UC San Diego

UC San Diego Previously Published Works

Title

Genome-Wide Association Study of Chronic Dizziness in the Elderly Identifies Loci Implicating MLLT10, BPTF, LINC01224, and ROS1

Permalink

https://escholarship.org/uc/item/6d21j9g0

Authors

Clifford, Royce

Munro, Daniel

Dochtermann, Daniel

et al.

Publication Date

2023-11-30

DOI

10.1007/s10162-023-00917-y

Peer reviewed

JARO

ORIGINAL ARTICLE

Genome-Wide Association Study of Chronic Dizziness in the Elderly Identifies Loci Implicating *MLLT10, BPTF, LINC01224*, and *ROS1*

Royce Clifford[1,2] · Daniel Munro[3,4] · Daniel Dochtermann[5] · Poornima Devineni[5] · Saiju Pyarajan[5] · Million Veteran Program · Francesca Telese[3] · Abraham A. Palmer[3,6] · Pejman Mohammadi[7,8] · Rick Friedman[1]

Abstract

Purpose Chronic age-related imbalance is a common cause of falls and subsequent death in the elderly and can arise from dysfunction of the vestibular system, an elegant neuroanatomical group of pathways that mediates human perception of acceleration, gravity, and angular head motion. Studies indicate that 27–46% of the risk of age-related chronic imbalance is genetic; nevertheless, the underlying genes remain unknown.

Methods The cohort consisted of 50,339 cases and 366,900 controls in the Million Veteran Program. The phenotype comprised cases with two ICD diagnoses of vertigo or dizziness at least 6 months apart, excluding acute or recurrent vertiginous syndromes and other non-vestibular disorders. Genome-wide association studies were performed as individual logistic regressions on European, African American, and Hispanic ancestries followed by trans-ancestry meta-analysis. Downstream analysis included case-case-GWAS, fine mapping, probabilistic colocalization of significant variants and genes with eQTLs, and functional analysis of significant hits.

Results Two significant loci were identified in Europeans, another in the Hispanic population, and two additional in trans-ancestry meta-analysis, including three novel loci. Fine mapping revealed credible sets of intronic single nucleotide polymorphisms (SNPs) in *MLLT10* - a histone methyl transferase cofactor, *BPTF* - a subunit of a nucleosome remodeling complex implicated in neurodevelopment, and *LINC01224* - a proto-oncogene receptor tyrosine kinase.

Conclusion Despite the difficulties of phenotyping the nature of chronic imbalance, we replicated two loci from previous vertigo GWAS studies and identified three novel loci. Findings suggest candidates for further study and ultimate treatment of this common elderly disorder.

Keywords Vertigo · Imbalance · Dizziness · Genome-wide association studies · Human

Introduction

Disorders of balance are a major source of falls, a leading cause of death in the elderly [1]. Epidemiological evidence suggests that over 6 million individuals in the USA experience chronic disequilibrium that can lead to falls [2]. The incidence of dizziness increases from 22% for adults between 65 and 69 years of age to over 40% for adults between the ages of 80 and 84 years [3–5] and there is an abundance of literature concerning associated

✉ Royce Clifford
r2clifford@health.ucsd.edu

1   Department of Otolaryngology-Head and Neck Surgery, University of California San Diego, La Jolla, CA 92093, USA

2   Research Dept, Veteran Administration Hospitals, San Diego, CA 92161, USA

3   Dept. of Psychiatry, University of California San Diego, La Jolla, CA 92093, USA

4   Dept. of Integrative Structural and Computational Biology, Scripps Research, La Jolla, CA 92093, USA

5   Veterans Administrations Hospitals, Million Veteran Program, Boston, MA 02130, USA

6   Institute for Genomic Medicine, University of California San Diego, La Jolla, CA 92093, USA

7   Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, WA 98101, USA

8   Department of Pediatrics, University of Washington School of Medicine, Seattle, WA 98195, USA

Springer

comorbidities related to dizziness in the elderly, including increased falls, depression, social isolation, fear, and overall functional decline [4, 6, 7]. Thus, while falls can have multiple causes, dizziness is a strong predictor of morbidity and mortality.

Dizziness can arise from dysfunction of the balance system, which requires an integration of sensory input in large part from the vestibular portion of the cochlea. This complex organ is a mechanosensory structure transmitting signals through an elegant and highly conserved neuroanatomical group of pathways that mediates our ability to perceive linear acceleration, gravity, and angular head motion. Significant age-related degeneration occurs in nearly all areas of the vestibule, including end organ hair cells, nerve fibers, and Scarpa ganglion cells [8–15]. Amalgamation of vestibular, visual, and proprioception input occurs in the cerebellum, with further processing in the cerebral cortex.

Evidence for a genetic component to balance dysfunction in the elderly includes a large Swedish population–based twin study, where self-estimated impaired balance was assessed as a risk for osteoporotic fractures and demonstrated an age- and sex-adjusted heritability of 27% [16]. An Australian twin study, using Lord's Balance test and Step test, found heritability of 46% for sensory balance modules and 30% for static and dynamic perturbations [17].

A recent large European GWAS used International Classification of Diseases (ICD) diagnoses of vertigo [18] for phenotyping. While their phenotype included acute and chronic diseases, we addressed chronicity by including only those participants with ICDs greater than 6 months apart, as electronic health record (EHR) studies show improved sensitivity and specificity for chronic syndromes with this algorithm [19, 20].

Dizziness itself is a broader term than "vertigo," creating a challenging phenotype for study. Younger patients present acutely with vertigo or a whirling sensation emanating from stimuli within the three semicircular canals, utricle, and saccule in the inner ear. As the vestibular system undergoes age-related degeneration [8, 11], complaints become vague, often described as an instability, dizziness, swaying, or an insecure gait, making diagnosis, treatment, and phenotyping more difficult [20, 21]. The recently coined term "presbyvestibulopathy" has stricter diagnostic criteria including objective clinical testing and is uncommon in patients hospitalized with complaints of vertigo and moderate impairment [22]. In contrast, the Million Veterans Program (MVP) cohort of the present study is predominantly ambulatory and may have milder impairment not requiring hospitalization, which nevertheless subjects them to stumbling and subsequent injury.

Chronic imbalance in the elderly may be associated with deficits in vision, cerebellar function, and proprioception,

thus dissecting out a purely cochleovestibular source is problematic at best. To address these challenges of chronicity while attempting to focus on the cochlea-vestibular system, we excluded those International Classification of Disease (ICD) diagnoses that were episodic or recurrent, i.e., Meniere's and benign paroxysmal positional vertigo (BPPV). While BPPV may be a consequence of the age-related degeneration of the vestibular system, it has multiple other causes [23, 24] and is episodic. Other exclusions included ataxias, i.e., Parkinson's Disease, cerebrovascular accidents, and traumatic brain injury (TBI). Thus, although imbalance in the elderly encompasses multiple systems, we endeavored to focus on the inner ear's sensory role in chronic imbalance and documented degeneration with age by selectively defining inclusion and exclusion criteria in the electronic health record (EHR).

Here, we report a genome-wide association study (GWAS) of imbalance/dizziness and vertigo with analyses of Europeans (EU), Americans of African descent (AA), and Hispanic ancestries (LA), then combine these ancestries in meta-analysis. Downstream analysis seeks to define genes and single nucleotide polymorphisms (SNPs) in regard to colocalization with expression quantitative trait loci (eQTLs) in specific tissues and identifies strong connections to the cerebellum despite the focus on the vestibular sensory apparatus. Case-case GWAS compares allele frequencies to the replication cohort to measure end results of phenotype definitions, and fine mapping suggesting plausible causal variants.

Methods

Participants

The Million Veteran Program (MVP) is a US Veterans Administration (VA) cohort with ongoing enrollment of those eligible for care in the nationwide system. After written informed consent, participants fill out basic and lifestyle health question surveys and have bloods drawn for genotyping. Questionnaires amass data in regard to demographics, lifestyle, work, military experience, and self-reported health conditions. Information including International Classification of Diseases (ICD) diagnostic codes, questionnaires, VA, and prior active-duty health records are linked to participants' de-identified ID#s behind the VA firewall.

Phenotyping and Replication

Of the 829,497 subjects in MVP v_20_1 release, 648,224 had been assigned ancestry according to a harmonized ancestry and race/ethnicity genomic algorithm (Fig. 1) [25]. Chronic imbalance was assessed as a binary variable with a case defined as

two diagnoses of dizziness or vertigo at least 6 months apart in the VA EHR, indicating chronicity. Although "dizziness" or the report of a whirling sensation is a symptom rather than a finding, we used International Classification of Diseases (ICD) codes which use the terms "vertigo" or "dizziness" rather than imbalance, since a health provider had performed a clinical examination on at least two occasions 6 months apart or more and the imbalance had continued beyond an acute episode. Subjects at the VA normally either undergo routine checkups with their healthcare provider or visit their provider for a complaint. In either case, an ICD in the electronic health record would indicate that the subject had a complaint of imbalance during the visit. ICDs included vertigo and dizziness identified within the Veteran's Administration Informatics and Computing Infrastructure (VINCI) space via SQL tables of diagnoses and descriptions. Participants assigned to the control group carried no diagnoses related to dizziness, imbalance, vertigo, or any of the excluded diagnoses (Table S1).

Excluded cases were those not genotyped, with no documentation of sex, of only one diagnosis of dizziness or vertigo, and those with diagnoses less than 6 months apart (Fig. 1). Since sample sizes for East and South Asian ancestry participants were too small to be statistically informative (< 10,000), these were excluded as well. In an attempt to focus on cochleovestibular causes of chronic imbalance, excluded diagnoses comprised the following: acute and episodic vestibular diseases including Meniere's Disease, primary cerebellar ataxias such as Parkinson's Disease, cerebrovascular syndromes, poisoning, and overdose leading to imbalance, syncope, and hypotension (Table S1). A total of 41,233 participants with the diagnosis of benign paroxysmal positional vertigo (BPPV) [24] were excluded, as the primary focus of our work was the aging vestibule. Migrainous vertigo was not eliminated because the ICD9

and ICD10 codes do not distinguish between types of aurae. Since a traumatic brain injury (TBI) can lead to vertigo, if the subject had a positive ICD for dizziness or vertigo AND had evidence of TBI, they were excluded.

The final sample after exclusions consisted of 417,239 participants (50,339 cases and 366,900 controls) with a mean age of 61.9 ± 14.3 (Table 1). The cohort consisted of 303,743 participants of European ancestry (72.80%); 79,388 participants of African American ancestry (19.02%); and 34,108 participants of Hispanic ancestry (8.17%). Males constituted 379,968 (91.06%), and 37,271 (8.93%) were female. Of the total, 255,559 were ≥ 50 years of age (61.25% of the cohort).

Since there was no directly applicable replication cohort, we compared our phenotype to a previously published meta-analysis with the diagnosis of vertigo [18]. Although this cannot be construed as a "replication" cohort, to our knowledge there were no chronic dizziness or imbalance GWAS studies available in the literature for direct comparison.

Consent

Prior written informed consent for use of health records and genomic data was obtained for all participants in this retrospective analysis within the Million Veteran Program. This study was approved by Central VA IRB reference E22-36, qualifying for exemption under category 4(ii).

Genetic and Statistical Analysis

Genotyping and Quality Control

Analyses were based on the version 3 release of the MVP unimputed genetic dataset and version 3 release of the MVP imputed genetic dataset. Details of genotyping, quality

Fig. 1 Details of cases and controls. Exclusion criteria and numbers in green ovals. Participants were recruited from 50 Veterans Administration Healthcare Centers across the USA, beginning in 2011. Final analyzed cohort included 79,388 participants of African American descent, 303,743 of European ancestry, and 34,108 with Hispanic ancestry. Note that some excluded participants may have fit into overlapping categories
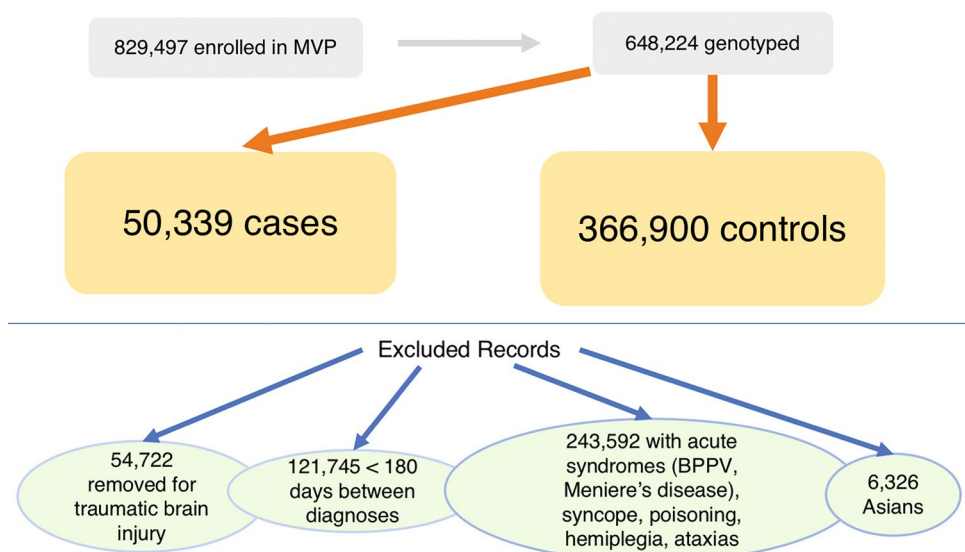
Table 1 Demographics

		Cases	(%)	Controls	(%)	Total	(%)	*p*-value[a]
TOTAL (%)		50,339	(12.06)	366,900	(87.94)	417,239		
Age (SD)		67	(11.73)	61	(14.62)	62	(14.28)	< 0.001
AFR	Male	8801	(12.93)	59,280	(87.07)	68,081	(85.75)	< 0.001
	Female	1334	(11.80)	9973	(88.20)	11,307	(14.24)	< 0.001
EUR	Male	33,605	(11.95)	247,696	(88.05)	281,301	(92.92)	< 0.001
	Female	2968	(13.23)	19,474	(86.77)	22,442	(7.08)	< 0.001
EUR ≥ 50		31,675	(12.39)	223,884	(87.60)	255,559	(100.00)	< 0.001
HIS	Male	3282	(10.73)	27,304	(89.27)	30,586	(89.67)	< 0.001
	Female	349	(9.91)	3173	(90.09)	3522	(10.33)	< 0.001

SD standard deviation, *AFR* African American ancestry, *EUR*, European ancestry, *EUR≥ 50* Europeans at least 50 years of age, *HIS* LatinX ancestry

[a]*p*-value of dizziness status predicted by demographic trait (linear regression)

control, and imputation of the MVP have been previously reported [26]. In brief, blood samples were gathered from subjects after written consent had been obtained, beginning in 2011. The Advanced Marker QC Pipeline excluded probe-set calls from batches that failed advanced QC tests, as described previously. Samples are stored at MAVERIC, VA Boston, MA. DNA was extracted and sent for genotyping at the Affymetrix Research Services Laboratory.

Genotyping had been carried out by the MVP Core via a 723,305–SNV biobank array (Axiom; Affymetrix), customized for MVP to include variants of interest in multiple diverse ancestries [26]. Imputation had been performed with Minimac 3 and the 1000 Genome Project phase 3 reference panel. Analyses used MVP Release 4 data (GrCh37). For further analysis, if only GrCh38 was available, loci and single nucleotide polymorphisms were lifted over to GrCh37 for analysis using Bioconductor in R. Final genotype data consisted of 96 million genetic variants. For each ancestry group, principal components of genotype were calculated using PLINK 2.0 alpha [27].

Subjects were genotyped on either the Affymetrix Axiom or UK BiLEVE Axiom arrays, with approximately 4700 samples per batch. Genotype calling was performed using a custom pipeline designed for biobank scale data. Routine quality control procedures were performed at each step of genotype extraction and calling. Quality control of marker genotypes included tests for batch and plate effects, deviations from Hardy–Weinberg equilibrium based on exact tests, sex effects, array effects, and discordance among technical replicates. Based on markers passing QC, subjects were removed for > 2% missing genotype rate, discrepancy between self-reported sex and genetically determined sex, or excessive heterozygosity. Phasing was performed using SHAPEIT3 in partially overlapping chunks of 15,000 markers, with the 1000 Genomes Phase

3 dataset used as a reference panel. Chunks were merged using hapfuse. Data was imputed with IMPUTE4, using the combination of the 1000 Genomes Phase 3, UK10K, and Haplotype Reference Consortium panels, where the latter was preferentially used as the imputation reference.

Although REGENIE compensates for relatedness, nevertheless, related individuals were removed. Relatedness was estimated using KING [28]. For analyses of unrelated subjects, from each pair with an estimated kinship coefficient > 0.0442 (3rd degree or closer), the individual with phenotype positive information was retained, and for pairs with identical symptoms, one was chosen at random.

For ancestral principal components calculations, SNPs were excluded that had a MAF < 5%, HWE *p-value* > 1×10^{-3}, call rate < 98%, were ambiguous (A/T, G/C), and located in the MHC region (chr6, 25–35 MB) or chromosome 8 inversion (chr8, 7–13 MB). SNPs were pairwise LD-pruned ($r^2 > 0.2$) and a random set of ~ 100 K markers was used for each subset to calculate PCs based on the smartPCA algorithm in EIGENSTRAT [29].

Power Calculations

We had previously conducted power calculations at 80% power based on a range of realistic assumptions for trait heritability, assuming 30% imbalance prevalence, a logadditive model, a type I error rate of 5.0×10^{-08}, and perfect linkage disequilibrium (LD) between marker and trait allele. A sample of 50,000 European Ancestry imbalance cases and 160,000 controls allowed us to detect a locus with ORs as low as 1.14 for allele frequencies between 0.05 and 0.20. Furthermore, there was excellent power to nominally replicate findings from the initial ancestry GWAS in other ancestry groups and European external replication studies.

Genome-Wide Association Study, Meta-analysis in MVP, and Comparison Cohort

GWAS was performed as a logistic regression testing imputed dosages for association with chronic balance under an additive model using REGENIE v2.2 [30], including the first 10 principal components of genotype, sex, and age as covariates. GWAS was analyzed separately on each of the three ancestries (EU, AA, and LA), with 913,319 SNPs filtered at MAF > 0.01, INFO > 0.60, and HWE > 1×10^{-15}. GWAS was calculated for two related phenotypes: two diagnoses ≥ 90 days apart on each ancestry—EU, AA, and LA ancestries ($N = 427,959$, cases = 55,404, controls = 372,555) and ≥ 180 days apart ($N = 417,230$, cases = 50,339, controls = 366,00). Since sample sizes including participants ≥ 50 years of age were sufficient only for EU, we then performed GWAS separately on this subset.

REGENIE runs as a 2-step process as described [30]. Step One consisted of analysis of MVP genotype array data only, where SNPs are partitioned into consecutive blocks consisting of 219 MB, with ridge regression predictions generated from each block. Predictions are then combined in a second ridge regression procedure and decomposed into separate chromosomes for leave-one-chromosome-out (LOCO) analysis. These LOCO predictions then became covariates in Step Two. In Step Two, Release 4 imputed data was used for cross validation. Firth logistic regression and saddle point approximation were used for this binary trait.

Summary statistics were combined in meta-analysis using METAL [31]. Results of each study were weighted proportional to the square root of each study's sample size. A total of 17,735,101 variants remained after filtering with INFO > 0.6, MAF ≥ 0.01, and inconsistent allele labels and strands adjusted or removed. Cochran's Q-test was implemented for each SNP.

Cohen's D was calculated to ascertain whether there was a difference in effect size between the European cohort and those over 50 years of age. Comparisons were computed for the lead SNP effect sizes in chromosomes 10 and 17, using beta for effect sizes, and converting standard error to standard deviation corrected for sample size in each cohort.

Quantile-quantile (QQ) plots of expected versus observed $-\log_{10}$ *p-values* included genotyped and imputed SNPs at MAF ≥ 1% (Fig. S1). We calculated h^2 values on GWAS summary data of MVP European participants ($N = 303,743$) with linkage disequilibrium score regression (LDSC) [32]. The proportion of inflation of test statistics due to the actual polygenic signal (rather than other causes such as population stratification) was estimated as 1 - (1 - LDSC intercept -1)/(mean observed chi-square - 1), using LD score regression [32, 33]. For primary analyses, genome-wide significance was declared at *p-value* < 5×10^{-08}. Regional association plots were generated using LocusZoom with 400 KB windows around the index variant with LD patterns calculated based on 1KGP EUR, AMR, or multiple reference populations for trans-ancestry analysis [33].

Methods of the Iceland/Finland/UK/US cohort GWAS are explained in detail in the original paper [18] and Supplementary Methods. For the Iceland/Finland/UKB/US comparison cohort ($n = 48,072$ and 894,541 controls), replication was considered significant at *p-value* < 0.00625, a Bonferroni adjustment for the eight significant lead SNPs from our GWAS.

The proportion of heritability explained by our GWAS was performed with linkage disequilibrium scores (LDSC) [34]. Pairwise genetic correlations (r_g) between results on chronic imbalance, hearing [35], and tinnitus [36] were calculated using bivariate LDSC regression, with publicly available GWAS summary statistics. LDSC regressions were computed with the Ice/Fin/UKB/US cohort for comparison.

Post-GWAS Analysis

Case-Case Genome-Wide Association Study

Case-case GWAS tests for differences between SNPs of two separate disorders, in this study to the Ice/Fin/UKB/US cohort, using CC-GWAS in R. Input consists of effect sizes, effective N's for cases and controls, Ps, population statistics and heritability for the disorders, and overlap between the two cohorts for significant SNPs (*p-value* = 5.0×10^{-08}) [37]. Output includes statistically significant effect sizes reflecting direction and magnitude of effect, discards false positives, and measures genetic distance between the two disorders.

Next, probabilistic colocalization between the GWAS associations (EU ancestry only) and expression quantitative trait loci analysis (eQTLs) was performed using fastENLOC to identify genes whose expression may mediate the GWAS association [38, 39]. We used the pre-computed eQTL annotations for GTEx v8 data from the fastENLOC repository, with SNPs labeled with rsIDs [40]. Approximately independent LD blocks for European ancestry from LDetect were also included as input [41]. fastENLOC was run separately on the eQTL data from each of the 49 tissues, and regional colocalization probability per eQTL per tissue was examined for evidence of colocalization. We used an RCP > 0.1 (regional colocalization probabilities) cutoff for putative causal gene contributors from the original fastEN-LOC study.

Fine Mapping

Fine mapping of European GWAS hits was computed with a "Sum of Single Effects" (susieR) for summary statistics model using default settings in R [42]. susieR v0.12.27 is a Bayesian method to assess uncertainty in which variables are

selected, involving highly correlated variables [43]. Originally formulated for linear regression, susieR has been verified for logistic regression. susieR provides "Credible Sets" (CS) of variables which include relevant rsIDs of SNPs and a posterior inclusion probability (PIP) set at 0.95 for each variable based on effect sizes and linkage disequilibrium (LD) between SNPs. Minimum LD is 0.5 within each CS, and indicates that within each CS, one of the variables has a non-zero effect coefficient.

An interval of 2.5 million base-pairs surrounding each significant locus was identified. This range was then downloaded from EnSembl in Data Slicer using these chromosome positions in GrCh38 and ancestries to match our lead SNPs from the 1000 genome. For example, since significant SNPs for the European cohort were identified in chromosomes 10 and 17, the LD matrix was fashioned from 1000 genomes of CEU individuals. Plink2 was used to convert the .vcf file to plink binary, then linkage disequilibrium matrices were fashioned for use in susieR [42]. A fitted regression with summary statistics was fashioned for each significant locus, which created credible sets of potentially causal variants, based on posterior probability of 0.95, with 100 iterations and refinement of the model. Increasing the iterations did not improve the model.

Functional Mapping and Annotation

Functional mapping and annotation of genetic associations (FUMA), version 1.4.2, was examined based on human genome assembly GRCh37 (hg19), with default settings unless stated otherwise [44]. Gene-based analysis was performed with the FUMA implementation of MAGMA version 1.08 [45]. 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 1KGP reference genotypes). SNPs in risk loci were mapped to protein coding genes with a 10 kb window.

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, Combined Annotation Dependent Depletion (CADD), RegulomeDB (RDB), and chromatin states in brain tissues/cell types [46]. Next, eQTL mapping was performed on significant (FDR < 0.05) SNP-gene pairs, mapping to GTEx v7 brain tissue, RNA-seq data from the CommonMind Consortium, and the BRAINEAC database. Chromatin interaction mapping was performed using built-in chromatin interaction data from the dorsolateral prefrontal cortex, hippocampus and neuronal progenitor cell line and adult and fetal cortex tissue. An FDR of $< 1 \times 10^{-5}$ defined significant interactions, based on previous recommendations, modified to account for the differences in cell lines used here.

Significance of genes was determined with a Bonferroni-corrected threshold of $p\text{-}value = 0.05/18,914$ protein coding genes $= 2.644 \times 10^{-06}$. eQTL gene mapping was included using GTEx v8 Brain with 13 tissues. 3D chromatin interaction mapping was performed to capture chromatin conformation with dorsolateral prefrontal cortex and hippocampus tissues (HiC GSE87112) [47].

Gene-Based and Gene Set Analysis with MAGMA

To ascertain specific biological pathways implicated, gene-based test statistics were used to perform a competitive set-based analysis of 10,894 pre-defined curated gene sets and GO terms obtained from MsigDB using MAGMA [45]. Significance of pathways was set at a Bonferroni-corrected threshold of $p\text{-}value = 0.05/10,894 = 4.6 \times 10^{-06}$. To test if tissue-specific gene expression was associated with dizziness, gene set–based analysis was also used with expression data from GTEx v7 RNA-seq and BrainSpan RNA-seq, where the expression of genes within specific tissues was used to define the gene properties used in the gene set analysis model.

Since the above platforms do not include cochleovestibular data, umgear.org [48] was explored for analogs to human tissue expression. Homologous genes in *mus musculus* were identified from Table 5 and explored in gene sets, i.e., utricular and crista ampullaris tissues. The following genes were explored: *Mfhas1*, *Mllt10*, *Bptf*, *1810010H24Rik* (homologue for C17orf58), *Kpna2*, *Zfp859* (ZNF91), *Skida1*, and *Gopc*. *LINC01224* is a primate-only non-coding RNA and is thus not represented in mice. Heat maps for identified genes were fashioned from utricular and crista ampullaris P7 studies (GSE172327, GSE168901, respectively).

PheWAS was performed using SNPs with the smallest *p-value* from fine mapping at the Atlas of GWAS Summary, with significance based on Bonferroni correction for the number of phenotypes engaged by each SNP [49].

Results

Demographics

Phenotype inclusions and exclusions are seen in Fig. 1 and Table S1. Overall case prevalence was 12.06%, with case mean age 67.21 ± 11.67 versus control mean age 61.15 ± 14.45, range (19 to 108), $p\text{-}value < 2.0 \times 10^{-16}$. Women had a prevalence of 12.48% versus 13.67% for men ($p\text{-}value < 2.0 \times 10^{-16}$). Those of African ancestry were less likely to have a EHR diagnosis of chronic dizziness than those of European or Hispanic ancestry (OR

0.728, 95% CI 0.756, 0.700, and OR 0.875, 95% CI 0.841, 0.909, *p-value* $< 2.0 \times 10^{-16}$ and 1.97×10^{-10}, respectively). Chronic imbalance was associated with increased age, *p-value* < 0.001 (Table 1).

The Ice/Fin/UKB/US cohort consisted of 48,072 cases and 894,541 controls aggregated from four European cohorts [18]. Case rate was reported as 5.10% [18, 50–52].

Heritability and Genetic Correlation of Chronic Imbalance with Other Traits

SNP heritability for those of European descent was h_s^{SNP} of 0.059 ± 0.0083 on the liability scale, assuming sample prevalence of 12% and population prevalence of 30% [4]. QQ plots identify a GCλ = 1.127, indicative of inflation of the test statistics (Fig. S1). LDSC intercept = 1.027, thus polygenic effects account for 94.87% of observed inflation.

LDSC genetic correlation compared genomic imbalance results to other cochlear disorders of the inner ear in EU ancestries (Table 2). Our phenotype was moderately correlated with hearing difficulties, $r_g = 0.314 \pm 0.048$ (*p-value* $= 5.77 \times 10^{-11}$), and tinnitus, $r_g = 0.20 \pm 0.8$ (*p-value* $= 1.51 \times 10^{-02}$). The Ice/Fin/UKB/US showed similar correlations for hearing, $r_g = 0.19 \pm 0.049$ (*p-value* $= 2.0 \times 10^{-4}$) and tinnitus, $r_g = 0.19 \pm 0.049$ (*p-value* $= 8.7 \times 10^{-05}$).

Genome-Wide Association Study and Replication

Results were not substantially different ($r_g = 1.00$, *p-value* $= 0.0$, Table 2) between diagnoses 90 days apart and 180 days apart. Comparison of GWAS on all Europeans to those of European ancestry ≥ 50 years of age yielded similar results, with no significant changes in Z-scores of the identified loci on chr 10 and 17 (data not shown, Cohen's $D = 0.001$ for both hits). Lead SNPs for EU ≥ 50 were in LD with both GWAS of all EU participants and with trans-ancestry GWAS results, with an identical lead SNP on chr10. On chromosome 17, rs7221651/rs4791046 $R^2 = 0.99$ for EU versus EU ≥ 50 years of age. Further, results for EU ≥ 50 were similar to GWAS on the entire cohort (Fig. S8). Thus, we report the results from the ≥ 180 days diagnosis and the entire cohort.

GWAS of chronic imbalance on EU ancestry indicated two significant loci, lead hits rs10828248 on chromosome 10 (*p-value* $= 3.17 \times 10^{-08}$) and rs7221651 on chromosome 17 (*p-value* $= 3.71 \times 10^{-10}$) (Fig. 2a and Table 3). GWAS of the Hispanic ancestry indicated a novel locus on chromosome 6, rs71717606 with *p-value* $= 4.84 \times 10^{-10}$ (Fig. 2b) with no significant hits in the AA ancestry (Fig. 2c). GWAS in trans-ancestry analysis identified four significant loci (Fig. 2d) at *p-value* $< 5.0 \times 10^{-08}$, including the two loci on chromosomes 10 and 17 identified in the EU GWAS.

Table 2 Linkage disequilibrium scores on cochlear disorders and replication cohort

p1	p2	rg	se	Z-score	p-value	h2 p1 liab	h2_p1 se	h2 p2 liab	h2 p2 se	h2_int	h2_int_se	gcov_int	gcov_int_se
Balance[a]	**I/F/UK/US**[d]	0.67	0.0734	9.16	5.3433e−20	0.0469	0.0059	0.0666	0.0074	1.0053	0.0076	0.0051	0.0052
Balance[a]	**Tinnitus**[b]	0.20	0.0833	2.43	0.0151	0.0462	0.0066	0.1026	0.0066	1.0116	0.0074	0.0028	0.0059
Balance[a]	**Hearing**[c]	0.31	0.0481	6.55	5.7721e−11	0.0461	0.0059	0.1199	0.0006	1.0273	0.0094	0.0002	0.0053
Tinnitus	**Hearing**[c]	0.52	0.0309	16.76	4.9136e−63	0.1071	0.0067	0.1193	0.0058	1.0282	0.0081	0.1544	0.0054
I/F/UK/US[d]	**Tinnitus**[b]	0.19	0.049	3.92	8.7001e−05	0.067	0.0072	0.1066	0.0068	1.0098	0.0073	0.0283	0.0048
I/F/UK/US[d]	**Hearing**[c]	0.19	0.0491	3.78	0.0002	0.067	0.0072	0.1208	0.006	1.0266	0.0087	0.0183	0.0052
Balance[a]	**Balance ≥ 50**	0.98	0.0063	154.25	0.0	0.0459	0.0058	0.0482	0.007	1.0217	0.0075	0.9827	0.0074
EU Bal_90	**EU Bal_180**	1.00	0.0035	288.08	0.0	0.0172	0.0024	.0172	.0024	1.0263	0.0100	0.9886	0.0098

p1 phenotype 1, *p2* phenotype 2, *rg* genomic correlation, *se* standard error, *h2* SNP heritability on the liability scale, *int* intercept, *gcov* genetic covariance, *I/F/UK/US* comparison cohort consisting of meta-analysis of Iceland/Finland/UKBiobank/US data, Skuladottir et al. [21], *EU Bal_90* European cohort for 2 diagnoses greater than 90 days apart, *EU Bal_180* European cohort for 2 diagnoses greater than 180 days apart

[a] Balance study – Trans-ancestry GWAS from Million Veterans Program (this study)

[b] Tinnitus study – European GWAS performed in UK Biobank [32]

[c] Hearing study – European GWAS performed in UK Biobank [31]

[d] Vertigo study – European meta-analysis GWAS on Iceland/Finland/UKBiobank/USA for diagnosis of vertigo [21]
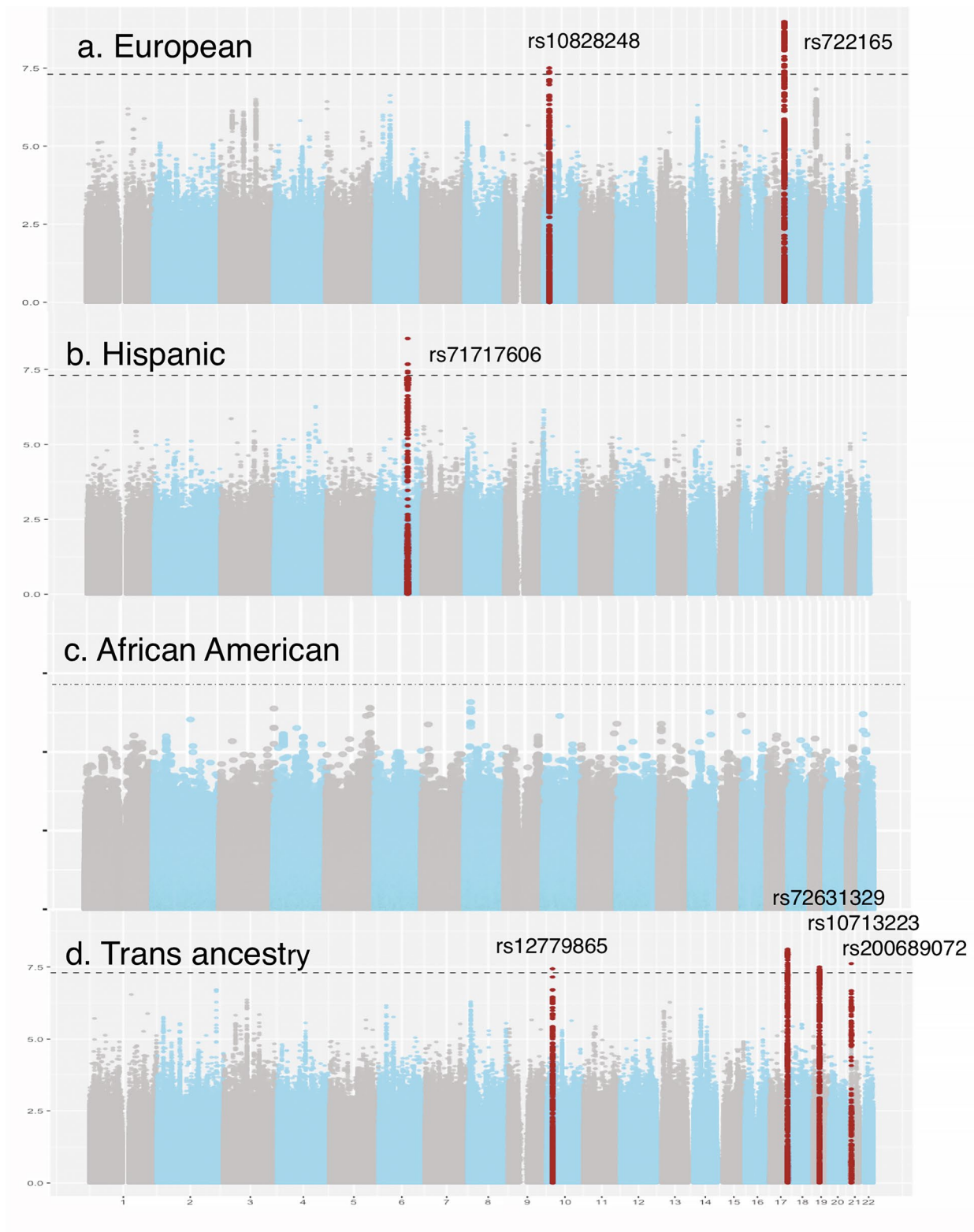
Fig. 2 Manhattan plot of GWAS for each ancestry, with lead SNPs indicated in each locus for which *p-value* < 5.0 × 10⁻⁰⁸: **a** GWAS of European cohort (EU). **b** GWAS of Hispanic ancestry (LA). **c** GWAS of African American ancestry (AA). **d** Meta-analysis of European, African American, and Hispanic populations. The horizontal line represents the Bonferroni-adjusted significance threshold *p-value* 5 × 10⁻⁰⁸. Red dots indicate single-nucleotide variations within the significant regional loci (See Supplemental Figs. S5–S12)

Table 3 Genome-wide significant loci associated with dizziness in the Million Veteran Program with replication in a European cohort

a. Discovery - Million Veteran Program (n = 417,239)

Cohort	Chr	Marker	BP[a]	A1	A2	A1Frq	INFO	BETA	SE	p	Z-score	Wt.	Dir[b]
European	10	rs10828248	21824619	A	G	0.638	0.9997	−0.0458	0.008	3.17E−08	5.532	303,743	
	17	rs7221651	65832390	A	G	0.795	0.9778	−0.0623	0.01	3.71E−10	6.266	303,743	
Europeans ≥ 50[h]	10	rs10828248	21824619	G	A	0.639	0.9997	−0.0481	0.008	1.49E−08	5.663	269,329	
	17	rs4791046	65851973	A	G	0.528	0.9778	−0.0598	0.01	5.24E−09	5.839	269,329	
Hispanic	6	rs71717606	117804875	C	CT	0.847	0.8871	−0.2138	0.036	4.84E−10	5.934	34,108	
Trans-ancestry[e]	10	rs12779865	21829162	T	C	0.289	>.60	0.0133	0.002	3.65E−08	5.507	417,239	+++
	17	rs72631329	65818374	T	C	0.129	>.60	0.0166	0.003	7.61E−09	5.777	417,239	+++
	19	rs10713223	23604022	T	TA	0.684	>.60	−0.0130	0.002	3.11E−08	−5.535	417,239	---
	21	rs200689072	18519809	A	C	0.553	>.60	−0.0123	0.002	2.43E−8	−5.578	417,239	---

b. Comparison - (n = 942,613)

Cohort	C Chr	Marker	BP[a]	A1	A2	A1frq	OR	SE	p[c]	Dir[d]	Predicted genes within loci
European	10	rs10828248	21824619	G	A	0.329	1.0284		**4.79E−04**	--	*CASC10; SKIDA1; MLLT10; DNAJC1; COMMD3 COMMD3-BMI1; BMI1; SPAG6*
	17	rs7221651	65832390	A	G	0.2	1.0213		2.59E−02	-+	*BPTF; C17orf58; KPNA17*
Europeans ≥ 50[h]	10	rs10828248	21824619	G	A	0.329	1.0284		**4.79E−04**	-+	*CASC10; SKIDA1; MLLT10; DNAJC1; COMMD3; COMMD3-BMI1; BMI1; SPAG6*
	17	rs4791046	65851973	A	G	0.2	1.02		3.06E−02	-+	*BPTF; C17orf58; KPNA17*
Hispanic[j]	6	rs71717606	117804875	CTTTTTT	CT	0.0382	1.0206		2.92E−01	-+	*GOPC; ROS1; DCBLD1*
Trans-ancestry[e]	10	rs12779865	21829162	C	T	0.3179	1.0237		**3.98E−03**	++	*CASC10; SKIDA1; MLLT10; DNAJC1; COMMD3; COMMD3-BMI1; BMI1; SPAG6*
	17	rs72631329	65818374	T	C	0.1278	1.0175		1.41E−01	++	*BPTF; C17orf58; KPNA17*
	19[f]	rs141771535	23601989	G	A	0.218	1.1031		**4.55E−27**	f	*ZNF91; LINC01224*
	21[g]	rs13047989	18530070	A	G	0.346	1.0045		3.69E−01	g	

A adenosine, *C* cytosine, *G* guanidine, *T* thymidine, *INFO* imputation information score, *Frq* frequency, *MVP* Million Veteran Program, *SE* standard error, *OR* odds ratio

[a]Indicates base pair position on chromosome GrCh37 Human Genome Build hg19

[b]Comparison of direction of effect in three ancestries in MVP, i.e., European, African American, and Hispanic

[c]Replicated loci in bold. Significance in replication cohort at p < 6.25E−03 after Bonferroni correction for 8 variants tested for replication

[d]Comparison of direction of effect between MVP and replication cohort

[e]Trans-ancestry consists of European (n = 303,743), African American (n = 79,388), and Hispanic (n = 34,108) meta-analysis

[f]Proxy variant for MVP Marker rs10713223, rs141771535 with r² = 0.82

[g]Proxy variant for MVP Marker rs200689072, rs13047989 with r² = 0.66

[h]MVP participants of European descent ≥ age of 50 (n = 258,533)

[j]Replication in European cohort

Comparing lead SNPs in chromosomes 10 and 17 between trans-ancestry and EU GWAS' demonstrated moderate to high LD. For chr10, R^2 of rs12779865/rs10828248 $= 0.87$ and for chr17, R^2 of rs72631329/rs7221651 $= 0.54$.

The MVP chronic imbalance phenotype and Ice/Fin/UKB/US vertigo phenotype demonstrated a genetic correlation of $r_g = 0.67 \pm 0.073$, $p\text{-}value = 5.34 \times 10^{-20}$ (Table 2). Despite differences in phenotypes, two loci were replicated in the Ice/Fin/UKB/US cohort (Table 3b), lead SNPs on chromosome 10, rs12779865/rs10828248 ($p\text{-}value = 3.98 \times 10^{-03}$ and 4.79×10^{-03}, respectively) and chromosome 19 rs10713223 ($p\text{-}value = 1.45 \times 10^{-27}$). Locus-Zoom plots are in Figs. S2–S9.

Gene set analysis identified seven mapped genes at $p\text{-}value < 0.05/19,093 = 2.619 \times 10^{-06}$ for 19,093 protein coding genes analyzed (Table 4). For the gene set analysis, two of the genes identified in our cohort were replicated in the Ice/Fin/UKB/US cohort: $ZNF19$ ($p\text{-}value = 6.34 \times 10^{-14}$) and $MLLT10$ ($p\text{-}value = 1.76 \times 10^{-03}$) [44].

Case-case GWAS (CC-GWAS) is a method to test for genetic differences between allele frequencies of two studies [37]. Comparison with Ice/Fin/UKB/US indicated a sole significant difference in alleles on the chromosome 19 locus (Table S2), with lead SNP rs17473980 ($p\text{-}value = 1.52 \times 10^{-23}$) (data for remainder of CC-GWAS upon request.).

Fine mapping was performed within significant loci using summary statistics and an LD matrix from 1000 genomes with ancestry-matched values [42]. Figure S10 shows a list of "credible sets" with a posterior probability > 0.95 for loci identified as significant. For the chromosome 6 sets, fine mapping identified SNPs within an intron of $ROS1$. The sets within chromosome 10 are within an intron of $MLLT10$ and in high LD with the lead SNP. In chromosome 17, the five identified SNPs are in equilibrium with the lead SNP, thus no conclusion as to "cause" can be made; however, they are all within introns of $BPTF$. In chromosome 19, both the lead SNP and the two identified SNPs are within a long non-coding RNA, $LINC01224$, in contrast to the Ice/Fin/UKB/US GWAS, where the lead SNP was within $ZNF19$. This finding is in agreement with CC-GWAS results, which indicated a significant difference in alleles within this locus.

Integration with Functional Genomic Data

We examined 54 body tissues for gene enrichment using sumstats from the EU GWAS and found that although ataxias and Parkinson's disorders were excluded from the phenotype, the cerebellum was nevertheless significantly enriched at $p\text{-}value = 9.4$ is 10^{-04} (Figs. 3a, S2, S3, S6−S8, and Tables S4 and S5). When reanalyzed with 3D chromatin

Table 4 Genome-wide significant genes associated with dizziness in Million Veteran Program

| ANCESTRY | GENE | CHR | Start[a] | Stop[b] | MVP discovery | | | Iceland/Finland/UKB/US replication | | | |
					# SNPs[c]	# params[d]	p	#SNPs[c]	Param[d]	p[e]	Dir Dis/Rep
TRANS-ANCESTRY	*MFHAS1*	8	8,640,864	8,751,155	1016	101	1.14E−06	517	29	0.74379	+-
	BPTF	17	65,821,640	65,980,494	832	58	4.84E−08	280	13	0.03156	++
	KPNA2	17	66,031,635	66,042,958	78	20	1.03E−06	39	5	0.33328	++
	ZNF91	19	23,487,793	23,578,362	584	51	4.98E−07	*304	12	6.14E-14	++
	LINC01224	19	23,582,035	23,598,876							
EUROPEAN	*MFHAS1*	8	8,640,864	8,751,155	550	28	1.02E−06	517	29	0.74379	++
	MLLT10	10	21,823,094	22,032,559	193	23	1.78E−06	*208	32	0.00176	**++**
	BPTF	17	65,821,640	65,980,494	305	9	5.04E−10	280	13	0.03156	++
	C17orf58	17	65,987,217	65,989,765	3	1	2.54E−07	3	1	0.02915	++
	KPNA2	17	66,031,635	66,042,958	39	4	6.76E−07	39	5	0.33328	++
	ZNF91	19	23,487,793	23,578,362	329	11	1.58E−06	*304	12	6.14E-14	++
	LINC01224	19	23,582,035	23,598,876							
EUR ≥ 50	*SKIDA1*	10	21,802,407	21,814,611	11	4	1.92E−06	11	5	0.00994	++
	MLLT10	10	21,823,094	22,032,559	193	23	4.97E−07	*208	32	0.00176	**++**
	BPTF	17	65,821,640	65,980,494	306	9	6.51E−09	280	13	0.03156	++
HIS	*GOPC*	6	117,639,374	117,923,691	1157	46	2.44E−07	1053	47	0.81116	++

MVP Million Veteran Program, *CHR* chromosome, *Dir Dis/Rep* direction of effect between discovery and replication

[a]Start position in GrCh37 Human Genome Build 19

[b]Stop position in GrCh37 Human Genome Build 19

[c]SNPs tested within gene

[d]Number of variants used to compute the gene-wise statistics

[e]Significance in replication cohort at $p < 6.25E{-}03$ for 8 genes tested

interaction mapping based on Hi-C and Chia-PET (Fig. 3b), frontal cortex and brain cortex genes became significantly over-expressed as well [44, 53]. It should be noted that this resource does not include the cochleovestibular organ.

We then looked at functional analysis of lead variants for all ancestries (Table S5, and Figs. S2–S9). All significant loci contained multiple SNPs with Combined Annotation Dependent Depletion (CADD) scores > 12.37 and Regu-lomeDB scores < 5, indicating that the SNP alteration is within the top 7% of deleterious substitutions and the locus contains a transcription factor binding site, a DNase peak, and/or eQTL [46].

We tested for colocalization between the European GWAS associations and eQTLs in 49 tissues assayed by the GTEx Project [40]. Regional colocalization probabilities were below 0.1 for all eQTLs, indicating lack of evidence for gene expression in the tested tissues mediating the genetic associations with chronic imbalance (Fig. 3c). The strongest colocalizations were with *MIR1915HG* (*CASC10*) in the aorta (RCP = 0.068) and *MLLT10* in sun-exposed skin (RCP = 0.064), while all other eQTLs had RCP < 0.03. Once again, these tissues do not include the cochleovestibular system.

In cells specific to the vestibular system examined in the murine model, several of the significant genes were over-expressed (Fig. 4). Particularly, *Bptf* was expressed in Type II hair cells of the utricle, and various cells of the crista amp-ullaris. *C17orf58*, an uncharacterized open reading frame, has been linked to longevity-promoting transcription factor *GATA2*, which has an effect on extracellular matrix mod-eling [54]. The fine-mapped genes identified in this study, i.e., *BPTF*, *ROS1*, and *MLLT10*, are strongly expressed in several tissues of the vestibule [55]. Although ROS1 is not represented in this particular single-cell study, it is identified in other murine studies, i.e., the vestibular sensory epithelium of the P0 mouse and within utricle and vestibular ganglion of the P16 mouse [56]. Long non-coding RNAs are not included in this dataset and therefore no conclusion can be made.

PheWAS of > 3300 traits and disorders for each signifi-cant SNP in the fine mapping analysis showed multiple cor-relations (Fig. S11 and Table S6). From the chr6 rs72965321 locus, the most significant hit was a correlation with phe-notypes for skin pigmentation (*p-value* = 1.87×10^{-05}). Although it did not maintain significance in colocalization with eQTLs, skin pigmentation was noted in association with *MLLT10*. Hippocampus anisotropy in diffusion tract imaging (DTI) indicates loss of white matter tracts in the hippocampus (*p-value* = 8.44×10^{-05}). The chr10 rs8904565 hit was significant for multiple areas of the brain, includ-ing anisotropy of the fornix (*p-value* = 3.39×10^{-11}), poste-rior limb of the internal capsule (*p-value* 4.87×10^{-09}), left accumbens, total brain volume, posterior corona radiata, and intra-cranial areas. Chr17 rs7221651 was significant for mul-tiple impedance measures and other measurements of body mass (impedance measurement of trunk fat mass, *p-value* 7.05×10^{-30}). Chr19 rs612285 showed minimal significance for frequent insomnia (*p-value* 1.40×10^{-04}) and menopause (*p-value* 2.89×10^{-04}).

Discussion

Good balance requires integration of sensory input from the vestibular system, proprioception, and the eye, with transla-tion to motor output of the eye and body muscles. Vestibular dysfunction, as measured by a modified Romberg, increases with age, and is abnormal in 35.4% of US adults over 40 years [2]. Imbalance carries a 12× increase in the odds of uninten-tional falling, leading to morbidity and mortality among older individuals [57]. Part of the reason for patients' lack of report of true vertigo rather than dizziness or imbalance is related to central compensation of vision and proprioception, and these systems degenerate with age as well [4].

"Minimal phenotyping" is defined as dependence on self-report, i.e., subjective "symptoms" rather than objec-tive "signs" for case definitions and can degrade the qual-ity and specificity of GWAS results [58]. Thus, we did not include self-report of imbalance or of falls. Further, it has been shown that even studies that use a single diagnosis in the electronic health record, implying an examination by a healthcare provider, can elicit a low sensitivity of 50% and specificity of 81%, leading to high rates of false positives in subsequent genome-wide association studies (GWAS), despite single examination by a clinician [19]. Two diagnoses in the EHR at least 6 months apart addressed both chronicity and allowed a second examination by a healthcare provider.

The non-specific report of dizziness can lead to a myriad of diagnoses, depending on testing and findings, including benign paroxysmal positional vertigo (BPPV), unilateral vestibulopathy, orthostatic hypotension, low vision, proprioceptive impairment, and extrapyramidal disorders, among others [57, 58]. Previous genome-wide association studies (GWAS) included reports on vestibular neuronitis [50] and motion sickness [51]. Benign paroxys-mal positional vertigo (BPPV) is the most common cause of dizziness in the elderly, and whether these diagnoses should be included in the phenotype is debatable. It is considered to be caused by free-floating debris attached to the cupula or within semicircular canals [24]; however, the source of the dislodgment of otoconia can occur in younger patients and have multiple other etiologies, such as traumatic brain injury. In addition, BPPV is treatable with repositioning techniques and is thus an intermittent disorder [59]. Here, we attempted to focus on chronic-ity, leading us to perhaps address one of its underlying causes, the source of the degeneration. Pathologic studies
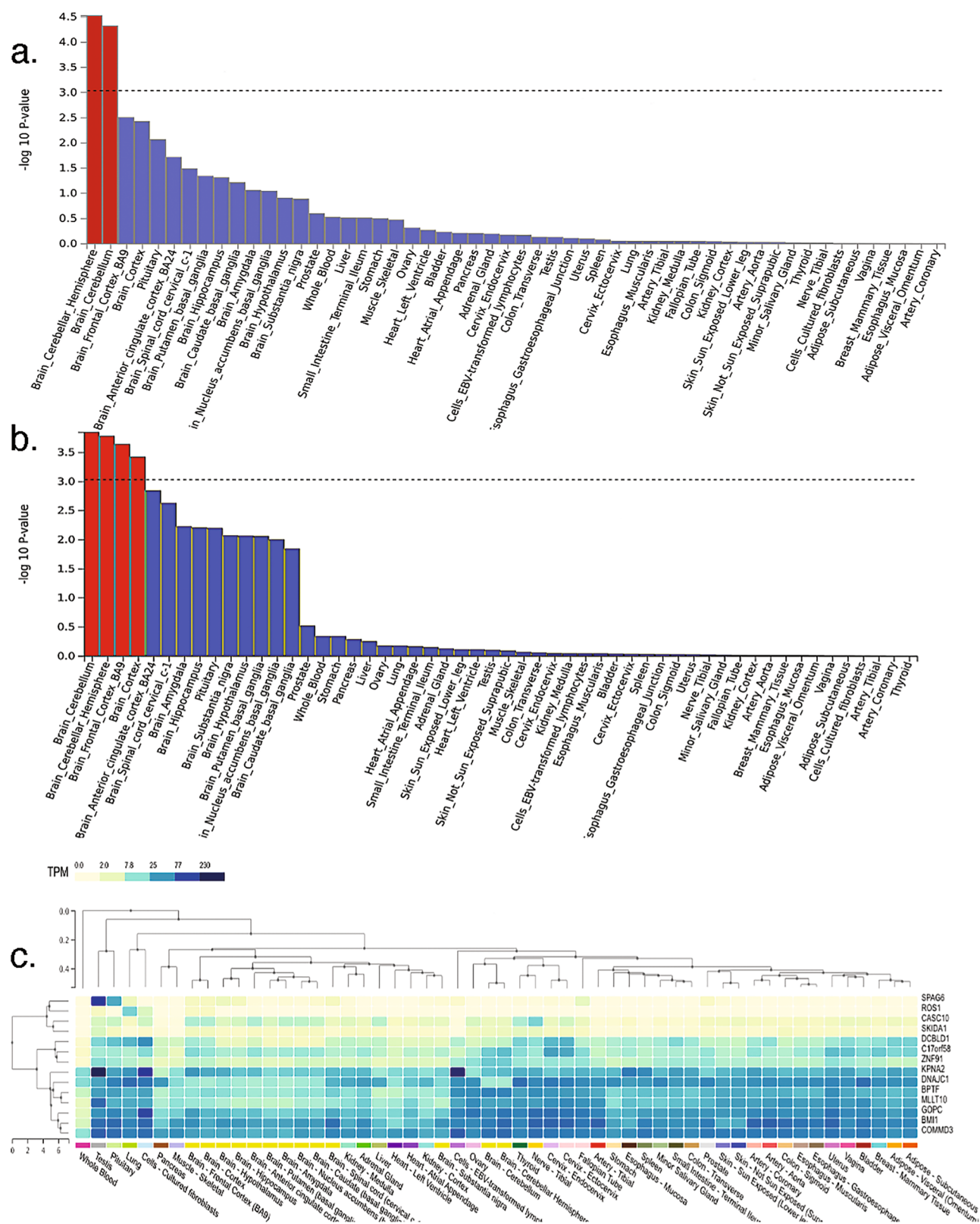
Fig. 3 Gene enrichment comparing body tissues. **a** With no chromatin enhancement, using the full distribution of SNP *p-value*s, only cerebellum shows significant expression. Dotted line denotes significant *P* corrected for 54 tissues. **b** With 3D chromatin analysis including Hi-C and Chia-PET interchromatin analysis, other aspects of the brain are enhanced, specifically frontal cortex. **c** Heatmap of gene expression for GWAS loci genes in each GTEx v8 tissue. Rows and columns are arranged by hierarchical clustering of the transcripts per million (TPM) values
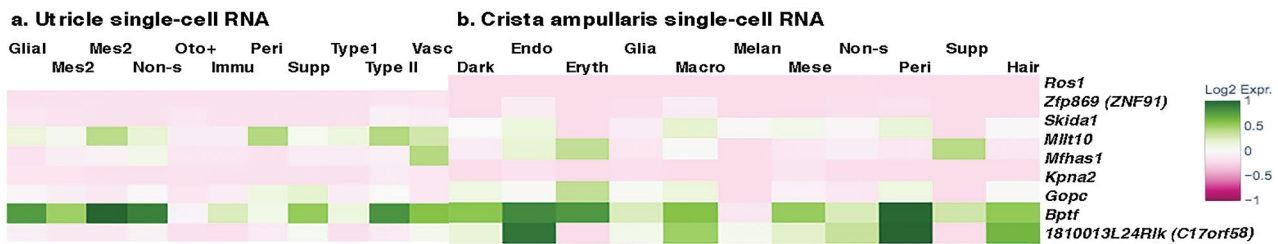
Fig. 4 Heatmaps of significant genes expressed in single-cell RNA studies of *M. muscularis* created at umgear.org/multigene-curator. html; **a** utricle P7. Cells are glial, mesenchymal1, mesenchymal2, non-sensory epithelial, oto+, immune, pericytes, supporting cells, Type 1 hair cells, Type 2 hair cells, vascular; **b** crista ampullaris P7. Cells are dark cells, endothelial, erythrocytes, glial, macrophages, melanocytes, mesenchymal, non-sensory, supporting cells, hair cells. Note: *Ros1* was not reported in the utricle. Derived from GSE172327 and GSE168901, respectively

demonstrate decreases in numbers, volume, and change in shape of otoconia during aging, and these irregularities may result in dislodgment and subsequent intermittent stimulus of the connected semicircular canals; however, the degeneration itself has not been shown to cause vertigo [24]. Previous GWAS have focused specifically on the diagnosis of vertigo, whether acute or chronic, excluding the complaint of "dizziness," which although it may be vague, may be more apt to capture the elderly complaint. In order to focus further on chronic imbalance, we excluded acute and intermittent vertiginous syndromes (i.e., Meniere's Disease, benign paroxysmal positional vertigo, and acute vestibular neuronitis). To focus on the chronic cochleovestibular disorder of aging, we also excluded non-vestibular system gait disorders such as Parkinson's and other ataxias. We also eliminated over 25,000 of those diagnosed with traumatic brain injury (TBI), since TBI is a strong predictor of vertigo that is presumed to have a separate etiology [60]. We used the ICD diagnosis of "dizziness" in an attempt to capture the most common complaint of chronic imbalance in the elderly, even though that may include some disorders originating outside the vestibule.

With the difficulty pinpointing this phenotype and excluding causes of imbalance outside the vestibular sensory organ, i.e., vision and proprioception, diagnoses are subject to misclassification. We first compared this study to the GWAS investigation that used a diagnosis of true vertigo [18]. Correlation between our "dizziness" study and the "vertigo" study was strong and highly significant at $r_g = 0.67$, p-value $= 5.34 \times 10^{-20}$, indicating that our "chronic imbalance" phenotype shared common genetic influences with other acute vestibular syndromes.

Some of the differences may reflect the removal of acute and recurrent causes of vertigo, as the findings in the study of vertigo was highly driven by BPPV [18]. To further explore the differences in phenotypes, we performed case-case GWAS, which compares individual allele differences from each study. Based on > 7 million SNPs compared between the two studies, the only locus that demonstrated significantly different alleles was on chromosome 19. While the Ice/Fin/UKB/US study had a strong hit suggesting *ZNF91*, our fine mapping analysis identified a credible set of SNPs within nearby *LINC01224* [60]. *LINC01224* affects the transcription of *ZNF91* by regulation of long-range interactions between the *ZNF91* enhancer and promoter [61]. *ZNF91* is a transcription factor specifically required to repress SINE-VNTR-Alu retrotransposons and, interestingly, is primate-specific [62]. The interplay of regulatory elements in this locus is yet to be determined. As further suggestion that this is a valid phenotype, this study replicated two of the loci in the Ice/Fin/UKB/US study, on chromosome 10 and chromosome 19. Thus, while our phenotype may not allow interpretation as a strictly vestibular organ description, and we cannot rule out other non-vestibular causes for the general term of "dizziness," comparison with another GWAS based on vertigo is favorable.

In post-GWAS analysis, we elicited potential causal SNPs and genes from significant loci. All lead SNPs and significant fine mapping SNPs identified were intronic or within a long non-coding RNA [63]. SNPs identified by fine mapping include intronic variants in the gene histone lysine methyltransferase *DOT1L* cofactor (*MLLT10*) within an enhancer region [64]. *MLLT10* is expressed in the vestibule and the cochlea [48] and is thought to prevent somatic cell reprogramming through regulation of *DOT1L*-mediated H3K79 methylation. A SNP in LD with the top hit rs10828248 has been associated with the genesis of meningiomas ($R^2 = 0.8403$, p-value < 0.001) [65].

Fine mapping identified a credible set of intronic SNPs in the gene bromodomain PHD finger transcription factor (*BPTF*) in chromosome 17. *BPTF* is the largest subunit of the nucleosome remodeling complex NURF and has been implicated in neurodevelopmental disorders, developmental delay, and autism [66]. While pheWAS identified predominant associations with body mass rather than

a neurological function, this gene is identified in Type I and Type II hair cells of the mature mouse utricle in the vestibular system in higher abundance than in the cochlea (Fig. 4) [48]. PheWAS also demonstrated a high significance with skin pigmentation, and *BPTF* has been identified in dark cells and melanocytes of the crista ampullaris (Fig. 4) in the mouse.

A novel variant in the Hispanic population, rs71717606 (*p-value* $= 4.84 \times 10^{-10}$) is an insertion/deletion within an intron of the gene *Discoidin*, *CUB*, and *LCCL* domain-containing protein (*DCBLD1*) on chromosome 6. *DCBLD1* is a transmembrane receptor involved in vertebrate development [67]. The top hit rs71717606 is located in a promoter site [64]. In contrast, fine mapping pointed to rs72965321, which is in an intron of an adjacent gene *ROS1*. eQTL analysis in FUMA reveals significant expression of *ROS1* in the frontal cortex, with significant HiC loops in the hippocampus (*p-value* $< 2.3 \times 10^{-11}$ after Bonferroni correction) [44, 45], indicative of enhancer-promoter and promoter-promoter interactions within *ROS1*. *ROS1* is expressed solely in the vestibular ganglion and has not been identified within the vestibular portion of the inner ear [48, 56].

Although our phenotype was crafted to eliminate causes of chronic dizziness outside the vestibular organ, nevertheless, gene enrichment identified the cerebellum as significantly involved. In addition, our GWAS-eQTL colocalization analysis using GTEx data did not identify any strong hits, and those data did not include expression data from the vestibule. However, cochleovestibular animal studies indicate that these genes are all expressed within the vestibule or vestibular ganglion. It is possible that further eQTL data from vestibular tissue or other relevant cell types could provide functional insight into the mechanisms underlying the genomic signal identified.

Note that this does not preclude expression quantitative trait loci in the vestibule as well but adds to the puzzle of chronic dizziness in the elderly. The general websites reporting functional annotations do not include information on the cochlea or vestibule. Thus, one limitation of this study is that lack of eQTLs documented within the cochlea or vestibule to compare to other body tissues. While there are excellent websites with vestibular and cochlear expression, data has not been integrated into other sites with brain or other body tissues information, making it difficult to compare these important tissues of interest in regard to expression data. On the other hand, examination of significant SNPs in linkage disequilibrium with lead SNPs points to multiple RegulomeDB scores > 5, indicating significant variants with a good likelihood of location in a functional region linked to expression, i.e., a transcription factor binding site, a DNAase footprint or peak, or transcription factor motif.

Besides demonstrating the intimate relationship of the vestibule with the cerebellum portion of the finely tuned balance system, this finding brings up other limitations of using ICDs for phenotyping. Future studies may engage natural language processing in the EHR to identify balance testing, i.e., Rombergs, VEMP testing, calorics, and others, to more objectively quantify balance in the elderly. Impairment of vision or proprioception is a sensory deficit that allow for mis-signaling emanating from the vestibular system. These systems were not addressed in this study.

Other limitations include the predominantly male aspect of the MVP cohort. Because of smaller numbers of females, leading to limited statistical power, we were unable to perform sex-stratified analysis, although sex and age were included as covariates. For the same reason, we did not include Asian ancestries. Other deficiencies include the MVP cohort that has unique battle-related environmental exposures that have poorly defined interactions with genetics and the aging process.

The use of electronic health records could lead to errors of omission for several reasons. First, some of the individuals in our study receive medical care outside the VA and second, patients with imbalance may perceive it to be normal for their age or be concerned with other health problems, consistent with a slightly lower prevalence of chronic dizziness identified in national surveys [2, 57].

Here, we have identified the first loci associated with chronic imbalance, including the first significant locus in a Hispanic population. Downstream analysis including correlation of eQTL through colocalization [38], CC-GWAS [37], fine mapping [42], and associated physiologic and anatomic information indicates variants and plausible genes of interest related to lead SNPs [68]. We have identified hits containing regulatory loci, i.e., *MLLT10* and *BPTF*, a gene identified in the vestibular ganglion (*ROS1*) and a non-coding RNA, *LINC01224*, that may control its downstream gene *ZNF19*, within the same locus. Ascertaining the physiologic/genetic architecture in the cochleovestibular system will aid in specific treatment of imbalance. Drug development can specifically target genes related to imbalance in the elderly. Those individuals with increased numbers of significant genetic variants can be identified to be at higher risk for imbalance at an earlier stage in order to provide more focused vestibular therapy. Future studies with larger and more gender-balanced cohorts will augment these results to clarify the role of the vestibular system in the genomics of chronic imbalance in the elderly.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Ekvall Hansson E, Magnusson M (2013) Vestibular asymmetry predicts falls among elderly patients with multi- sensory dizziness. BMC Geriatr 13. https://doi.org/10.1186/1471-2318-13-77

2. Harun A, Semenov YR, Agrawal Y (2015) Vestibular function and activities of daily living: analysis of the 1999 to 2004 National Health and Nutrition Examination Surveys. Gerontol Geriatr Med 1:2333721415607124. https://doi.org/10.1177/2333721415607124. Epub 2015 Sep 21. PMID: 26753170; PMCID: PMC4706363. https://doi.org/10.1177/2333721415607124

3. Kippenbrock T, Soja ME (2007) Preventing falls in the elderly: interviewing patients who have fallen. Researchers attempt to identify fall-risk factors from the patients' point of view. Geriatr Nurs 14:205–9

4. Fernández L, Breinbauer HA, Delano PH (2015) Vertigo and dizziness in the elderly. Front Neurol 6:144

5. Cutson TM (1994) Falls in the elderly. Am Fam Physician 49:149–156

6. Teixeira AR, Wender MH, Gonçalves AK, Freitas CDLR, dos Santos AMPV, Soldera CLC (2016) Dizziness, physical exercise, falls, and depression in adults and the elderly. Int Arch Otorhinolaryngol 20:124–131

7. Mitchell AJ, Vaze A, Rao S (2009) Clinical diagnosis of depression in primary care: a meta-analysis. Lancet 374:609–619

8. Furman JM, Redfern MS (2001) Effect of aging on the otolith-ocular reflex. J Vestib Res 11:91–103

9. Baloh RW, Jacobson KM, Socotch TM (1993) The effect of aging on visual-vestibuloocular responses. Exp Brain Res 95:509–516

10. Matheson AJ, Darlington CL, Smith PF (1999) Further evidence for age-related deficits in human postural function. J Vestib Res 9:261–264

11. Rauch SD, Velazquez-Villaseñor L, Dimitri PS, Merchant SN (2001) Decreasing hair cell counts in aging humans. Ann N Y Acad Sci 942:220–227

12. Jang YS, Hwang CH, Shin JY, Bae WY, Kim LS (2006) Age-related changes on the morphology of the otoconia. Laryngoscope 116:996–1001

13. Walther LE, Westhofen M (2007) Presbyvertigo-aging of otoconia and vestibular sensory cells. J Vestib Res 17:89–92

14. Maes L, Dhooge I, D'haenens W, Bockstael A, Keppler H, Philips B, Swinnen F, Vinck BM (2010) The effect of age on the sinusoidal harmonic acceleration test, pseudorandom rotation test, velocity step test, caloric test, and vestibular-evoked myogenic potential test. Ear Hear 31:84–94

15. Tanioka H, Tanioka S, Kaga K (2020) Vestibular aging process from 3D physiological imaging of the membranous labyrinth. Published Online First. https://doi.org/10.1038/s41598-020-66520-w

16. Wagner H, Melhus H, Pedersen NL, Michaëlsson K (2009) Heritability of impaired balance: a nationwide cohort study in twins. Osteoporosis Int 20:577–583

17. El Haber N, Hill KD, Cassano A-MT, Paton LM, Macinnis RJ, Cui JS, Hopper JL, Wark JD (2006) Genetic and environmental influences on variation in balance performance among female twin pairs aged 21–82 years. Am J Epidemiol 164:246–56

18. Skuladottir ATh, Bjornsdottir G, Nawaz MS, Petersen H, Rognvaldsson S, Moore KHS, Olafsson PI, Magnusson SH, Bjornsdottir A, Sveinsson OA, Sigurdardottir GR, Saevarsdottir S, Ivarsdottir EV, Stefansdottir L, Gunnarsson B, Muhlestein JB, Knowlton KU, Jones DA, Nadauld LD, Hartmann AM, Rujescu D, Strupp M, Walters GB, Thorgeirsson TE, Jonsdottir I, Holm H, Thorleifsson G, Gudbjartsson DF, Sulem P, Stefansson H, Stefansson K (2021) A genome-wide meta-analysis uncovers six sequence variants conferring risk of vertigo. Commun Biol 4:1148

19. Xiao AY, Tan ML, Plana MN, Yadav D, Zamora J, Petrov MS (2018) The use of international classification of diseases codes to identify patients with pancreatitis: a systematic review and meta-analysis of diagnostic accuracy studies. Clin Transl Gastroenterol 9. https://doi.org/10.1038/s41424-018-0060-1

20. Foley HE, Knight JC, Ploughman M, Asghari S, Audas R (2020) Identifying cases of chronic pain using health administrative data: a validation study. Can J Pain 4:252–267

21. Wassermann A, Finn S, Axer H (2022) Age-associated characteristics of patients with chronic dizziness and vertigo. J Geriatr Psychiatry Neurol 35:580–585

22. Müller KJ, Becker-Bense S, Strobl R, Grill E, Dieterich M (2022) Chronic vestibular syndromes in the elderly: presbyvestibulopathy—an isolated clinical entity? Eur J Neurol 29:1825–1835

23. Ji L, Zhai S (2018) Aging and the peripheral vestibular system. J Otol 13:138–140

24. Yetiser S (2020) Review of the pathology underlying benign paroxysmal positional vertigo. J Int Med Res 48:300060519892370

25. Fang H, Hui Q, Lynch J, Honerlaw J, Assimes TL, Huang J, Vujkovic M, Damrauer SM, Pyarajan S, Gaziano JM, DuVall SL, O'Donnell CJ, Cho K, Chang K-M, Wilson PWF, Tsao PS, Sun YV, Tang H, Gaziano JM, Ramoni R, Breeling J, Chang K-M, Huang G, Muralidhar S, O'Donnell CJ, Tsao PS, Muralidhar S, Moser J, Whitbourne SB, Brewer J V., Concato J, Warren S, Argyres DP, Stephens B, Brophy MT, Humphries DE, Do N, Shayan S, Nguyen X-MT, Pyarajan S, Cho K, Hauser E, Sun Y, Zhao H, Wilson P, McArdle R, Dellitalia L, Harley J, Whittle J, Beckham J, Wells J, Gutierrez S, Gibson G, Kaminsky L, Villareal G, Kinlay S, Xu J, Hamner M, Haddock KS, Bhushan S, Iruvanti P, Godschalk M, Ballas Z, Buford M, Mastorides S, Klein J, Ratcliffe N, Florez H, Swann A, Murdoch M, Sriram P, Yeh SS, Washburn R, Jhala D, Aguayo S, Cohen D, Sharma S, Callaghan J, Oursler KA, Whooley M, Ahuja S, Gutierrez A, Schifman R, Greco J, Rauchman M, Servatius R, Oehlert M, Wallbom A, Fernando R, Morgan T, Stapley T, Sherman S, Anderson G, Sonel E, Boyko E, Meyer L, Gupta S, Fayad J, Hung A, Lichy J, Hurley R, Robey B, Striker R (2019) Harmonizing genetic ancestry and self-identified race/ethnicity in genome-wide association studies. Am J Human Gen 105:763–72

26. Hunter-Zinck H, Shi Y, Li M, Gorman BR, Ji SG, Sun N, Webster T, Liem A, Hsieh P, Devineni P, Karnam P, Gong X, Radhakrishnan L, Schmidt J, Assimes TL, Huang J, Pan C, Humphries D, Brophy M, Moser J, Muralidhar S, Huang GD, Przygodzki R, Concato J, Gaziano JM, Gelernter J, O'Donnell CJ, Hauser ER, Zhao H, O'Leary TJ, Tsao PS, Pyarajan S (2020) Genotyping array design

and data quality control in the Million Veteran Program. Am J Hum Genet 106:535–548

27. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4:1–16

28. Zhou W, Nielsen JB, Fritsche LG, Dey R, Maiken E, Wolford BN, Lefaive J, Vandehaar P, Sarah A, Gifford A, Bastarache LA, Wei W, Denny JC, Hveem K, Kang HM, Abecasis GR, Willer CJ (2019) Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat Genetics 50:1335–1341

29. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. PLoS Genet 2:190

30. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, Benner C, O'Dushlaine C, Barber M, Boutkov B, Habegger L, Ferreira M, Baras A, Reid J, Abecasis G, Maxwell E, Marchini J (2021) Computationally efficient whole-genome regression for quantitative and binary traits. Nat Genet 53. https://doi.org/10.1038/s41588-021-00870-7

31. Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26:2190–2191

32. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47:291–295

33. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ (2010) LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26(18):2336–2337. https://doi.org/10.1093/bioinformatics/btq419

34. Finucane, H, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y et al (2015) Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 47:1228–35

35. Wells HRR, Freidin MB, Zainul Abidin FN et al (2019) GWAS identifies 44 independent associated genomic loci for self-reported adult hearing difficulty in UK Biobank. Am J Hum Genet 105(4):788–802. https://doi.org/10.1016/j.ajhg.2019.09.008

36. Clifford RE, Maihofer AX, Stein MB, Ryan AF, Nievergelt CM (2020) Novel risk loci in tinnitus and causal inference with neuropsychiatric disorders among adults of European ancestry. JAMA Otolaryngol Head Neck Surg 146:1015–1025

37. Peyrot WJ, Price AL (2021) Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. Nat Genet 53:445–454

38. Hukku A, Pividori M, Luca F, Pique-Regi R, Im HK, Wen X (2021) Probabilistic colocalization of genetic variants from complex and molecular traits: promise and limitations. Am J Hum Genet 108:25–35

39. Pividori M, Rajagopal PS, Barbeira A, Liang Y, Melia O, Bastarache L, Park Y, GTEx Consortium, Wen X, Im HK (2020) PhenomeXcan: mapping the genome to the phenome through the transcriptome. Sci Adv 6. https://doi.org/10.1126/sciadv.aba2083

40. GTEx Consortium (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science (1979) 369:1318–30

41. Berisa T, Pickrell JK (2016) Approximately independent linkage disequilibrium blocks in human populations. Bioinformatics 32(2):283–285. https://doi.org/10.1093/bioinformatics/btv546

42. Zou Y, Carbonetto P, Wang G, Stephens M (2022) Fine-mapping from summary data with the 'Sum of Single Effects' model. Published Online First. https://doi.org/10.1371/journal.pgen.1010299

43. Wang G, Sarkar A, Carbonetto P, Stephens M (2020) A simple new approach to variable selection in regression, with application to genetic fine mapping. J R Stat Soc Series B Stat Methodol 82(5):1273–1300. https://doi.org/10.1111/rssb.12388

44. Watanabe K, Taskesen E, van Bochoven A, Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8:1826

45. de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015) MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 11:e1004219

46. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M (2019) CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res 47. https://doi.org/10.1093/nar/gky1016

47. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (1979) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009(326):289–293

48. Orvis J, Gottfried B, Kancherla J, Adkins RS, Song Y, Dror AA, Olley D, Rose K, Chrysostomou E, Kelly MC, Milon B, Matern MS, Azaiez H, Herb B, Colantuoni C, Carter RL, Ament SA, Kelley MW, White O, Bravo HC, Mahurkar A, Hertzano R (2021) gEAR: Gene Expression Analysis Resource portal for community-driven, multi-omic data exploration. Nat Methods 18:843–844

49. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, van der Sluis S, Andreassen OA, Neale BM, Posthuma D (2019) A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet 51:1339–1348

50. Hromatka BS, Tung JY, Kiefer AK, Do CB, Hinds DA, Eriksson N (2015) Genetic variants associated with motion sickness point to roles for inner ear development, neurological processes and glucose homeostasis. Hum Mol Genet 24:2700–2708

51. Rujescu D, Hartmann AM, Giegling I, Konte B, Herrling M, Himmelein S, Strupp M (2018) Genome-wide association study in vestibular neuritis: involvement of the host factor for HSV-1 replication. Front Neurol 9:1–9

52. Strupp M, Maul S, Konte B, Hartmann AM, Giegling I, Wollenteit S, Feil K, Rujescu D (2020) A variation in FGF14 is associated with downbeat nystagmus in a genome-wide association study. Cerebellum 19:348–357

53. Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S (2015) Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep 10:1297–1309

54. Vidović T, Ewald CY (2022) Longevity-promoting pathways and transcription factors respond to and control extracellular matrix dynamics during aging and disease. Front Aging 3. https://doi.org/10.3389/fragi.2022.935220

55. Rudnicki A, Isakov O, Ushakov K, Shivatzki S, Weiss I, Friedman LM, Shomron N, Avraham KB (2014) Next-generation sequencing of small RNAs from inner ear sensory epithelium identifies micro-RNAs and defines regulatory pathways. BMC Genomics 15:1–12

56. Shen J, Scheffer DI, Kwan KY, Corey DP (2015) SHIELD: an integrative gene expression database for inner ear research. Database 2015:1–9

57. Cai N, Revez JA, Adams MJ et al (2020) Minimal phenotyping yields genome-wide association signals of low specificity for major depression. Nat Genet 52(4):437–447. https://doi.org/10.1038/s41588-020-0594-5

58. Agrawal Y, Carey JP, della Santina CC, Schubert MC, Minor LB. (2009) Disorders of balance and vestibular function in US adults: data from the National Health and Nutrition Examination Survey, 2001–2004. Arch Intern Med 169:938–944

59. Agrawal Y, van de Berg R, Wuyts F, Walther L, Magnusson M, Oh E, Sharpe M, Strupp M (2019) Presbyvestibulopathy: diagnostic criteria consensus document of the classification committee of the Bárány Society. J Vestib Res 29:161–170

60. Kalmanson O, Foster CA (2023) Cupulolithiasis: a critical reappraisal. OTO Open 7. https://doi.org/10.1002/oto2.38

61. Sessoms PH, Gottshall KR, Sturdy J, Viirre E (2015) Head stabilization measurements as a potential evaluation tool for comparison of persons with TBI and vestibular dysfunction with healthy controls. Mil Med 180:135–142

62. Fu W, Zhao J, Hu W, Dai L, Jiang Z, Zhong S, Deng B, Huang Y, Wu W, Yin J (2021) LINC01224/ZNF91 promote stem cell-like properties and drive radioresistance in non-small cell lung cancer. Cancer Manag Res 13:5671–5681

63. Edwards SL, Beesley J, French JD, Dunning AM (2013) Beyond GWASs: illuminating the dark road from association to function. Am J Hum Genet 93(5):779–97. https://doi.org/10.1016/j.ajhg.2013.10.012. PMID: 24210251; PMCID: PMC3824120

64. He D, Wu D, Muller S, Wang L, Saha P, Ahanger SH, Liu SJ, Cui M, Hong SJ, Jain M, Olson HE, Akeson M, Costello JF, Diaz A, Lim DA (2021) miRNA-independent function of long noncoding pri-miRNA loci. Proc Natl Acad Sci 118. https://doi.org/10.1073/pnas.2017562118

65. Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, Parton A, Armean IM, Trevanion SJ, Flicek P, Cunningham F (2018) Ensembl variation resources. Database (Oxford) 2018:bay119. https://doi.org/10.1093/database/bay119

66. Walsh KM, Zhang C, Calvocoressi L, Hansen HM, Berchuck A, Schildkraut JM, Bondy ML, Wrensch M, Wiemels JL, Claus EB (2022) Pleiotropic MLLT10 variation confers risk of meningioma and estrogen-mediated cancers. Neurooncol Adv 4. https://doi.org/10.1093/noajnl/vdac044

67. Stankiewicz P, Khan TN, Szafranski P, Slattery L, Streff H, Vetrini F, Bernstein JA, Brown CW, Rosenfeld JA, Rednam S, Scollon S, Bergstrom KL, Parsons DW, Plon SE, Vieira MW, Quaio CRDC, Baratela WAR, Acosta Guio JC, Armstrong R, Mehta SG, Rump P, Pfundt R, Lewandowski R, Fernandes EM, Shinde DN, Tang S, Hoyer J, Zweier C, Reis A, Bacino CA, Xiao R, Breman AM, Smith JL, Katsanis N, Bostwick B, Popp B, Davis EE, Yang Y (2017) Haploinsufficiency of the chromatin remodeler BPTF causes syndromic developmental and speech delay, postnatal microcephaly, and dysmorphic features. Am J Hum Genet 101:503–515

68. Schmoker AM, Weinert JL, Markwood JM, Albretsen KS, Lunde ML, Weir ME, Ebert AM, Hinkle KL, Ballif BA (2020) FYN and ABL regulate the interaction networks of the DCBLD receptor family. Mol Cell Proteomics 19:1586–1601