Functional & Integrative Genomics (2026) 26:130
https://doi.org/10.1007/s10142-026-01900-5

RESEARCH

# Peripheral and central immune features in insomnia: Integrative bulk and single-cell transcriptomic analyses reveal core hub genes and candidate compounds

**Wenwen Zhu[1,2] · Zhijuan Wen[1,2] · Yimeng Gao[1] · Bijuan He[1,2] · Ruitong Chen[1,2] · Haosen Pan[1,2] · Wei Bin[1,2] · Zhimin Yang[1,2] · Fei Tan[1]**

Received: 1 January 2026 / Revised: 25 March 2026 / Accepted: 25 May 2026


## Abstract

Insomnia is closely associated with immune dysregulation, yet the overall pattern of peripheral-central immune disequilibrium and its underlying molecular basis remains incompletely understood. To characterize the peripheral-central immune features associated with insomnia, identify key immune cell populations and core molecular programs, and prioritize candidate therapeutic compounds with preliminary experimental validation. Peripheral blood bulk transcriptomic dataset GSE208668 was analyzed using differential expression analysis, weighted gene co-expression network analysis (WGCNA), and functional enrichment analysis to identify insomnia-associated genes. Protein–protein interaction network analysis and machine learning models were then applied in independent peripheral blood datasets to refine core genes. Immune deconvolution and peripheral blood single-cell transcriptomic dataset GSE213496 were used to determine the immune-cell context, cell-type localization, and intercellular communication features of these genes. The brain single-cell transcriptomic dataset GSE137665 was further analyzed to assess central alterations. Drug prediction, molecular docking, and molecular dynamics simulations were performed to prioritize candidate compounds, followed by in vitro validation of resveratrol in an LPS-induced THP-1 macrophage model. A total of 5,321 differentially expressed genes associated with insomnia were identified, and weighted gene co-expression network analysis highlighted the turquoise and blue modules as key insomnia-related modules. Integrative analysis yielded 390 intersecting genes enriched mainly in immune, inflammatory, and oxidative stress-related pathways. Protein interaction analysis and machine learning further identified six refined core genes: FN1, HMOX1, HSP90AA1, IL10, MYD88, and NFE2L2. Because IL10 was not stably detected in the single-cell datasets, the remaining five genes were used for downstream single-cell analyses. Immune deconvolution suggested selective peripheral immune remodeling in insomnia, characterized by increased resting CD4 memory T cells and M2 macrophages, together with reduced activated NK cells. Peripheral single-cell analysis showed that HMOX1, HSP90AA1, MYD88, and NFE2L2 were mainly enriched in neutrophils, inflammatory macrophages, conventional dendritic cells, and selected lymphocyte populations, whereas FN1 showed a more restricted distribution pattern. CellChat analysis indicated enhanced intercellular communication under the sleep deprivation-related condition. In contrast, brain single-cell analysis revealed comparatively modest but detectable central alterations, including enrichment of ependymal

Wenwen Zhu, Zhijuan Wen and Yimeng Gao contributed equally to this work.

Wenwen Zhu, Zhijuan Wen and Yimeng Gao should be considered co-first authors.

✉ Zhimin Yang
yangzhimin1966@gzucm.edu.cn

✉ Fei Tan
tanfei2021@gzucm.edu.cn

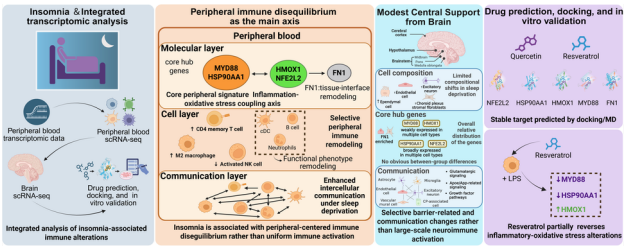
[1] State Key Laboratory of Traditional Chinese Medicine Syndrome/The Second Clinical School, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 510006, China

[2] Guangdong Laboratory of Chinese Medicine (Hengqin Laboratory), Hengqin, Guangdong Macao Deep Cooperation Zone 519031, China

Published online: 10 June 2026

cells, slight increases in excitatory neurons, brain endothelial cells, and choroid plexus stromal fibroblasts, together with heterogeneous expression of the core hub genes and selective rewiring of intracerebral communication networks. Drug prediction consistently prioritized quercetin and resveratrol, and structural analyses supported stable interactions with key targets. In THP-1 macrophages, resveratrol downregulated MYD88 and HSP90AA1 while further upregulating HMOX1. Insomnia appears to be associated predominantly with a peripheral-centered immune disequilibrium pattern, characterized by selective remodeling of innate immune-related populations, enhanced inflammatory and oxidative stress programs, and increased intercellular communication. The MYD88-HMOX1-HSP90AA1-NFE2L2 axis may represent a key molecular program linking inflammatory activation and oxidative stress adaptation. Resveratrol was identified as a potential compound and has been validated through preliminary in vitro experiments.

**Graphical abstract**




**Keywords** Insomnia · Sleep deprivation · Immune homeostasis imbalance · Oxidative stress · MYD88

## Introduction

Insomnia is one of the most common sleep disorders and is characterized by difficulty initiating sleep, impaired sleep maintenance, or early morning awakening (Riemann et al. 2022). Chronic insomnia is associated with impaired attention, emotional disturbances, and cognitive dysfunction, and is also linked to depression, anxiety, cardiovascular disease, metabolic syndrome, and neurodegenerative disorders (Khan and Aouad 2022; Riemann et al. 2022; Wu et al. 2023). Recent epidemiological studies have estimated that the global prevalence of insomnia is approximately 16.2% (Benjafield et al. 2025). Clinically, insomnia is mainly assessed using subjective questionnaires, such as the Pittsburgh Sleep Quality Index (PSQI) and the Insomnia Severity Index (ISI), in combination with objective measures including polysomnography (PSG). Current treatments mainly include cognitive behavioral therapy and pharmacological interventions such as benzodiazepines, melatonin, and GABAergic agents (De Crescenzo et al. 2022; Riemann et al. 2023). Although these approaches can improve symptoms, they primarily target neurotransmission and arousal regulation, and their long-term use may be limited by adverse effects and dependence. These limitations indicate the need for a broader biological framework to better understand the pathophysiology of insomnia.

In recent years, advances in neuroimmunology have revised the traditional view of the brain as an immune-privileged organ. As pointed out by Rustenhoven and Kipnis, the central nervous system (CNS) communicates with the peripheral immune system through a neuroimmune interface that includes the meninges, the blood–brain barrier (BBB), the choroid plexus, and meningeal lymphatic pathways (Castellani et al. 2023; Rustenhoven and Kipnis 2022). These structures enable continuous molecular exchange and functional interaction between peripheral immunity and the CNS (Castellani et al. 2023; Rustenhoven and Kipnis 2022). Neurons can regulate immune responses through cytokine signaling pathways, including IL-13-related pathways (Alves de Lima et al. 2020a, b; Herz et al. 2021; Li et al. 2023), whereas immune cells are also capable of sensing neurotransmitters and neuropeptides (Boahen et al. 2023; Castellani et al. 2023; Rustenhoven and Kipnis 2022). In addition, the meninges and choroid plexus contain diverse immune cell populations, including T cells, B cells, dendritic cells, and macrophages, which can sense signals in the cerebrospinal fluid and relay information related to tissue homeostasis or immune activation (Alves de Lima et al. 2020a, b; Cugurra et al. 2021; Rustenhoven and Kipnis 2022; Sankowski et al. 2024; Xu et al. 2024). Together, these findings support the concept of a "neuroimmune ecosystem" and provide a theoretical framework for investigating insomnia beyond a purely neuronal model.

Evidence from both central and peripheral levels further suggests that immune dysregulation may participate in sleep-related pathological processes. At the central level, single-cell and organoid studies have shown that microglia can respond to peripheral inflammatory signals and influence neural network remodeling (Jha et al. 2022; Schafer et al. 2023; Zhang et al. 2023a, b). Experimental sleep

deprivation has also been associated with TREM2-dependent microglial activation and increased Aβ deposition, thereby linking neuroimmune responses to sleep-related neuropathology (Parhizkar et al. 2023; Zhao et al. 2018). At the peripheral level, chronic insomnia is often associated with elevated inflammatory mediators, including IL-6, TNF-α, and IL-1β, together with reduced antioxidant capacity (Ballesio et al. 2026; Davinelli et al. 2024; Zhang et al. 2023a, b). Peripheral immune signals can affect the CNS through humoral routes or neural reflex pathways (Boahen et al. 2023; Garbarino et al. 2021), whereas central immune activation may in turn influence peripheral immune regulation (Castellani et al. 2023; Garbarino et al. 2021; Rustenhoven and Kipnis 2022). These observations suggest that insomnia may involve coordinated alterations across peripheral and central immune levels.

Although increasing attention has been paid to the immunobiology of insomnia, previous studies have largely focused on only a single aspect of this process, such as circulating cytokines, oxidative stress markers, or microglial responses, without systematically elucidating the molecular networks and cellular programs linking peripheral immune alterations to central neuroimmune changes (Tang et al. 2025). In the present study, we performed an integrative analysis of bulk and single-cell transcriptomic data to characterize the immunological features of insomnia at both peripheral and central levels. Using peripheral blood transcriptomic datasets, we identified insomnia-associated transcriptional alterations and prioritized candidate genes through differential expression analysis, weighted gene co-expression network analysis, protein–protein interaction analysis, and machine learning. We then used peripheral blood single-cell data to define relevant immune cell populations, cell states, and intercellular communication programs, and further analyzed brain single-cell data to assess the distribution of these signals across barrier-associated, immune-related, and neural cell populations. Finally, based on the resulting gene set, we predicted candidate compounds and further evaluated them through molecular docking, molecular dynamics simulation, and in vitro validation in an LPS-stimulated THP-1 macrophage model. Through this integrative analytical framework spanning peripheral and central levels, we aimed to clarify the immunological features of insomnia and provide clues for future mechanistic studies and therapeutic exploration.

## Materials and methods

### Study design and analytical framework

We first used a human peripheral blood transcriptome dataset (GSE208668) as the discovery cohort to identify insomnia-associated transcriptional changes and hub genes through differential expression analysis, weighted gene co-expression network analysis (WGCNA), functional enrichment analysis, and protein–protein interaction (PPI) network analysis. We then used independent human peripheral blood datasets (GSE56931 and GSE37667) as the training and validation cohorts to further prioritize the refined core genes from these hub genes through machine learning models and SHAP-based interpretation. To define the immune cell context of these genes, we next combined bulk immune deconvolution with mouse peripheral blood single-cell transcriptomic analysis to identify associated immune cell subsets, intercellular communication patterns, and differentiation-related states. The expression data for GSE208668, GSE56931, and GSE37667 were downloaded from the Gene Expression Omnibus (GEO) database as series matrix files. As these GEO series matrix files contained preprocessed expression matrices provided by the submitters, subsequent analyses were performed on the processed data. Probe annotation was conducted according to the corresponding platform annotation files and converted to gene symbols. When multiple probes mapped to the same gene, the median value was used as the representative expression level.

Among the six refined core genes, five genes (FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2) were subsequently retained as the core hub genes based on their clear cell type-specific expression patterns in the peripheral blood single cell dataset (GSE213496). We further examined the distribution of the core hub genes in a mouse brain tissue single cell dataset (GSE137665) to assess their cell type-specific expression across barrier-associated, immune-related, and neuroglial populations. Finally, based on the refined core genes and the core hub genes, we performed drug prediction and structural analyses to prioritize candidate compounds for subsequent experimental validation.

### Identification of insomnia-associated peripheral transcriptional changes in the discovery cohort

#### Peripheral blood transcriptome dataset

The GSE208668 dataset was downloaded from the Gene Expression Omnibus (GEO) database and used as the discovery cohort. This microarray dataset contains 17 insomnia samples and 25 control samples derived from human peripheral blood and profiled on the GPL10904 platform.

#### Differential gene expression analysis

To identify insomnia-associated transcriptional alterations in peripheral blood, differential expression analysis was

performed using the limma package (version 3.40.6) in R. Gene expression values were fitted using linear models, and empirical Bayes moderation was applied to estimate moderated t-statistics, F statistics, and log odds values for differential expression. Genes with fold change ≥ 1.5 and $P < 0.05$ were defined as differentially expressed genes (DEGs).

### Weighted gene co-expression network analysis

Weighted gene co-expression network analysis (WGCNA) was performed using the WGCNA R package (version 1.72.1) in R 4.3.3 to identify co-expression modules associated with insomnia. After filtering out low-expression genes, the top 5,000 most variable genes ranked by standard deviation were retained for network construction. A signed co-expression network was established using biweight mid-correlation (bicor). The soft-thresholding power was evaluated from 1 to 20 using the pickSoftThreshold() function, and power = 12 was selected to achieve approximate scale-free topology.

Modules were identified using the blockwiseModules() function with the following parameters: networkType = "signed", TOMType = "signed", minModuleSize = 30, deepSplit = 4, and mergeCutHeight = 0.10. Module eigengenes were calculated for each module and correlated with the binary phenotype (Insomnia = 1, Control = 0) to identify insomnia-related modules. Based on the module–trait relationship analysis, the turquoise and blue modules, which showed the strongest positive and negative correlations with insomnia, respectively, were selected for downstream analysis. Gene significance (GS) and module membership (MM) were then calculated within these two modules to evaluate the association between individual genes and the insomnia phenotype. Although functional enrichment was not performed on the full WGCNA modules, the downstream intersecting genes derived in part from the key insomnia-associated modules were mainly enriched in immune-, inflammation-, and oxidative stress-related pathways, suggesting that these modules may capture co-expression patterns relevant to inflammatory and stress-response processes in insomnia.

### Integration with prior insomnia-related genes

To focus on genes most relevant to insomnia, insomnia-associated genes were retrieved from the GeneCards database (https://www.genecards.org/) using "Insomnia" as the search term. Subsequently, the intersection was taken among three gene sets: (1) DEGs identified in GSE208668; (2) genes from the turquoise and blue modules identified by WGCNA as the modules most strongly associated with insomnia; and (3) insomnia-related genes retrieved from GeneCards.

### PPI network construction, functional annotation, and hub gene screening

To characterize the biological relevance and network properties of the intersecting genes, the gene set was first submitted to the STRING database (http://string-db.org/), with the species restricted to Homo sapiens and the minimum interaction confidence set to the default medium threshold (confidence score > 0.4). Based on this input, a protein–protein interaction (PPI) network was generated, and functional enrichment analyses were performed on the same STRING platform, including Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Disease Ontology, tissue-specific expression, Human Phenotype Ontology (HPO), and UniProt keyword enrichment. These analyses were used to define the major biological processes, pathways, and phenotype associations represented by the intersecting genes.

The resulting PPI network was then imported into Cytoscape (version 3.9.1) for visualization and topological analysis. Hub gene screening was performed using the CytoHubba plugin. Four algorithms, namely MCC, MNC, EPC, and Degree, were applied to rank node importance, and the top 20 genes identified by each algorithm were extracted. The intersection of these four gene lists was defined as the hub genes.

## Identification of the refined core genes by machine learning

### Training and validation cohorts

To further prioritize the most relevant genes from the hub genes, two independent peripheral blood microarray datasets were downloaded from GEO. GSE56931 (platform GPL10379) was used as the training cohort. In this dataset, samples labeled "day1, 8 h, baseline" were defined as controls, whereas samples labeled "day3, 8 h, sleepdep" were assigned to the sleep deprivation-related group. GSE37667 (platform GPL570) was used as the validation cohort, with "baseline" samples defined as controls and "Prolonged_wakefulness" samples assigned to the sleep deprivation-related group. Sleep deprivation datasets were used to capture sleep-loss-related immune/transcriptional programs that may overlap with insomnia, rather than to serve as strictly disease-matched validation cohorts. For clarity of presentation, the sleep deprivation/prolonged wakefulness group in the external datasets was labeled as "Insomnia" in the relevant figures.

### Machine learning model construction and feature prioritization

To further improve the robustness of hub gene identification, a comprehensive machine learning framework was constructed (Liu et al. 2022). The GSE56931 dataset was used as the training set, and GSE37667 was used as an independent validation set. A total of 20 machine learning models were implemented, including glmBoost, Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Naïve Bayes, Ridge regression, and stepwise generalized linear models (forward, backward, and both directions). Model performance was evaluated using the area under the receiver operating characteristic curve (AUC). Genes selected by each model were recorded, and their selection frequency across models was calculated. Genes that were consistently identified in multiple high-performing models were defined as refined core genes, representing the most stable and reliable biomarkers.

To further enhance model interpretability, SHapley Additive exPlanations (SHAP) analysis was applied to quantify the contribution of each gene to model predictions. SHAP values were used to rank gene importance and to validate the stability of refined core genes across different models.

### Single gene GSEA of the refined core genes

To explore the potential functional context of each refined core gene, single-gene GSEA was performed separately in the GSE208668 cohort using the normalized expression matrix (Subramanian et al. 2005). For each refined core gene, samples were dichotomized into high- and low-expression groups according to the median expression value of that gene. Genome-wide expression differences between the two groups were then analyzed by GSEA using GSEA software (version 3.0; Broad Institute). The reference gene sets were obtained from MSigDB, and the c2.cp.kegg.v7.5.1.symbols.gmt collection was used as the background gene set (Liberzon et al. 2011). The main parameters were set as follows: minGSSize = 5, maxGSSize = 5000, and nPerm = 1000. Pathways with nominal $P < 0.05$ and FDR < 0.25 were considered significantly enriched.

### Expression validation of the refined core genes

The expression levels of the refined core genes (FN1, HMOX1, HSP90AA1, IL10, MYD88, and NFE2L2) were further validated in the GSE208668 discovery cohort using the normalized expression matrix. For each gene, expression levels were compared between the control and insomnia groups, and the results were visualized using boxplots. Statistical significance was assessed using the Wilcoxon rank-sum test, and $P < 0.05$ was considered statistically significant.

## Peripheral immune cell context of the refined core genes

### Immune cell deconvolution of peripheral blood transcriptomes

To characterize the peripheral immune landscape associated with insomnia, the gene symbol-annotated GSE208668 peripheral blood expression matrix was analyzed using CIBERSORT in microarray mode to estimate the relative proportions of 22 immune cell types (Newman et al. 2015). Immune cell subsets with a mean estimated proportion of < 1% or a zero proportion in ≥ 70% of samples were excluded to avoid unstable estimates caused by sparse distributions. The remaining cell types were used for between-group comparisons and correlation analyses with the refined core genes. Differences in immune cell proportions between the control and insomnia groups were assessed using the Wilcoxon rank-sum test. Associations between immune cell proportions and refined core gene expression were evaluated using Spearman correlation analysis, and correlations with $|r| \geq 0.6$ and $P < 0.001$ were considered significant.

### Peripheral blood single cell transcriptomic dataset and quality control

The mouse peripheral blood single-cell transcriptomic dataset GSE213496 was downloaded from GEO and analyzed using Seurat (version 5.0.1) in R (version 4.3.1). Samples labeled "control" were assigned to the control group, whereas samples labeled "SD" were assigned to the sleep deprivation group. The sleep deprivation dataset was used to capture sleep-loss-related immune/transcriptional programs that may overlap with insomnia, rather than to serve as strictly disease-matched validation cohorts. For clarity of presentation, the sleep deprivation group was labeled as "Insomnia" in the relevant figures. Raw count matrices were imported using Read10X() and converted into Seurat objects with CreateSeuratObject() using min.cells = 3 and min.features = 300. The percentages of mitochondrial transcripts and hemoglobin-related transcripts were calculated using PercentageFeatureSet(), with mitochondrial genes identified by the prefix "mt-" and hemoglobin genes defined as Hba-a1, Hba-a2, Hba-x, Hbb-bs, Hbb-bt, Hbb-bh1, and Hbb-y. Cells were retained according to the following criteria: nFeature_RNA > 300, nFeature_RNA < 5000, nCount_RNA > 500, nCount_RNA < quantile (nCount_RNA, 0.97),

mt_percent < 10, and HB_percent < 5. After filtering, samples were merged for downstream analysis. Cell cycle scoring was performed using the CellCycleScoring function, and cell cycle effects were regressed out during data scaling using S.Score and G2M.Score.

#### Dimensionality reduction, clustering, and cell annotation

After quality control, the merged dataset was normalized using the LogNormalize method. Highly variable genes were identified using FindVariableFeatures() with nfeatures = 2000, and the data were scaled using ScaleData(). Principal component analysis (PCA) was then performed, and the first 12 principal components were selected according to the elbow plot for downstream analysis. Graph-based clustering was performed using FindNeighbors() and FindClusters() with a clustering resolution of 0.2, and the cell clusters were visualized by UMAP and t-SNE. Cluster marker genes were identified using FindAllMarkers() with the Wilcoxon rank-sum test, with only.pos = TRUE, logfc.threshold = 0.25, and min.pct = 0.25. Statistical significance was evaluated using Benjamini–Hochberg-adjusted P values. Cell identities were assigned based on canonical marker genes curated from PanglaoDB, CellMarker 2.0, ImmGen, and published literature. UMAP plots, heatmaps, and stacked bar plots were generated to visualize cell distribution, marker expression, and cell-type composition across groups. Batch effects were assessed by comparing sample distributions across clusters, and no additional batch correction was applied because the observed differences were considered to primarily reflect biological variation rather than technical bias.

#### Mapping of the refined core genes to peripheral immune cell subsets

To determine the cellular context of the refined core genes, their expression patterns were visualized across annotated peripheral immune cell populations using UMAP feature plots and dot plots based on the normalized Seurat object. Gene symbols were matched to the mouse single-cell dataset in a case-insensitive manner before visualization. Group-specific expression patterns were further examined using split UMAP feature plots according to the control and sleep deprivation groups. Because IL10 was not consistently detected in the peripheral blood single-cell dataset, it was excluded from subsequent analyses. The remaining five genes—FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2—were retained as the core hub genes for downstream evaluation.

#### Peripheral cell communication analysis

To investigate intercellular communication associated with sleep deprivation in peripheral blood, CellChat analysis was performed using the annotated single-cell dataset and CellChatDB.mouse. CellChat objects were constructed separately for the control and sleep deprivation groups based on the normalized RNA expression matrix and cell-type labels, and then merged for comparative analysis. The analysis workflow included subsetData(), identifyOverExpressedGenes(), identifyOverExpressedInteractions(), computeCommunProb(population.size = TRUE), filterCommunication(min.cells = 10), computeCommunProbPathway(), and aggregateNet(). Communication number, communication strength, differential interaction networks, sender-receiver relationships, and pathway-level signaling patterns were analyzed to identify major communication hubs and altered signaling axes associated with sleep deprivation. In addition, macrophage-related signaling pathways were further ranked according to pathway-associated communication strength.

#### Differentiation-related state analysis

To evaluate differentiation-related cell states in peripheral immune cells, CytoTRACE2 was applied in mouse mode to the annotated Seurat object. CytoTRACE2 scores were calculated for individual cells and projected onto the UMAP embedding to visualize the distribution of differentiation-related states. Group-wise differences in CytoTRACE2 scores were further compared between the control and sleep deprivation groups to assess whether sleep deprivation was associated with shifts in relative differentiation potential.

### Assessment of the core hub genes in brain tissue single-cell transcriptomic data

#### Brain tissue single cell dataset and quality control

The mouse brain tissue single-cell transcriptomic dataset GSE137665, including the cerebral cortex, hypothalamus, and brainstem, was downloaded from GEO and analyzed using Seurat in R. Six samples, including control1-3 and sleep deprivation1-3, were included in the analysis. Samples labeled Normal Sleep were assigned to the control group, whereas samples labeled Sleep Deprived were assigned to the sleep deprivation group. The sleep deprivation dataset was used to capture sleep-loss-related immune/transcriptional programs that may overlap with insomnia, rather than to serve as strictly disease-matched validation cohorts.

For clarity of presentation, the sleep deprivation group was labeled as "Insomnia" in the relevant figures. Raw count matrices were imported using Read10X() and converted into Seurat objects using CreateSeuratObject() with min.cells = 3 and min.features = 300. The percentage of mitochondrial transcripts was calculated using PercentageFeatureSet() with the pattern "^mt-". Cells were retained according to the following criteria: nFeature_RNA > 300, nFeature_RNA < 4000, mt_percent < 12, nCount_RNA > 500, and nCount_RNA < quantile (nCount_RNA, 0.99). After quality control, the filtered samples were merged for downstream analysis.

#### Data integration, clustering, and cell annotation

After quality control, the merged dataset was normalized using NormalizeData(), and 2,000 highly variable genes were identified using FindVariableFeatures(). The data were then scaled with ScaleData() and subjected to principal component analysis using RunPCA() with 30 principal components. Because sample-driven clustering was observed before integration, batch correction was performed using Harmony with original identity as the batch variable. Downstream analysis was based on the Harmony reduction. Specifically, FindNeighbors() and RunUMAP() were performed using the first 20 Harmony components, and clustering was carried out using FindClusters() at a resolution of 0.1. Marker genes for each cluster were identified using FindAllMarkers() with test.use = "wilcox", only.pos = TRUE, logfc.threshold = 0.25, min.pct = 0.25, and layer = "data". Statistical significance was evaluated using Benjamini–Hochberg-adjusted P values. Cell identities were manually assigned based on marker gene expression patterns and published reference resources, resulting in 11 annotated brain cell populations.

#### Mapping of the core hub genes in brain cell populations

To assess the cellular distribution of the core hub genes in the central compartment, their expression patterns were visualized across annotated brain cell populations using UMAP feature plots and dot plots based on the normalized Seurat object. Gene symbols were matched to the mouse dataset in a case-insensitive manner before visualization. Group-specific expression patterns were further examined using split UMAP feature plots and split dot plots according to the control and sleep deprivation groups. Particular attention was given to barrier-associated, immune-related, and neuroglial cell populations, including brain endothelial cells, fenestrated endothelial cells, choroid plexus epithelial cells, choroid plexus stromal fibroblasts, astrocytes, microglia, and border-associated macrophages.

#### Intracerebral cell–cell communication analysis

To characterize the intracerebral communication landscape, CellChat analysis was performed using the annotated brain single-cell dataset and CellChatDB.mouse. A CellChat object was constructed from the normalized RNA expression matrix using cell-type labels as grouping information. The analysis workflow included subsetData(), identifyOverExpressedGenes(do.fast = FALSE), identifyOverExpressedInteractions(), computeCommunProb(population.size = TRUE), filterCommunication(min.cells = 10), computeCommunProbPathway(), and aggregateNet(). Communication number, pathway-level communication strength, sender-receiver relationships, and representative ligand-receptor pathways were analyzed to identify major intracerebral communication hubs and signaling axes.

### Drug prediction and structural evaluation of candidate compounds

#### Drug prediction

Potential therapeutic compounds were predicted separately based on the refined core genes and the core hub genes using a previously described drug-repurposing framework (Peng et al. 2020). Candidate compounds were screened against integrated drug–target interaction resources, and the output compounds were annotated using DrugBank identifiers in the final ranked results. For each gene set, compounds were prioritized according to distance score, nominal P value, and false discovery rate (FDR). To improve robustness, we focused on compounds that were reproducibly identified in both the refined core gene model and the core hub gene model. Based on this overlapping strategy, quercetin and resveratrol were retained as shared candidate compounds for downstream structural analyses because both showed significant enrichment in the two independent screening models. Supplementary Excel S3 provides the ranked compound lists for both analyses.

#### Molecular docking

To evaluate the binding potential between the prioritized compounds and the core hub proteins, the two-dimensional structures of quercetin and resveratrol were obtained from the PubChem database and converted into three-dimensional conformations using ChemOffice. Protein structures were downloaded from the RCSB PDB database, and the following crystal structures were selected for docking: FN1 (PDB ID: 2CG7), HMOX1 (PDB ID: 1N45), HSP90AA1 (PDB ID: 6TN5), MYD88 (PDB ID: 4EO7), and NFE2L2

(PDB ID: 2FLU). Protein preprocessing was performed in PyMOL 2.6, including removal of water molecules and irrelevant heteroatoms. Ligand and receptor preparation, including hydrogen addition, charge assignment, rotatable bond definition, and docking box setup, was conducted using AutoDockTools 1.5.6. Molecular docking was then performed using AutoDock Vina, and the optimal binding conformations were selected according to binding energy scores. Docking poses were visualized using Discovery Studio 2019 and PyMOL 2.6. In this study, a binding energy below −5.0 kcal/mol was considered indicative of favorable binding activity, whereas a value below −7.0 kcal/mol suggested relatively strong binding affinity.

#### Molecular dynamics simulation

To further assess the stability of the docked complexes, molecular dynamics (MD) simulations were performed using GROMACS 2022. Protein topologies were generated with the AMBER14SB force field, and ligand parameters were assigned using GAFF2. Ligand topologies were prepared with Antechamber, and atomic charges were calculated using the RESP method. Each protein–ligand complex was solvated in a cubic box containing TIP3P water molecules, with a minimum distance of 1.0 nm from the solute to the box edge. Na and Cl ions were added to neutralize the system and achieve a final salt concentration of 0.15 M. Long-range electrostatic interactions were treated using the particle mesh Ewald method, and hydrogen-containing bonds were constrained using the LINCS algorithm.

Energy minimization was carried out by 3000 steps of steepest descent followed by 2000 steps of conjugate gradient minimization. The minimized systems were then equilibrated and subjected to a 100-ns production run under the NPT ensemble at 310 K and 1 bar, using the Nosé–Hoover thermostat and Parrinello–Rahman barostat. Structural stability was evaluated by analyzing the root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), and hydrogen bonds throughout the trajectories.

### In vitro *validation in THP-1 macrophage model*

#### *Cells, reagents, and* in vitro *model establishment*

The human monocytic cell line THP-1 was obtained from Procell and cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS). Primary antibodies against GAPDH (60004–1-Ig), MYD88 (67969–1-Ig), HSP90AA1 (13171–1-AP), and HMOX1 (10701–1-AP) were purchased from Proteintech. FBS was purchased from Gibco Inc. Phorbol 12-myristate 13-acetate (PMA) was obtained from MedChemExpress, lipopolysaccharide (LPS) from Aladdin, and resveratrol (RSV) from Macklin. The Cell Counting Kit-8 (CCK-8) was purchased from Beijing Solarbio Technology Co., Ltd. The qRT-PCR kit was obtained from Vazyme, and primers were synthesized by Sangon Biotech. The primer sequences were as follows:

MYD88-F: 5′ CGGATGGTGGTGGTTGTCTCTG 3′
MYD88-R: 5′ GCTTCTGATGGGCACCTGGAG 3′
HSP90AA1-F: 5′ CCTCCTCGCCGCCGTTTC 3′
HSP90AA1-R: 5′ CCAGACACCATCAGATGCCAGAG 3′
HMOX1-F: 5′ CTACCGCTCCCGCATGAACTC 3′
HMOX1-R: 5′ AGCAGGAACGCAGTCTTGGC 3′
ACTB-F: 5′ CATGTACGTTGCTATCCAGGC 3′
ACTB-R: 5′ CTCCTTAATGTCACGCACGAT 3′

#### Cell differentiation, inflammatory stimulation, and drug treatment

To induce the differentiation of THP-1 cells into a macrophage-like phenotype, THP-1 cells in the logarithmic growth phase were seeded into culture plates and treated with 100 nM PMA for 24 h. After the PMA-containing medium was removed, the cells were gently washed with PBS and then cultured in fresh complete medium for an additional 24 h to minimize the residual effects of PMA on subsequent inflammatory stimulation. Differentiated THP-1 macrophage-like cells were then stimulated with 1 μg/mL LPS for 24 h to induce an inflammatory response. Resveratrol (RSV) was dissolved in DMSO to prepare a stock solution and diluted with low-serum medium to the desired final concentrations. The final concentration of DMSO was kept consistent among groups and did not exceed 0.1%. Cells were treated for 24 h.

#### CCK-8 assay for cell viability and determination of the optimal resveratrol concentration

To determine the non-cytotoxic dose range and optimal treatment concentration of resveratrol, THP-1 cells were seeded into 96-well plates and induced into macrophage-like cells. The cells were then treated with different concentrations of resveratrol (5, 10, 25, 50, and 100 μM) for 24 h. Subsequently, fresh culture medium containing 10% CCK-8 working solution was added to each well, and the cells were incubated for an additional 2 h. The absorbance at 450 nm was measured using a microplate reader.

#### Western blot analysis

Three groups were included in the Western blot analysis: control group, LPS group, and LPS + optimal concentration

of resveratrol group. After treatment, the culture medium was discarded, and the cells were washed twice with pre-cooled PBS. Total protein was extracted using RIPA lysis buffer on ice for 30 min, followed by centrifugation at 12,000 rpm and 4 °C for 15 min. The supernatants were collected, and protein concentrations were determined using the BCA method. Equal amounts of total protein were mixed with 5 × loading buffer at a ratio of 4:1 and denatured at 100 °C for 10 min. Proteins were then separated by SDS-PAGE and transferred onto PVDF membranes. After blocking with 5% non-fat milk at room temperature for 1.5 h, the membranes were incubated with primary antibodies overnight at 4 °C. After washing with TBST, the membranes were incubated with HRP-conjugated secondary antibodies for 1 h at room temperature. Protein bands were visualized using an ECL chemiluminescence reagent and captured with a gel imaging system. Band intensities were quantified using ImageJ software.

### qRT-PCR analysis

Three groups were included for qRT-PCR analysis: control group, LPS group, and LPS + optimal concentration of resveratrol group. After treatment, total RNA was extracted according to the manufacturer's instructions. Qualified RNA samples were reverse-transcribed into cDNA. Quantitative real-time PCR was subsequently performed using the SYBR Green method. The PCR reaction system and cycling conditions were set according to the kit instructions. Each 20 μL reaction contained 10 μL SYBR Green Mix, 0.8 μL forward and reverse primers, 1 μL cDNA template, and 8.2 μL RNase-free water. The amplification conditions were as follows: initial denaturation at 95 °C for 5 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s. GAPDH was used as the internal reference gene, and the relative expression levels of target genes were calculated using the 2^-ΔΔCt method.

### Statistical analysis

All experiments were performed with at least three independent biological replicates. For the CCK-8 assay, each condition was tested in triplicate wells. For qRT-PCR, each sample was analyzed in technical triplicate. Data are presented as mean ± SD. Statistical analysis was conducted using GraphPad Prism 10 software. Comparisons between two groups were performed using the independent-samples t-test, while comparisons among three or more groups were analyzed by one-way analysis of variance (one-way ANOVA), followed by Tukey's post hoc test for multiple comparisons when appropriate. A p-value < 0.05 was considered statistically significant.

## Results

### Identification of insomnia-associated transcriptional alterations in peripheral blood

To characterize peripheral transcriptional alterations associated with insomnia, we first analyzed the peripheral blood transcriptome dataset GSE208668. Differential expression analysis identified 5,321 differentially expressed genes (DEGs) between the insomnia and control groups (Fig. 1A; Supplementary Excel S1), indicating substantial peripheral transcriptional alterations associated with insomnia.

To identify gene co-expression modules associated with insomnia, WGCNA was performed using the top 5,000 most variable genes. As shown in Fig. 1B, a soft-thresholding power of 12 was selected to ensure approximate scale-free topology. Based on this threshold, a signed co-expression network was constructed, and multiple gene modules were identified (Fig. 1C). Module-trait relationship analysis showed that the turquoise module was most positively correlated with insomnia, whereas the blue module was most negatively correlated with insomnia (Fig. 1D). Consistently, module eigengene analysis demonstrated that the turquoise module eigengene was significantly increased in the insomnia group, while the blue module eigengene was significantly decreased (Fig. 1E, Supplementary Excel S1). To further assess the biological relevance of genes within these key modules, the relationships between gene significance (GS) and module membership (MM) were examined. Strong GS–MM correlations were observed in both the blue and turquoise modules, indicating that genes highly connected within these modules were also strongly associated with the insomnia phenotype (Fig. 1F). These findings identified the turquoise and blue modules as the key insomnia-related modules for subsequent analysis.

Integrative analysis yielded 390 intersecting genes associated with insomnia (Fig. 2A; Supplementary Excel S1). Functional annotation of this gene set showed enrichment in immune, inflammatory, and oxidative stress-related pathways, including NF-κB, Toll-like receptor, and NOD-like receptor signaling (Fig. 3A-H; Supplementary Excel S1), indicating that peripheral transcriptional alterations in insomnia are largely centered on stress and immune regulation. As these intersecting genes were derived in part from the key insomnia-associated WGCNA modules, these enrichment results also provide a functional interpretation of the major co-expression patterns identified above.

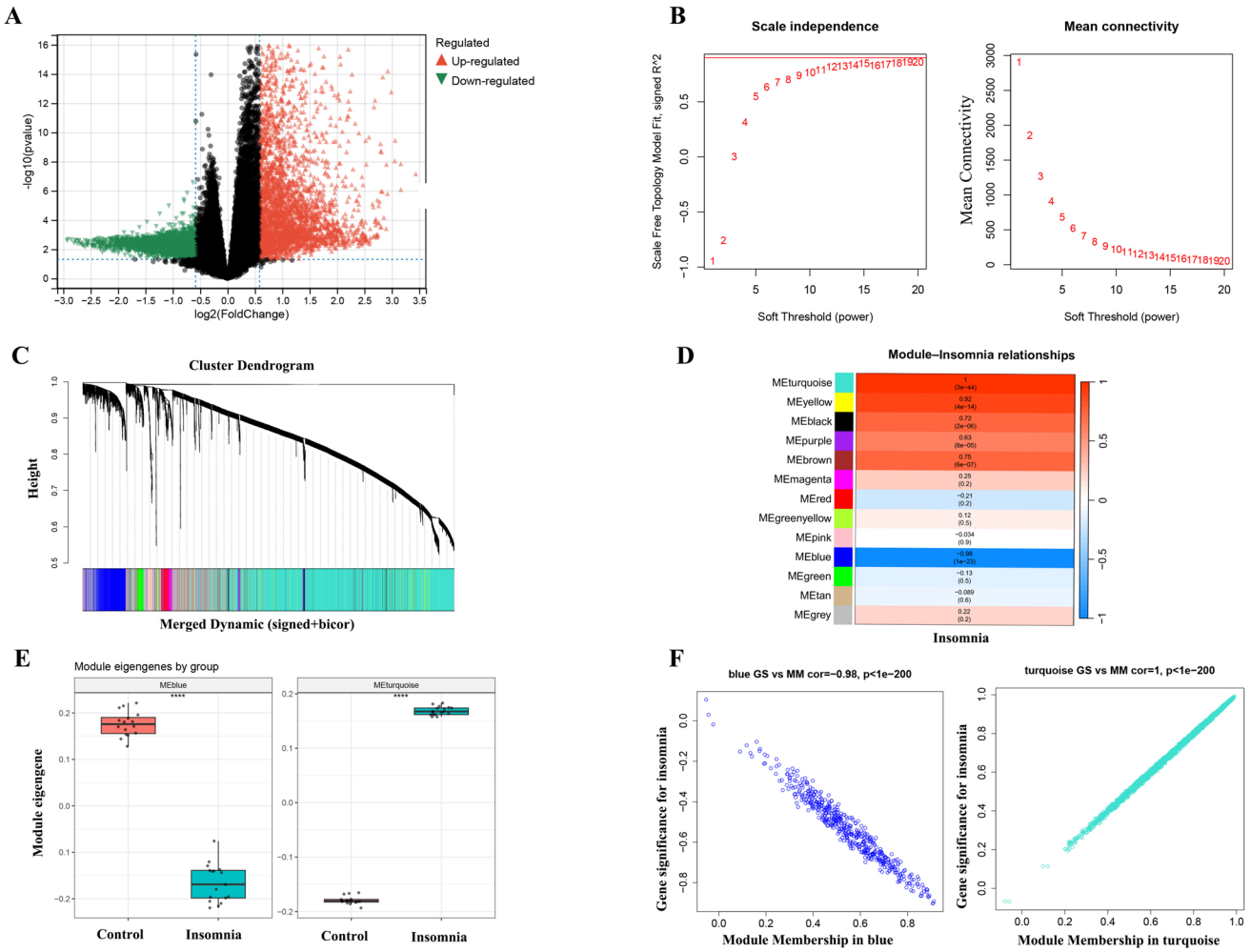



**Fig. 1** Differentially expressed gene identification and WGCNA co-expression network analysis in insomnia. (**A**) Volcano plot of differentially expressed genes in the GSE208668 dataset. (**B**) Scale-free topology analysis of soft-thresholding powers for WGCNA network construction. (**C**) Sample clustering dendrogram used to assess overall data quality and detect outliers. (**D**) Module–trait relationship heatmap showing correlations between module eigengenes and the insomnia phenotype. (**E**) Boxplots show the distribution of module eigengene values in the Control and Insomnia groups, with individual samples overlaid as jittered points. Group differences were evaluated using the Wilcoxon rank-sum test. Significance is indicated as ns, not significant; *$P<0.05$; **$P<0.01$; ***$P<0.001$; ****$P<0.0001$. (**F**) Scatter plots of gene significance versus module membership in the MEturquoise and MEblue modules

## Network and machine learning based prioritization of refined core genes associated with insomnia

To identify hub genes among the insomnia-associated intersecting genes, we first constructed a protein–protein interaction network based on the 390 intersecting genes. The resulting network showed significant connectivity, supporting the presence of a biologically coherent interaction structure (Fig. 2B; Supplementary Excel S1). In Cytoscape, the intersection of the top 20 genes identified by four topological algorithms (Degree, EPC, MCC, and MNC) yielded nine hub genes: FN1, MYD88, NFE2L2, HMOX1, VCAM1, HSP90AA1, HSPD1, MAPK3, and IL10. These genes were therefore considered key nodes in the topology-based analysis (Fig. 2C-E; Supplementary Excel S1), suggesting that they may represent central components of the peripheral molecular alterations associated with insomnia.

To identify the refined core genes, we next applied machine learning models in two independent peripheral blood datasets. Among the tested models, glmBoost+RF showed the highest overall predictive performance, with AUCs of 0.974 in the training cohort and 0.716 in the validation cohort (mean AUC = 0.845) (Fig. 4A). However, this model combination showed a relatively low recurrence frequency in the model-screening framework, indicating that it should be regarded as the best-performing representative model, rather than the sole or most stable basis for feature selection. In contrast, several other models, including LDA and Stepglm[forward], were retained more frequently across the screening process (Fig. 4B). Therefore, the final

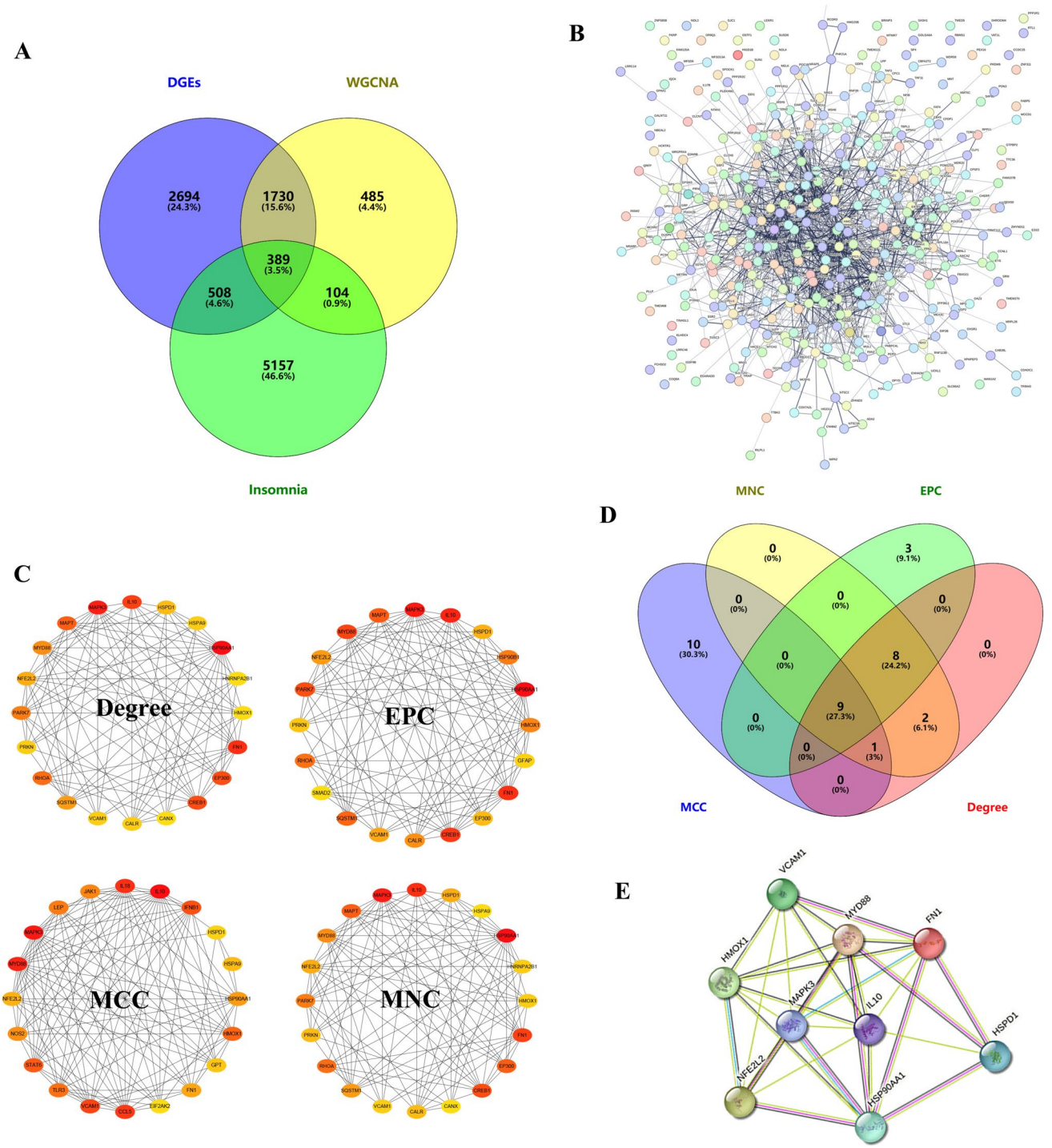



**Fig. 2** Identification of hub genes associated with insomnia. (**A**) Venn diagram showing the intersecting genes among DEGs, WGCNA module genes, and GeneCards-derived insomnia-related genes. (**B**) STRING-derived protein–protein interaction (PPI) network of the 390 intersecting genes. (**C**) Ranking of top 20 genes by four CytoHubba algorithms (MCC, MNC, EPC, Degree). (**D**) Venn diagram displaying the intersection of the four algorithm-specific top 20 gene sets. (**E**) Network view highlighting the nine selected hub genes within the PPI network

refined core genes were determined by integrating model performance, cross-model recurrence, and SHAP-based interpretability, rather than relying on a single classifier. At the gene level, FN1, HMOX1, HSP90AA1, IL10, MYD88, and NFE2L2 showed relatively high selection frequencies across machine-learning models, whereas HSPD1, MAPK3, and VCAM1 were retained less frequently (Fig. 4C). SHAP analysis of the representative glmBoost + RF model showed

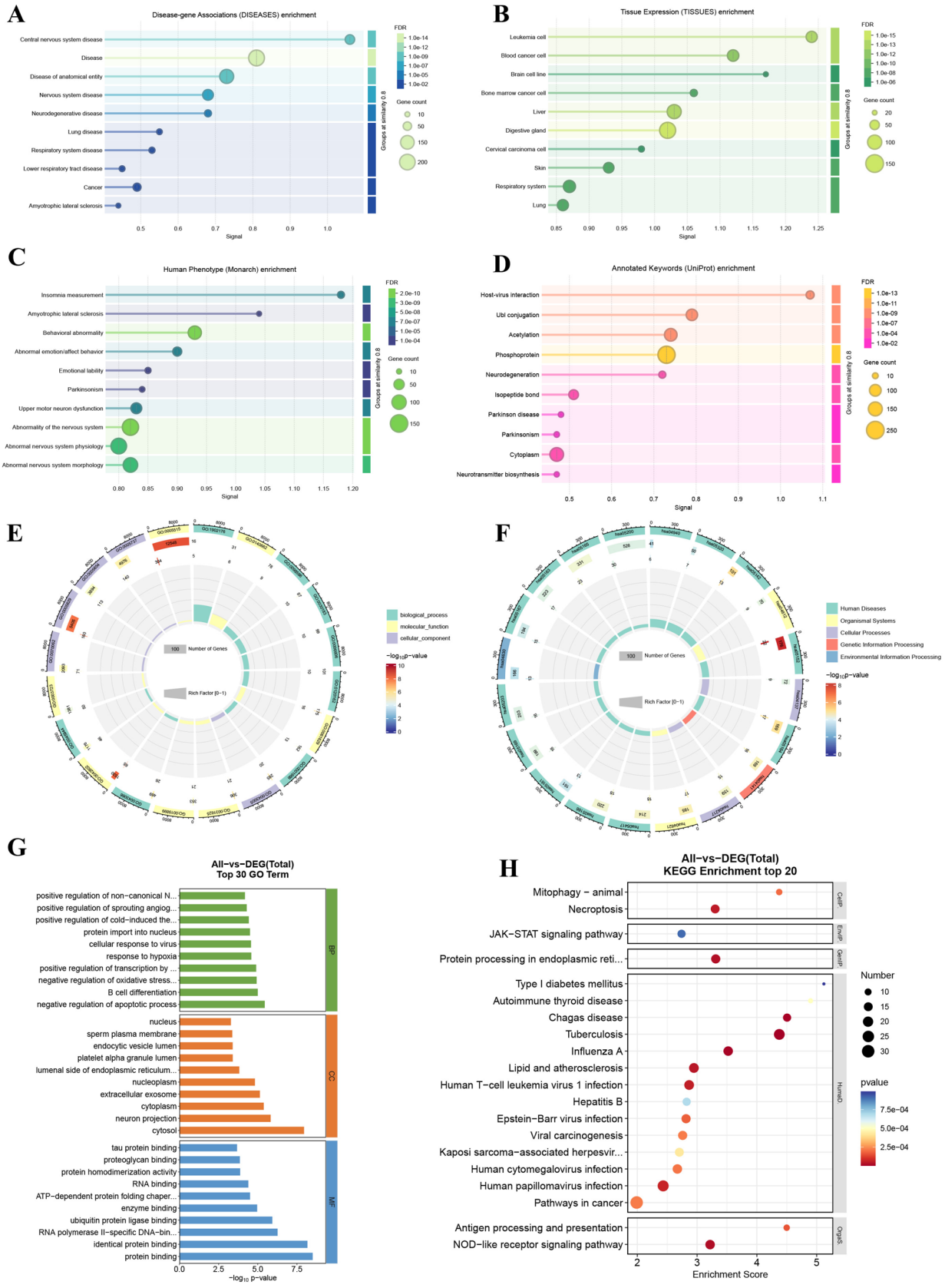

A
Disease-gene Associations (DISEASES) enrichment
Central nervous system disease
Disease
Disease of anatomical entity
Nervous system disease
Neurodegenerative disease
Lung disease
Respiratory system disease
Lower respiratory tract disease
Cancer
Amyotrophic lateral sclerosis
Signal
FDR
Gene count
Groups at similarity 0.8
B
Tissue Expression (TISSUES) enrichment
Leukemia cell
Blood cancer cell
Brain cell line
Bone marrow cancer cell
Liver
Digestive gland
Cervical carcinoma cell
Skin
Respiratory system
Lung
Signal
C
Human Phenotype (Monarch) enrichment
Insomnia measurement
Amyotrophic lateral sclerosis
Behavioral abnormality
Abnormal emotion/affect behavior
Emotional lability
Parkinsonism
Upper motor neuron dysfunction
Abnormality of the nervous system
Abnormal nervous system physiology
Abnormal nervous system morphology
Signal
D
Annotated Keywords (UniProt) enrichment
Host-virus interaction
Ubl conjugation
Acetylation
Phosphoprotein
Neurodegeneration
Isopeptide bond
Parkinson disease
Parkinsonism
Cytoplasm
Neurotransmitter biosynthesis
Signal
E
biological_process
molecular_component
Number of Genes
Rich Factor [0–1]
F
Human Diseases
Organismal Systems
Cellular Processes
Genetic Information Processing
Environmental Information Processing
Number of Genes
Rich Factor [0–1]
G
All−vs−DEG(Total)
Top 30 GO Term
positive regulation of non−canonical N...
positive regulation of sprouting angiog...
positive regulation of cold−induced the...
protein import into nucleus
cellular response to virus
response to hypoxia
positive regulation of transcription by ...
negative regulation of oxidative stress...
B cell differentiation
negative regulation of apoptotic process
nucleus
sperm plasma membrane
endocytic vesicle lumen
platelet alpha granule lumen
lumenal side of endoplasmic reticulum...
nucleoplasm
extracellular exosome
cytoplasm
neuron projection
cytosol
tau protein binding
proteoglycan binding
protein homodimerization activity
RNA binding
ATP−dependent protein folding chaper...
enzyme binding
ubiquitin protein ligase binding
RNA polymerase II−specific DNA−bin...
identical protein binding
protein binding
BP
CC
MF
−log10 p−value
H
All−vs−DEG(Total)
KEGG Enrichment top 20
Mitophagy − animal
Necroptosis
JAK−STAT signaling pathway
Protein processing in endoplasmic reti...
Type I diabetes mellitus
Autoimmune thyroid disease
Chagas disease
Tuberculosis
Influenza A
Lipid and atherosclerosis
Human T−cell leukemia virus 1 infection
Hepatitis B
Epstein−Barr virus infection
Viral carcinogenesis
Kaposi sarcoma−associated herpesvir...
Human cytomegalovirus infection
Human papillomavirus infection
Pathways in cancer
Antigen processing and presentation
NOD−like receptor signaling pathway
CellP
EnvIP
GenIP
HumD
OrgS
Number
pvalue
Enrichment Score

**Fig. 3** Functional enrichment analysis of the 390 intersection genes. (**A**) Bar plot of disease–gene association analysis. (**B**) Tissue expression profiling across major human tissues and cell lines. (**C**) Human phenotype enrichment analysis. (**D**) UniProt annotated keyword enrichment analysis. (**E**) GO circular plot displaying enriched Gene Ontology terms. (**F**) KEGG circular plot displaying enriched KEGG pathways. (**G**) GO chord plot illustrating representative GO terms and their associated genes. (**H**) KEGG chord plot illustrating representative KEGG pathways and their associated genes

that MYD88 made the strongest positive contribution in the displayed prediction, while FN1, HMOX1, HSP90AA1, IL10, and NFE2L2 also showed interpretable contributions in the SHAP-based analyses. Based on this integrated evaluation, FN1, HMOX1, HSP90AA1, IL10, MYD88, and NFE2L2 were retained as the refined core genes (Fig. 4D–E; Supplementary Figure S1A–C). These genes are broadly involved in extracellular matrix regulation, innate immune signaling, oxidative stress responses, protein homeostasis, and immunomodulation, supporting their potential relevance to insomnia-related immune dysregulation. Single gene GSEA suggested that these genes were mainly associated with immune activation and stress response pathways (Fig. 4F). Consistently, expression analysis in the discovery cohort GSE208668 showed that HMOX1, HSP90AA1, IL10, MYD88, and NFE2L2 were significantly upregulated in the insomnia group, whereas FN1 showed a downward trend without reaching statistical significance (Fig. 4G; Supplementary Excel S1). Together, these findings support a refined six-gene signature associated with peripheral immune dysregulation in insomnia.

## Peripheral immune remodeling links the refined core gene signature to specific immune cell populations and enhanced intercellular communication across insomnia- and sleep-loss-related conditions

To define the immune context of the refined core genes, we first assessed peripheral immune cell composition using immune deconvolution analysis. After excluding unstable cell populations, 16 immune cell types were retained for downstream analysis (Fig. 5A; Supplementary Excel S2). Overall, both groups were dominated by multiple T cell subsets, together with moderate proportions of plasma cells and resting dendritic cells. Correlation analysis revealed distinct relationships between the refined core genes and specific immune cell populations (Fig. 5B), with FN1 showing an opposite correlation pattern to HMOX1, HSP90AA1, MYD88, and NFE2L2, whereas IL10 showed no significant associations. Group-wise comparisons further suggested selective remodeling of the peripheral immune landscape in the insomnia cohort, characterized by increased resting CD4 memory T cells and M2 macrophages, together with reduced activated NK cells (Fig. 5C-D). Immune cell correlation analysis also revealed coordinated and antagonistic relationships among several immune subsets (Fig. 5E), indicating that insomnia is associated with altered peripheral immune organization rather than uniform immune activation.

To further examine whether these peripheral signals could also be observed at single-cell resolution under a related sleep-loss condition, we next analyzed the peripheral blood single-cell transcriptomic dataset GSE213496. Because cells from different samples were already well mixed before batch correction and showed no obvious sample-specific segregation on the UMAP embedding, with only minimal changes observed after correction, batch correction was not applied in downstream analyses to avoid overcorrection and preserve potential biological differences (Fig. 6A; Supplementary Figure S1D-H). Clustering analysis identified 11 immune cell populations, including activated B cells, neutrophils, $CD4^+$ helper T cells, tissue-resident $CD8^+$ T cells, resident macrophages, NK cells, cDCs, inflammatory macrophages, naïve $CD8^+$ T cells, follicular B cells, and naïve $CD4^+$ T cells (Fig. 6B-C), and marker gene analysis confirmed their transcriptional identities (Fig. 6D-E). Compared with controls, the sleep deprivation group showed changes in multiple immune cell populations, including relatively higher proportions of resident macrophages, tissue-resident $CD8^+$ T cells, neutrophils, naïve $CD4^+$ T cells, inflammatory macrophages, follicular B cells, and $CD4^+$ helper T cells, together with reduced cDCs (Fig. 6F). However, these compositional differences exhibited considerable variability across individual samples, suggesting that they may reflect a combination of biological heterogeneity and condition-associated effects rather than uniform group-level shifts. Among the six refined core genes, five, FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2, showed clear cell type-specific expression patterns (Fig. 6G). Because IL10 was not consistently detected in the single-cell dataset, the remaining five genes—FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2—were hereafter defined as the core hub genes for downstream single-cell analyses. FN1 was mainly localized to resident macrophages, whereas the other four genes were preferentially expressed in macrophage, neutrophil, cDC, and selected lymphocyte-related populations. Group-split UMAP and DotPlot visualizations further suggested that HMOX1, HSP90AA1, MYD88, and NFE2L2 displayed a broader distribution across neutrophils and inflammatory macrophages in the sleep deprivation-related condition, whereas cDCs showed an opposite distribution pattern (Fig. 7A-B). To determine whether these peripheral immune changes were accompanied by altered intercellular signaling, we next performed CellChat analysis. The global communication network identified activated B cells, tissue-resident $CD8^+$ T cells, resident macrophages,

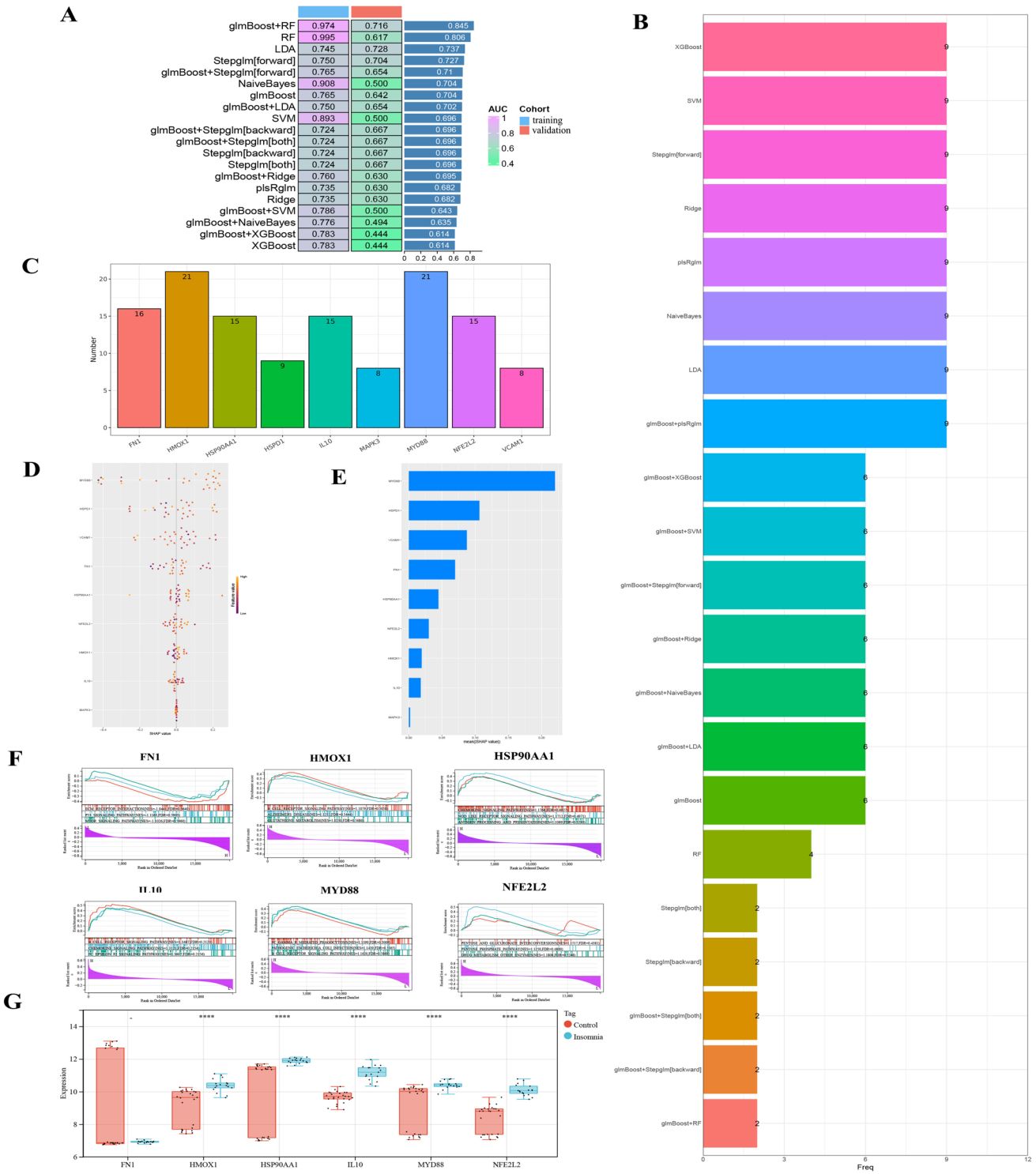

A
glmBoost+RF 0.974 0.716 0.845
RF 0.995 0.617 0.806
LDA 0.745 0.728 0.737
Stepglm[forward] 0.750 0.704 0.727
glmBoost+Stepglm[forward] 0.765 0.654 0.71
NaiveBayes 0.908 0.500 0.704
glmBoost 0.765 0.642 0.704
glmBoost+LDA 0.750 0.654 0.702
SVM 0.893 0.500 0.696
glmBoost+Stepglm[backward] 0.724 0.667 0.696
glmBoost+Stepglm[both] 0.724 0.667 0.696
Stepglm[backward] 0.724 0.667 0.696
Stepglm[both] 0.724 0.667 0.696
glmBoost+Ridge 0.760 0.630 0.695
plsRglm 0.735 0.630 0.682
Ridge 0.735 0.630 0.682
glmBoost+SVM 0.786 0.500 0.643
glmBoost+NaiveBayes 0.776 0.494 0.635
glmBoost+XGBoost 0.783 0.444 0.614
XGBoost 0.783 0.444 0.614
0 0.2 0.4 0.6 0.8
AUC
1
0.8
0.6
0.4
Cohort
training
validation
B
XGBoost 9
SVM 9
Stepglm[forward] 9
Ridge 9
plsRglm 9
NaiveBayes 9
LDA 9
glmBoost+plsRglm 9
glmBoost+XGBoost 6
glmBoost+SVM 6
glmBoost+Stepglm[forward] 6
glmBoost+Ridge 6
glmBoost+NaiveBayes 6
glmBoost+LDA 6
glmBoost 6
RF 4
Stepglm[both] 2
Stepglm[backward] 2
glmBoost+Stepglm[both] 2
glmBoost+Stepglm[backward] 2
glmBoost+RF 2
0 3 6 9 12
Freq
C
Number
0 5 10 15 20
16 21 15 9 15 8 21 15 8
FN1 HMOX1 HSP90AA1 HSPD1 IL10 MAPK3 MYD88 NFE2L2 VCAM1
D
SHAP value
Feature value
High
Low
E
mean(|SHAP value|)
F
FN1
HMOX1
HSP90AA1
IL10
MYD88
NFE2L2
Enrichment score
Ranked list metric
Rank in Ordered DataSet
G
Expression
6 8 10 12 14
****
FN1 HMOX1 HSP90AA1 IL10 MYD88 NFE2L2
Tag
Control
Insomnia

**Fig. 4** Machine learning- and SHAP-based identification and functional characterization of refined core genes in peripheral blood datasets relevant to insomnia. (**A**) ROC curves of 20 machine learning model combinations, shown by AUC values and mean AUC. glmBoost+RF achieved the highest overall predictive performance and was selected as the representative model for SHAP interpretation. (**B**) Recurrence frequency of different machine-learning models or model combinations in the screening framework, showing that the top-performing model was not necessarily the most recurrent one. (**C**) Selection frequency of candidate genes across machine-learning models, used to prioritize reproducible biomarkers. (**D**) Representative SHAP local explanation plot for the selected model, illustrating the contribution of individual genes to the predicted outcome. (**E**) SHAP summary plots showing the overall contribution and ranking of candidate genes in the representative model. MYD88 showed the strongest overall contribution. Additional SHAP dependence plots are shown in Supplementary Figure S1A–C. (**F**) Single-gene GSEA plots for the six core hub genes based on the GSE208668 dataset. (**G**) Expression levels of the refined core genes in the Control and Insomnia groups. Statistical significance was assessed using the Wilcoxon rank-sum test, with $P<0.05$ considered statistically significant. Data are presented as boxplots with overlaid individual data points, **** $P<0.0001$

and neutrophils as the major communication hubs (Fig. 7C). Analysis of the strongest ligand-receptor interactions highlighted multiple inflammatory and antigen presentation-related pathways, including Thy1-Adgre5, Sirp-Cd47, Selplg-Sell, Ppbp-Cxcr2, and MHC I/CD8-associated interactions (Fig. 7D). Macrophage-centered analyses further showed prominent signaling between resident or inflammatory macrophages and neutrophils, activated B cells, and tissue-resident $CD8^+$ T cells (Fig. 7E-F; Supplementary Figure S2A-L; Supplementary Excel S3). Compared with controls, the sleep deprivation group showed both a higher number and greater strength of inferred ligand-receptor interactions, resulting in a denser communication network (Fig. 7G-H). Sleep deprivation predominantly strengthened immune cell communication, with neutrophils emerging as a major communication hub (Fig. 7I-J). In contrast, CytoTRACE2 analysis showed substantial overlap in differentiation-related scores between groups (Fig. 7K-L; Supplementary Figure S3A-C), suggesting that the major peripheral single-cell differences under the sleep deprivation-related condition are more consistent with altered cell composition, cell state, and communication activity than with large shifts in developmental hierarchy.

## Brain single-cell analysis in a mouse sleep deprivation model reveals modest central support for the core hub gene signature and limited neuroimmune remodeling

To assess whether the core hub genes also showed relevant patterns in the central compartment, we analyzed the brain tissue single-cell transcriptomic dataset GSE137665. Because cells clustered primarily by sample rather than by cell type in the uncorrected PCA and UMAP spaces (Supplementary Figure S3D-E), while showing broadly similar cell cycle distributions across samples(Supplementary Figure S4A-B), the observed separation was considered mainly attributable to technical batch effects, and Harmony was therefore applied for batch correction. After batch correction, cells from different samples were well mixed on the UMAP embedding, suggesting effective integration and minimal batch effects (Fig. 8A), and 11 CNS cell populations were identified, including brain endothelial cells, pituitary endocrine cells, excitatory neurons, microglia, astrocytes, vascular mural cells, ependymal cells, choroid plexus stromal fibroblasts, border-associated macrophages, choroid plexus epithelial cells, and fenestrated endothelial cells (Fig. 8B-E). Group-wise comparison of cellular composition suggested changes in several brain cell populations, particularly in barrier-associated cell types, including enrichment of ependymal cells and modest increases in excitatory neurons, brain endothelial cells, and choroid plexus stromal fibroblasts (Fig. 8F). Notably, these compositional patterns also showed substantial variability across individual samples, indicating that inter-sample heterogeneity may contribute to the observed differences.

The five core hub genes showed distinct but heterogeneous cell type-specific expression patterns in brain tissue. FN1 was mainly enriched in endothelial and choroid plexus-related compartments, HSP90AA1 and NFE2L2 showed broader expression across CNS cell types, whereas HMOX1 and MYD88 remained relatively weak overall (Fig. 8G). Group-wise visualization showed that most core hub genes did not exhibit obvious between-group differences across brain cell populations (Fig. 9A-B). These findings indicate that central alterations were comparatively modest under the current sleep deprivation-related condition. To further characterize the CNS microenvironment under the analyzed sleep deprivation condition, we next examined intracerebral cell–cell communication. The CellChat network identified brain endothelial cells and excitatory neurons as major communication hubs, with dense interactions involving astrocytes, microglia, vascular mural cells, and choroid plexus-associated cells (Fig. 9C), thereby outlining the dominant intracerebral communication architecture under the analyzed condition. The ligand–receptor matrix further highlighted prominent communication axes involving glutamatergic signaling, Apoe/App-related signaling, and multiple growth factor-associated pathways, suggesting that sleep deprivation reshaped CNS intercellular communication through coordinated alterations in neuronal, glial, and vascular signaling networks (Fig. 9D–E).

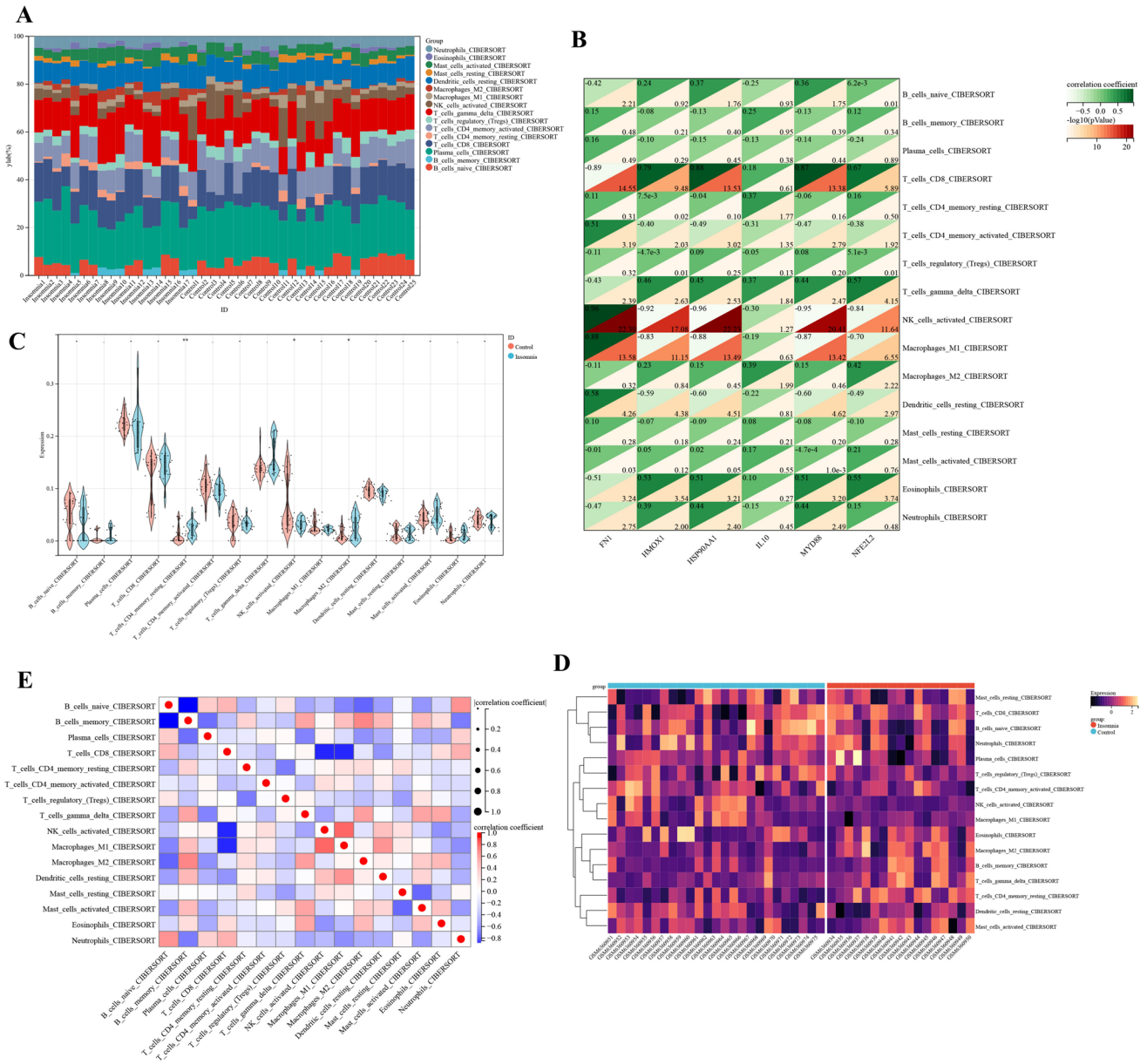



**Fig. 5** Peripheral immune infiltration landscape and immune cell interaction patterns in insomnia. (**A**) Relative proportions of the 16 retained peripheral immune cell types in individual samples from the control and insomnia groups, as estimated by immune deconvolution after exclusion of unstable cell populations. (**B**) Spearman correlation analysis between the proportions of retained immune cell types and the expression levels of the refined core genes. The upper triangle indicates correlation coefficients, and the lower triangle indicates significance levels as −log10(P value). Correlations with $|r| \geq 0.6$ and $P < 0.001$ were considered significant. (**C**) Comparison of immune cell proportions between the control and insomnia groups. Differences were assessed using the Wilcoxon rank-sum test. Significance is indicated as ns, not significant; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$. (**D**) Heatmap summarizing group-wise distribution patterns of peripheral immune cell composition. (E) Correlation matrix depicting pairwise associations among peripheral immune cell subsets

Together, these findings suggest that CNS alterations under the current sleep deprivation condition were modest, with changes involving glutamatergic signaling, Apoe/App-associated pathways, and growth factor-related interactions, and with little evidence of large-scale shifts in overall cell abundance.

## Drug prediction and structural analyses prioritize quercetin and resveratrol for further evaluation

Drug prediction analyses were performed independently using the refined core gene signature and the core hub gene signature. As shown in Supplementary Figure S4C-D,

candidate compounds displayed clear shifts in distance score distributions relative to the reference background, indicating potential relevance to the transcriptional features identified in this study. The full ranked results are provided in Supplementary Excel S3. Among the predicted compounds, quercetin and resveratrol were consistently identified in both models and showed strong statistical support. Specifically, in the refined core gene-based analysis, quercetin had a distance score of 1.0797 ($P = 2.76 \times 10^{-18}$, FDR = $1.50 \times 10^{-14}$), whereas resveratrol had a distance score of 1.3077 ($P = 3.33 \times 10^{-10}$, FDR = $1.80 \times 10^{-6}$). In the core hub gene-based analysis, quercetin yielded a distance score of 1.1197 ($P = 7.15 \times 10^{-17}$, FDR = $3.88 \times 10^{-13}$), and resveratrol yielded a distance score of 1.3846 ($P = 2.15 \times 10^{-8}$, FDR = $1.16 \times 10^{-4}$) (Supplementary Excel S3). These results prioritized quercetin and resveratrol as two shared candidate compounds for subsequent structural analyses.

Molecular docking was then performed to evaluate the binding potential of quercetin and resveratrol with the five core hub proteins, FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2. The overall docking results are summarized in Fig. 10A, and representative docking conformations are shown in Fig. 10B-C. As shown in Fig. 10D-M, both compounds exhibited favorable binding energies across all five targets. For quercetin, the binding energies with FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2 were −7.7, −7.9, −9.0, −7.5, and −10.0 kcal/mol, respectively. For resveratrol, the corresponding values were −7.2, −7.2, −7.5, −6.5, and −8.5 kcal/mol. Notably, both compounds showed relatively stronger predicted binding to NFE2L2 and HSP90AA1, with quercetin exhibiting the lowest docking scores among all tested interactions. These findings support quercetin and resveratrol as candidate compounds with potential relevance to oxidative stress- and immune inflammation-related pathways implicated in insomnia-related immune dysregulation.

MD simulations were performed for the NFE2L2-quercetin, NFE2L2-resveratrol, HSP90AA1-quercetin, and HSP90AA1-resveratrol complexes. RMSD analysis showed that all four systems reached dynamic equilibrium during the simulation and remained stable thereafter, with only moderate fluctuations. The Rg and SASA profiles were also stable across the trajectories, indicating that ligand binding did not markedly alter global protein compactness or solvent exposure. In addition, all complexes maintained persistent hydrogen bond interactions throughout the simulation, and RMSF analysis showed limited residue-level fluctuations (Fig. 11A-E). Together, these findings indicate that both quercetin and resveratrol can form relatively stable complexes with NFE2L2 and HSP90AA1, providing computational support for the docking results.

Although quercetin showed overall stronger docking scores, resveratrol was also consistently prioritized by both drug prediction models and demonstrated stable target engagement in the subsequent structural analyses, thereby justifying its selection for downstream experimental validation.

### Resveratrol modulated LPS-induced alterations in MYD88, HMOX1, and HSP90AA1 expression in THP-1 macrophages

CCK-8 assay showed that resveratrol caused a mild dose-dependent reduction in cell viability. Although slight decreases were observed at 5, 10, and 25 μM, cell viability remained at a relatively high level, whereas 50 and 100 μM induced more pronounced cytotoxicity. Therefore, 25 μM was selected as the working concentration for subsequent experiments because it maintained high cell viability without overt cytotoxicity (Fig. 12A). In the LPS-induced THP-1 macrophage model, Western blot analysis showed that MYD88 and HSP90AA1 (HSP90α) protein expression was significantly upregulated in the model group compared with the control group, whereas resveratrol treatment markedly attenuated these increases. In contrast, HMOX1 (HO-1) protein expression was elevated in the model group and was further enhanced after resveratrol intervention (Fig. 12B-C). Consistent with the protein results, qRT-PCR analysis showed that MYD88 and HSP90AA1 mRNA levels were upregulated after LPS stimulation and downregulated following resveratrol treatment, whereas HMOX1 mRNA expression was increased in the model group and further enhanced by resveratrol (Fig. 12D). These findings suggest that resveratrol modulates LPS-induced inflammatory and oxidative stress-related responses in THP-1 macrophages, partly consistent with the predicted regulation of the inflammation-oxidative stress axis.

## Discussion

Insomnia is increasingly recognized as an immune dysregulation-related disorder that is not confined to the central nervous system (Garbarino et al. 2021; Liu et al. 2021). By integrating peripheral blood bulk transcriptomic data, peripheral and brain single-cell transcriptomic datasets, candidate compound prediction, structural analyses, and macrophage-based experimental validation, the present study suggests that insomnia-related alterations are more consistent with a peripheral-centered pattern of immune disequilibrium. Overall, the most prominent and robust abnormalities were observed at the peripheral immune

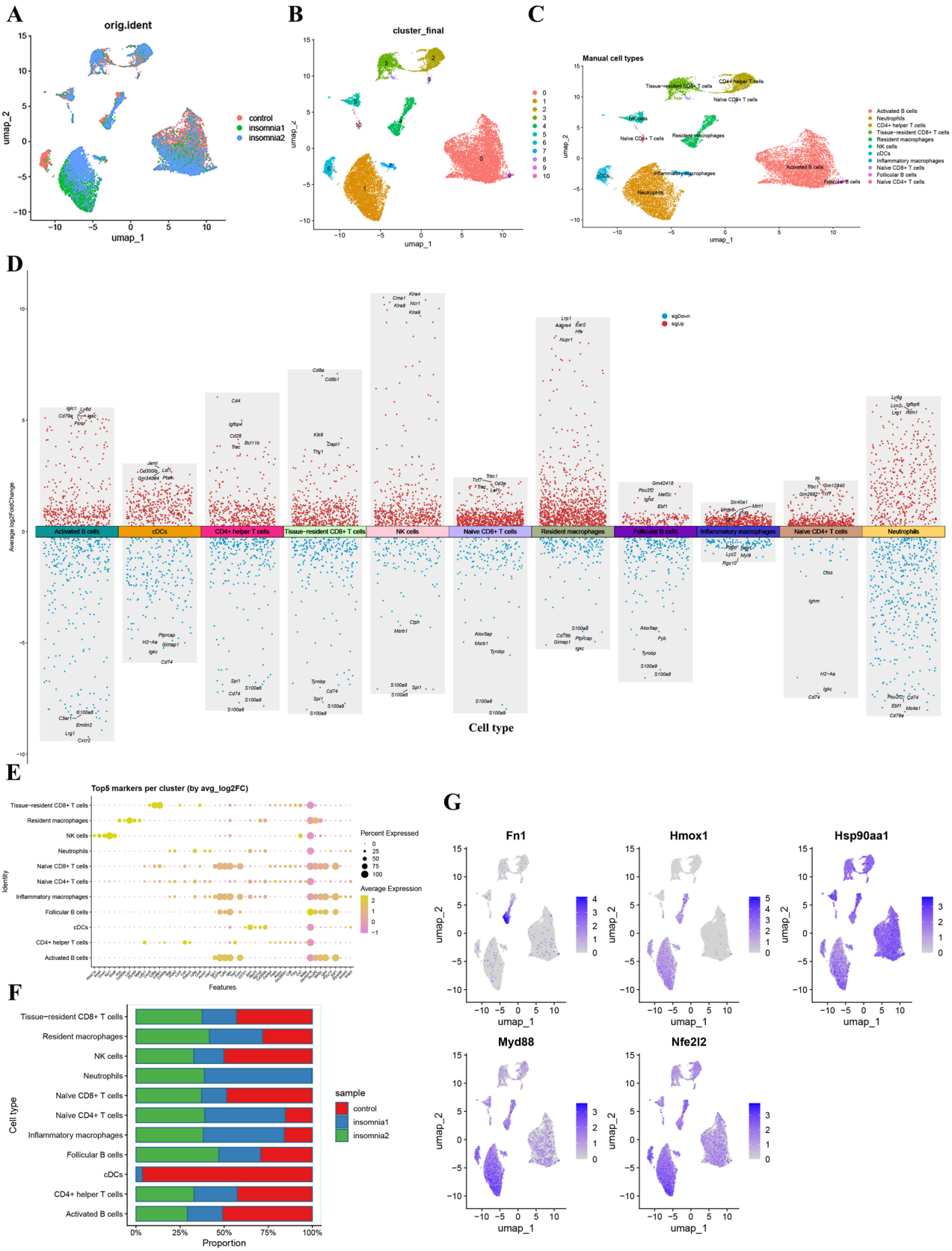

A
orig.ident
control
insomnia1
insomnia2
umap_1
umap_2
B
cluster_final
C
Manual cell types
Activated B cells
Neutrophils
CD4+ helper T cells
Tissue−resident CD8+ T cells
Resident macrophages
NK cells
cDCs
Inflammatory macrophages
Naïve CD8+ T cells
Follicular B cells
Naïve CD4+ T cells
D
sigDown
sigUp
Average log2FoldChange
Activated B cells
cDCs
CD4+ helper T cells
Tissue−resident CD8+ T cells
NK cells
Naïve CD8+ T cells
Resident macrophages
Follicular B cells
Inflammatory macrophages
Naïve CD4+ T cells
Neutrophils
Cell type
E
Top5 markers per cluster (by avg_log2FC)
Identity
Features
Percent Expressed
Average Expression
F
Cell type
Proportion
sample
control
insomnia1
insomnia2
G
Fn1
Hmox1
Hsp90aa1
Myd88
Nfe2l2

**Fig. 6** Single-cell profiling of immune cell composition and core hub genes expression in peripheral blood datasets relevant to insomnia. (**A**) UMAP plot showing the distribution of peripheral immune cells after quality control. (**B**) UMAP clustering of peripheral immune cells into 11 distinct clusters. (**C**) UMAP plot colored by annotated cell types, including Activated B cells, Neutrophils, $CD4^+$ helper T cells, Tissue-resident $CD8^+$ T cells, Resident macrophages, NK cells, cDCs, Inflammatory macrophages, Naïve $CD8^+$ T cells, Follicular B cells, and Naïve $CD4^+$ T cells. (**D**) Overview of cell type specific differentially expressed genes in the peripheral blood single-cell dataset. Red and blue dots indicate significantly upregulated and downregulated genes, respectively, as identified by Seurat FindAllMarkers using adjusted p-values. (**E**) The top five marker genes for each cluster supporting cell-type annotation. (**F**) Comparing the proportions of major immune cell types between Control and Insomnia groups. (**G**) Showing the cell type–specific expression patterns of five core hub genes (FN1, HMOX1, HSP90AA1, MYD88, NFE2L2)

level, whereas changes in the brain appeared comparatively modest, indicating that disruption of peripheral immune homeostasis may represent an important component of the pathophysiology of insomnia.

At the systems level, the strongest and most consistent signals identified in this study were mainly enriched in immune-, inflammatory-, and oxidative stress-related pathways, including the NF-κB, Toll-like receptor, and NOD-like receptor signaling pathways. These findings indicate that insomnia-related alterations are unlikely to be restricted to isolated abnormalities in a small number of genes, but may instead reflect broad peripheral transcriptional remodeling. Further immune deconvolution and single-cell analyses suggest that this disequilibrium should not be simply interpreted as generalized activation across all immune cell types, but is more consistent with selective remodeling of specific immune cell populations. The observed changes were mainly concentrated in macrophage-, neutrophil-, and conventional dendritic cell-related levels, together with alterations in specific T-cell states. This pattern is broadly consistent with previous studies showing that sleep loss can remodel the peripheral immune cell landscape, weaken NK cell-related features, and enhance CD4+T-/B-cell-related signals (Alhamawi et al. 2025; Liu et al. 2021). Collectively, these findings favor a model of biased immune reorganization rather than global immune activation.

At the molecular level, MYD88, HMOX1, HSP90AA1, and NFE2L2 constituted the most coherent core gene set, supporting the presence of an inflammation-oxidative stress coupling axis in insomnia. Among them, MYD88 is a key adaptor molecule that mediates innate immune activation downstream of Toll-like receptor signaling; HSP90AA1 is involved in protein homeostasis maintenance and stabilization of signaling complexes under stress conditions; and the NFE2L2-HMOX1 pathway represents a classical antioxidant defense module (Davinelli et al. 2024; Kawasaki and Kawai 2014; Moran Luengo et al. 2019). These findings suggest that the immune abnormalities associated with insomnia do not merely reflect a simple pro-inflammatory state, but more likely represent a coupled condition in which inflammatory signaling and oxidative stress responses are activated in parallel yet remain dynamically imbalanced (Davinelli et al. 2024; Garbarino et al. 2021; Kawasaki and Kawai 2014). The THP-1 cell experiments further support this interpretation: resveratrol downregulated MYD88 and HSP90AA1 while upregulating HMOX1, suggesting that it may exert a dual effect by suppressing inflammatory activation and enhancing antioxidant defense. Because the external machine-learning and single-cell datasets were derived from sleep deprivation models rather than clinically diagnosed insomnia cohorts, these signatures may reflect sleep-loss-related molecular features that partially overlap with insomnia.

Beyond alterations in cellular composition, the peripheral immune phenotype associated with insomnia was also characterized by changes in cell states and enhanced intercellular communication. Although several immune cell populations exhibited proportional changes, these differences showed substantial inter-individual heterogeneity. In contrast, CellChat analysis more consistently demonstrated a denser and stronger ligand-receptor communication network under sleep deprivation-related conditions, whereas CytoTRACE2 did not indicate obvious differences in developmental hierarchy. These findings suggest that, rather than being understood as a simple increase or decrease in the abundance of a particular leukocyte subset, insomnia may be better regarded as a state characterized by immune functional reprogramming and enhanced intercellular communication. This interpretation is also consistent with previous findings showing that prolonged sleep deprivation can induce systemic inflammatory amplification and neutrophil accumulation (Sang et al. 2023).

Notably, FN1 appeared to represent a relatively independent biological dimension. Although it was retained in the refined core gene set, it was not fully integrated into the dominant inflammation-oxidative stress axis. Its direction of change in the bulk transcriptomic analysis was not entirely consistent with that of the other core molecules, and its statistical robustness was comparatively weaker. In the single-cell analyses, FN1 was mainly distributed in peripheral resident macrophages and in endothelial/choroid plexus-related cell populations in the brain. Given its established roles in extracellular matrix organization, cell adhesion, tissue repair, and barrier-associated microenvironment remodeling, FN1 is more likely to reflect structural adaptation or tissue-interface remodeling rather than acting as a major inflammatory driver (Dzobo and Dandara 2023; Xu et al. 2024).

The relatively modest central signals observed in this study do not indicate that the central nervous system is

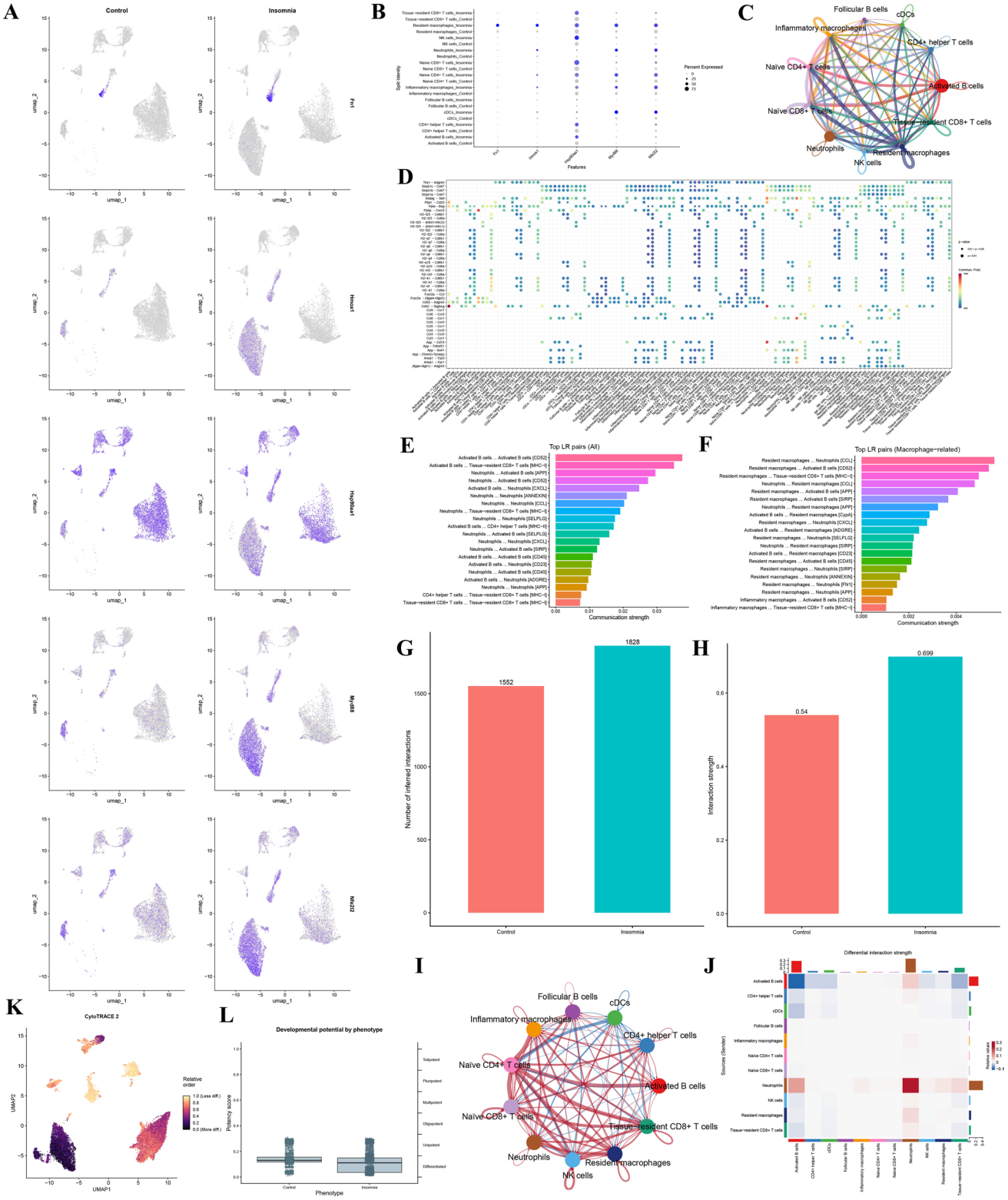

A
Control
Insomnia
Fn1
Hmox1
Hspb8as1
Myd88
Nfkbiz
umap_1
umap_2
B
Features
Split Identity
Percent Expressed
C
Follicular B cells
cDCs
Inflammatory macrophages
CD4+ helper T cells
Naïve CD4+ T cells
Activated B cells
Naïve CD8+ T cells
Tissue−resident CD8+ T cells
Neutrophils
Resident macrophages
NK cells
D
E
Top LR pairs (All)
Communication strength
F
Top LR pairs (Macrophage−related)
Communication strength
G
Number of inferred interactions
1552
1828
Control
Insomnia
H
Interaction strength
0.54
0.699
Control
Insomnia
I
Follicular B cells
cDCs
Inflammatory macrophages
CD4+ helper T cells
Naïve CD4+ T cells
Activated B cells
Naïve CD8+ T cells
Tissue−resident CD8+ T cells
Neutrophils
Resident macrophages
NK cells
J
Differential interaction strength
Sources (Sender)
Relative values
K
CytoTRACE 2
UMAP1
UMAP2
Relative order
1.0 (Less diff.)
0.0 (More diff.)
L
Developmental potential by phenotype
Potency score
Totipotent
Pluripotent
Multipotent
Oligopotent
Unipotent
Differentiated
Control
Insomnia
Phenotype

**Fig. 7** Group-wise expression of core hub genes, immune communication network, and differentiation trajectory in peripheral blood datasets relevant to insomnia. (**A**) Split UMAP plots showing the distribution of peripheral immune cell types in Control and Insomnia groups. (**B**) Comparison of the cell type–specific distribution of the core hub genes (FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2) between the Control and Insomnia groups. (**C**) Global cell–cell communication network depicting ligand–receptor interactions among peripheral immune cell populations. (**D**) Dot plot showing the top 12 signaling pathways and their corresponding ligand–receptor interactions across immune cell types, with dot size indicating significance and color reflecting communication strength. (**E**) Top 20 ligand–receptor pairs with the highest communication strength across all cell types. (**F**) Top 20 macrophage-related ligand–receptor pairs, highlighting dominant interactions involving resident and inflammatory macrophages, mainly with neutrophils, activated B cells, and tissue-resident $CD8^+$ T cells. (**G**) Bar plot showing the total number of inferred cell–cell interactions in the Control and Insomnia groups. (**H**) Bar plot showing the overall interaction strength in the Control and Insomnia groups. (**I**) Differential cell–cell communication network comparing the Control and Insomnia groups. Red lines indicate interactions increased in the Insomnia group, whereas blue lines indicate interactions decreased in the Insomnia group. Line thickness represents the magnitude of change. Supplementary Figure S2 shows a communication network focused on macrophages. (**J**) Heatmap showing changes in interaction strength between cell-type pairs in the Insomnia group relative to the Control group. (**K**) CytoTRACE2 analysis projected onto the UMAP embedding, showing the relative developmental state of cells. (**L**) Boxplot showing the developmental potential scores of cells in the Control and Insomnia groups

uninvolved in insomnia-related pathology. Rather, they suggest that neuroimmune alterations associated with insomnia may preferentially localize to brain-border niches and neurovascular interfaces, instead of manifesting as widespread parenchymal remodeling (Alves de Lima et al. 2020a, b; Castellani et al. 2023; Cugurra et al. 2021; Rustenhoven and Kipnis 2022; Sankowski et al. 2024; Xu et al. 2024). It should also be noted that the brain single-cell dataset used in this study reflected an acute sleep deprivation state of approximately 12 h, whereas the peripheral blood dataset corresponded to 24/48 h of sleep deprivation. This difference in sampling time points suggests that insomnia-related alterations may exhibit a degree of time-dependent evolution, such that peripheral abnormalities may precede, or be more pronounced than, detectable changes at the central level (Sang et al. 2023; Vanrobaeys et al. 2023). Accordingly, our findings to some extent support a "periphery-brain border-central nervous system" pattern of alteration, in which peripheral inflammatory and oxidative stress-related signals may initially act on more vulnerable border-associated microenvironments and, under sustained conditions, subsequently extend to broader intracerebral changes.

Overall, the present study supports a peripheral-centered immune disequilibrium model of insomnia, characterized by selective remodeling of innate immune-related cell populations, coordinated abnormalities in inflammatory and oxidative stress responses, and enhanced intercellular communication. MYD88, HMOX1, HSP90AA1, and NFE2L2 define a relatively coherent core molecular program, whereas FN1 may represent another dimension related to tissue-interface remodeling. These findings provide an integrated framework for understanding insomnia from a peripheral-central neuroimmune perspective and also offer a basis for future mechanistic studies and the exploration of potential therapeutic strategies.

## Limitations and future perspectives

Several limitations of the present study should be acknowledged. First, the peripheral bulk transcriptomic data were derived from human samples, whereas both single-cell datasets were obtained from mouse sleep deprivation models. Accordingly, the inferred peripheral-central relationship should be interpreted as an integrative framework spanning analytical levels and species, rather than as strictly matched direct evidence. Second, the temporal windows differed across datasets. The brain single-cell dataset reflected approximately 12 h of acute sleep deprivation, whereas the peripheral blood dataset corresponded to 24/48 h of sleep deprivation. This discrepancy is likely to be important, as recent studies have shown that acute sleep deprivation induces highly region-specific and directionally heterogeneous transcriptional responses in the brain (Vanrobaeys et al. 2023), whereas peripheral inflammatory readouts in humans tend to become more stable only after more sustained or repeated sleep restriction (Ballesio et al. 2026). In addition, the brain dataset used in this study primarily represented signals from the brain parenchyma, which may have diluted focal alterations arising from the meninges, choroid plexus, neurovascular interfaces, or subarachnoid structures (Alves de Lima et al. 2020a, b; Cugurra et al. 2021; Rustenhoven and Kipnis 2022; Sankowski et al. 2024; Xu et al. 2024). Therefore, the relatively modest central alterations observed here likely reflect, at least in part, the combined influence of species background, model type, duration of sleep deprivation, and anatomical resolution. Another limitation that warrants explicit mention concerns IL10. Although IL10 was retained as a refined core gene

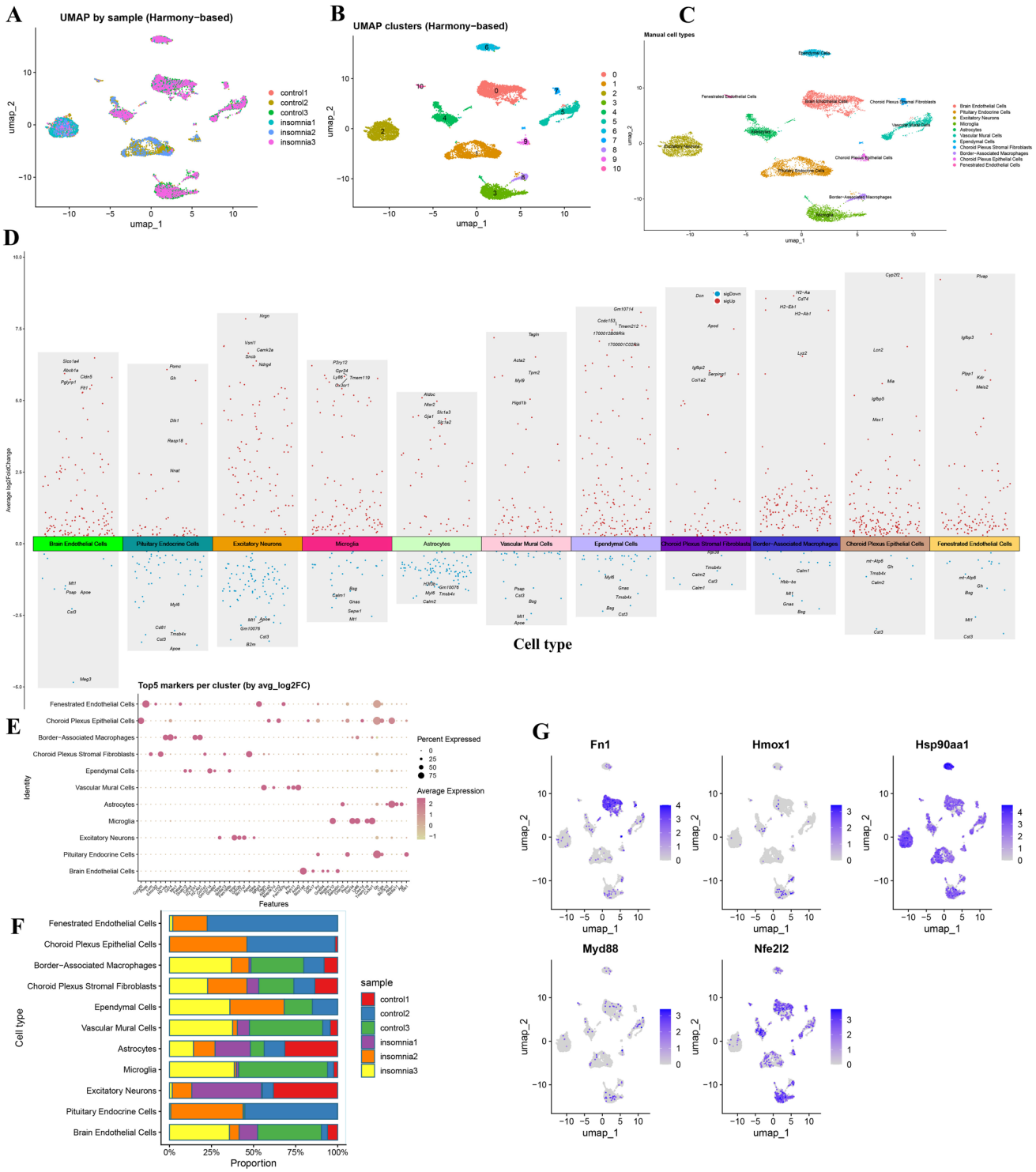



**Fig. 8** Single-cell atlas in brain tissue datasets relevant to insomnia. (**A**) Harmony-corrected UMAP plot showing mixed distribution of cells from different samples. (**B**) UMAP clustering of brain cells into 11 distinct clusters. (**C**) UMAP plot colored by annotated cell types. (**D**) Overview of cluster-specific differentially expressed genes in the brain single-cell dataset. Red and blue dots indicate significantly upregulated and downregulated genes, respectively, as identified by Seurat FindAllMarkers using adjusted p-values. (**E**) Heatmap of cluster-specific marker genes showing distinct transcriptional profiles. (**F**) Stacked bar plots comparing proportions of major brain cell types between Control and Insomnia groups. (**G**) UMAP feature plots showing cell type–specific expression patterns of the five core hub genes (FN1, HMOX1, HSP90AA1, MYD88, NFE2L2)

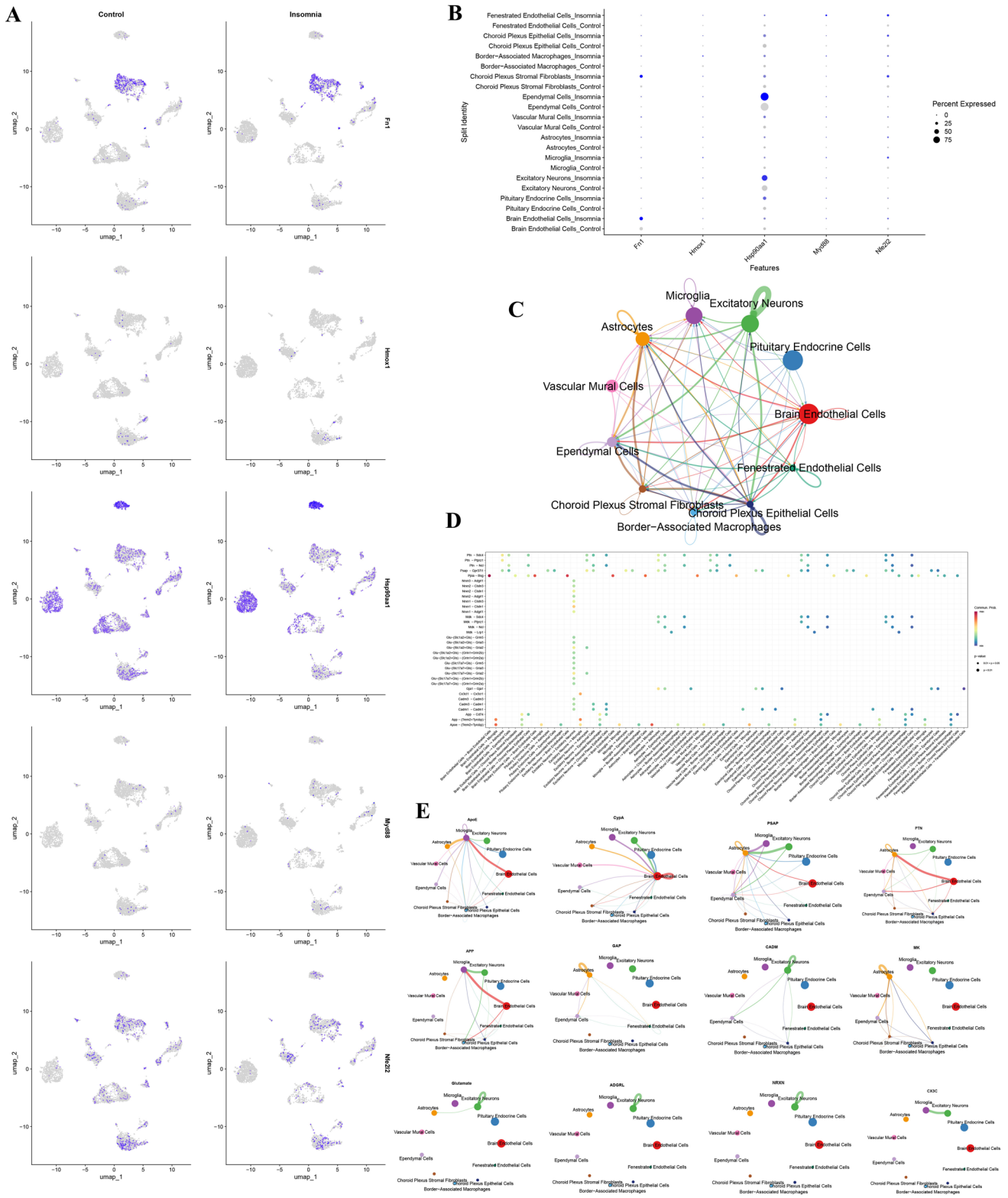



**Fig. 9** Group-wise expression of core hub genes and intracerebral cell–cell communication network in brain tissue datasets relevant to insomnia. (**A**) Split UMAP feature plots showing FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2 cell type distribution in Control and Insomnia groups. (**B**) Comparison of the cell type–specific distribution of the core hub genes (FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2) between the Control and Insomnia groups. (**C**) CellChat-derived intracerebral cell–cell communication network highlighting major sender and receiver brain cell types. (**D**) Ligand–receptor interaction matrix summarizing pathway-level communication strengths among brain cell populations. (**E**) Network visualizations of 12 representative signaling pathways (CADM, CX3C, CypA, GAP, MK, NRXN, ApoE, APP, PSAP, ADGRL, Glutamate, PTN) and their dominant sender–receiver cell pairs in the brain

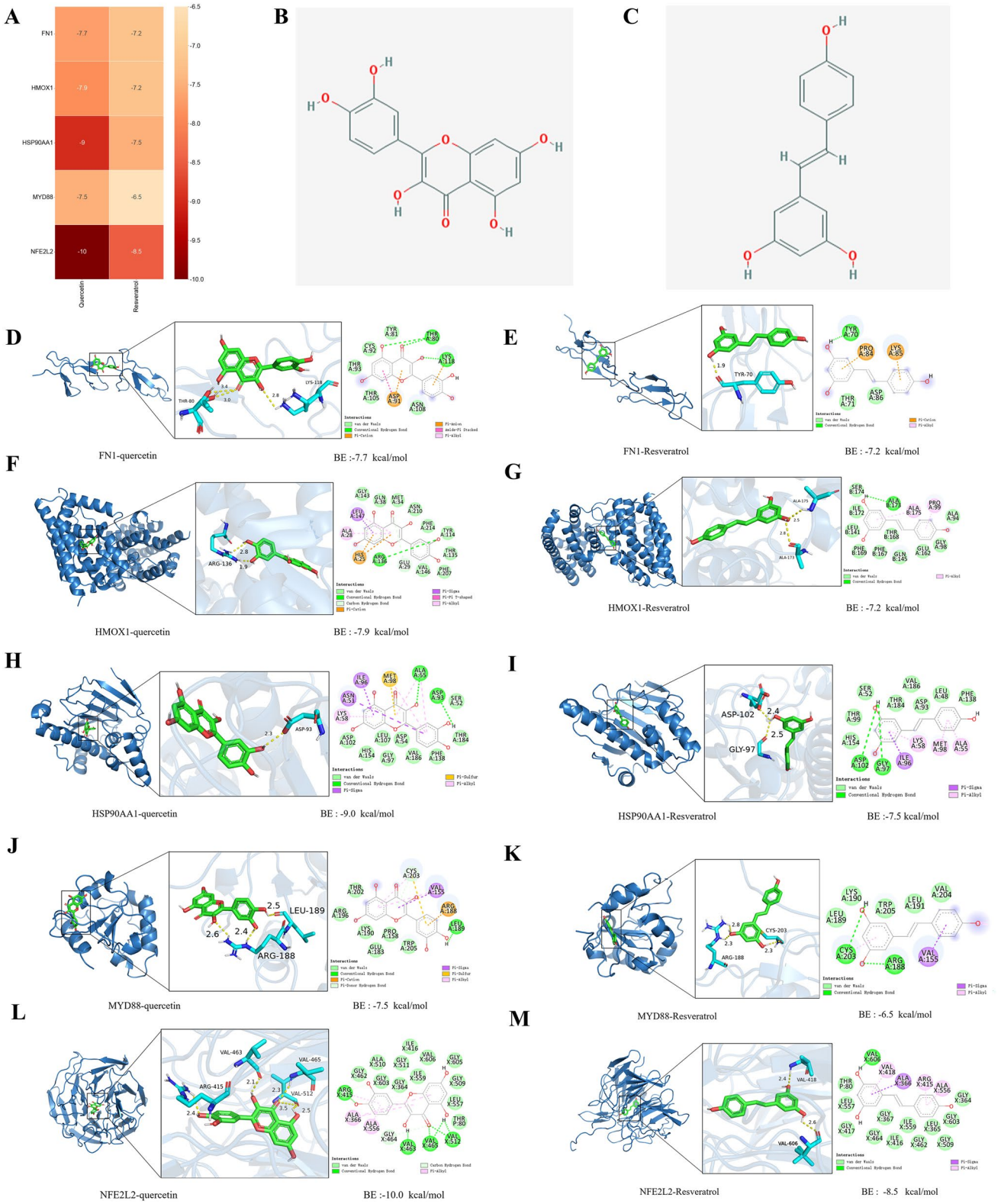



**Fig. 10** Molecular docking of quercetin and resveratrol with core hub proteins. (**A**) Summary of docking scores for quercetin and resveratrol with FN1, HMOX1, HSP90AA1, MYD88, and NFE2L2. (**B**) Chemical structure of quercetin. (**C**) Chemical structure of resveratrol. (**D**) Quercetin–FN1 binding pose. (**E**) Quercetin–HMOX1 binding pose. (**F**) Quercetin–HSP90AA1 binding pose. (**G**) Quercetin–MYD88 binding pose. (**H**) Quercetin–NFE2L2 binding pose. (**I**) Resveratrol–FN1 binding pose. (**J**) Resveratrol–HMOX1 binding pose. (**K**) Resveratrol–HSP90AA1 binding pose. (**L**) Resveratrol–MYD88 binding pose. (**M**) Resveratrol–NFE2L2 binding pose

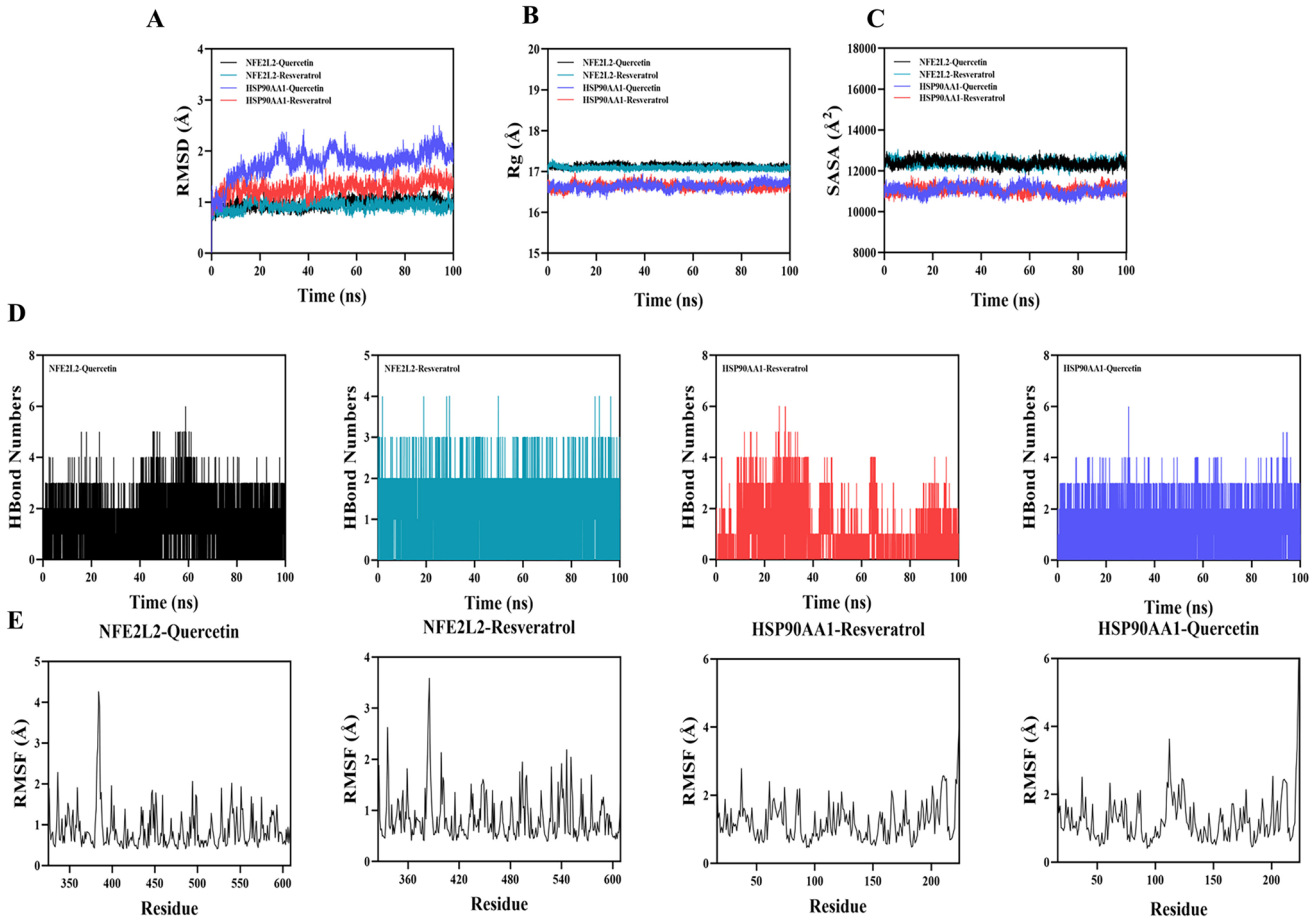



**Fig. 11** Molecular dynamics simulations of quercetin and resveratrol complexes with NFE2L2 and HSP90AA1. (**A**) Time evolution of backbone RMSD for NFE2L2–quercetin, NFE2L2–resveratrol, HSP90AA1–quercetin, and HSP90AA1–resveratrol complexes. (**B**) Radius of gyration (Rg) profiles for the four protein–ligand complexes over 100 ns simulations. (**C**) Solvent-accessible surface area (SASA) trajectories of the four complexes. (**D**) Time-dependent changes in the number of hydrogen bonds formed between proteins and ligands. (**E**) RMSF plots showing residue-level fluctuations of NFE2L2 and HSP90AA1 in complex with quercetin and resveratrol

at the bulk transcriptomic and machine-learning levels, it was not stably detected in the single-cell datasets. Given that IL10 is a low-abundance, transient, and highly state-dependent secreted cytokine, this result most likely reflects the limited sensitivity of single-cell platforms for capturing sparsely expressed transcripts, rather than indicating a lack of biological relevance (Glass et al. 2022; Qiu 2020).

Despite these limitations, the present study also suggests two potentially meaningful translational directions. First, the refined six-gene signature, together with the more condensed core hub gene set, may serve as candidate peripheral blood biomarkers for identifying an immune-dysregulated subtype of insomnia. Second, computational drug prediction, molecular docking, and molecular dynamics analyses consistently highlighted quercetin and resveratrol as candidate compounds. Although quercetin showed stronger overall docking performance, resveratrol was also robustly prioritized and demonstrated stable target engagement; moreover, its in vitro effects in THP-1 macrophages supported both anti-inflammatory and antioxidant regulatory properties. These findings suggest that targeting the inflammation-oxidative stress coupling axis may represent a feasible therapeutic strategy for insomnia-related immune disequilibrium.

Future studies would benefit from matched peripheral-central sampling designs, longer observation windows, and single-cell or spatial omics analyses with finer anatomical resolution. Particular attention should be directed toward brain-border compartments, neurovascular interfaces, and immune-cell state transitions (Alves de Lima et al. 2020a, b; Cugurra et al. 2021; Rustenhoven and Kipnis 2022; Sankowski et al. 2024; Xu et al. 2024), in order to more precisely define the temporal dynamics, key cellular interfaces, and potential therapeutic targets underlying insomnia-related immune disequilibrium.

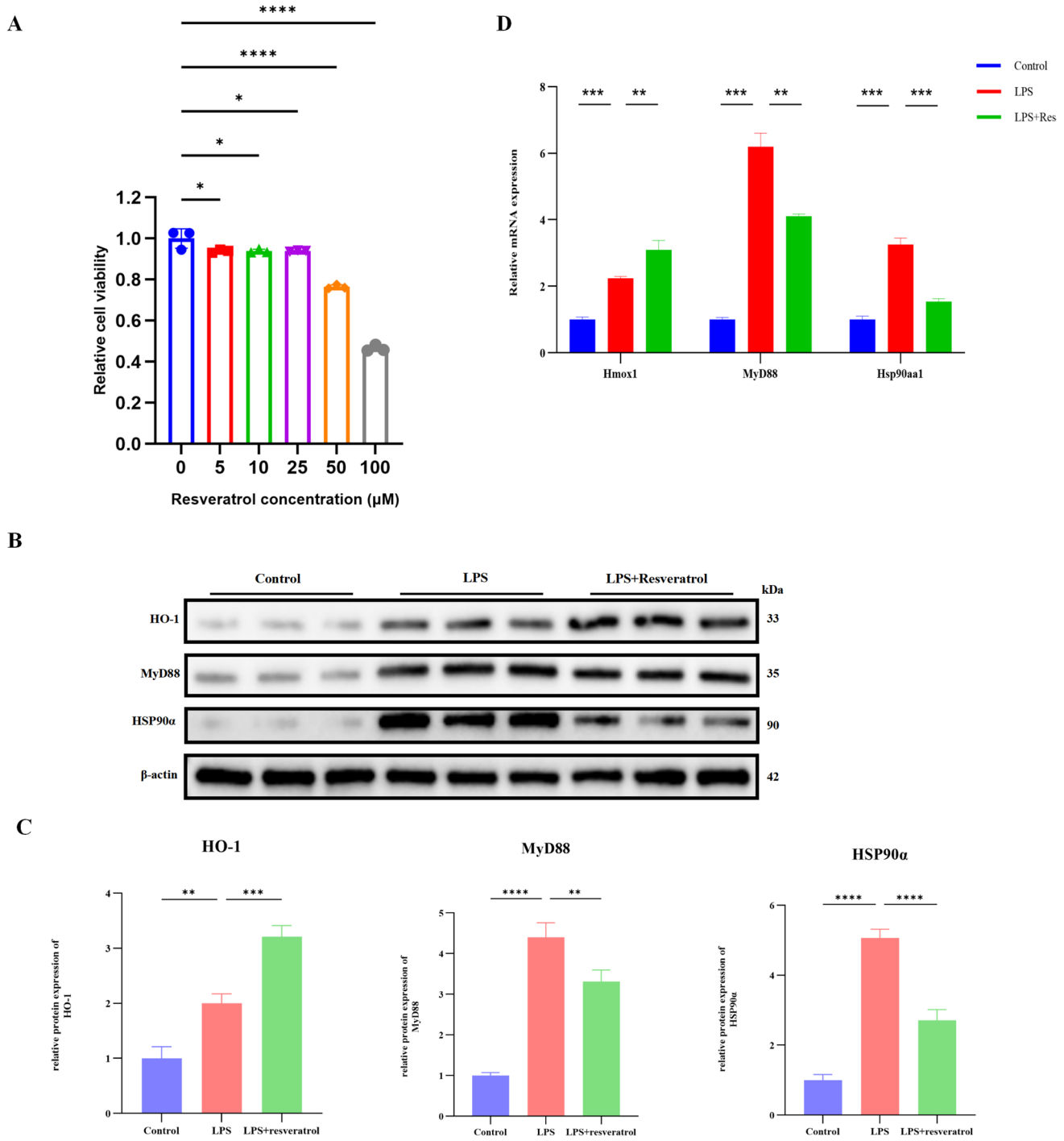



**Fig. 12** Resveratrol modulated LPS-induced alterations in MYD88, HMOX1, and HSP90AA1 expression in THP-1 macrophages. (**A**) Cell viability of THP-1 macrophages treated with different concentrations of resveratrol (0, 5, 10, 25, 50, and 100 μM), as determined by the CCK-8 assay. (**B**) Representative Western blot bands showing the protein expression of HMOX1 (HO-1), MYD88, and HSP90AA1 (HSP90α) in the Control, LPS, and LPS + resveratrol groups. (**C**) Quantitative analysis of HMOX1, MYD88, and HSP90AA1 protein expression. (**D**) qRT-PCR analysis of HMOX1, MYD88, and HSP90AA1 mRNA expression in the Control, LPS, and LPS + resveratrol groups. Data are presented as mean ± SD from at least three independent experiments. Statistical analysis was performed using one-way ANOVA followed by Tukey's post hoc test. Significance is indicated as ns, not significant; *$P<0.05$; **$P<0.01$; ***$P<0.001$; ****$P<0.0001$

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10142-026-01900-5.

**Acknowledgements** We are especially grateful to Mr. Xiang Zeng from the Second Affiliated Hospital of Guangzhou University of Chinese Medicine for his guidance. The graphical abstract was created in BioRender. Zeng, S. (2026) https://BioRender.com/v1q3mqn.

**Authors' contributions** WZ conducted all analyses and drafted the original manuscript. ZW and YG performed molecular docking and molecular dynamics simulations. BH, RC and HP contributed to figure layout, data interpretation, and academic discussion. BW provided conceptual guidance for the machine learning component of the study. ZY and FT provided overall project supervision and critical revision of the manuscript. All authors reviewed and approved the final version of the manuscript.

**Funding** This work was supported by the Science and Technology Projects of Guangzhou (2024A03J0034), the Science and Technology Research Project of Guangdong Provincial Hospital of Traditional Chinese Medicine (YN2022QN21), the Research Project of the Guangdong Provincial Department of Traditional Chinese Medicine (20261134) and the Chinese Medicine Guangdong Laboratory (Hengqin Laboratory) Reserach Program Project (HQL2024PZ034).

**Data availability** The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethics approval** Not applicable.

**Clinical trial number** Not applicable.

**Competing interest** The authors declare no competing interests.



## References

Alhamawi RM, Halawani FA, Hakeem SF, Alslimi HA, Alhamawi EM, Aljohani AM, Mohammed IN, Zahid HM, Almutawif YA (2025) Insomnia and anxiety: exploring their hidden effect on natural killer cells among young female adults. Front Immunol 16:1698155. https://doi.org/10.3389/fimmu.2025.1698155

Alves de Lima K, Rustenhoven J, Da Mesquita S, Wall M, Salvador AF, Smirnov I, Martelossi Cebinelli G, Mamuladze T, Baker W, Papadopoulos Z, Lopes MB, Cao WS, Xie XS, Herz J, Kipnis J (2020a) Meningeal gammadelta T cells regulate anxiety-like behavior via IL-17a signaling in neurons. Nat Immunol 21:1421–1429. https://doi.org/10.1038/s41590-020-0776-4

Alves de Lima K, Rustenhoven J, Kipnis J (2020b) Meningeal immunity and its function in maintenance of the central nervous system in health and disease. Annu Rev Immunol 38:597–620. https://doi.org/10.1146/annurev-immunol-102319-103410

Ballesio A, Fiori V, Lombardo C (2026) Effects of experimental sleep deprivation on peripheral inflammation: an updated meta-analysis of human studies. J Sleep Res 35:e70099. https://doi.org/10.1111/jsr.70099

Benjafield AV, Sert Kuniyoshi FH, Malhotra A, Martin JL, Morin CM, Maurer LF, Cistulli PA, Pepin JL, Wickwire EM (2025) Estimation of the global prevalence and burden of insomnia: a systematic literature review-based analysis. Sleep Med Rev 82:102121. https://doi.org/10.1016/j.smrv.2025.102121

Boahen A, Hu D, Adams MJ, Nicholls PK, Greene WK, Ma B (2023) Bidirectional crosstalk between the peripheral nervous system and lymphoid tissues/organs. Front Immunol 14:1254054. https://doi.org/10.3389/fimmu.2023.1254054

Castellani G, Croese T, Peralta Ramos JM, Schwartz M (2023) Transforming the understanding of brain immunity. Science 380:eabo7649. https://doi.org/10.1126/science.abo7649

Cugurra A, Mamuladze T, Rustenhoven J, Dykstra T, Beroshvili G, Greenberg ZJ, Baker W, Papadopoulos Z, Drieu A, Blackburn S, Kanamori M, Brioschi S, Herz J, Schuettpelz LG, Colonna M, Smirnov I, Kipnis J (2021) Skull and vertebral bone marrow are myeloid cell reservoirs for the meninges and CNS parenchyma. Science 373. https://doi.org/10.1126/science.abf7844

Davinelli S, Medoro A, Savino R, Scapagnini G (2024) Sleep and oxidative stress: current perspectives on the role of NRF2. Cell Mol Neurobiol 44:52. https://doi.org/10.1007/s10571-024-01487-0

De Crescenzo F, D'Alo GL, Ostinelli EG, Ciabattini M, Di Franco V, Watanabe N, Kurtulmus A, Tomlinson A, Mitrova Z, Foti F, Del Giovane C, Quested DJ, Cowen PJ, Barbui C, Amato L, Efthimiou O, Cipriani A (2022) Comparative effects of pharmacological interventions for the acute and long-term management of insomnia disorder in adults: a systematic review and network meta-analysis. Lancet 400:170–184. https://doi.org/10.1016/s0140-6736(22)00878-9

Dzobo K, Dandara C (2023) The extracellular matrix: its composition, function, remodeling, and role in Tumorigenesis. Biomimetics (Basel) 8. https://doi.org/10.3390/biomimetics8020146

Garbarino S, Lanteri P, Bragazzi NL, Magnavita N, Scoditti E (2021) Role of sleep deprivation in immune-related disease risk and outcomes. Commun Biol 4:1304. https://doi.org/10.1038/s42003-021-02825-4

Glass MC, Glass DR, Oliveria JP, Mbiribindi B, Esquivel CO, Krams SM, Bendall SC, Martinez OM (2022) Human IL-10-producing B cells have diverse states that are induced from multiple B cell subsets. Cell Rep 39:110728. https://doi.org/10.1016/j.celrep.2022.110728

Herz J, Fu Z, Kim K, Dykstra T, Wall M, Li H, Salvador AF, Zou B, Yan N, Blackburn SM, Andrews PH, Goldman DH, Papadopoulos Z, Smirnov I, Xie XS, Kipnis J (2021) GABAergic neuronal IL-4R mediates T cell effect on memory. Neuron 109:3609-3618.e3609. https://doi.org/10.1016/j.neuron.2021.10.022

Jha PK, Valekunja UK, Ray S, Nollet M, Reddy AB (2022) Single-cell transcriptomics and cell-specific proteomics reveals molecular signatures of sleep. Commun Biol 5:846. https://doi.org/10.1038/s42003-022-03800-3

Kawasaki T, Kawai T (2014) Toll-like receptor signaling pathways. Front Immunol 5:461. https://doi.org/10.3389/fimmu.2014.00461

Khan MS, Aouad R (2022) The effects of insomnia and sleep loss on cardiovascular disease. Sleep Med Clin 17:193–203. https://doi.org/10.1016/j.jsmc.2022.02.008

Li S, Olde Heuvel F, Rehman R, Aousji O, Froehlich A, Li Z, Jark R, Zhang W, Conquest A, Woelfle S, Schoen M, Cc OM, Reinhardt RL, Voehringer D, Kassubek J, Ludolph A, Huber-Lang M, Knoll B, Morganti-Kossmann MC, Brockmann MM, Boeckers T, Roselli F (2023) Interleukin-13 and its receptor are synaptic proteins involved in plasticity and neuroprotection. Nat Commun 14:200. https://doi.org/10.1038/s41467-023-35806-8

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP (2011) Molecular signatures database (MSigDB) 3.0. Bioinformatics 27:1739–1740. https://doi.org/10.1093/bioinformatics/btr260

Liu X, Chen B, Huang Z, Duan R, Li H, Xie L, Wang R, Li Z, Gao Y, Zheng Y, Su W (2021) Effects of poor sleep on the immune cell landscape as assessed by single-cell analysis. Commun Biol 4:1325. https://doi.org/10.1038/s42003-021-02859-8

Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, Wang L, Lu T, Zhang Y, Sun Z, Han X (2022) Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. Nat Commun 13:816. https://doi.org/10.1038/s41467-022-28421-6

Moran Luengo T, Mayer MP, Rudiger SGD (2019) The Hsp70-Hsp90 chaperone cascade in protein folding. Trends Cell Biol 29:164–177. https://doi.org/10.1016/j.tcb.2018.10.004

Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA (2015) Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12:453–457. https://doi.org/10.1038/nmeth.3337

Parhizkar S, Gent G, Chen Y, Rensing N, Gratuze M, Strout G, Sviben S, Tycksen E, Zhang Q, Gilmore PE, Sprung R, Malone J, Chen W, Remolina Serrano J, Bao X, Lee C, Wang C, Landsness E, Fitzpatrick J, Wong M, Townsend R, Colonna M, Schmidt RE, Holtzman DM (2023) Sleep deprivation exacerbates microglial reactivity and Aβ deposition in a TREM2-dependent manner in mice. Sci Transl Med 15:eade6285. https://doi.org/10.1126/scitranslmed.ade6285

Peng Y, Yuan M, Xin J, Liu X, Wang J (2020) Screening novel drug candidates for Alzheimer's disease by an integrated network and transcriptome analysis. Bioinformatics 36:4626–4632. https://doi.org/10.1093/bioinformatics/btaa563

Qiu P (2020) Embracing the dropouts in single-cell RNA-seq analysis. Nat Commun 11:1169. https://doi.org/10.1038/s41467-020-14976-9

Riemann D, Benz F, Dressle RJ, Espie CA, Johann AF, Blanken TF, Leerssen J, Wassing R, Henry AL, Kyle SD, Spiegelhalder K, Van Someren EJW (2022) Insomnia disorder: state of the science and challenges for the future. J Sleep Res 31:e13604. https://doi.org/10.1111/jsr.13604

Riemann D, Espie CA, Altena E, Arnardottir ES, Baglioni C, Bassetti CLA, Bastien C, Berzina N, Bjorvatn B, Dikeos D, Dolenc Groselj L, Ellis JG, Garcia-Borreguero D, Geoffroy PA, Gjerstad M, Goncalves M, Hertenstein E, Hoedlmoser K, Hion T, Holzinger B, Janku K, Jansson-Frojmark M, Jarnefelt H, Jernelov S, Jennum PJ, Khachatryan S, Krone L, Kyle SD, Lancee J, Leger D, Lupusor A, Marques DR, Nissen C, Palagini L, Paunio T, Perogamvros L, Pevernagie D, Schabus M, Shochat T, Szentkiralyi A, Van Someren E, van Straten A, Wichniak A, Verbraecken J, Spiegelhalder K (2023) The European Insomnia Guideline: an update on the diagnosis and treatment of insomnia 2023. J Sleep Res 32:e14035. https://doi.org/10.1111/jsr.14035

Rustenhoven J, Kipnis J (2022) Brain borders at the central stage of neuroimmunology. Nature 612:417–429. https://doi.org/10.1038/s41586-022-05474-7

Sang D, Lin K, Yang Y, Ran G, Li B, Chen C, Li Q, Ma Y, Lu L, Cui XY, Liu Z, Lv SQ, Luo M, Liu Q, Li Y, Zhang EE (2023) Prolonged sleep deprivation induces a cytokine-storm-like syndrome in mammals. Cell 186:5500-5516.e5521. https://doi.org/10.1016/j.cell.2023.10.025

Sankowski R, Suss P, Benkendorff A, Bottcher C, Fernandez-Zapata C, Chhatbar C, Cahueau J, Monaco G, Gasull AD, Khavaran A, Grauvogel J, Scheiwe C, Shah MJ, Heiland DH, Schnell O, Markfeld-Erol F, Kunze M, Zeiser R, Priller J, Prinz M (2024) Multiomic spatial landscape of innate immune cells at human central nervous system borders. Nat Med 30:186–198. https://doi.org/10.1038/s41591-023-02673-1

Schafer ST, Mansour AA, Schlachetzki JCM, Pena M, Ghassemzadeh S, Mitchell L, Mar A, Quang D, Stumpf S, Ortiz IS, Lana AJ, Baek C, Zaghal R, Glass CK, Nimmerjahn A, Gage FH (2023) An in vivo neuroimmune organoid model to study human microglia phenotypes. Cell 186:2111-2126.e2120. https://doi.org/10.1016/j.cell.2023.04.022

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550. https://doi.org/10.1073/pnas.0506580102

Tang N, Zeng Y, He G, Chen S (2025) Interference between immune cells and insomnia: a bibliometric analysis from 2000 to 2023. Front Neurol 16:1486548. https://doi.org/10.3389/fneur.2025.1486548

Vanrobaeys Y, Peterson ZJ, Walsh EN, Chatterjee S, Lin LC, Lyons LC, Nickl-Jockschat T, Abel T (2023) Spatial transcriptomics reveals unique gene expression changes in different brain regions after sleep deprivation. Nat Commun 14:7095. https://doi.org/10.1038/s41467-023-42751-z

Wu TT, Zou YL, Xu KD, Jiang XR, Zhou MM, Zhang SB, Song CH (2023) Insomnia and multiple health outcomes: umbrella review of meta-analyses of prospective cohort studies. Public Health 215:66–74. https://doi.org/10.1016/j.puhe.2022.11.021

Xu H, Lotfy P, Gelb S, Pragana A, Hehnly C, Byer LIJ, Shipley FB, Zawadzki ME, Cui J, Deng L, Taylor M, Webb M, Lidov HGW, Andermann ML, Chiu IM, Ordovas-Montanes J, Lehtinen MK (2024) The choroid plexus synergizes with immune cells during neuroinflammation. Cell 187:4946-4963.e4917. https://doi.org/10.1016/j.cell.2024.07.002

Zhang W, Jiang J, Xu Z, Yan H, Tang B, Liu C, Chen C, Meng Q (2023a) Microglia-containing human brain organoids for the study of brain development and pathology. Mol Psychiatry 28:96–107. https://doi.org/10.1038/s41380-022-01892-1

Zhang Y, Zhao W, Liu K, Chen Z, Fei Q, Ahmad N, Yi M (2023b) The causal associations of altered inflammatory proteins with sleep duration, insomnia and daytime sleepiness. Sleep 46. https://doi.org/10.1093/sleep/zsad207

Zhao Y, Wu X, Li X, Jiang LL, Gui X, Liu Y, Sun Y, Zhu B, Pina-Crespo JC, Zhang M, Zhang N, Chen X, Bu G, An Z, Huang TY, Xu H (2018) TREM2 is a receptor for beta-amyloid that mediates microglial function. Neuron 97:1023-1031.e1027. https://doi.org/10.1016/j.neuron.2018.01.031