Integrated analysis of proteome-wide and transcriptome-wide association studies identified novel genes and chemicals for vertigo

Bolun Cheng,1,2,3 Peilin Meng,1,2,3 Xuena Yang,1,2,3 Shiqiang Cheng,1,2,3 Li Liu,1,2,3 Yumeng Jia,1,2,3 Yan Wen1,2,3 and Feng Zhang1,2,3

Vertigo is a leading symptom of various peripheral and central vestibular disorders. Although genome-wide association studies (GWASs) have identified multiple risk variants for vertigo, how these risk variants contribute to the risk of vertigo remains unknown. Discovery proteome-wide association study (PWAS) was first performed by integrating the protein quantitative trait loci from the dorsolateral prefrontal cortex (DLPFC) in the Banner Sun Health Research Institute dataset (n = 152) and GWAS summary of vertigo (n = 942,613), followed by replication PWAS using the protein quantitative trait loci from the DLPFC in Religious Orders Study or the Rush Memory and Aging Project dataset (n = 376). Transcriptome-wide association studies (TWASs) were then performed by integrating the same GWAS datasets of vertigo (n = 942,613) with mRNA expression reference from human fetal brain, and DLPFC. Chemical-related gene set enrichment analysis (GSEA) and Gene ontology/Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses were finally conducted to further reveal the pathogenesis of vertigo. Permutation-based empirical P values were calculated in PWAS, TWAS, and GSEA. By integrating the GWAS of vertigo and two independent brain proteomes from human DLPFC, three genes were identified to genetically regulate protein abundance levels in vertigo, and were not previously implicated by GWAS, including MTERFD2 (P_Banner = 0.045, P_ROSMAP = 0.031), MGST1 (P_Banner = 0.014, P_ROSMAP = 0.018), and RAB3B (P_Banner = 0.045, P_ROSMAP = 0.035). Compared with TWAS results, we identified overlapping genes RAB3B (P_TWAS = 0.017) and MTERFD2 (P_TWAS = 0.003) that showed significant associations with vertigo at both proteome-wide and transcriptome-wide levels. Chemical-related GSEA identified multiple chemicals that might be associated with vertigo, such as nickel (P = 0.007), glycidamide (P = 0.005), and proanthocyanidins (P = 0.015). Our study provides novel clues for understanding the biological mechanism of vertigo, and highlights several possible risks and therapeutic chemicals for vertigo.

1  Key Laboratory of Trace Elements and Endemic Diseases, Collaborative Innovation Center of Endemic Disease and Health Promotion for Silk Road Region, School of Public Health, Health Science Center, Xi’an Jiaotong University, No. 76 Yan Ta West Road, Xi’an 710061, China
2  Key Laboratory for Disease Prevention and Control and Health Promotion of Shaanxi Province, Health Science Center, Xi’an Jiaotong University, Xi’an 710061, China
3  Key Laboratory of Environment and Genes Related to Diseases of Ministry of Education of China, Health Science Center, Xi’an Jiaotong University, Xi’an 710061, China

Correspondence to: Feng Zhang
School of Public Health, Health Science Center
Xi’an Jiaotong University
No. 76 Yan Ta West Road
Xi’an 710061, China
E-mail: fzhxjtu@mail.xjtu.edu.cn
Keywords: vertigo; proteome-wide association study; transcriptome-wide association study; chemical; gene set enrichment analysis

Abbreviations: 3-NP = 3-nitropropionic acid; Banner dataset = Banner Sun Health Research Institute dataset; BSLMM = Bayesian sparse linear mixed model; CI = cerebral ischaemia; CMC = CommonMind Consortium; CNS = central nervous system; CTD = Comparative Toxicogenomics Database; DLPFC = dorsolateral prefrontal cortex; eQTL = expression quantitative trait loci; ESs = enrichment scores; GO = gene ontology; GSEA = gene set enrichment analysis; GWAS = genome-wide association study; LD = linkage disequilibrium; LSPC = lotus seedpod proanthocyanidins; NMDA = N-methyl-D-aspartate; PAC = proanthocyanidins; pQTL = protein quantitative trait loci; PWAS = proteome-wide association study; QC = quality control; ROSMAP = Religious Orders Study or the Rush Memory and Aging Project; s = standard deviation; sQTL = splicing quantitative trait loci; TWAS = transcriptome-wide association study; WGS = whole-genome sequencing

Graphical Abstract

Introduction

Vertigo is a subtype of dizziness defined as the illusion of motion caused by the asymmetrical intervention of the vestibular system.1 Vestibular central lesions affecting the pons, medulla, or cerebellum may cause vertigo, severe ataxia, vomiting, nausea, and other neurological signs.1 Conventionally, the diseases causing vertigo can be classified into three broad categories: otological vertigo, central vertigo, and psychogenic dizziness.2,3 Common causes of vertigo include benign paroxysmal postural vertigo, migraine, vestibular neuritis, Ménière’s syndrome, adverse drug effects, and disturbed blood pressure regulation.4 The prevalence of vertigo is 6.5% and increases with age, with approximately 65% of patients being female5 and a lifetime prevalence of about 20–30%.6 Vertigo can come on suddenly and last for a few seconds or maybe constant for several days, which is a major risk factor for falls and bone fractures, placing a huge burden on the healthcare system.7 Several common vertigo syndromes are known to be genetically heterogeneous and have been identified in patients with isolated recurrent attacks of vertigo, genetic deafness...
syndromes, and neurological disorders. Genetic studies have identified several mutations in the KCNA1 and CACNA1A genes associated with recurrent vertigo, suggesting that voltage-gated channels and solute carriers in the plasma membrane of neurons play a key role in recurrent vertigo. A genome-wide association study (GWAS) also uncovered six sequence variations associated with vertigo risk. In addition, the central nervous system (CNS) plays an important role in the pathogenesis of vertigo. Interference with protein–protein interaction networks in neurons, glia, and other cell types has also been found to be associated with multifactorial neurological disorders. Although previous research works have focused on transcription as a central regulator of protein expression, it is now increasingly recognized that the CNS relies on efficient updating of the protein landscape. However, the characteristics of brain proteins and genetics of vertigo remain unclear.

GWAS seeks to strongly link genetic loci to diseases and other heritable traits, but the statistical power is limited. Linkage disequilibrium (LD) and population stratification make it convoluted for GWAS to identify exact causal variants. The transcriptome-wide association study (TWAS) tests whether the phenotype being studied correlates with the level of gene expression predicted from genetic variants, usually requiring a target tissue for association tests. As a new method to detect gene–phenotype associations mediated by changes in protein abundance, PWAS aggregates all variants’ signals that jointly affect protein-coding genes, reducing the burden of multiple testing and providing more interpretable discoveries. For example, Liu et al. performed PWAS to identify genes associated with changes in cis-regulatory protein abundance in the human brain and found several high-confidence risk proteins (including CNNM2 and CTNND1) for schizophrenia and depression. Another study performed PWAS by integrating GWAS and human brain proteomics, followed by Mendelian randomization, and identified 19 genes consistent with depression causality.

Human brain development is disrupted by chemical exposure, leading to irreversible damage to the nervous system. Conversely, some chemicals have positive effects on the CNS. Recent studies have found that chemical exposure also affects epigenetic inheritance. For example, prenatal exposure to endocrine-disrupting chemicals might lead to poor behaviour and cognitive dysfunction in children. The contribution of chemicals to brain development is multifaceted and complex, and chemicals associated with the endocrine activity are suspected to interfere with neurodevelopment. In addition, neural tube and axial defects were frequently found when evaluating the developmental effects of chemicals in laboratory animal studies. Therefore, it is necessary to identify the chemicals related to the pathogenesis of vertigo.

In this study, PWASs were performed on vertigo by integrating a large-scale GWAS and two independent human brain protein quantitative trait loci (pQTL) datasets. Discovery PWAS was first conducted using the pQTL from the dorsolateral prefrontal cortex (DLPFC) in the Banner dataset. Replication PWAS was then performed using the pQTL from DLPFC in the ROSMAP dataset. To validate the results of PWASs, we also performed TWASs by integrating the same GWAS datasets and mRNA expression reference from the human fetal brain, and DLPFC in which gene expression was quantified with RNA-seq and RNA-seq splicing, respectively. Chemical-related gene set enrichment analysis (GSEA) and Gene ontology (GO)/Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were finally conducted to further reveal the pathogenesis of vertigo. Our study identified risk genes whose protein abundance is associated with vertigo, providing novel insights into further mechanistic studies and the development of new therapies.

Materials and methods

GWAS summary data of vertigo

Genome-wide association summary level data used in this study was derived from a recent publicly available large-scale GWAS of human vertigo. Briefly, this dataset had 48,072 vertigo cases and 894,541 controls in the vertigo meta-analysis, all of whom were verified as being of white origin. The deCODE was used to perform the sample preparation and the whole-genome sequencing (WGS) in Iceland samples. The Affymetrix chip UK BiLEVE Axiom and the Affymetrix UK Biobank Axiom array were used to genotype the UK Biobank samples. NovaSeq Illumina was used for WGS in US samples, and Illumina Global Screening Array chips were used for genotyping. FinnGen ThermoFisher Axiom array was used to genotype FinnGen samples. The logistic regression assuming an additive model was applied and tested for association between sequence variants and vertigo using deCODE software. The fixed-effects inverse variance method based on effect estimates and standard errors was used to combine vertigo GWAS summary results from Iceland, the UK, Finland, and the US. This method assumes that all study groups have a common OR, but allows for different population frequencies for alleles and genotypes. In total, 62,056,310 variants in the meta-analysis had imputation information above 0.8 and MAF > 0.01%. The genotyping, quality control, imputation, and association analysis of the dataset have been described in detail elsewhere.

Human brain pQTL datasets for protein reference weights

Human brain proteomes datasets were obtained from a recent publicly available study. Briefly, a proteome analysis was performed using brain DLPFC of 376 European descent subjects collected by the Religious Orders Study (ROS) or the Rush Memory and Aging Project (ROSMAP dataset). After quality control (QC), 8356 proteins were
included in proteomic profiles for pQTL analysis. By integrating proteomic data and SNP genotypes, 1475 proteins were significantly heritable to genetic variation, and their protein weights were used in our PWAS. Besides, another proteome analysis was performed using brain DLPCF of 152 European descent subjects from the Banner Sun Health Research Institute (Banner dataset). After QC, 8168 proteins were included in proteomic profiles for pQTL analysis. By integrating proteomic data and SNP genotypes, 1139 proteins were significantly heritable to genetic variation, and their protein weights were used in our PWAS. The brain protein weights from the ROSMAP and Banner brain proteome can be downloaded from https://doi.org/10.7303/syn23627957. Detailed information for the brain protein weights has been described in the study of Wingo et al.

**PWAS for vertigo**

The PWAS analysis of vertigo was performed by using FUSION software (http://gusevlab.org/projects/fusion/). Using the pre-computed reference weights of the ROSMAP and Banner brain proteome together with GWAS summary data of vertigo, FUSION is capable to estimate the associations of each protein abundance gene with vertigo. Firstly, the reference expression weights were calculated using the prediction models in FUSION. The calculated expression weights were then combined with GWAS results to impute association statistics between protein abundance gene expression levels and vertigo. Bayesian sparse linear mixed model (BSLMM) was used to compute the SNP-expression weights in the 1-Mb cis loci of the gene for a given gene. The association test statistics between predicted gene expression and vertigo were calculated as \( Z_{PWAS} = W^T Z (W^T W)^{-1/2} \). \( Z \) denotes the scores of vertigo, while \( W \) denotes the weights. \( S \) denotes the SNP-correlation covariance matrix. In this study, a total of 5000 permutation tests in FUSION were implemented to control the potential impact of multiple test problems. The permutation-based empirical \( P \) value was calculated for each gene within the ROSMAP and Banner brain proteome. Significant protein abundance genes were considered at permuted \( P < 0.05 \).

**TWAS for vertigo**

The TWAS analyses were performed by using FUSION software (http://gusevlab.org/projects/fusion/). Using the pre-computed gene expression weights of different tissues together with GWAS summary data of vertigo, FUSION is capable to estimate the associations of each gene with vertigo in different tissues. In this study, TWAS analysis was performed based on the GWAS summary statistics of vertigo and pre-computed expression reference panel in TWAS included the splicing-level (splicing QTL, sQTL) and mRNA-level (eQTL) weights from CommonMind Consortium (CMC) in DLPCF profiles, and fetal brain expression (eQTL) weights. The only difference between PWASs and TWASs is the different external presentation reference panels. PWAS is based on protein abundance, while TWAS is based on mRNA expression level. Briefly, the gene expression weights were calculated using the prediction models of FUSION, respectively. The calculated expression weights were then combined with GWAS summary statistics of vertigo. BSLMM was used to compute the SNP-expression weights in the 1-Mb cis loci of the gene for a given gene. Multiple testing problem of each gene was computed using 5000 permutation tests. Genes with significant correlation signals were identified at permuted \( P < 0.05 \).

**Colocalization analysis**

Colocalization analysis was performed using Fusion software with parameter—coloc \( P \) 0.05, which indicated that only genes with \( P < 0.05 \) were included to perform colocalization analysis. Colocalization of GWAS and pQTL (ROSMAP and Banner datasets), eQTL (CMC and fetal brain dataset), and sQTL (CMC dataset) signals were performed to explore the same risk variants in vertigo.

**Statistical analysis**

The chemical-related gene set was obtained from the public Comparative Toxicogenomics Database (CTD) which consisted of 1 379 105 chemical–gene interactions (http://ctdbase.org/). CTD is an innovative digital ecosystem that relates toxicological information for chemicals, genes, phenotypes, diseases, and exposures to advance the understanding of human health. The GSEA (4.2.3) was implemented to evaluate the functional relevance between 103 chemicals and vertigo. Briefly, GSEA was first applied to the PWAS-level data of vertigo to determine whether vertigo-associated protein abundance genes were significantly enriched in the chemical-related gene sets. Following the standard GSEA approach, we set \( GS_m \) as the PWAS statistic. All genes were then ordered by ranking \( GS_m \) from the largest to the smallest, set as \( GS_m = (GS_{m1}, GS_{m2}, \ldots, GS_{mN}) \). We set \( G_i \) as the \( i \)th gene in chemical-related gene set \( C \) with \( N_C \) genes. A weighted Kolmogorov-Smirnov-like running sum statistic was used to calculate the enrichment scores (ES) of each analyzed chemical.

\[
ES_s = \max_{1 \leq s \leq N} \left\{ \sum_{G_i \in C, s \geq j} |GS_{mi}|^H \frac{N_R}{N} - \frac{1}{N - N_C} \right\}
\]

where \( N_R \) defined as \( N_R = \sum_{G_i \in C} |GS_{mi}|^H \), \( N \) as the total number of genes; \( H \) as the parameter that gives higher weights to genes with extreme statistics. We followed the recommendation that using \( H = 1 \) for the original GSEA algorithm. The TWAS-level data were then utilized to perform the same GSEA algorithm.
Following the normalization procedure, statistical test was performed to calculate the null distribution of ES, with SNP labels of vertigo GWAS randomly shuffled and then used for ES calculation. The null distribution of ES (ES(0), ES(1), ES(2), ..., ES(n)) was obtained after t time permutations. To adjust the bias caused by different sizes of chemical-related gene sets, the observed ES was normalized by the standard deviation (SD) and mean of permuted ES:

$$NES = \frac{ES - mean(ES_{50})}{SD(ES_{50})}$$

Each enrichment analysis implemented 20000 permutations. The permutation-based empirical P values were finally calculated using the NES for each chemical-related gene set. Significant enrichment was identified at P < 0.05. Detailed information of permutation and statistical analysis were described elsewhere. 36

**Gene ontology and pathway enrichment analysis**

To determine biological features and enriched pathways for vertigo, the gene ontology (GO) and pathway enrichment analysis were performed on DAVID (Dec.2021), STRING (11.5), and Reactome (Version 81) using the significant genes identified by PWAS and TWAS (P_TWAS < 0.05 or P_TWAS < 0.05), including pQTL PWAS using Banner datasets, eQTL TWAS using CMC and fetal brain dataset, and sQTL TWAS using CMC dataset. All enrichment analyses were false discovery rate (FDR) controlled using the Benjamini-Hochberg method.

**Data availability**

Detailed description of the datasets used in our study are shown in Supplementary Material. The datasets generated and/or analyzed in the current study are available from the corresponding authors upon reasonable request.

**Results**

**PWASs identified three genes that regulate protein levels in the brain and were linked to risk of vertigo**

Discovery PWAS was first performed by integrating the Banner human brain pQTL data and GWAS of vertigo and identified 22 significant genes for vertigo (Table 1). To validate these results, replication PWAS was then conducted using the ROSMAP human brain pQTL dataset and the same GWAS of vertigo and identified 13 significant genes for vertigo (Table 1). Importantly, three common vertigo risk genes showed proteome-wide significant signals in both discovery and replication stages, including MTERFD2 (P_Banner = 0.045, P_ROSMAP = 0.031), MGST1 (P_Banner = 0.014, P_ROSMAP = 0.018), and RAB3B (P_Banner = 0.045, P_ROSMAP = 0.035) (Table 1). The original PWAS results based on the Banner and ROSMAP human brain proteomes are presented in Fig. 1.

**TWASs identified eight genes that regulate gene levels in the brain and were linked to risk of vertigo**

The eQTL-based TWAS identified 99 cis-regulated expression genes associated with vertigo, such as RAB3B (P_TWAS = 0.017), TMEM132E (P_TWAS = 0.002), and TCEA3 (P_TWAS = 0.037) (Supplementary Table 1). The sQTL-based TWAS identified 148 genes associated with vertigo, such as MTERFD2 (P_TWAS = 0.003), C12orf73 (P_TWAS = 0.027), and SNRPC (P_TWAS = 0.003) (Supplementary Table 2). Notably, eight candidate common genes were shared in eQTL-based and sQTL-based TWAS, such as PIGG (P_eQTL = 0.011, P_sQTL = 0.099), ZSWIM7 (P_eQTL = 0.023, P_sQTL = 0.017), and PMS2P5 (P_eQTL = 0.033, P_sQTL = 0.031) (Supplementary Tables 1 and 2).

### Table 1: Proteome-wide significant genes for vertigo

| Gene   | CHR | Z_TWAS | P_Banner | Z_TWAS | P_Banner | Z_TWAS | P_Banner |
|--------|-----|--------|----------|--------|----------|--------|----------|
| ASPHD1 | 16  | 4.967  | 2.00 × 10^{-4} | -2.321 | 4.00 × 10^{-4} |
| PPP2R4 | 9   | -2.427 | 0.005    |         |          |         |          |
| PTGR1  | 9   | -2.285 | 0.007    |         |          |         |          |
| TMEM132E| 17  | -3.173 | 0.007    |         |          |         |          |
| SHISA7 | 19  | -2.441 | 0.010    |         |          |         |          |
| SLCT7A1| 19  | 2.491  | 0.010    |         |          |         |          |
| RIM53  | 1   | -1.97  | 0.011    |         |          |         |          |
| MSTO1  | 1   | 2.393  | 0.019    |         |          |         |          |
| NOSTRIN| 2   | 2.838  | 0.013    |         |          |         |          |
| CCDC91 | 12  | 7.986  | 0.013    |         |          |         |          |
| CDH7   | 18  | 2.279  | 0.014    |         |          |         |          |
| MGST1a | 12  | 2.526  | 0.018    |         |          |         |          |
| S100A1 | 1   | 2.779  | 0.019    |         |          |         |          |
| C12orf73| 12  | -2.90  | 0.019    |         |          |         |          |
| DHR54  | 14  | 2.638  | 0.020    |         |          |         |          |
| SPROD4 | 12  | 1.971  | 0.023    |         |          |         |          |
| AGA    | 4   | -2.357 | 0.025    |         |          |         |          |
| C2orf76| 2   | 2.181  | 0.026    |         |          |         |          |
| MTERFD2a| 2   | -2.999 | 0.045    | -2.999 | 0.031   |         |          |
| PAK4   | 19  | 2.147  | 0.032    |         |          |         |          |
| GPX1   | 3   | -2.267 | 0.035    |         |          |         |          |
| Rab3Bb | 1   | 2.184  | 0.045    |         |          |         |          |
| COQ10A | 12  | 4.018  | 0.039    |         |          |         |          |
| EXOG   | 3   | 2.067  | 0.0393   |         |          |         |          |
| IPEC1  | 6   | -2.335 | 0.0396   |         |          |         |          |
| PRMA3  | 19  | -2.583 | 0.0396   |         |          |         |          |
| RT2    | 10  | -2.044 | 0.042    |         |          |         |          |
| AD5L   | 22  | 2.165  | 0.044    |         |          |         |          |

CHR = chromosome; PERM = permutation; PWAS = proteome-wide association study; ROSMAP = Religious Orders Study and Rush Memory and Aging Project.  
Genes that reached proteome-wide significant level in both the Banner and ROSMAP brain proteomes.

[57x314]Table 1 Proteome-wide significant genes for vertigo

**Results**

**PWASs identified three genes that regulate protein levels in the brain and were linked to risk of vertigo**

Discovery PWAS was first performed by integrating the Banner human brain pQTL data and GWAS of vertigo and identified 22 significant genes for vertigo (Table 1). To validate these results, replication PWAS was then conducted using the ROSMAP human brain pQTL dataset and the same GWAS of vertigo and identified 13 significant genes for vertigo (Table 1). Importantly, three common vertigo risk genes showed proteome-wide significant signals in both discovery and replication stages, including MTERFD2 (P_Banner = 0.045, P_ROSMAP = 0.031), MGST1 (P_Banner = 0.014, P_ROSMAP = 0.018), and RAB3B (P_Banner = 0.045, P_ROSMAP = 0.035) (Table 1). The original PWAS results based on the Banner and ROSMAP human brain proteomes are presented in Fig. 1.
Fetal brain-based TWAS identified 20 cis-regulated expression genes associated with vertigo, such as MXRA7 ($P_{\text{TWAS}} = 0.006$), MSH3 ($P_{\text{TWAS}} = 0.018$), and DHFR ($P_{\text{TWAS}} = 0.038$) (Supplementary Table 3). Interestingly, no common gene in fetal brain-based TWAS was shared with eQTL and sQTL-based TWAS, which might highlight the differences for vertigo mechanism in RNA level between adult and fetus.

The comparison of PWAS and TWAS highlighted two risk genes for vertigo

PWAS and TWAS highlighted several promising candidate genes for vertigo, and a series of comparisons thus to be performed between PWAS and TWAS results to explore the overlapped risk genes at the protein and RNA level, including eQTL-based TWAS, sQTL-based TWAS, and fetal brain-based TWAS. Proteome-wide candidate risk gene RAB3B was significantly overlapped in eQTL-based TWAS ($P_{\text{TWAS}} = 0.017$). MTERFD2 was statistically significant between PWAS and sQTL-based TWAS result ($P_{\text{TWAS}} = 0.003$). However, no proteome-wide significant genes were supported by fetal brain-based TWAS results.

Colocalization analysis of common vertigo risk genes between PWAS and TWAS

Colocalization analysis based on the PP4 hypothesis which represents the posterior probability for coloc colocalization analysis hypothesis 4, that is, eQTL and GWAS signals are driven by the same common causal variant. The results indicated that the GWAS and pQTL signals were driven by the same risk variants, including RAB3B ($PP4_{\text{banner}} = 0.007$, $PP4_{\text{ROSMAP}} = 0.004$).
In particular, for vertigo based on PWAS results

**Table 2 Chemical-related gene set enrichment analysis**

The pathogenesis of vertigo. A total of 15 GO annotation and pathway analysis were accomplished to GO and pathway enrichment analysis. Seven candidate chemicals were observed, such as nickel (P = 0.007), proanthocyanidins (PAC; P = 0.015), and irinotecan (P = 0.025) (Table 2). For the TWAS gene list, 22, 32, and 31 candidate chemicals were significantly enriched in fetal brain, RNA-seq and splicing data, respectively. Notably, three common chemicals overlapped in RNA-seq and splicing TWAS results, including fipronil, phenylpropanolamine, and fenbucloazole (Supplementary Table 4).

**GO and pathway enrichment analysis**

GO annotation and pathway analysis were accomplished to further reveal the pathogenesis of vertigo. A total of 15 significant terms were identified by using significant genes from PWAS and TWAS. For example, CNS (P_{FDR} = 2.84 × 10^{-7}) and brain stem (P_{FDR} = 0.039) in CMC splicing dataset-based TWAS, mitochondrion (P_{FDR} = 3.9 × 10^{-4}) in Banner dataset-based PWAS, cognitive function measurement (P_{FDR} = 0.023) and serine metabolism (P_{FDR} = 0.012) in fetal brain dataset-based TWAS (Supplementary Table 5).

**Discussion**

In this study, we performed PWASs for vertigo by integrating human brain pQTL data and genome-wide associations. We identified three genes that regulate protein levels in the human brain that were associated with risk of vertigo by combining the results from the discovery and replication PWASs. Among the three candidate genes, RAB3B and MTERFD2 were also validated in TWAS, strongly emphasizing the consistent findings at mRNA and protein levels. Chemical-related GSEA and GO/pathway enrichment analysis highlights that the pathogenesis of vertigo may be related to neuronal exocytosis and CNS development.

Our study suggesting that RAB3B and MTERFD2 are highly risk genes whose expression and protein abundance are significantly associated with vertigo. RAB3B encodes a low molecular weight GTP-binding protein, which is involved in the exocytosis of synaptic vesicles and secretory granules in the CNS and the anterior pituitary cells. In addition, RAB3B may be involved in polarized transport of basolateral and tight junctional membrane proteins to the plasma membrane and the regulation of synaptic plasticity. RAB3B has been confirmed to be involved in neurotransmitters and synapses, pituitary exocytosis, and immune function. In mouse models, Rab3b is required for long-term depression of inhibitory synapses in the hippocampus, short-term plasticity, and normal reverse learning. As a key intracellular signalling molecule, RAB3A is highly expressed in the brain and can control the downstream exocrine secretion of other calcium-dependent processes in anterior pituitary cells. RAB3B immunoreactivity plays an important role in human pituitary adenoma. Possibly reflecting the number of secretory granules and exocytosis activity. In addition, GO and KEGG pathway analysis also found that vertigo was involved in the biological function of neurons. For example, regulation of short-term neuronal synaptic plasticity in Banner PWAS, cell proliferation in the midbrain, commissural neuron axon guidance in RNA-seq TWAS, neurotransmitter secretion, and GABAergic synapse in splicing TWAS.

MTERFD2, also known as MTERF4, is another vertigo risk gene identified in this study that belongs to the component of the mitochondrial transcription termination factor (MTERF) family. Colocalization of EGFP-MTERFD2 fusion protein suggested that MTERFD2 targeting mitochondria exhibit a dynamic expression pattern during embryogenesis and might play an important role in organ differentiation. MTERFD2 has been shown to recruit...
It was found that MTERFD2 also plays a vital role in mitochondrial regulation and neurodegenerative disease. Ye et al. found that the overexpression of MTERFD2 in SH-SY5Y cells partially alleviated 1-methyl-4-phenylpyridinium induced mitochondrial dysfunction, and proposed that MTERFD2 might be the trigger factor of the pathogenesis of Parkinson’s disease induced by environmental toxins. Wang et al. emphasized that MTERFD2 plays a key role in the pathogenesis of Alzheimer’s disease by inhibiting ADAM10 in HEK293-APPswe cells and promoting amyloidogenic processing of β-amyloid precursor protein. Besides, GO annotation also found that vertigo is involved in the biological function of mitochondrial regulation, such as mitochondrial in Banner PWAS. Among the three genes overlapped in Banner and ROSMAP PWAS, RAB3B and MTERFD2 were also verified by TWAS to be associated with vertigo at mRNA-level, emphasizing that genetic risk variants likely confer risk of vertigo by regulating messenger RNA expression and protein abundance of these genes.

In particular, MGST1 and SPRYD4 were only recognized by the PWAS rather than TWAS, highlighting the additional insights provided by focusing brain proteins directly. MGST1 protein is ubiquitous in human tissues and cell lines and is located in the endoplasmic reticulum and outer mitochondrial membrane, playing a protective role against membrane oxidative stress. It was found that Mgst1 deletion in mice resulted in embryonic lethal and impaired haematopoietic function, emphasizing that Mgst1 is essential for embryonic development and haematopoietic function in vertebrates. SPRYD4 has not been widely explored, but it has been recently found to have anti-cancer effects, such as acute myeloid leukemia and hepatocellular carcinoma. In our GO annotation, glutathione peroxidase activity was also found to involve in vertigo by Banner PWAS. Notably, all three proteome-wide significant genes were not specified by the original vertigo GWAS, highlighting the ability of PWAS in exploring vertigo risk genes.

Chemical-related GSEA identified multiple neurotoxic chemicals that may be associated with vertigo risk, such as nickel, glycidamide, fipronil. Nickel is a potential risk chemical for vertigo identified by our ROSMAP PWAS and has also been shown to cause vertigo, with toxicity manifesting in affecting T cell systems and inhibiting the activity of natural killer cells. Thus, environmental and occupational exposure to nickel is a potential risk factor for human brain dysfunction and neurological symptoms. In addition, GSEA based on TWAS results of RNA-seq and splicing jointly identified a potential risk chemical, fipronil, which is a member of the phenylpyrazole class of insecticide and has broad spectrum activity. Fipronil interferes with the GABAergic system, affecting brain growth and neural development. Lassiter et al. found that fipronil caused oxidative stress in undifferentiated neuronal PC12 cells and subsequently inhibited DNA and protein synthesis. Ki et al. subsequently indicated that fipronil-induced neuronal apoptosis was mediated by reactive oxygen species production.

Conversely, our chemical-related GSEA have also identified some chemicals that might play a protective role in vertigo, such as PAC and piroxicam. PAC was a potential therapeutic chemical for vertigo identified by our ROSMAP PWAS and was also widely distributed in common foods including cereals, fruits, vegetables, and wines, and belonged to a universal group of plant polyphenols with their powerful antioxidant capacity and possible protective effects on human health. Chen et al. suggested that ethanol causes cognitive impairment and that the protective effect of PAC on ethanol-induced cognitive impairment may be due to its antioxidant and anti-inflammatory activities. Gong et al. found that lotus seedpod proanthocyanidins (LSPC) could prevent hippocampal neuron damage and reduce the expression of P53 protein in the hippocampus, suggesting that LSPC might be used to treat Alzheimer’s disease. Piroxicam was detected to be significantly enriched in fetal brain-based TWAS for vertigo, which could be converted into novel CNS stimulants or inhibitors through desulphurization, methylation, dehydrogenation, carboxylation, and carbonylation. Animal toxicity studies showed that Piroxicam significantly reversed oxidative stress induced by the mycotoxin 3-nitropropionic acid (3-NP) in mice, suggesting that Piroxicam protected mice from 3-NP-induced oxidative stress and behavioural changes in the brain. Mazumder et al. recommended the use of Piroxicam as the gold standard for the prevention of cerebral ischaemia (CI) neurodegeneration, since Piroxicam inhibits N-methyl-D-aspartate (NMDA) receptor, which are also associated with CI-induced neurodegeneration.

Our study has several strengths. First, this study examined the mRNA and protein levels in vertigo through PWAS and TWAS. Second, GSEA provided potential risk chemicals and possible treatments for vertigo. There are several limitations of our study. First, the pre-computed reference weights of the ROSMAP and Banner brain proteome included both older and neurological disorder individuals, which may result in slight bias in our results. Besides, although tissue differences were controlled in the TWAS analysis, the weak anatomical correlation between vertigo and DLPFC could still lead to potential bias. Third, our method is a theoretical analysis of the sequencing data of European ancestors. Further studies of other genetic backgrounds and annotation analyses are needed to confirm our findings and reveal the underlying biological mechanisms of identified genes and chemicals in the development of vertigo.

In conclusion, we identified three high-confidence genes (RAB3B, MTERFD2, and MGST1) showed proteome-wide significant associations in the human brain with vertigo in two independent brain proteomes, strongly highlighting the underlying pathogenesis of these proteins in vertigo. Our study provides new insights into the genetically regulated protein abundance in vertigo and also highlights promising chemicals for further pathogenesis investigations and therapeutics research.
Funding

This study was supported by the Natural Science Basic Research Plan in Shaanxi Province of China (2021JCW-08), and the Fundamental Research Funds for the Central Universities.

Competing interests

The authors report no conflicts of interest.

Supplementary material

Supplementary material is available at Brain Communications online.

References

1. Karatas M. Central vertigo and dizziness: Epidemiology, differential diagnosis, and common causes. Neurolog. 2008;14(6):355-364.
2. Bisdorff A, Von Brevern M, Lempert T, Newman-Toker DE. Classification of vestibular symptoms: Towards an international classification of vestibular disorders. J Vestib Res. 2009;19(1-2):1-13.
3. Hain TC, Uddin M. Pharmacological treatment of vertigo. CNS Drugs. 2003;17(2):85-100.
4. Baloh RW. Vertigo. Lancet (London, England). 1998;352(9143):1841-1846.
5. Hulse R, Biesdorf A, Hörmann K, et al. Peripheral vestibular disorders: An epidemiologic survey in 70 million individuals. Otol Neurotol. 2019;40(4):88-95.
6. Feil K, Strobl R, Schindler A, et al. What is behind cerebellar vertigo and dizziness? Cerebellum (London, England). 2019;18(3):320-332.
7. Neuhauser HK, Radtke A, von Brevern M, Lezisis F, Feldmann M, Lempert T. Burden of dizziness and vertigo in the community. Arch Int Med. 2008;168(19):2118-2124.
8. Kerber KA. The genetics of vertigo. Sem Neurol. 2006;26(5):484-491.
9. Gazquez I, Lopez-Escamez JA. Genetics of recurrent vertigo and vestibular disorders. Curr Genomics. 2011;12(6):443-450.
10. Skuladottir AT, Bjornsdottir G, Nawaz MS, et al. A genome-wide meta-analysis uncovers six sequence variants conferring risk of vertigo. Commun Biol. 2021;4(1):1148.
11. Saha K. Vertigo related to central nervous system disorders. Continuum (Minneapolis, Minn). 2021;27(2):447-467.
12. Basu A, Ash PE, Woizolin B, Emili A. Protein interaction network biology in neuroscience. Proteomics. 2021;21(3-4):e1900311.
13. Lagusse S, Ron D. Protein translation and psychiatric disorders. Neuroscientist. 2020;26(1):21-42.
14. Lewis CM, Knight J. Introduction to genetic association studies. Cold Spring Harb Protoc. 2012;2012(3):297-306.
15. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012;90(1):7-24.
16. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016;48(3):245-252.
17. Brandes N, Linial N, Linial M. PWAS: Proteome-wide association study-linking genes and phenotypes by functional variation in proteins. Genome Biol. 2020;21(1):173.
18. Liu J, Li X, Luo XJ. Proteome-wide association study provides insights into the genetic component of protein abundance in psychiatric disorders. Biol Psychiatry. 2021;90(11):781-789.
19. Wingo TS, Liu Y, Gerasimov ES, et al. Brain proteome-wide association study implicates novel proteins in depression pathogenesis. Nat Neurosci. 2021;24(6):810-817.
20. Fritsche L, Barenys M, Klose J, et al. Development of the concept for stem cell-based developmental neurotoxicity evaluation. Toxicol Sci. 2018;165(1):14-20.
21. Santamaría A, Díaz-Muñoz M. Chemical agents positively and negatively affecting the central nervous system. Cent Nerv Syst Agents Med Chem. 2010;10(4):258-269.
22. Beyer JN, Raniszewski NR, Burslem GM. Advances and opportunities in epigenetic chemical biology. Chembiochem. 2021;22(1):17-42.
23. Weng JC, Hong CI, Tasi JD, Shen CY, Su PH, Wang SL. The association between prenatal endocrine-disrupting chemical exposure and altered resting-state brain fMRI in teenagers. Brain Struct Funct. 2020;225(5):1669-1684.
24. Knudsen TB, Martin MT, Kavlock RJ, Judson RS, Dix DJ, Singh AV. Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA’s ToxRefDB. Reprod Toxicol. 2009;28(2):209-219.
25. Gudbjartsson DF, Helgason H, Gudjonsson SA, et al. Large-scale whole-genome sequencing of the Icelandic population. Nat Genet. 2015;47(5):435-444.
26. Welsh S, Peakman T, Sheard S, Almond R. Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK biobank cohort. BMC Genomics. 2017;18(1):26.
27. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719-748.
28. Wingo AP, Liu Y, Gerasimov ES, et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer’s disease pathogenesis. Nat Genet. 2021;53(2):143-146.
29. Wingo AP, Fan W, Duong DM, Gerasimov ES, Dammer EB. Shared proteomic effects of cerebral atherosclerosis and Alzheimer’s disease on the human brain. Nat Neurosci. 2020;23(6):696-700.
30. Zhou X, Carbonetto P, Stephens M. Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genet. 2013;9(2):e1003264.
31. Fromer M, Roussos P, Sieberts SK, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci. 2016;19(11):1442-1453.
32. Hall LS, Pain O, O’Brien HE, Anney R, Walters JTR, Owen MJ. eQTL-effects on gene expression in the human prenatal brain associated with genetic risk for neuropsychiatric disorders. Mol Psychiatry. 2021;26(6):2082-2088.
33. Davis AP, Grondin CJ, Johnson RJ, et al. The comparative toxicogenomics database: Update 2017. Nucleic Acids Res. 2017;45(D1):D972-d978.
34. Davis AP, Grondin CJ, Johnson RJ, et al. Comparative toxicogenomics database (CTD): Update 2021. Nucleic Acids Res. 2021;49(D1):D1138-d1143.
35. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet. 2007;81(6):1278-1283.
36. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genomewide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545-15550.
37. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57.
38. Szklarczyk D, Gable AL, Nasu KC, et al. The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(D1):D605-d612.
39. Wright AJ, Orlic-Milacic M, Rothfels K, et al. Evaluating the predictive accuracy of curated biological pathways in a public knowledgebase. Database. 2022;2022:baac009.

40. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. J Clin Epidemiol. 2014;67(8):850-857.

41. Matsuno A, Itoh J, Takekoshi S, et al. Dynamics of subcellular organelles, growth hormone, Rab3B, SNAP-25, and syntaxin in rat pituitary cells caused by growth hormone releasing hormone and somatostatin. Microsc Res Tech. 2003;62(3):232-239.

42. Schlüter OM, Basu J, Sudhof TC, Rosenmund C. Rab3 superprimers synaptic vesicles for release: Implications for short-term synaptic plasticity. J Neurosci. 2006;26(4):1239-1246.

43. Yamamoto Y, Nishimura N, Morimoto S, et al. Distinct roles of Rab3B and Rab13 in the polarized transport of apical, basolateral, and tight junctional membrane proteins to the plasma membrane. Biochem Biophys Res Commun. 2003;308(2):270-275.

44. Tsitsenis T, Younts TJ, Chiu CQ, Kaeser PS, Castillo PE, Südhof TC. Rab3B protein is required for long-term depression of hippocampal inhibitory synapses and for Normal reversal learning. Proc Natl Acad Sci USA. 2011;108(34):14300-5.

45. Lledo PM, Vernier P, Vincent JD, Mason WT, Zorec R. Inhibition of Rab3B expression attenuates ca(2+)-dependent exocytosis in rat anterior pituitary cells. Nature. 1995;364(6437):540-544.

46. Rotondo F, Scheithauer BW, Kovacs K, Bell DC. Rab3B immunoreexpression in human pituitary adenomas. Appl Immunohistochem Mol Morphol. 2009;17(3):185-188.

47. Nishioka H, Haraoka K. Significance of immunohistochemical expression of Rab3B and SNAP-25 in growth hormone-producing pituitary adenomas. Acta Neuropathol. 2005;109(6):598-602.

48. Xu Q, Zhang F, He H, et al. Expression profile of mouse Mterfd2, a novel component of the mitochondrial transcription termination factor (MTERF) family. Genes Genet Syst. 2011;86(4):269-275.

49. Spähl H, Habermann B, Gustafsson CM, Larsson NG, Hallberg BM. Structure of the human MTERF4-NSUN4 protein complex that regulates mitochondrial ribosome biogenesis. Proc Natl Acad Sci USA. 2012;109(38):15253-8.

50. Ye X, Han Y, Zhang L, Liu W, Zuo J. MTERF4 Regulates the mitochondrial dysfunction induced by MPP(+) in SH-SY5Y cells. Biochem Biophys Res Commun. 2015;464(1):214-220.

51. Wang XL, Liu Q, Chen GJ, Li ML, Ding YH. Overexpression of MTERF4 promotes the amyloidogenic processing of APP by inhibiting ADAM10. Biochem Biophys Res Commun. 2017;482(4):928-934.

52. Prall WC, Czibere A, Jäger M, et al. Age-related transcription levels of KU70, MGST1 and BIK in CD34 + hematopoietic stem and progenitor cells. Mech Ageing Dev. 2007;128(9):503-510.

53. Bräutigam L, Zhang J, Dreij K, et al. MGST1, A GSH transferase/peroxidase essential for development and hematopoietic stem cell differentiation. Redox Biol. 2018;17:171-179.