Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identifying core genes via plasma metabolite mediation during the progression from bacterial pneumonia to secondary bacterial meningitis

Hexiang Jiang

jianghx91@zafu.edu.cn

Zhejiang A&F University https://orcid.org/0009-0005-5342-0257

Xinyu Cheng

Zhejiang A & F University

Jingyan Fan

Zhejiang A & F University

Houhui Song

Zhejiang A & F University

Yang Yang

Zhejiang A & F University

Research Article

Keywords: Bacterial pneumonia; Bacterial meningitis; Lung-brain axis; Metabolic genes; Mendelian randomization; multi-omics

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11817044/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Neurological damage secondary to acute pneumonia from diverse pathogens implies a shared underlying pathological mechanism, which remains poorly defined. In this study, we established a "bacterial pneumonia (BP)–metabolite–bacterial meningitis (BM)" mediation pathway and identified seven core bridging metabolites via Mendelian randomization. Subsequent gene mapping of metabolite-associated single nucleotide polymorphisms, integrated with differential gene expression analysis, weighted gene co-expression network analysis, and receiver operating characteristic curve analysis, pinpointed five BM-associated genes: *OLAH*, *LTB4R*, *PYGL*, *PPP1R3B*, and *MMP25*. Immune infiltration analysis and single-cell RNA sequencing revealed that *PYGL*, *PPP1R3B*, and *MMP25* are predominantly expressed in neutrophils, with their expression levels strongly correlated with the extent of neutrophil infiltration. During BM progression, these genes were markedly upregulated in the early phase alongside neutrophil differentiation and downregulated during recovery; this dynamic pattern was consistent in both cerebrospinal fluid and peripheral blood. Upstream regulatory analysis showed their corresponding transcription factors were also initially upregulated. We further constructed competing endogenous RNA networks, and used network pharmacology and molecular docking to screen potential gene-targeting compounds. Collectively, our findings delineate a metabolite-gene mediated pathway linking BP to BM, providing a theoretical basis for lung–brain axis-targeted interventions.

Introduction

Keywords: Bacterial pneumonia; Bacterial meningitis; Lung-brain axis; Metabolic genes; Mendelian randomization; multi-omics

Bacterial meningitis (BM), a life-threatening central nervous system (CNS) infection, causes approximately 318, 000 global deaths yearly and disproportionately threatens infants, the elderly, and immunocompromised individuals (van de Beek et al., 2021). Notably, zoonotic BM carries risks of animal-to-human transmission (van Samkar et al., 2016; Ma et al., 2025). Current prevention and therapeutic strategies based on vaccination and antibiotics face growing challenges from bacterial resistance and emerging serotypes, underscoring the urgent need for new therapeutics (van de Beek et al., 2021).

Conserved pathogenic pathways have inspired broad-spectrum drug discovery across multiple diseases (Langenberg et al., 2023; Grotzinger et al., 2025; You et al., 2025). Clinically, bacterial pneumonia (BP) and BM frequently co-occur: 17% of BM patients present with concurrent BP (Figueiredo et al., 2020). Neurological sequelae after severe pneumonia further suggest overlapping pathological processes linking respiratory infection to cerebral injury (Douaud et al., 2022; Ma et al., 2023). A landmark 2022 study described the “lung-brain axis” and verified that lipopolysaccharide (LPS) signaling enables pulmonary microbiota to shape cerebral immunity (Hosang et al., 2022). In LPS-triggered severe acute BP, disruption of both the pulmonary capillary barrier and blood–brain barrier (BBB) facilitates endogenous bacterial translocation from the lung to the brain, explaining lung-brain-axis-dependent neurological injury after severe pneumonia (Ma et al., 2023). Still, LPS is exclusive to Gram-negative

bacteria, and the conserved pathways driving BM caused by diverse invasive respiratory pathogens remain poorly defined.

Progress in computational and genomic tools has enabled the generation and public availability of large genetic datasets across diseases. Coupled with refined causal-inference frameworks, these resources have advanced molecular multimorbidity research (Langenberg et al., 2023). Here, we leverage the lung-brain axis to systematically explore its role as a common pathway linking BP and BM using genetic and multi-omics approaches. We built a Mendelian randomization (MR)-based BP–metabolite–BM mediation model and identified seven metabolites with putative mediating effects within this axis. SNP loci for these core metabolites were mapped to neighboring genes. We then integrated differential gene expression (DGE), weighted gene co-expression network analysis (WGCNA), and receiver operating characteristic (ROC) analysis, with validation across independent clinical datasets, to uncover five core lung-brain-axis genes associated with BM. Finally, network pharmacology and molecular docking were applied to screen small-molecule compounds targeting these genes. Collectively, this study illuminates conserved pathogenic pathways underlying pathogen-induced BM and provides solid evidence for precision interventions targeting the lung-brain axis.

Methods

Data Download and Processing

The datasets GSE40586, GSE163194, GSE163195, GSE163196, and GSE261482 were downloaded from the GEO database(Edgar et al., 2002). The GSE40586 dataset (platform GPL6244) contained peripheral blood transcriptomic data from 21 BM patients and 18 healthy controls, which were utilized for DGE analysis, WGCNA, and immune infiltration analysis (Lill et al., 2013; Jiang et al., 2025). The GSE163194 dataset comprised single-cell RNA sequencing (scRNA-seq) data of cerebrospinal fluid (CSF) from BM patients at different disease stages and was used for single-cell and pseudotime analyses. The GSE163195 and GSE163196 datasets (both based on platform GPL20795) contained distinct transcriptomic profiles: GSE163195 included data from the CSF of 12 pediatric BM patients across different phases; GSE163196 comprised data from peripheral blood leukocytes of 7 controls, 6 septic pediatric patients with BM in the acute phase, and 6 in the recovery phase. These datasets were employed to validate the expression of key genes. Sample annotation details for S1 to S9 in datasets GSE163194, GSE163195, and GSE163196 are provided in references (Xiao et al., 2022; Jiang et al., 2025). The GSE261482 dataset contains a total of 177 samples, including 38 whole blood samples from healthy individuals and 139 whole blood samples from children with pneumonia. The standardized gene expression matrix was downloaded to verify the expression levels of core genes in the whole blood of pneumonia patients.

Mendelian Randomization Analysis

Genome-wide association study (GWAS) summary statistics for BM and BP were obtained from the FinnGen R12 database (Kurki et al., 2023). The BM dataset (finngen_R12_G6_MENINGBACT) included 1,258 cases and 497,087 controls (total n = 498,345). The BP dataset (J10_PNEUMOBACT) included 21,582 cases and 415,538 controls (total n = 437,120). GWAS data for plasma metabolites were sourced from the GWAS Catalog(Cerezo et al., 2025), encompassing summary statistics for 1,400 metabolites (accession IDs: GCST90199621–GCST90201020)(Shin et al., 2014; Yin et al., 2022).

Using these GWAS datasets for BP, plasma metabolites, and BM, we constructed a "BP–metabolite–BM" mediation model. First, an MR analysis was performed with plasma metabolites as the exposure and BM as the outcome to identify metabolites showing significant causal associations with BM. Subsequently, MR analysis was conducted with BP as the exposure and the identified metabolites as outcomes to determine which metabolites were causally influenced by BP. This process systematically evaluated whether BP influences BM risk by regulating these metabolites.

All analytical data were based on European populations. The instrumental variables met the following criteria: $P < 1\times10^{-5}$, $R^2 < 0.001$, and distance > 10,000 kb. The harmonise_data() function in the TwoSampleMR package (version 0.5.7) of R was used to align the SNP effects between exposure and outcome datasets (Hemani et al., 2018). MR analysis was then performed using the mr() function, which incorporates five complementary methods: MR Egger, weighted median, inverse variance weighted (IVW), simple mode, and weighted mode. The IVW method served as the primary reference for causal inference. Horizontal pleiotropy and heterogeneity were assessed using the mr_pleiotropy_test() and mr_heterogeneity() functions, respectively. A *P*-value > 0.05 was considered to indicate no significant evidence for pleiotropy or heterogeneity (Hemani et al., 2018). Directionality of the causal relationship was evaluated via the directionality_test() function. A result of correct_causal_direction = TRUE supported the hypothesized direction. If correct_causal_direction = FALSE, the exposure and outcome variables were swapped, and the MR analysis was repeated to determine the correct causal direction.

Differential Gene Expression Analysis and Screening

DGE analysis of the microarray dataset GSE40586 was conducted using the limma package (version 3.58.1) (Ritchie et al., 2015). Genes with a false discovery rate (FDR) < 0.05 and an absolute $\log_2$ fold change ($|\log_2 FC|$) > 0.5 were considered as differentially expressed. DGE boxplots were generated using the ggpubr package (version 0.6.0) and further visualized with the ggplot2 package (version 3.5.0).

Weighted Gene Co-expression Network Analysis

A weighted gene co-expression network was constructed using the R package WGCNA (version 1.72-5) (Langfelder and Horvath, 2008). First, sample differences and the presence of outliers were examined by constructing a sample clustering heatmap. Subsequently, to ensure the co-expression network conformed to the scale-free network property, the optimal soft-thresholding power (β) was selected.

Next, the expression matrix was converted into an adjacency matrix, which was further transformed into a topological overlap matrix (TOM). Based on the TOM, hierarchical clustering of genes was performed, and modules were identified using the dynamic tree cutting algorithm. The minimum module size was set to 200 genes, and module clustering was conducted with MEDissThres = 0.2 to merge similar modules. Finally, BM occurrence was defined as a trait, and Pearson correlation coefficients between each module and BM were calculated. Module genes that were significantly correlated with BM ($P < 0.05$) and exhibited the strongest correlation were selected. These selected genes, termed WGCNA_GENEs, were subjected to further analysis.

Gene Mapping of Key Metabolite-associated SNPs

First, instrumental SNPs corresponding to the metabolites identified in the mediation MR analysis were extracted. Subsequently, all genes within ± 250 kb of each SNP were obtained using the R package vautils (version 0.1.0). Based on this mapping, candidate genes potentially associated with BM were defined as MR_GENEs for subsequent analysis.

ROC Analysis

ROC analysis of the core genes was performed using the R package pROC (version 1.18.4) (Robin et al., 2011). The area under the ROC curve (AUC) was calculated to evaluate diagnostic performance, where a value closer to 1.0 indicates higher accuracy. Based on conventional thresholds, an AUC > 0.9 represents high diagnostic accuracy, an AUC of 0.7–0.9 indicates moderate utility, an AUC of 0.5–0.7 suggests low accuracy, and an AUC = 0.5 denotes no discriminative value.

Immune Infiltration Analysis

A total of 28 immune cells were used to calculate the level of immune infiltration, and their relative enrichment scores were assessed via single-sample gene set enrichment analysis using the R package GSVA (version 1.50.5) (Hänzelmann et al., 2013). The gene sets for 28 types of immune cells were sourced from reference (Charoentong et al., 2017). Differential infiltration between BM patients and controls was assessed using the Wilcoxon rank-sum test ($P < 0.05$), and results were visualized with boxplots generated by the ggplot2 package (version 3.4.1) (Hänzelmann et al., 2013). The Spearman correlation between differentially infiltrating immune cells and core genes was analyzed based on the enrichment scores of these immune cells and the expression levels of the core genes, and the correlation coefficients (r values) between them were calculated and visualized as a heatmap (Jiang et al., 2025).

scRNA-seq Analysis

scRNA-seq data from single-cell samples of BM at the same disease stages were combined and analyzed. Cells with > 50 detected genes and < 40% mitochondrial gene expression were selected for subsequent analyses. We obtained single-cell transcriptomes of 2,524 cells at stage S1, 2,896 cells at stage S3, 2,246 cells at stage S4, 1,930 cells at stage S6, 3,854 cells at stage S7, and 2,537 cells at stage S8.

First, we used the R package Seurat (version 5.0.6) to perform t-SNE clustering and marker gene identification for CSF cells at each stage, and manually annotated the clustering results (Stuart et al., 2019). Then, we removed apoptotic cells and conducted integrated clustering of normal immune cells. Subsequently, the NormalizeData() function was used to normalize the data, and the FindVariableFeatures() function was applied to identify highly variable genes (HVGs) between cells (feature selection) for subsequent cell type annotation. We used default parameters, i.e., using the "vst" method to select 2,000 HVGs. Batch effect correction was then performed using canonical correlation analysis (CCA) implemented in Seurat, followed by principal component analysis (PCA) for dimensionality reduction. An elbow plot was generated to identify usable data dimensions, and principal components preceding the elbow point were selected for subsequent analyses. In this study, the top 15 principal components were chosen for further clustering.

After completing PCA dimensionality reduction, we performed unsupervised clustering analysis on the cells using the FindNeighbors() and FindClusters() functions in the Seurat package, obtaining 15 clusters at a default resolution of 1.0. Marker genes were retrieved by referencing the CellMarker database and the original literature to facilitate manual annotation of different clusters (Xiao et al., 2022). After manual annotation, a total of 8 cell subsets were identified, namely plasmacytoid dendritic cells (pDCs), B cells, myeloid dendritic cells (mDCs), natural killer cells (NKs), T cells, neutrophils, monocytes, and macrophages. Then, we generated t-SNE clustering plots to visualize the distribution of cell subsets and annotated the marker genes used. The proportion of each cell subset was calculated and visualized using bar charts generated with the ggplot2 package.

The expression distribution of core genes across different cell clusters and disease stages in the single-cell data was analyzed using the R package Seurat (version 5.0.6). Furthermore, violin plots were generated to show the expression of core genes at different stages, and intergroup statistical significance testing was performed using the ggpubr package (version 0.6.0).

Pseudotime Analysis

Pseudotime analysis of key cell types was performed using the Monocle package (version 2.32.0) in R (Qiu et al., 2017). The input data required to create a CellDataSet (cds) object via the newCellDataSet() function included: (1) the gene-cell matrix of raw count data from single-cell data processed by Seurat; (2) the meta.data from the single-cell data, which was converted into an S4 object using the new() function; (3) the data frame storing gene names, which was converted into an S4 object using the same new() function. Subsequently, the expression data within the cds object was normalized, and genes

expressed in fewer than 10 cells were excluded. For trajectory analysis, cells were ordered by pseudotime using genes that met the Monocle 2 identification thresholds: mean_expression ≥ 0.1 and dispersion_empirical ≥ 1 × dispersion_fit. Dimensionality reduction was then performed using the reduceDimension() function with the parameters "reduction_method = "DDRTree" and max_components = 2", followed by cell ordering and visualization using the plot_cell_trajectory() function.

Transcription Factor Prediction and Expression Analysis

To investigate the regulatory role of transcription factors (TFs) in neutrophil cells with high expression of key genes at the single-cell level, we constructed a regulatory network using the Python package pyscenic (version 0.12.1) based on the cells and genes after quality control (Aibar et al., 2017). First, the “GRNBoost2” algorithm was applied to infer the co-expression regulatory networks between TFs and their candidate target genes from scRNA-seq data (GSE163194); second, the DNA-binding motifs of TFs were identified using the RcisTarget database (Aibar et al., 2017), thereby screening out regulons (i.e., sets of TFs and their high-confidence target genes) with regulatory evidence; finally, the activity of these regulons was evaluated in each cell using the "AUCell" scoring method, enabling the quantification of differential expression and activity states of these regulons across different cells.

Construction of ceRNA Regulatory Network

Target miRNAs of the biomarkers were predicted using the DIANA-microT database (Maragkakis et al., 2009; Tastsoglou et al., 2023), and upstream lncRNAs were predicted via the starBase database (Li et al., 2013). Based on these predictions, the ceRNA regulatory network (lncRNA–miRNA–mRNA) was constructed and visualized using the Sankey diagram. Finally, Cytoscape was used to visualize the miRNA–mRNA–TFs molecular regulatory network (Shannon et al., 2003).

Network Pharmacology Analysis and Molecular Docking

Core gene names were used to retrieve and download related compounds from the HERB database (Fang et al., 2021). Subsequently, we constructed a disease–gene–compound network using Cytoscape software (version 3.10.3).

Relevant compound information was retrieved from the TCMSP database (Ru et al., 2014). The screening criteria for active ingredients were set as follows: oral bioavailability ≥ 30% and drug-likeness ≥ 0.18, and the effective active ingredients were finally identified. AutoDock 4 was used for molecular docking between the active compounds and core gene targets, followed by analysis of their binding affinities. A binding free energy < -5.0 kcal/mol was set as the screening threshold. The docking results were visualized with Discovery Studio. The three-dimensional structures of active ingredients were downloaded from PubChem (Kim et al., 2025), and the three-dimensional structures of target proteins were downloaded from the RCSB PDB database(Berman et al., 2000) and the GeneCards database (Stelzer et al., 2016).

Statistical Analysis

All statistical analyses were performed using R software (version 4.3.2). The Wilcoxon rank-sum test was applied to compare gene expression and immune cell infiltration between the different groups. A *P*-value < 0.05 was considered statistically significant.

Results

BP Affects BM via Core Metabolites

MR analysis was performed with plasma metabolites as exposure factors and BM as the outcome. A total of 57 metabolites showed significant causal associations with BM ($P < 0.05$), including 32 protective factors (odds ratio [OR] < 1) and 25 risk factors (OR > 1) (Fig. 1A). All results passed the heterogeneity, horizontal pleiotropy, and directional tests, thus ensuring their reliability (Suppl. Tables 1–3).

Further MR analysis was conducted with BP as the exposure and aforementioned metabolites as outcomes, which revealed that BP had significant causal relationships with 8 of these metabolites (Fig. 2B). Results of heterogeneity, horizontal pleiotropy, and directional tests (Suppl. Tables 4–6) showed that the metabolite dibutyl sulfosuccinate exhibited horizontal pleiotropy ($P < 0.05$) with a "FALSE" result in the directional test. Reverse MR analysis was performed by swapping the exposure and outcome variables, which confirmed that a reverse causal relationship existed only between BP and dibutyl sulfosuccinate, but not between BP and the remaining seven metabolites (Suppl. Figure 1).

Sensitivity analyses of each exposure factor were visualized using scatter plots (Suppl. Figures 2, 6), forest plots (Suppl. Figures 3, 7), funnel plots (Suppl. Figures 4, 8), and leave-one-out validation plots (Suppl. Figures 5, 9). In summary, BP regulates the development and progression of BM mainly by affecting 7 metabolites, including the adenosine 3',5'-cyclic monophosphate (cAMP)-to-adenosine 5'-monophosphate (AMP) ratio, bilirubin (Z, Z)-to-etiocholanolone glucuronide ratio, C12H22O4 (1) glucuronide levels, inosine 5'-monophosphate (IMP)-to-urate ratio, pyridoxate levels, sphinganine levels, and unknown metabolite X-13844 levels.

Screening and Validation of Core Genes in the Lung-Brain Axis

Based on the BM transcriptome dataset (GSE40586), we performed DGE analysis and visualized the results (Fig. 2A). In WGCNA, sample clustering revealed small differences among samples with no outliers detected (Suppl. Figure 10A). To ensure that the constructed network conformed to a scale-free topology, the optimal soft-thresholding power (β) was set at 25 (Suppl. Figure 10B). Gene clustering was conducted based on the TOM, and gene modules were identified using the dynamic tree-cutting method (Fig. 2B). A total of 32 gene modules were obtained via module eigengene clustering (Fig. 2C), and 12 modules were finally retained after merging similar modules (Figs. 2B and 2D).

After calculating the correlation between each module and disease grouping, we found that the pink module exhibited the strongest positive correlation with the disease status (correlation coefficient = 0.67, $P < 0.05$) and a negative correlation with the control status (Fig. 2D). There was a significant positive correlation between gene significance (GS) and module membership (MM) in the pink module, further indicating that the genes in this module were highly associated with BM (Fig. 2E). Finally, a Venn diagram was used to integrate DEGs, metabolite SNP-associated genes (MR_GENEs), and pink module genes (WGCNA_GENEs), leading to the identification of 27 candidate core genes of the "lung–brain axis" (Fig. 2F).

We performed validation by integrating CSF transcriptome data from patients at different disease stages (GSE163195). The results showed that the expression levels of 7 of these 27 genes were significantly upregulated in the initial phase and significantly decreased in the remission and recovery phases (Fig. 2G), while no significant differences were observed in the remaining genes (Suppl. Figure 11A). This finding was consistent with the validation results from the peripheral blood transcriptome dataset of BM patients (GSE40586) (Suppl. Figure 11B). To avoid the limitations of a single dataset, validation was further conducted using another peripheral blood transcriptome dataset of BM patients (GSE163196), which finally identified 5 core genes, namely *OLAH*, *LTB4R*, *PYGL*, *PPP1R3B*, and *MMP25*, whose expression dynamics were consistent with the CSF data (Fig. 2H). Notably, these five core genes were also significantly upregulated in whole blood specimens from patients with pneumonia (Suppl. Figure 12). ROC analysis confirmed that these genes had excellent diagnostic efficacy for BM (AUC > 0.8), suggesting their potential as early diagnostic biomarkers (Fig. 2I).

Expression of Core Genes in Immune Cells

To investigate the relationship between core genes and immune cells, immune infiltration analysis was performed based on the peripheral blood transcriptome dataset (GSE40586) of BM patients. The results showed that the infiltration of neutrophils and activated dendritic cells was significantly increased in patients (Fig. 3A). Further correlation analysis indicated that *MMP25*, *PPP1R3B*, *PYGL*, and *OLAH* were strongly positively correlated with neutrophils, while *OLAH* and *LTB4R* were also positively correlated with activated dendritic cells (Fig. 3B).

Further analysis using the CSF single-cell dataset (GSE163194) of BM patients revealed that cell clustering yielded 15 cell clusters (Fig. 3C). After annotation, 8 major cell types were identified, including pDCs, B cells, mDCs, NKs, T cells, neutrophils, monocytes, and macrophages (Fig. 3D), with the marker genes used illustrated in Fig. 3E. Cell proportion analysis demonstrated that the number of neutrophils decreased significantly during the disease's remission and recovery phases, and the number of monocytes also decreased during the recovery phase (Fig. 3F). Gene expression analysis revealed that core genes including *MMP25*, *PPP1R3B*, *PYGL*, *OLAH*, and *LTB4R* were mainly expressed in neutrophils and monocytes (Fig. 3G), and their expression was significantly downregulated during the remission and recovery phases (Fig. 3H), which was consistent with the aforementioned results. Collectively, these findings suggest that core genes are closely associated with neutrophil infiltration.

Expression of Core Genes During Neutrophil Differentiation

Pseudotime analysis of neutrophils revealed a distinct differentiation trajectory over time, which could be divided into three stages. Undifferentiated cells were predominantly concentrated in Stage 1 (Figs. 4A, B). The initial phase of meningitis was dominated by Stage 2 neutrophils, whereas the recovery phase was characterized by Stage 1 neutrophils (Fig. 4C). These findings indicate that the disease induces neutrophil differentiation, and the cells gradually reverted to their original undifferentiated state during convalescence.

Previous results suggested that *MMP25*, *PYGL*, and *PPP1R3B* were positively correlated with neutrophils. Further analysis demonstrated that as cells differentiated, the expression levels of *MMP25* and *PYGL* exhibited a marked trend of increasing first and then decreasing, while the expression variation of *PPP1R3B* was relatively moderate (Fig. 4D). Across different disease stages, neutrophils progressively returned to the undifferentiated state as the disease progressed from the initial phase to the remission phase and then to the recovery phase, and the expression levels of these three genes also basically decreased to their baseline levels (Fig. 4E).

In conclusion, neutrophils undergo differentiation and then revert to their original state during the progression from the initial phase to the recovery phase of meningitis, and the expression levels of *MMP25*, *PYGL*, and *PPP1R3B* correspondingly restored to baseline after an initial elevation.

Molecular Regulatory Network of Core Genes

Using the expression data of neutrophils, we employed pyscenic to predict the TFs of core genes. A total of 6 TFs were identified: *MMP25* was regulated by CEBPB, RELB, RUNX1, and STAT2; *PPP1R3B* was regulated by ETS2, and RELB; and *PYGL* was regulated by FOS.

Analysis of TF expression levels in neutrophils across different samples revealed that all TFs exhibited higher transcriptional activity during the initial phase of the disease (Fig. 5A). To further explore the regulatory mechanisms of core genes, we utilized the DIANA_microT database to predict miRNAs targeting these core genes, with 22 miRNAs identified in total. Subsequently, the starBase database was used to predict upstream lncRNAs, resulting in the identification of 7 lncRNAs; accordingly, a core gene-related lncRNA–miRNA–mRNA regulatory network was constructed (Fig. 5B). Finally, the relationships among mRNA, miRNA and TFs were visualized as an interaction network (Fig. 5C).

Screening of Targeted Drugs for Core Genes

Using the core genes *MMP25*, *PPP1R3B*, *PYGL*, *OLAH*, and *LTB4R* as queries, we retrieved related compounds and constructed a disease–gene–compound network (Fig. 6A). By searching the TCMSP database and conducting subsequent screening, we finally identified Quercetin/Meletin, Stepholidine, and D-coclaurine as potential active ingredients.

These small molecules were used as ligands and the core genes as receptors for molecular docking analysis. Molecular docking results with the ligand–receptor binding free energy < -5.0 kcal/mol were visualized, including *PYGL*–D-coclaurine (Fig. 6B), *OLAH*–Stepholidine (Fig. 6C), *PPP1R3B*–Stepholidine (Fig. 6D) and *PYGL*–Stepholidine (Fig. 6E). Detailed information was provided in Suppl. Table 7.

Discussion

Pneumonia is a major cause of morbidity and mortality, frequently driven by bacterial pathogens, and can reshape plasma metabolite profiles (More et al., 2022). Blood metabolomics reliably distinguishes BP from viral pneumonia (More et al., 2022; Rischke et al., 2025). Animal studies further show that exposure to diverse airborne bacteria alters metabolite profiles in plasma and bronchoalveolar lavage fluid, with detectable links between pulmonary metabolites, gut microbiota and circulating metabolites (Song et al., 2025). Collectively, these observations support a model whereby BP remodels systemic plasma metabolism. Brain metabolism relies heavily on circulating substrates: the brain takes up blood-borne metabolites and releases metabolic products into the circulation. Prior work has mapped cerebral metabolite exchange with peripheral blood, and tryptophan metabolism is known to modify outcomes in tuberculous meningitis (Ardiansyah et al., 2023; Wang et al., 2025). Together, these data raise the possibility that plasma metabolites cross into the brain and contribute to meningitis pathogenesis. Even so, causal relationships connecting BP, plasma metabolites and BM remain uncharacterized. To address this gap, we applied two-sample MR, setting BP as exposure, BM as outcome and plasma metabolites as mediators, to dissect causal pathways and mediating effects. Our MR analysis identified seven plasma metabolites that mediate the BP–BM relationship: the adenosine 3',5'-cAMP/adenosine 5'-AMP ratio, the (Z,Z)-bilirubin/etiocholanolone glucuronide ratio, C12H22O4 (1) glucuronide, the IMP/urate ratio, pyridoxate, sphinganine, and unknown metabolite X-13844.

Among these candidates, sphinganine and cAMP are functionally well characterized in infection and neurological disease. Sphinganine is required for TLR4 signaling and amplifies inflammatory responses in sepsis and melanoma models (Hering et al., 2024). Its circulating concentration declines in community-acquired pneumonia, whereas CSF sphinganine rises in patients with neurological sequelae after SARS-CoV-2 infection (Ning et al., 2018; Chen et al., 2024). Our MR results indicate BP exerts a protective effect on sphinganine levels, while sphinganine itself increases BM risk. This pattern points to sphinganine as a plausible mediator of pneumonia-associated meningitis, and the underlying mechanism awaits dedicated experimental dissection. cAMP, by contrast, regulates immune evasion during *Streptococcus pneumoniae* brain invasion, yet whether systemic cAMP levels influence BM susceptibility remains unclear (Pinho-Ribeiro et al., 2023). Pyridoxate, a vitamin B6 metabolic marker, has previously been linked to pulmonary arterial hypertension (Costeira et al., 2023). The (Z,Z)-bilirubin/etiocholanolone glucuronide ratio correlates with gut microbiota composition, immune thrombocytopenia and ischemic stroke (He et al., 2024; Hong et al., 2024; Zeng and Zhong, 2024). Functional roles for C12H22O4 (1) glucuronide and the IMP/urate ratio in pulmonary or cerebral infection have not yet been reported. Notably, metabolite X-13844 remains structurally unannotated; as a statistically supported mediator in the BP–BM cascade, it represents a high-priority target for future

structural identification and functional validation, which may uncover previously unrecognized small molecules mediators bridging pulmonary infection and meningeal injury.

Building on the MR-identified metabolite mediators, integrated SNP mapping, DGE, and WGCNA prioritized five core genes: *OLAH*, *LTB4R*, *PYGL*, *PPP1R3B*, and *MMP25*. Rather than acting in isolation, these genes converge into three functionally interrelated modules that collectively drive the transition from systemic metabolic perturbation to central nervous system pathology: an inflammation-immunity module encompassing *LTB4R* and *OLAH*, a metabolism-regulated cell death module formed by *PPP1R3B* and *PYGL*, and a barrier integrity module represented by *MMP25*.

The LTB4–LTB4R axis is a cornerstone of leukocyte recruitment and antimicrobial defense. *LTB4R* signaling drives neutrophil infiltration (Jala et al., 2017) and can amplify pathological inflammation in non-immune cells (He et al., 2020). Consistent with these reports, we observed elevated *LTB4R* expression in serum and CSF at BM onset, implicating this gene in neutrophil-dependent host defense within the central nervous compartment. Given its central role in inflammatory regulation, *LTB4R* constitutes an attractive candidate for anti-inflammatory therapeutic targeting in meningitis (Tager and Luster, 2003). *OLAH* encodes an enzyme involved in oleic acid synthesis; its expression correlates with disease severity across multiple respiratory viral infections and is upregulated in severe influenza, COVID-19, and RSV infection (Jia et al., 2024; Palanki et al., 2024). The dynamic upregulation of *OLAH* in serum and CSF during early BM, followed by downregulation during recovery, supports its participation in inflammatory cascades. Since oleic acid serves as a key membrane lipid, we hypothesize that *OLAH* may modulate lipid raft organization and thereby tune *LTB4R* signaling or the activity of other immune receptors. Collectively, these two genes likely synergize to amplify neuroinflammation: *LTB4R* directs leukocyte trafficking, while *OLAH*-dependent lipid metabolism provides structural and signaling substrates for sustained immune activation.

For the metabolism-regulated cell death module, *PPP1R3B* orchestrates glycolipid metabolic reprogramming and immune dysregulation via p-STAT signaling to promote macrophage M2 polarization and glycogenolysis, and published data from the same work also place *PPP1R3B* upstream of PYGL through the mitochondrial p-GSK-3β/p-PYGL/p-GYS2 axis (Shen et al., 2025). Its marked upregulation in peripheral blood and CSF during acute BM positions *PPP1R3B* as a central metabolic switch. *PYGL*, a vitamin B6-dependent glycogen phosphorylase, has been linked to necroptosis in pulmonary sepsis (Choi et al., 2025). Studies have indicated that inhibition of PYGL may represent an effective strategy to restrain inflammatory responses (Martínez-Navarro et al., 2020). We found *PYGL* increased in serum and CSF from patients with early BM, supporting a model whereby *PPP1R3B*-driven metabolic rewiring activates *PYGL*, leading to glycogen breakdown and necroptotic cell death. This cascade may trigger a localized energy crisis and release damage-associated molecular patterns (DAMPs) that further recruit inflammatory cells. The interplay between these two genes thus connects systemic metabolic alterations to programmed cell death within the meninges.

Finally, the barrier integrity module centers on *MMP25*. *MMP25*, a granulocyte-enriched matrix metalloproteinase, has been proposed as a biomarker for allergy-related respiratory inflammation (Blumenthal et al., 2010). It drives endothelial glycocalyx breakdown via succinate-triggered membrane remodeling (Abdullah et al., 2023). The cerebral endothelial glycocalyx is essential for maintaining BBB integrity, and its disruption exacerbates barrier leakage (Shi et al., 2025). We detected upregulated *MMP25* during acute BM, leading us to hypothesize that *MMP25* impairs BBB function by cleaving glycocalyx components. Its expression may be secondarily induced by inflammatory cues emanating from the *LTB4R* axis or metabolic stress originating from the *PPP1R3B–PYGL* cascade, such as succinate accumulation.

In summary, the five core genes assemble into a functional network that transduces circulating metabolic signals into localized brain inflammation, regulated cell death, and barrier disruption. This modular framework not only deepens our understanding of the lung–brain axis but also suggests that targeting specific modules (e.g., *LTB4R* inhibition for inflammation control, *MMP25* blockade for barrier protection) could interrupt the progression from BP to BM. Building on this five-gene signature, we further performed network pharmacology and molecular docking to screen small molecules capable of modulating this pathway. L-Stepholidine, an alkaloid derived from *Stephania intermedia*, readily crosses the BBB (Sun et al., 2009). It protects mice against methamphetamine-associated memory impairment (Zhou et al., 2019), improves prefrontal cortical function and memory by upregulating glutamate receptors (Hu et al., 2024), and has been characterized as a candidate antipsychotic compound (Meade et al., 2015). Our network pharmacology and molecular docking results predicted that L-stepholidine may target *PYGL*, *OLAH* and *PPP1R3B* among the five core genes to limit BM pathology. This computational prediction requires functional validation, ideally in experimental models that recapitulate the BP-initiated meningeal inflammatory cascade.

This two-sample MR mediation analysis has inherent limitations. Despite sensitivity testing, residual horizontal pleiotropy cannot be completely ruled out. Linkage disequilibrium and inter-metabolite correlations hinder separation of independent mediating effects. All GWAS data originate from European populations, restricting generalizability to other ancestries. MR captures lifelong genetic effects and cannot reflect dynamic, dose-dependent changes during acute infection. Moreover, this analysis only supports statistical mediation; direct biological connections between metabolites and genes await independent experimental validation.

Integrating genetic and multi-omic evidence, this study delineates a molecular pathway within the lung–brain axis: BP promotes BM through sequential changes in circulating metabolites and downstream core gene programs. This work extends the lung–brain axis from an epidemiological association toward a testable “BP–metabolite–gene–BM” cascade, offering a new framework to dissect the pathogenesis of meningitis.

Abbreviations

AMP
monophosphate
AUC
area under the ROC curve
BBB
blood-brain barrier
BM
bacterial meningitis
BP
bacterial pneumonia
cAMP
cyclic monophosphate
CCA
canonical correlation analysis
CNS
central nervous system
CSF
cerebrospinal fluid
DGE
differential gene expression
GS
gene significance
GWAS
genome-wide association study
HVGs
highly variable genes
IMP
inosine 5'-monophosphate
IVW
inverse variance weighted
LPS
lipopolysaccharide
LTB4
leukotriene B4
LTB4R
leukotriene B4 receptor
MMP25
matrix metallopeptidase 25
MM
module membership

MR
mendelian randomization
mDCs
myeloid dendritic cells
NKs
natural killer cells
OLAH
oleoyl-acyl carrier protein hydrolase
OR
odds ratio
PCA
principal component analysis
pDCs
plasmacytoid dendritic cells
PPP1R3B
protein phosphatase 1 regulatory subunit 3B
PYGL
glycogen phosphorylase L
ROC
receiver operating characteristic
scRNA-seq
single-cell RNA sequencing
SNP
single nucleotide polymorphism
TOM
topological overlap matrix
TFs
transcription factors
WGCNA
weighted gene co-expression network analysis

Declarations

Competing Interests

The authors affirm that there are no financial, personal, or other relationships with other people or organizations that could influence the work inappropriately. All the authors have read, edited, and approved the manuscript, and they have consented to the publication.

Ethics and Consent to Participate declarations

The data used in this study were all from GEO and the IEU OPEN GWAS databases, which are publicly accessible. The datasets have received ethical approval, are de-identified, and can be downloaded for free for non-commercial research and publishing relevant articles. Therefore, there are no ethical issues or conflicts of interest in our study.

Consent for publication

All authors consent to publication.

Funding

Declaration

This study was supported by the National Natural Science Foundation of China (32302877; 32272951) and Zhejiang A&F University Research and Development Fund (2024LFR078).

Authors' contributions

H.J. performed and designed the research; H.J. and J.F. analyzed the data and drafted the manuscript; X.C. polished the language; H.S. and Y.Y. supervised the whole project and revised the manuscript. All the authors read, edited, and approved the manuscript.

Acknowledgements

We thank the participants and investigators of the FinnGen study.

Availability of data and materials

The data that support the findings of this study are available within the article and its supplementary material. The datasets analyzed during the current study are available in the GenBank of National Center for Biotechnology Information, accession number GSE40586, GSE163194, GSE163195, GSE163196, and GSE261482.

References

1. Abdullah S, Ghio M, Cotton-Betteridge A, Vinjamuri A, Drury R, Packer J, Aras O, Friedman J, Karim M, Engelhardt D et al (2023) Succinate metabolism and membrane reorganization drives the endotheliopathy and coagulopathy of traumatic hemorrhage. Sci Adv 9(24):eadf6600. 10.1126/sciadv.adf6600

2. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J et al (2017) SCENIC: single-cell regulatory network inference and clustering. Nat Methods 14(11):1083–1086. 10.1038/nmeth.4463
3. Ardiansyah E, Avila-Pacheco J, Nhat LTH, Dian S, Vinh DN, Hai HT, Bullock K, Alisjahbana B, Netea MG, Estiasari R et al (2023) Tryptophan metabolism determines outcome in tuberculous meningitis: a targeted metabolomic analysis. Elife 12:e85307. 10.7554/eLife.85307
4. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28(1):235–242. 10.1093/nar/28.1.235
5. Blumenthal MN, Zhong W, Miller M, Wendt C, Connett JE, Pei D (2010) Serum metalloproteinase leukolysin (MMP-25/MT-6): a potential metabolic marker for atopy-associated inflammation. Clin Experimental Allergy 40:859–866. 10.1111/j.1365-2222.2010.03475.x
6. Cerezo M, Sollis E, Ji Y, Lewis E, Abid A, Bircan KO, Hall P, Hayhurst J, John S, Mosaku A et al (2025) The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. Nucleic Acids Res 53(D1):D998–D1005. 10.1093/nar/gkae1070
7. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z (2017) Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep 18(1):248–262. 10.1016/j.celrep.2016.12.019
8. Chen S, Liang J, Chen D, Huang Q, Sun K, Zhong Y, Lin B, Kong J, Sun J, Gong C et al (2024) Cerebrospinal fluid metabolomic and proteomic characterization of neurologic post-acute sequelae of SARS-CoV-2 infection. Brain Behav Immun 115:209–222. 10.1016/j.bbi.2023.10.016
9. Choi H, Lee JY, Yoo H, Jeon K (2025) Identification and Verification of Necroptosis-Related Genes in Patients with Sepsis by Bioinformatic Analysis and Molecular Experiments. J Cell Mol Med 29(9):e70582. 10.1111/jcmm.70582
10. Costeira R, Evangelista L, Wilson R, Yan X, Hellbach F, Sinke L, Christiansen C, Villicaña S, Masachs OM, Tsai P et al (2023) Metabolomic biomarkers of habitual B vitamin intakes unveil novel differentially methylated positions in the human epigenome. Clin Epigenetics 15(1):166. 10.1186/s13148-023-01578-7
11. Douaud G, Lee S, Alfaro-Almagro F, Arthofer C, Wang C, McCarthy P, Lange F, Andersson JLR, Griffanti L, Duff E et al (2022) SARS-CoV-2 is associated with changes in brain structure in UK Biobank. Nature 604(7907):697–707. 10.1038/s41586-022-04569-5
12. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30(1):207–210. 10.1093/nar/30.1.207
13. Fang S, Dong L, Liu L, Guo J, Zhao L, Zhang J, Bu D, Liu X, Huo P, Cao W et al (2021) HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine. Nucleic Acids Res 49(D1):D1197–D1206. 10.1093/nar/gkaa1063
14. Figueiredo AHA, Brouwer MC, Bijlsma MW, van der Ende A, van de Beek D (2020) Community-acquired pneumonia in patients with bacterial meningitis: a prospective nationwide cohort study.

Clin Microbiol Infect 26(4):513. e7-513.ell

15. Grotzinger AD, Werme J, Peyrot WJ, Frei O, de Leeuw C, Bicks LK, Guo Q, Margolis MP, Coombes BJ, Batzler A et al (2026) Mapping the genetic landscape across 14 psychiatric disorders. Nature 649(8096):406–415. 10.1038/s41586-025-09820-3
16. Hänzelmann S, Castelo R, Guinney J (2013) BMC Bioinformatics 14:7. 10.1186/1471-2105-14-7. GSVA: gene set variation analysis for microarray and RNA-Seq data
17. He M, Xu C, Yang R, Liu L, Zhou D, Yan S (2024) Causal relationship between human blood metabolites and risk of ischemic stroke: a Mendelian randomization study. Front Genet 15:1333454. 10.3389/fgene.2024.1333454
18. He R, Chen Y, Cai Q (2020) The role of the LTB4-BLT1 axis in health and disease. Pharmacol Res 158:104857. 10.1016/j.phrs.2020.104857
19. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R et al (2018) The MR-Base platform supports systematic causal inference across the human phenome. Elife 7:e34408. 10.7554/eLife
20. Hering M, Madi A, Sandhoff R, Ma S, Wu J, Mieg A, Richter K, Mohr K, Knabe N, Stichling D et al (2024) Sphinganine recruits TLR4 adaptors in macrophages and promotes inflammation in murine models of sepsis and melanoma. Nat Commun 15(1):6067. 10.1038/s41467-024-50341-w
21. Hong Y, Zhang C, Shen K, Dong X, Chen B (2024) Genetically predicted plasma metabolites mediate the causal relationship between gut microbiota and primary immune thrombocytopenia (ITP). Front Microbiol 15:1447729. 10.3389/fmicb.2024.1447729
22. Hosang L, Canals RC, van der Flier FJ, Hollensteiner J, Daniel R, Flügel A, Odoardi F (2022) The lung microbiome regulates brain autoimmunity. Nature 603:138–144. 10.1038/s41586-022-04427-4
23. Hu Z, Zhu X, Liang Y, Zhang Y, Zheng P, Zhang X (2024) Levo-Stepholidine as a Potential Cognitive Enhancer: Insights into Executive Function and Memory Improvements. Biomedicines 12(12):2680. 10.3390/biomedicines12122680
24. Jala VR, Bodduluri SR, Satpathy SR, Chheda Z, Sharma RK, Haribabu B (2017) The yin and yang of leukotriene B 4 mediated inflammation in cancer. Semin Immunol 33:58–64. 10.1016/j.smim.2017.09.005
25. Jia X, Crawford JC, Gebregzabher D, Monson EA, Mettelman RC, Wan Y, Ren Y, Chou J, Novak T, McQuilten HA et al High expression of oleoyl-ACP hydrolase underpins life-threatening respiratory viral diseases. Cell 2024, 187(17):4586–4604e20. 10.1016/j.cell.2024.07.026
26. Jiang H, Yu X, Fan J, Song H, Yang Y (2025) Identification of predictors for bacterial meningitis diagnosis based on transcriptomics and genetic analysis. Amb Express 15(1):84. 10.1186/s13568-025-01893-7
27. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B et al (2025) PubChem 2025 update. Nucleic Acids Res 53(D1):D1516–D1525. 10.1093/nar/gkae1059
28. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, Reeve MP, Laivuori H, Aavikko M, Kaunisto MA et al (2023) FinnGen provides genetic insights from a well-phenotyped isolated

population. Nature 613(7952):508–518. 10.1038/s41586-023-05837-8

29. Langenberg C, Hingorani AD, Whitty CJM (2023) Biological and functional multimorbidity—from mechanisms to management. Nat Med 29(7):1649–1657. 10.1038/s41591-023-02420-6
30. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559. 10.1186/1471-2105-9-559
31. Li JH, Liu S, Zhou H, Qu LH, Yang JH starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014(Database issue), 42:D92–D97. 10.1093/nar/gkt1248
32. Lill M, Kõks S, Soomets U, Schalkwyk LC, Fernandes C, Lutsar I, Taba P (2013) Peripheral blood RNA gene expression profiling in patients with bacterial meningitis. Front Neurosci 7:33. 10.3389/fnins.2013.00033
33. Ma J, Wu H, Ma Z, Wu Z (2025) Bacterial and host factors involved in zoonotic Streptococcal meningitis. Microbes Infect 27(1):105335. 10.1016/j.micinf.2024.105335
34. Ma Q, Yao C, Wu Y, Wang H, Fan Q, Yang Q, Xu J, Dai H, Zhang Y, Xu F et al (2023) Neurological disorders after severe pneumonia are associated with translocation of endogenous bacteria from the lung to the brain. Sci Adv 9(42):eadi699. 10.1126/sciadv.adi0699
35. Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K et al (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Res W273–W276 37(Web Server issue). 10.1093/nar/gkp292
36. Martínez-Navarro FJ, Martínez-Morcillo FJ, López-Muñoz A, Pardo-Sánchez I, Martínez-Menchón T, Corbalán-Vélez R, Cayuela ML, Pérez-Oliva AB, García-Moreno D, Mulero V (2020) The vitamin B6-regulated enzymes PYGL and G6PD fuel NADPH oxidases to promote skin inflammation. Dev Comp Immunol 108:103666. 10.1016/j.dci.2020.103666
37. Meade JA, Free RB, Miller NR, Chun LS, Doyle TB, Moritz AE, Conroy JL, Watts VJ, Sibley DR (2015) (-)-Stepholidine is a potent pan-dopamine receptor antagonist of both G protein- and β-arrestin-mediated signaling. Psychopharmacology 232(5):917–930. 10.1007/s00213-014-3726-8
38. More TH, Mozafari B, Märtens A, Herr C, Lepper PM, Danziger G, Volk T, Hoersch S, Krawczyk M, Guenther K et al (2022) Plasma Metabolome Alterations Discriminate between COVID-19 and Non-COVID-19 Pneumonia. Metabolites 12(11):1058. 10.3390/metabo12111058
39. Ning P, Zheng Y, Luo Q, Liu X, Kang Y, Zhang Y, Zhang R, Xu Y, Yang D, Xi W et al (2018) Metabolic profiles in community-acquired pneumonia: developing assessment tools for disease severity. Crit Care 22(1):130. 10.1186/s13054-018-2049-2
40. Palanki R, Yamagata H, Mitchell MJ (2024) OLAH connects fatty acid metabolism to the severity of respiratory viral disease. Cell 187(17):4549–4551. 10.1016/j.cell.2024.07.032
41. Pinho-Ribeiro FA, Deng L, Neel DV, Erdogan O, Basu H, Yang D, Choi S, Walker AJ, Carneiro-Nascimento S, He K et al (2023) Bacteria hijack a meningeal neuroimmune axis to facilitate brain invasion. Nature 615(7952):472–481. 10.1038/s41586-023-05753-x

42. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C (2017) Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 14(10):979–982. 10.1038/nmeth.4402
43. Rischke S, Gurke R, Zielbauer A, Ziegler N, Hahnefeld L, Köhm M, Kannt A, Vehreschild MJ, Geisslinger G, Rohde G et al (2025) Proteomic, metabolomic and lipidomic profiles in community acquired pneumonia for differentiating viral and bacterial infections. Sci Rep 15(1):1922. 10.1038/s41598-025-85229-2
44. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43(7):e47. 10.1093/nar/gkv007
45. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinformatics 12:77. 10.1186/1471-2105-12-77
46. Ru J, Li P, Wang J, Zhou W, Li B, Huang C, Li P, Guo Z, Tao W, Yang Y et al (2014) TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. J Cheminform 6:13. 10.1186/1758-2946-6-13
47. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504. 10.1101/gr.1239303
48. Shen L, Yu J, Chen W, Bi Y, Yang Z, Lu C, Jiang C, Yang Y, Chen M, Zou J et al (2025) PPP1R3B Suppresses Atherosclerosis by Promoting the M2 Polarization of Macrophages Through Glycogen Metabolic Reprogramming. Adv Sci 12(41):e06345. 10.1002/advs.202506345
49. Shi SM, Suh RJ, Shon DJ, Garcia FJ, Buff JK, Atkins M, Li L, Lu N, Sun B, Luo J et al (2025) Glycocalyx dysregulation impairs blood–brain barrier in ageing and disease. Nature 639(8056):985–994. 10.1038/s41586-025-08589-9
50. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP et al (2014) An atlas of genetic influences on human blood metabolites. Nat Genet 46(6):543–550. 10.1038/ng.2982
51. Song C, Zhou L, Zhao L, Guo Z, Zhang L, Yang H, Lei Y, Han Y, Xu Y, Xiang Z et al (2025) Exposure to diverse airborne bacterial communities following Klebsiella pneumoniae infection elicits distinct host response mechanisms in mice. Ecotox Environ Safe 303:118792. 10.1016/j.ecoenv.2025.118792
52. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y et al (2016) The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinform 54(1301–). 10.1002/cpbi.5
53. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R (2019) Comprehensive Integration of Single-Cell Data. Cell 177(7):1888–1902e21. 10.1016/j.cell.2019.05.031

54. Sun Y, Dai J, Hu Z, Du F, Niu W, Wang F, Liu F, Jin G, Li C (2009) Oral bioavailability and brain penetration of (-)-stepholidine, a tetrahydroprotoberberine agonist at dopamine D (1) and antagonist at D (2) receptors, in rats. Br J Pharmacol 158(5):1302–1312. 10.1111/j.1476-5381.2009.00393.x
55. Tager AM, Luster AD (2003) BLT1 and BLT2: the leukotriene B4 receptors. Prostaglandins Leukot Essent Fat Acids 69(2–3):123–134. 10.1016/s0952-3278(03)00073-5
56. Tastsoglou S, Alexiou A, Karagkouni D, Skoufos G, Zacharopoulou E, Hatzigeorgiou AG (2023) DIANA-microT 2023: including predicted targets of virally encoded miRNAs. Nucleic Acids Res 51(W1):W148–W153. 10.1093/nar/gkad283
57. van de Beek D, Brouwer MC, Koedel U, Wall EC (2021) Community-acquired bacterial meningitis. Lancet 398(10306):1171–1183. 10.1016/S0140-6736(21)00883-7
58. van Samkar A, Brouwer MC, van der Ende A, van de Beek D (2016) Zoonotic bacterial meningitis in human adults. Neurology 87(11):1171–1179. 10.1212/WNL.0000000000003101
59. Wang Y, Zhou L, Wang N, Qiu B, Yao D, Yu J, He M, Li T, Xie Y, Yu X et al (2025) Comprehensive characterization of metabolic consumption and production by the human brain. Neuron 113(11):1708–1722e5. 10.1016/j.neuron.2025.03.003
60. Xiao H, Xiao H, Zhang Y, Guo L, Dou Z, Liu L, Zhu L, Feng W, Liu B, Hu B et al (2022) High-throughput sequencing unravels the cell heterogeneity of cerebrospinal fluid in the bacterial meningitis of children. Front Immunol 13:872832. 10.3389/fimmu.2022.872832
61. Yin X, Chan LS, Bose D, Jackson AU, VandeHaar P, Locke AE, Fuchsberger C, Stringham HM, Welch R, Yu K et al (2022) Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. Nat Commun 13:1644. 10.1038/s41467-022-29143-5
62. You D, Wu Y, Lu M, Shao F, Tang Y, Liu S, Liu L, Zhou Z, Zhang R, Shen S et al (2025) A genome-wide cross-trait analysis characterizes the shared genetic architecture between lung and gastrointestinal diseases. Nat Commun 16(1):3032. 10.1038/s41467-025-58248-w
63. Zeng X, Zhong B Investigating the causal relationship between human blood metabolites and pulmonary hypertension: a two-sample Mendelian randomization study. Front Cardiovasc Med 2024, 11:1304986. 10.3389/fcvm.2024.1304986
64. Zhou M, Gong X, Ru Q, Xiong Q, Chen L, Si Y, Xiao H, Li C (2019) The Neuroprotective Effect of L-Stepholidine on Methamphetamine-Induced Memory Deficits in Mice. Neurotox Res 36(2):376–386. 10.1007/s12640-019-00069-z

Figures

A

exposure	outcome	nsnp	method	pval	OR(95% CI)
1−(1−enyl−palmitoyl)−GPE (p−16:0) levels	BM	23	Inverse variance weighted	**0.015**	0.752 (0.597 to 0.946)
1−(1−enyl−stearoyl)−GPE (p−18:0) levels	BM	24	Inverse variance weighted	**0.011**	0.761 (0.615 to 0.940)
1−oleoylglycerol (18:1) levels	BM	30	Inverse variance weighted	**0.003**	1.308 (1.097 to 1.558)
2−palmitoleoyl−GPC (16:1) levels	BM	25	Inverse variance weighted	**0.032**	0.805 (0.660 to 0.982)
2−stearoyl−GPE (18:0) levels	BM	32	Inverse variance weighted	**0.046**	0.836 (0.701 to 0.997)
2,4−di−tert−butylphenol levels	BM	19	Inverse variance weighted	**0.027**	0.792 (0.644 to 0.973)
3−formylindole levels	BM	29	Inverse variance weighted	**0.005**	0.816 (0.708 to 0.940)
4−acetylcatechol sulfate (1) levels	BM	18	Inverse variance weighted	**0.047**	0.892 (0.796 to 0.998)
5−oxoproline to citrate ratio	BM	15	Inverse variance weighted	**0.024**	0.828 (0.703 to 0.975)
5alpha−pregnan−diol disulfate levels	BM	22	Inverse variance weighted	**0.038**	1.166 (1.008 to 1.347)
7−methylguanine levels	BM	23	Inverse variance weighted	**0.037**	0.796 (0.642 to 0.987)
Adenosine 3',5'−cyclic monophosphate (cAMP) to adenosine 5'−monophosphate (AMP) ratio	BM	13	Inverse variance weighted	**0.003**	0.657 (0.497 to 0.869)
Adenosine 5'−diphosphate (ADP) to choline phosphate ratio	BM	22	Inverse variance weighted	**0.010**	1.267 (1.057 to 1.518)
Adenosine 5'−diphosphate (ADP) to gluconate ratio	BM	23	Inverse variance weighted	**0.010**	1.262 (1.058 to 1.505)
Adenosine 5'−diphosphate (ADP) to glycerol 3−phosphate ratio	BM	23	Inverse variance weighted	**0.042**	1.176 (1.006 to 1.375)
Alpha−ketoglutarate levels	BM	26	Inverse variance weighted	**0.047**	1.244 (1.003 to 1.544)
Alpha−ketoglutarate to proline ratio	BM	24	Inverse variance weighted	**0.041**	1.250 (1.009 to 1.549)
Bilirubin (Z,Z) to etiocholanolone glucuronide ratio	BM	28	Inverse variance weighted	**0.042**	1.138 (1.005 to 1.288)
Bilirubin degradation product, C16H18N2O5 (1) levels	BM	23	Inverse variance weighted	**0.031**	1.178 (1.015 to 1.369)
Bilirubin degradation product, C16H18N2O5 (3) levels	BM	29	Inverse variance weighted	**0.038**	1.131 (1.007 to 1.270)
Carnitine to propionylcarnitine (C3) ratio	BM	24	Inverse variance weighted	**0.030**	1.227 (1.020 to 1.475)
Cholate to phosphate ratio	BM	18	Inverse variance weighted	**0.046**	0.800 (0.643 to 0.996)
Cholesterol to oleoyl−linoleoyl−glycerol (18:1 to 18:2) [2] ratio	BM	35	Inverse variance weighted	**0.025**	0.837 (0.716 to 0.978)
Citrate to oxalate (ethanedioate) ratio	BM	16	Inverse variance weighted	**0.024**	0.745 (0.576 to 0.963)
Cys−gly, oxidized levels	BM	24	Inverse variance weighted	**0.004**	1.236 (1.068 to 1.429)
Dibutyl sulfosuccinate levels	BM	36	Inverse variance weighted	**0.008**	0.919 (0.864 to 0.979)
Dihydroferulic acid sulfate levels	BM	28	Inverse variance weighted	**0.026**	1.115 (1.013 to 1.228)
Dodecadienoate (12:2) levels	BM	24	Inverse variance weighted	**0.029**	1.094 (1.009 to 1.187)
Glucuronide of C12H22O4 (1) levels	BM	18	Inverse variance weighted	**0.013**	1.274 (1.052 to 1.543)
Glycolithocholate levels	BM	23	Inverse variance weighted	**0.024**	0.808 (0.672 to 0.972)
Iminodiacetate (IDA) levels	BM	17	Inverse variance weighted	**0.047**	1.342 (1.004 to 1.794)
Inosine 5'−monophosphate (IMP) to urate ratio	BM	20	Inverse variance weighted	**0.036**	0.829 (0.697 to 0.988)
Linoleate (18:2n6) levels	BM	21	Inverse variance weighted	**0.048**	1.301 (1.002 to 1.688)
N,N−dimethylalanine levels	BM	25	Inverse variance weighted	**0.003**	1.309 (1.093 to 1.567)
Perfluorooctanesulfonate (PFOS) levels	BM	24	Inverse variance weighted	**0.013**	0.785 (0.648 to 0.951)
Phenylpyruvate to citrate ratio	BM	29	Inverse variance weighted	**0.020**	1.243 (1.034 to 1.495)
Pyridoxate levels	BM	24	Inverse variance weighted	**0.002**	1.395 (1.130 to 1.722)
Ribitol levels	BM	34	Inverse variance weighted	**0.020**	0.880 (0.790 to 0.980)
Sphinganine levels	BM	22	Inverse variance weighted	**0.042**	1.262 (1.009 to 1.580)
Stachydrine levels	BM	24	Inverse variance weighted	**0.036**	0.747 (0.568 to 0.981)
Suberate (C8−DC) levels	BM	27	Inverse variance weighted	**0.020**	0.837 (0.721 to 0.972)
Taurodeoxycholic acid 3−sulfate levels	BM	22	Inverse variance weighted	**0.029**	0.872 (0.771 to 0.986)
Trans−urocanate levels	BM	19	Inverse variance weighted	**0.031**	1.217 (1.018 to 1.455)
Undecenoylcarnitine (C11:1) levels	BM	29	Inverse variance weighted	**0.026**	0.839 (0.719 to 0.979)
X−10458 levels	BM	24	Inverse variance weighted	**0.014**	0.778 (0.637 to 0.950)
X−11470 levels	BM	34	Inverse variance weighted	**0.042**	1.179 (1.006 to 1.383)
X−12216 levels	BM	23	Inverse variance weighted	**0.027**	0.789 (0.640 to 0.974)
X−12707 levels	BM	30	Inverse variance weighted	**0.039**	1.213 (1.010 to 1.456)
X−12730 levels	BM	18	Inverse variance weighted	**0.028**	0.778 (0.621 to 0.973)
X−13844 levels	BM	17	Inverse variance weighted	**0.032**	0.722 (0.535 to 0.972)
X−15503 levels	BM	20	Inverse variance weighted	**0.008**	0.728 (0.575 to 0.921)
X−21796 levels	BM	24	Inverse variance weighted	**0.033**	1.314 (1.023 to 1.687)
X−22509 levels	BM	20	Inverse variance weighted	**0.014**	0.862 (0.766 to 0.971)
X−25109 levels	BM	29	Inverse variance weighted	**0.042**	0.814 (0.668 to 0.993)
X−25265 levels	BM	28	Inverse variance weighted	**0.034**	0.891 (0.801 to 0.991)
X−25433 levels	BM	26	Inverse variance weighted	**0.011**	0.829 (0.718 to 0.958)
X−25520 levels	BM	16	Inverse variance weighted	**0.036**	0.769 (0.602 to 0.983)

B

exposure	outcome	nsnp	method	pval	OR(95% CI)
BP	Adenosine 3',5'−cyclic monophosphate (cAMP) to adenosine 5'−monophosphate (AMP) ratio	30	Inverse variance weighted	**0.047**	1.134 (1.002 to 1.283)
BP	Bilirubin (Z,Z) to etiocholanolone glucuronide ratio	29	Inverse variance weighted	**0.023**	0.868 (0.768 to 0.980)
BP	Dibutyl sulfosuccinate levels	29	MR Egger	**0.003**	0.634 (0.482 to 0.836)
BP	Glucuronide of C12H22O4 (1) levels	29	MR Egger	**0.044**	0.724 (0.536 to 0.976)
BP	Glucuronide of C12H22O4 (1) levels	29	Inverse variance weighted	**0.039**	0.866 (0.756 to 0.993)
BP	Inosine 5'−monophosphate (IMP) to urate ratio	29	Inverse variance weighted	**0.003**	0.796 (0.685 to 0.925)
BP	Pyridoxate levels	29	Inverse variance weighted	**0.036**	1.136 (1.008 to 1.279)
BP	Sphinganine levels	29	MR Egger	**0.015**	0.697 (0.531 to 0.914)
BP	Sphinganine levels	29	Inverse variance weighted	**0.011**	0.852 (0.753 to 0.963)
BP	X−13844 levels	30	Inverse variance weighted	**0.048**	1.130 (1.001 to 1.274)

Figure 1

Mendelian randomization results of bacterial pneumonia (BP)−plasma metabolites−bacterial meningitis (BM). A: Forest plot showing the effect of plasma metabolites on BM; B: Forest plot showing the effect of BP on plasma metabolites. An odds ratio (OR) > 1 indicates a risk factor; an OR < 1 indicates a protective factor.

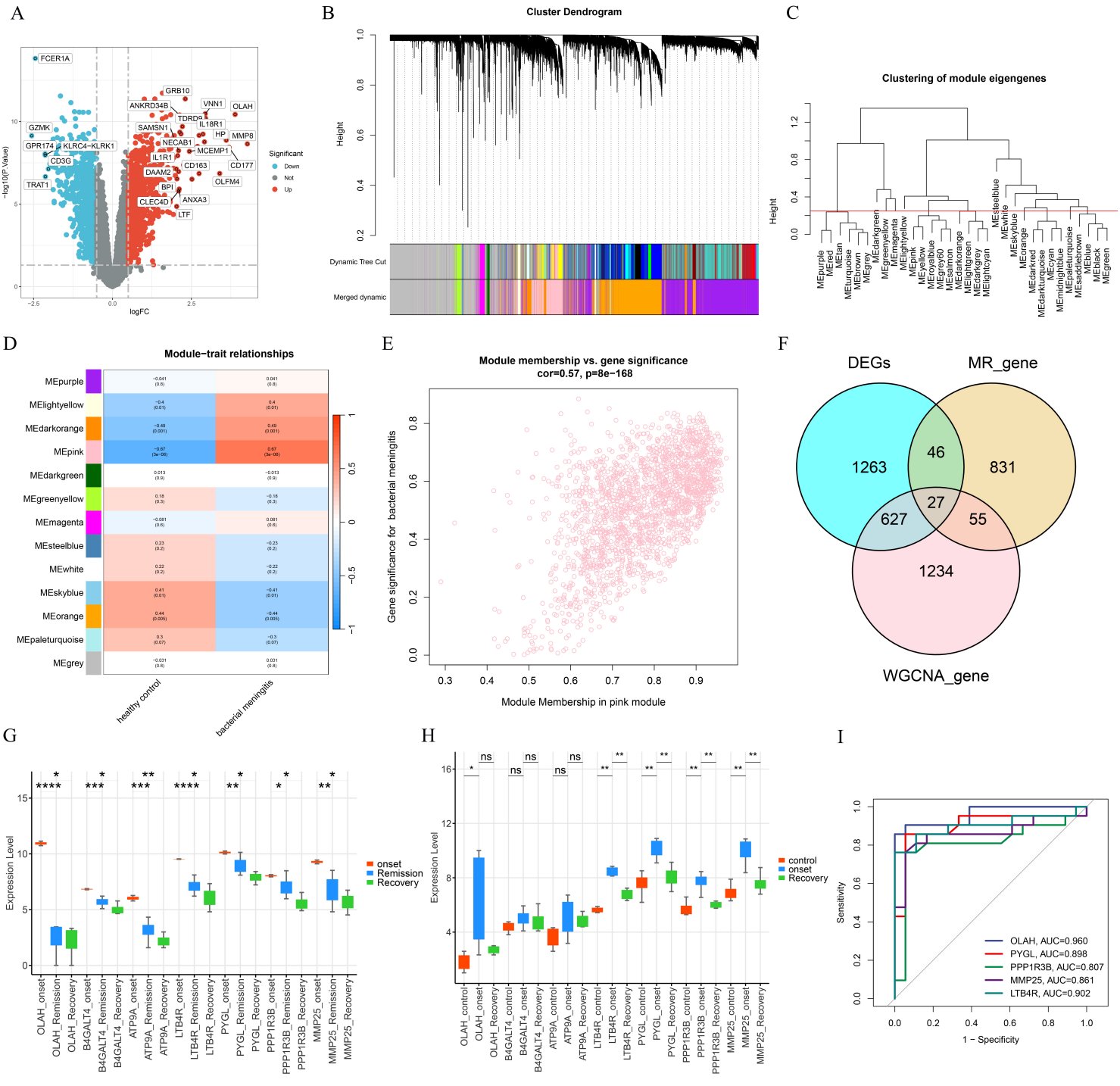

Figure 2

Screening and Validation of Core Genes. A: Volcano plot visualizing differentially expressed genes (DEGs); B-E: Construction and analysis of a weighted gene co-expression network; B: Cluster dendrogram. Gene clustering based on TOM-based dissimilarity (Top); Co-expression modules assigned distinct colors for each gene in the hierarchical clustering tree. Dynamic tree cut denotes the original modules, while merged dynamic corresponds to the finally identified merged modules (Bottom); C: Clustering of module eigengenes; D: Heatmap showing module-trait correlations. Numbers indicate correlation coefficients, with P-value in parentheses; E: Scatter plot of Gene Significance (GS) and

Module Membership (MM) in the pink module; F: Venn diagram of DEGs, metabolite SNP-associated genes (MR_GENEs), and pink module genes (WGCNA_GENEs); G–H: Validation of core gene expression in BM at different stages using datasets GSE163195 (G) and GSE163196 (H); I: ROC curves of core genes.

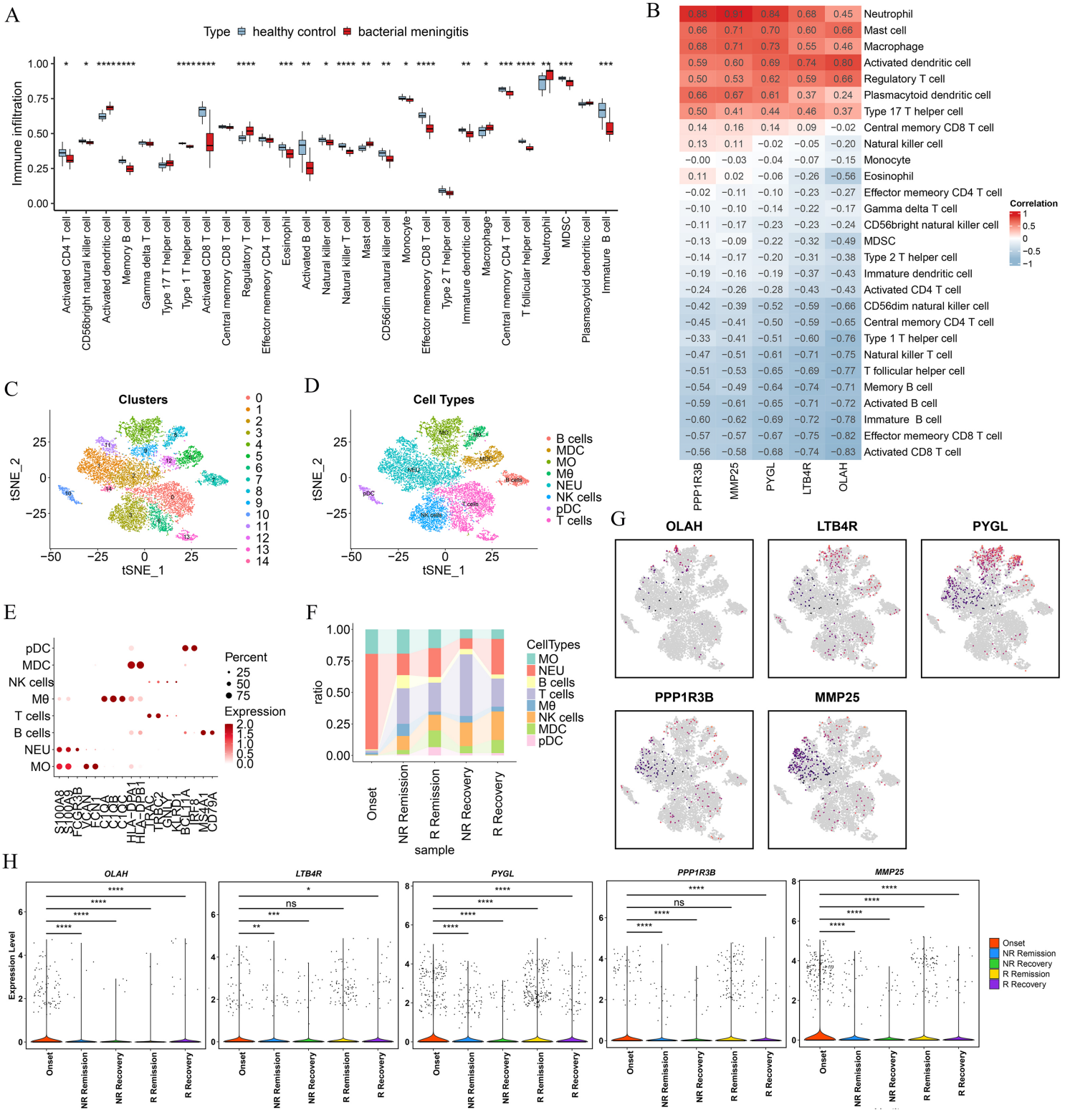

Figure 3

Expression of Core Genes in Immune Cells. A: Violin plot illustrating immune cell infiltration scores; B: Heatmap of the correlation between core genes and immune cells; C: t-SNE map of cell subsets; D: Single-cell annotation; E: Annotated marker genes; F: Bar plot of cell subset proportions; G: Expression of core genes in different cell subsets; H: Expression of core genes in the single-cell dataset (GSE163194).

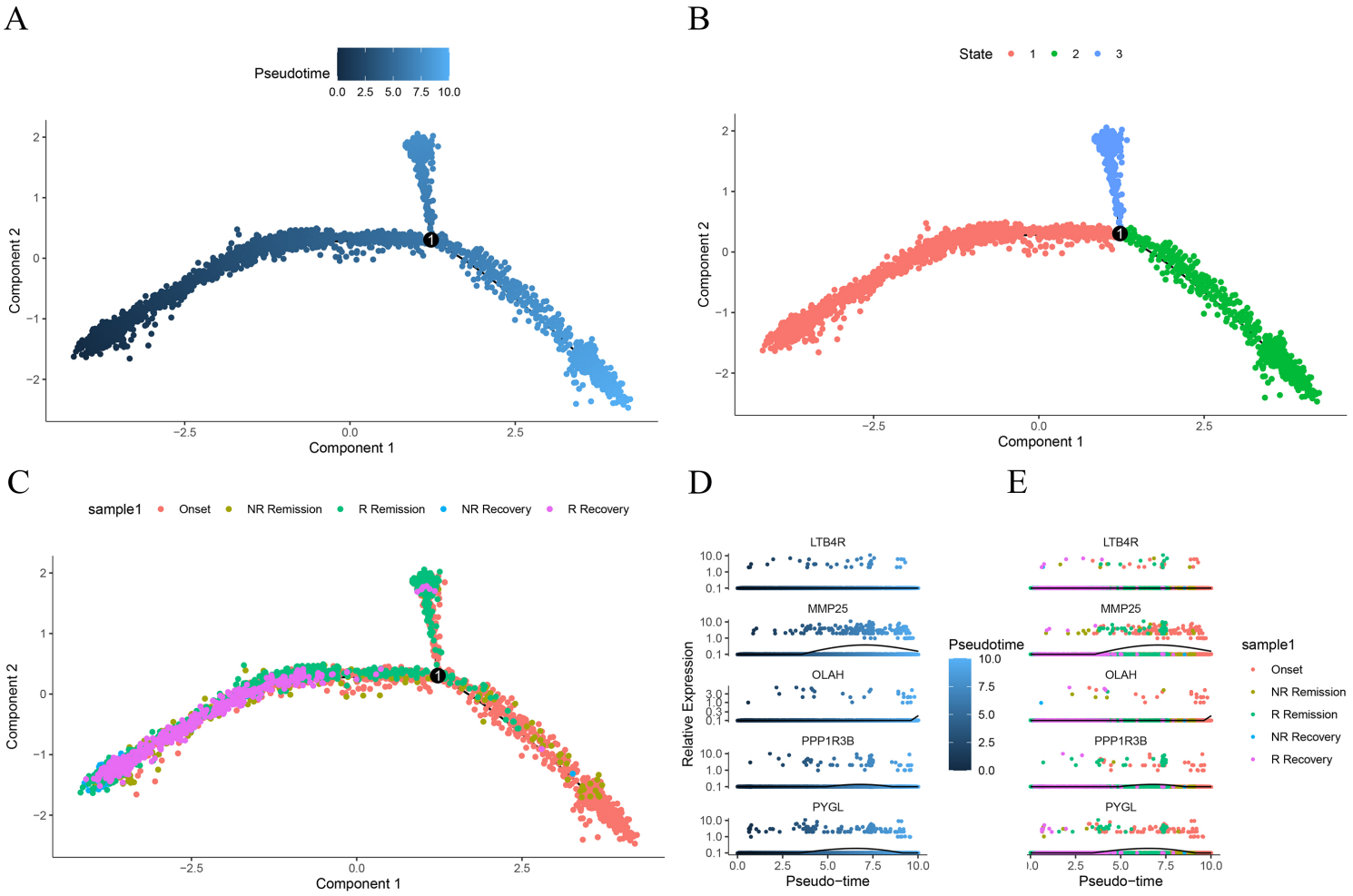

Figure 4

Expression profiles of core genes during neutrophil differentiation. A: Pseudotime trajectory over time; B: Pseudotime trajectory across stages; C: Pseudotime trajectory among samples; D: Expression dynamics of core genes during pseudotime; E: Expression patterns of core genes across differentiation stages.

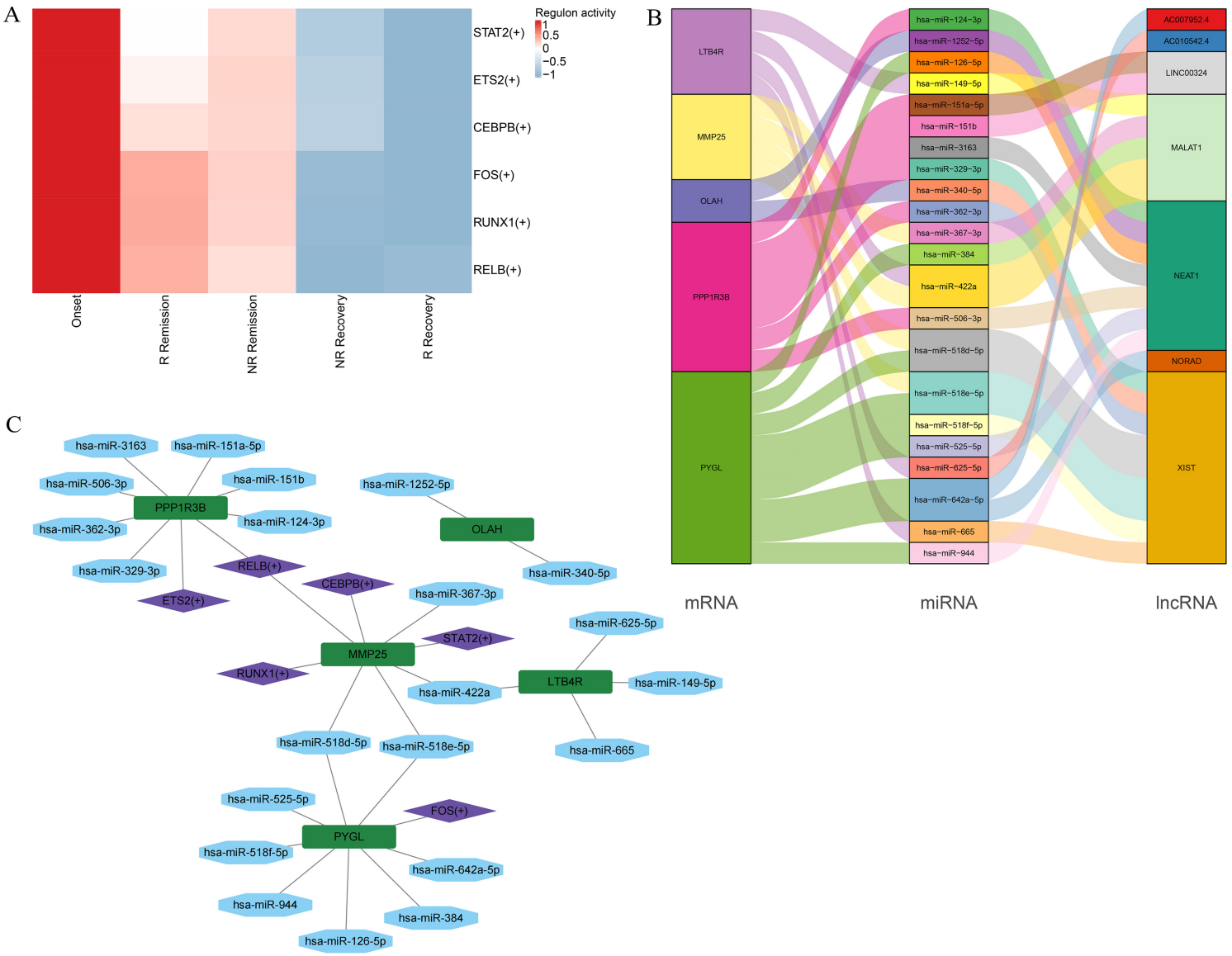

Figure 5

Molecular regulatory network of core genes. A: Expression profiles of transcription factors (TFs) in distinct neutrophil populations; B: Sankey diagram for the ceRNA (lncRNA–miRNA–mRNA) regulatory network; C: Molecular regulatory network of mRNA–miRNA–TFs.

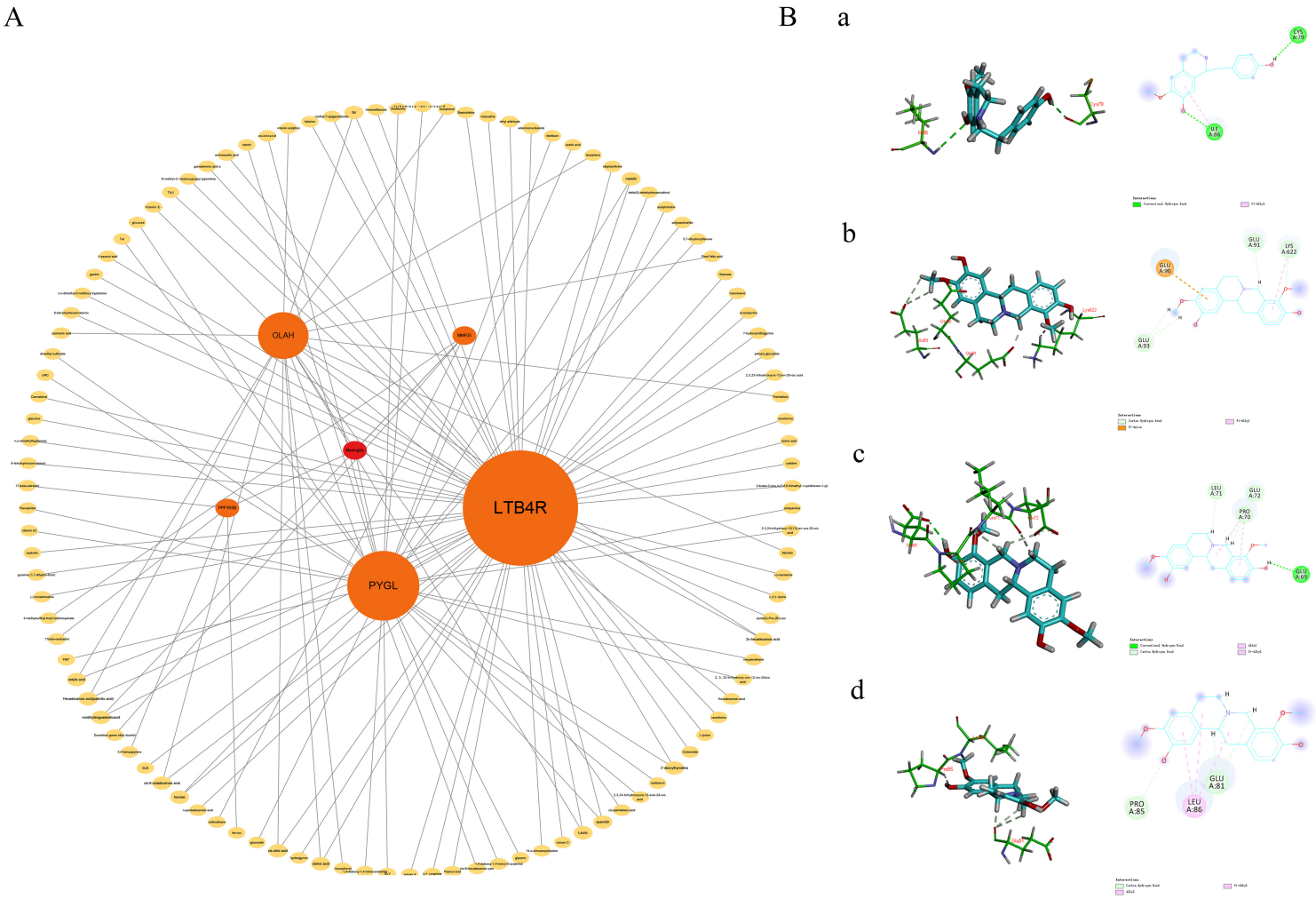

Figure 6

Screening of targeted drugs for core genes. A: Disease–gene–compound network based on core genes; B: three- and two-dimensional molecular docking diagrams for: (a) PYGL–D-coclaurine; (b) PYGL–Stepholidine; (c) OLAH–Stepholidine; (d) PPP1R3B–Stepholidine.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.jpg
- FigureS2.jpg
- FigureS3.jpg
- FigureS4.jpg
- FigureS5.jpg
- FigureS6.jpg
- FigureS7.jpg

- FigureS8.jpg
- FigureS9.jpg
- FigureS10.jpg
- FigureS11.jpg
- FigureS12.jpg
- TableS1.docx
- TableS2.docx
- TableS3.docx
- TableS4.docx
- TableS5.docx
- TableS6.docx
- TableS7.docx