K.E. & Ramachandran, V. Hearing aid use among older U.S. adults; the national health and nutrition examination survey, 2005-2006 and 2009-2010. *Ear Hear* **35**, 289-94 (2014). PMID 24521924
8. Duan, H. *et al.* Heritability of Age-Related Hearing Loss in Middle-Aged and Elderly Chinese: A Population-Based Twin Study. *Ear Hear* **40**, 253-259 (2019). PMID 29794565
9. Bogo, R. *et al.* The role of genetic factors for hearing deterioration across 20 years: a twin study. *J Gerontol A Biol Sci Med Sci* **70**, 647-53 (2015). PMID 25665831
10. Wingfield, A. *et al.* A twin-study of genetic contributions to hearing acuity in late middle age. *J Gerontol A Biol Sci Med Sci* **62**, 1294-9 (2007). PMID 18000151
11. Fransen, E. *et al.* Genome-wide association analysis demonstrates the highly polygenic character of age-related hearing impairment. *Eur J Hum Genet* **23**, 110-5 (2015). PMID 24939585
12. Trpchevska, N. *et al.* Genome-wide association meta-analysis identifies 48 risk variants and highlights the role of the stria vascularis in hearing loss. *Am J Hum Genet* **109**, 1077-1091 (2022). PMID 35580588
13. Nagtegaal, A.P. *et al.* Genome-wide association meta-analysis identifies five novel loci for age-related hearing impairment. *Sci Rep* **9**, 15192 (2019). PMID 31645637
14. Ivarsdottir, E.V. *et al.* The genetic architecture of age-related hearing impairment revealed by genome-wide association analysis. *Commun Biol* **4**, 706 (2021). PMID 34108613
15. Praveen, K. *et al.* Population-scale analysis of common and rare genetic variation associated with hearing loss in adults. *Commun Biol* **5**, 540 (2022). PMID 35661827
16. De Angelis, F. *et al.* Sex differences in the polygenic architecture of hearing problems in adults. *Genome Med* **15**, 36 (2023). PMID 37165447
17. Hoffmann, T.J. *et al.* A Large Genome-Wide Association Study of Age-Related Hearing Impairment Using Electronic Health Records. *PLoS Genet* **12**, e1006371 (2016). PMID 27764096
18. Vuckovic, D. *et al.* Genome-wide association analysis on normal hearing function identifies PCDH20 and SLC28A3 as candidates for hearing function and loss. *Hum Mol Genet* **24**, 5655-64 (2015). PMID 26188009
19. Wolber, L.E. *et al.* Salt-inducible kinase 3, SIK3, is a new gene associated with hearing. *Hum Mol Genet* **23**, 6407-18 (2014). PMID 25060954

20. Friedman, R.A. *et al.* GRM7 variants confer susceptibility to age-related hearing impairment. *Hum Mol Genet* **18**, 785-96 (2009). PMID 19047183
21. Jiang, L., Wang, D., He, Y. & Shu, Y. Advances in gene therapy hold promise for treating hereditary hearing loss. *Mol Ther* **31**, 934-950 (2023). PMID 36755494
22. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat Genet* **52**, 437-447 (2020). PMID 32231276
23. Cherny, S.S. *et al.* Self-reported hearing loss questions provide a good measure for genetic studies: a polygenic risk score analysis from UK Biobank. *Eur J Hum Genet* **28**, 1056-1065 (2020). PMID 32203203
24. McCullagh, M.C., Raymond, D., Kerr, M.J. & Lusk, S.L. Prevalence of hearing loss and accuracy of self-report among factory workers. *Noise Health* **13**, 340-7 (2011). PMID 21959114
25. Brennan-Jones, C.G. *et al.* Self-reported hearing loss and manual audiometry: A rural versus urban comparison. *Aust J Rural Health* **24**, 130-5 (2016). PMID 26311193
26. Wells, H.R.R. *et al.* GWAS Identifies 44 Independent Associated Genomic Loci for Self-Reported Adult Hearing Difficulty in UK Biobank. *Am J Hum Genet* **105**, 788-802 (2019). PMID 31564434
27. Johnson, K.R., Zheng, Q.Y. & Noben-Trauth, K. Strain background effects and genetic modifiers of hearing in mice. *Brain Res* **1091**, 79-88 (2006). PMID 16579977
28. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015). PMID 25885710
29. Jean, P. *et al.* Single-cell transcriptomic profiling of the mouse cochlea: An atlas for targeted therapies. *Proc Natl Acad Sci U S A* **120**, e2221744120 (2023). PMID 37339214
30. Hoa, M. *et al.* Characterizing Adult Cochlear Supporting Cell Transcriptional Diversity Using Single-Cell RNA-Seq: Validation in the Adult Mouse and Translational Implications for the Adult Human Cochlea. *Front Mol Neurosci* **13**, 13 (2020). PMID 32116546
31. Wang, J. *et al.* Polygenic Risk Scores and Hearing Loss Phenotypes in Children. *JAMA Otolaryngol Head Neck Surg* (2024). PMID 39509092
32. Khela, H. & Kenna, M.A. Genetics of pediatric hearing loss: A functional perspective. *Laryngoscope Investig Otolaryngol* **5**, 511-519 (2020). PMID 32596495
33. Rad, A. *et al.* Aberrant COL11A1 splicing causes prelingual autosomal dominant nonsyndromic hearing loss in the DFNA37 locus. *Hum Mutat* **42**, 25-30 (2021). PMID 33169910
34. Ahmadmehrabi, S. *et al.* Genome-first approach to rare EYA4 variants and cardio-auditory phenotypes in adults. *Hum Genet* **140**, 957-967 (2021). PMID 33745059
35. Katsuno, T. *et al.* TRIOBP-5 sculpts stereocilia rootlets and stiffens supporting cells enabling hearing. *JCI Insight* **4**(2019). PMID 31217345
36. Fasquelle, L. *et al.* Tmprss3, a transmembrane serine protease deficient in human DFNB8/10 deafness, is critical for cochlear hair cell survival at the onset of hearing. *J Biol Chem* **286**, 17383-97 (2011). PMID 21454591
37. Li, X. *et al.* PON2 and ATP2B2 gene polymorphisms with noise-induced hearing loss. *J Thorac Dis* **8**, 430-8 (2016). PMID 27076938
38. Stover, E.H. *et al.* Novel ATP6V1B1 and ATP6V0A4 mutations in autosomal recessive distal renal tubular acidosis with new evidence for hearing loss. *J Med Genet* **39**, 796-803 (2002). PMID 12414817
39. Sekerkova, G., Zheng, L., Loomis, P.A., Mugnaini, E. & Bartles, J.R. Espins and the actin cytoskeleton of hair cell stereocilia and sensory cell microvilli. *Cell Mol Life Sci* **63**, 2329-41 (2006). PMID 16909209
40. Dunbar, L.A. *et al.* Clarin-2 is essential for hearing by maintaining stereocilia integrity and function. *EMBO Mol Med* **11**, e10288 (2019). PMID 31448880

41. Liu, W. *et al.* Expression of trans-membrane serine protease 3 (TMPRSS3) in the human organ of Corti. *Cell Tissue Res* **372**, 445-456 (2018). PMID 29460002
42. Pacentine, I., Chatterjee, P. & Barr-Gillespie, P.G. Stereocilia Rootlets: Actin-Based Structures That Are Essential for Structural Stability of the Hair Bundle. *Int J Mol Sci* **21**(2020). PMID 31947734
43. Cooper, N.P., Vavakou, A. & van der Heijden, M. Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea. *Nat Commun* **9**, 3054 (2018). PMID 30076297
44. Gagnon, L.H. *et al.* The chloride intracellular channel protein CLIC5 is expressed at high levels in hair cell stereocilia and is essential for normal inner ear function. *J Neurosci* **26**, 10188-98 (2006). PMID 17021174
45. Silverstein, R.S. & Tempel, B.L. Atp2b2, encoding plasma membrane Ca2+-ATPase type 2, (PMCA2) exhibits tissue-specific first exon usage in hair cells, neurons, and mammary glands of mice. *Neuroscience* **141**, 245-57 (2006). PMID 16675132
46. Cousin, M.A. *et al.* Pathogenic SPTBN1 variants cause an autosomal dominant neurodevelopmental syndrome. *Nat Genet* **53**, 1006-1021 (2021). PMID 34211179
47. Wang, P. *et al.* LOXHD1 is indispensable for coupling auditory mechanosensitive channels to the site of force transmission. *Res Sq* (2024). PMID 38260480
48. Shpargel, K.B., Makishima, T. & Griffith, A.J. Col11a1 and Col11a2 mRNA expression in the developing mouse cochlea: implications for the correlation of hearing loss phenotype with mutant type XI collagen genotype. *Acta Otolaryngol* **124**, 242-8 (2004). PMID 15141750
49. Chen, T. *et al.* Cochlear supporting cells require GAS2 for cytoskeletal architecture and hearing. *Dev Cell* **56**, 1526-1540 e7 (2021). PMID 33964205
50. Szeto, I.Y.Y. *et al.* SOX9 and SOX10 control fluid homeostasis in the inner ear for hearing through independent and cooperative mechanisms. *Proc Natl Acad Sci U S A* **119**, e2122121119 (2022). PMID 36343245
51. Fernandez, K.A. *et al.* Noise-induced Cochlear Synaptopathy with and Without Sensory Cell Loss. *Neuroscience* **427**, 43-57 (2020). PMID 31887361
52. Kujawa, S.G. & Liberman, M.C. Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss. *Hear Res* **330**, 191-9 (2015). PMID 25769437
53. Jean, P., Ozcete, O.D., Tarchini, B. & Moser, T. Intrinsic planar polarity mechanisms influence the position-dependent regulation of synapse properties in inner hair cells. *Proc Natl Acad Sci U S A* **116**, 9084-9093 (2019). PMID 30975754
54. Aoki, H., Hara, A. & Kunisada, T. White spotting phenotype induced by targeted REST disruption during neural crest specification to a melanocyte cell lineage. *Genes Cells* **20**, 439-49 (2015). PMID 25818501
55. Ridley, C.L., Kopun, J.G., Neely, S.T., Gorga, M.P. & Rasetshwane, D.M. Using Thresholds in Noise to Identify Hidden Hearing Loss in Humans. *Ear Hear* **39**, 829-844 (2018). PMID 29337760
56. Yankaskas, K. Prelude: noise-induced tinnitus and hearing loss in the military. *Hear Res* **295**, 3-8 (2013). PMID 22575206
57. Chen, J.X. *et al.* Systematic review of hearing loss after traumatic brain injury without associated temporal bone fracture. *Am J Otolaryngol* **39**, 338-344 (2018). PMID 29506762
58. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol* **186**, 1026-1034 (2017). PMID 28641372
59. Lutman, M.E. & Spencer, H.S. Occupational noise and demographic factors in hearing. *Acta Otolaryngol Suppl* **476**, 74-84 (1990). PMID 2087983

60. Gaziano, J.M. *et al.* Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol* **70**, 214-23 (2016). PMID 26441289
61. Loh, P.R., Palamara, P.F. & Price, A.L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet* **48**, 811-6 (2016). PMID 27270109
62. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat Genet* **48**, 1284-1287 (2016). PMID 27571263
63. Chen, C.Y. *et al.* Improved ancestry inference using weights from external reference panels. *Bioinformatics* **29**, 1399-406 (2013). PMID 23539302
64. Nievergelt, C.M. *et al.* International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nat Commun* **10**, 4558 (2019). PMID 31594949
65. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015). PMID 25722852
66. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867-73 (2010). PMID 20926424
67. Abraham, G., Qiu, Y. & Inouye, M. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. *Bioinformatics* **33**, 2776-2778 (2017). PMID 28475694
68. Cook, J.P., Mahajan, A. & Morris, A.P. Guidance for the utility of linear models in meta-analysis of genetic association studies of binary phenotypes. *Eur J Hum Genet* **25**, 240-245 (2017). PMID 27848946
69. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010). PMID 20616382
70. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017). PMID 29184056
71. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010). PMID 20601685
72. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* **47**, D886-D894 (2019). PMID 30371827
73. Dong, S. *et al.* Annotating and prioritizing human non-coding variants with RegulomeDB v.2. *Nat Genet* **55**, 724-726 (2023). PMID 37173523
74. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010). PMID 20634204
75. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015). PMID 26771021
76. Blake, J.A. *et al.* Mouse Genome Database (MGD): Knowledgebase for mouse-human comparative biology. *Nucleic Acids Res* **49**, D981-D987 (2021). PMID 33231642
77. Lee, S.H., Wray, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* **88**, 294-305 (2011). PMID 21376301
78. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015). PMID 25642630
79. Holland, D. *et al.* Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genet* **16**, e1008612 (2020). PMID 32427991
80. Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat Commun* **10**, 2417 (2019). PMID 31160569

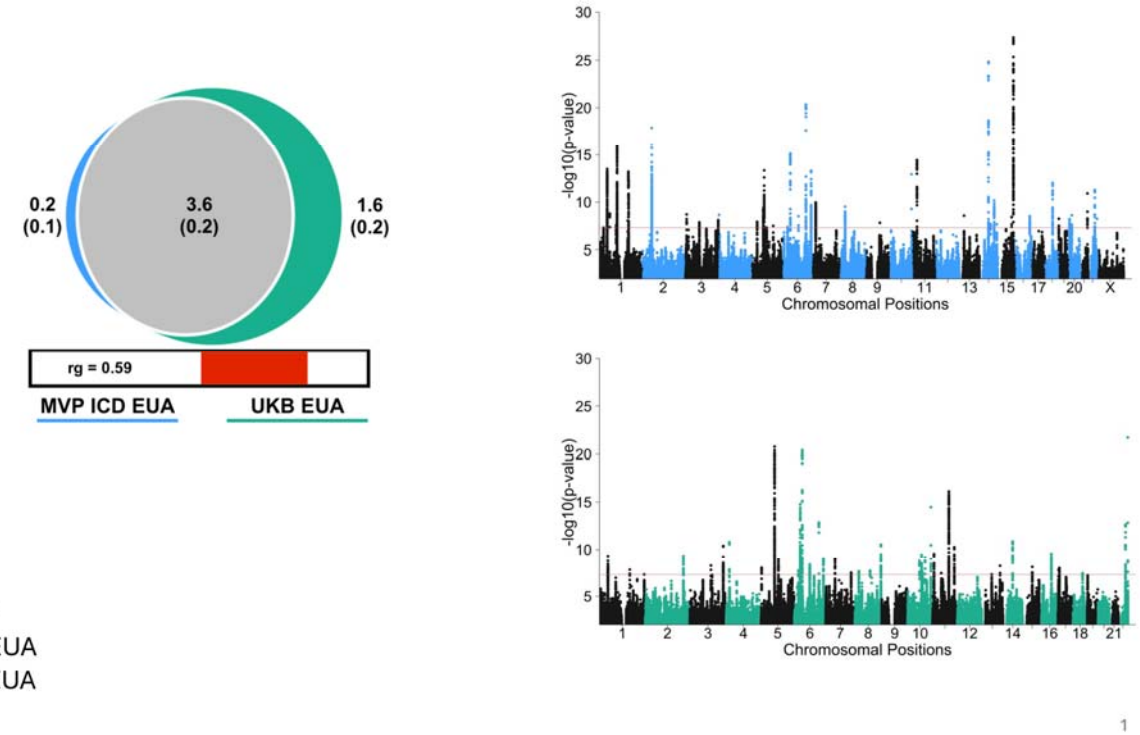

Fig. 1
A. Mixer
B. MVP EUA
C. UKB EUA

1

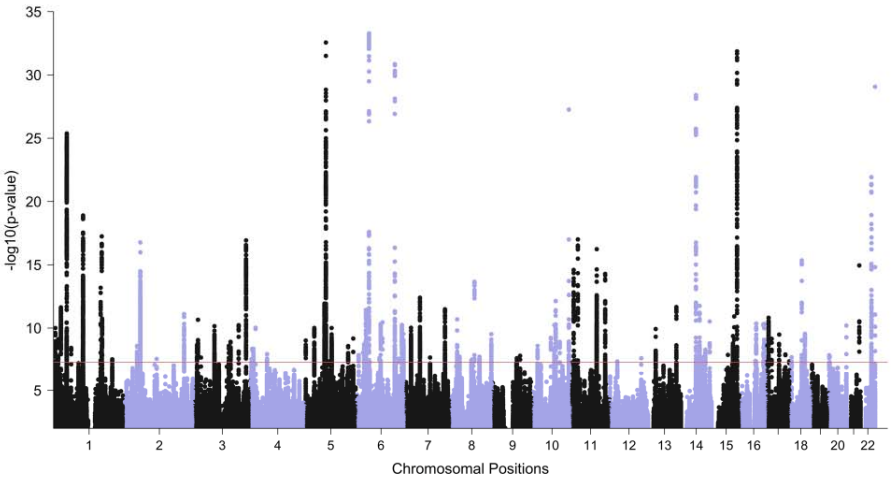

Fig. 2a
MVP Multi-Ancestry + UKB EUA Meta Analysis
Loci Plot

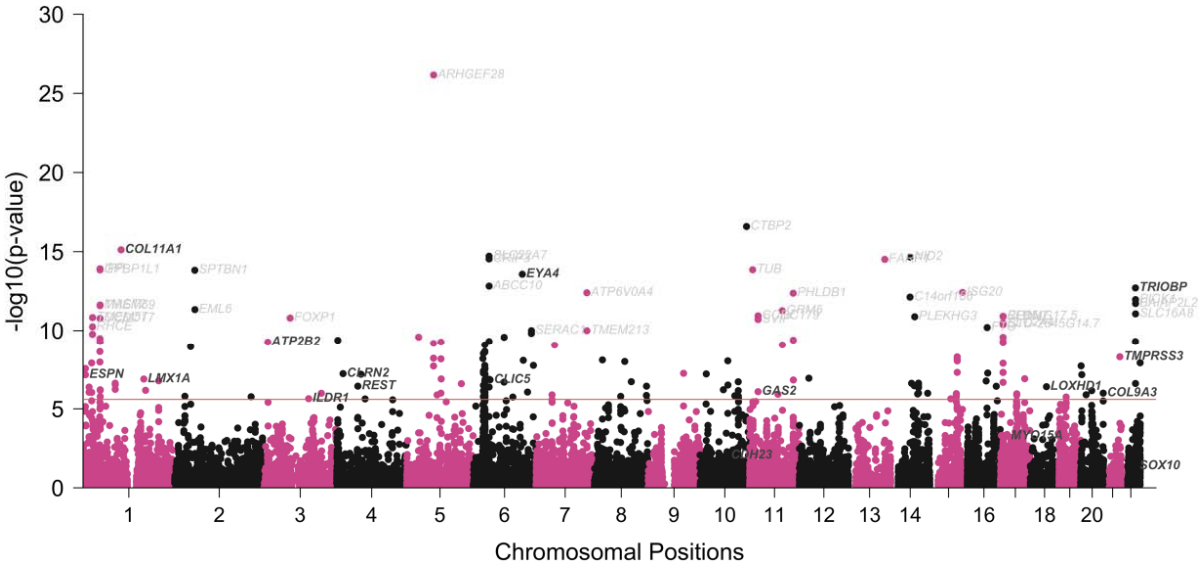

Fig. 2b
Top 40 + sig. HereditaryHearingLoss.org genes
MVP Multi-Ancestry + UKB EUA Meta Analysis
Gene Plot

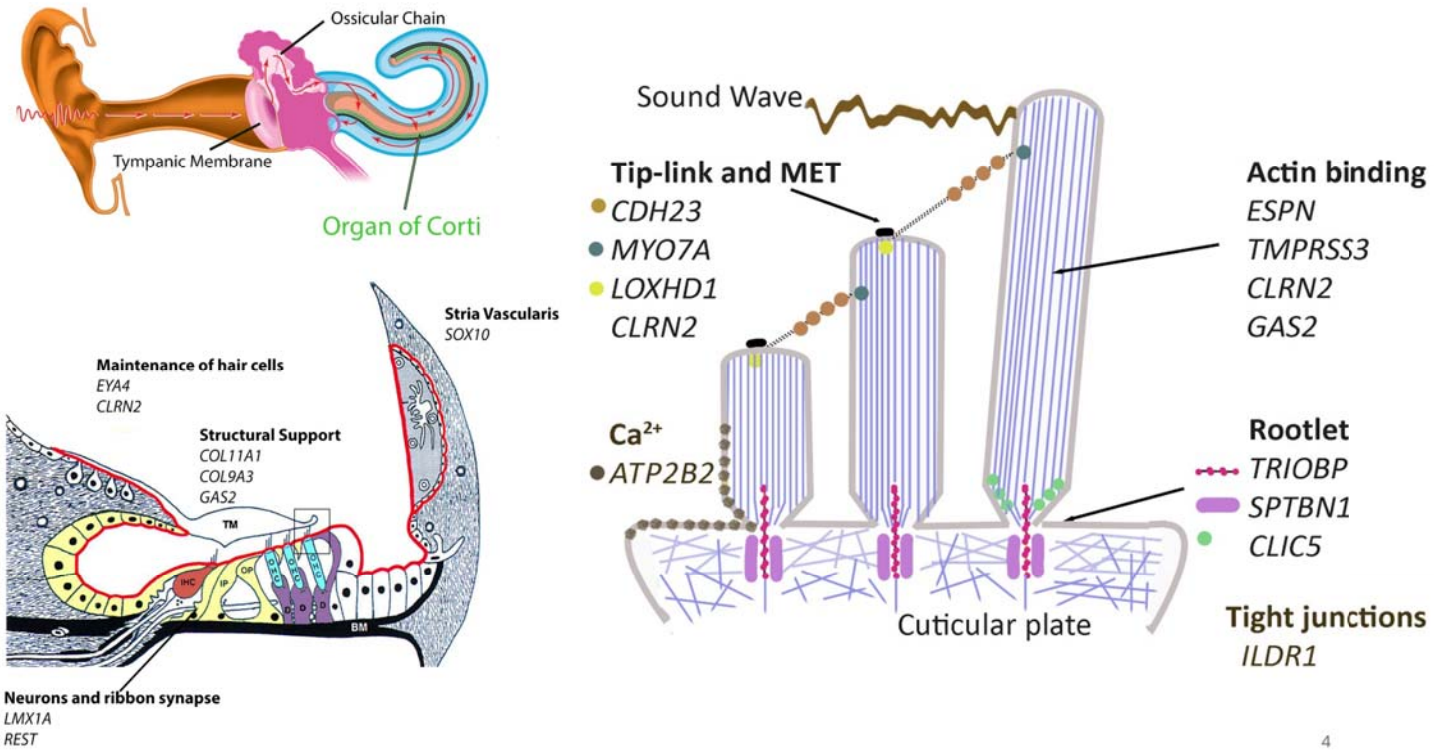

Figure 3.

A. Cross-section of outer, middle, and inner ear. Sound waves travel through the external auditory canal, vibrating the tympanic membrane and ossicular chain, which then transfers a fluid wave into the cochlea or inner ear. The Organ of Corti (OC), arraarranged tonotopically, is suspended within the center of the cochlea (green) where this fluid wave travels from base to apex.

B. Cross-section of the Organ of Corti (OC). The OC is supported by the basilar membrane below (black) and movement is restricted by gelatinous tectorial membrane above (TM), where tips of longest OHC stereocilia are embedded. The shearing motion of the fluid endolymph wave motion against the OHC stereocilia induces a centrally-based dampening reflex movement, redirecting and focusing the wave towards the locally juxtaposed inner hair cells (IHCs). In response, IHCs transduce this mechanical movement into an electric signal, releasing neurotransmitter at the ribbon synapse. Post-synaptic neurons of the cochlear nerve transmit signals through the auditory pathway to the brain's auditory cortex. The stria vascularis maintains the electric potential of endolymph that bathes the apices of hair cells.

C. Expanded area of apical hair cells from B., showing stereociliary movement that applies to both OHCs and IHCs. The fluid wave movement pushes stereocilia laterally, stretching tip-links attaching a taller stereocilium to its adjacent shorter one. This movement opens a mechanoelectric transduction (MET) channel at its tip (black), allowing influx of Ca++ and K+ and triggering hypopolarization at the base of the cell. The stereocilium's parallel actin fibers are critical for stiffness and rigidity of the stereocilia, while allowing for bending at the

reinforced rootlet area. Tight junctions between hair cells and supporting cells maintain the specificity of endolymph that bathes the stereocilia of IHCs and OHCs.

Abbr: OC – Organ of Corti, TM – tectorial membrane, OHC – outer hair cell, IHC – inner hair cells, MET – mechanicoelectrical transduction channel, IP – inner pillar, OP – outer pillar, D – Deiter's cell.