Check for updates

Investigating the causal relationship and potential shared diagnostic genes between primary biliary cholangitis and systemic lupus erythematosus using bidirectional Mendelian randomization and transcriptomic analyses

Tian Tao[1†], Anqi Tang[1†], Lizeyu Lv[1], Jianhua Yuan[2], Ling Wu[1], Liangbin Zhao[1*] and Jun Chen[3*]

[1]Department of Nephrology, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China, [2]Department of Cardiovascular Medicine, Chengdu Second People's Hospital, Chengdu, Chengdu, China, [3]Department of Intensive Care Medicine, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China

Background: The co-occurrence of primary biliary cholangitis (PBC) and systemic lupus erythematosus (SLE) has been consistently reported in observational studies. Nevertheless, the underlying causal correlation between these two conditions still needs to be established.

Methods: We performed a bidirectional two-sample Mendelian randomization (MR) study to assess their causal association. Five MR analysis methods were utilized for causal inference, with inverse-variance weighted (IVW) selected as the primary method. The Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) and the IVW Radial method were applied to exclude outlying SNPs. To assess the robustness of the MR results, five sensitivity analyses were carried out. Multivariable MR (MVMR) analysis was also employed to evaluate the effect of possible confounders. In addition, we integrated transcriptomic data from PBC and SLE, employing Weighted Gene Co-expression Network Analysis (WGCNA) to explore shared genes between the two diseases. Then, we used Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment methods to perform on the shared genes. The Least Absolute Shrinkage and Selection Operator (LASSO) regression algorithm was utilized to identify potential shared diagnostic genes. Finally, we verified the potential shared diagnostic genes in peripheral blood mononuclear cells (PBMCs)-specific cell populations of SLE patients by single-cell analysis.

Results: Our MR study provided evidence that PBC had a causal relationship with SLE (IVW, OR: 1.347, 95% CI: 1.276 - 1.422, P < 0.001) after removing outliers (MR-PRESSO, rs35464393, rs3771317; IVW Radial, rs11065987, rs12924729, rs3745516). Conversely, SLE also had a causal association with PBC (IVW, OR:

1.225, 95% CI: 1.141 - 1.315, P < 0.001) after outlier correction (MR-PRESSO, rs11065987, rs3763295, rs7774434; IVW Radial, rs2297067). Sensitivity analyses confirmed the robustness of the MR findings. MVMR analysis indicated that body mass index (BMI), smoking and drinking were not confounding factors. Moreover, bioinformatic analysis identified PARP9, ABCA1, CEACAM1, and DDX60L as promising diagnostic biomarkers for PBC and SLE. These four genes are highly expressed in CD14+ monocytes in PBMCs of SLE patients and potentially associated with innate immune responses and immune activation.

Conclusion: Our study confirmed the bidirectional causal relationship between PBC and SLE and identified PARP9, ABCA1, CEACAM1, and DDX60L genes as the most potentially shared diagnostic genes between the two diseases, providing insights for the exploration of the underlying mechanisms of these disorders.

Introduction

Systemic lupus erythematosus (SLE) is a complex autoimmune disease that affects multiple organs and systems (1, 2). Genetic factors are widely acknowledged to participate in the development of SLE, even though the specific cause of the disease remains elusive (2–4). Abnormal liver enzymes can be observed in approximately 50% of patients with SLE. Lupus hepatitis and drug-induced hepatitis are common causes of liver enzyme abnormalities, while autoimmune liver diseases such as primary biliary cholangitis (PBC) can also lead to such abnormalities (5–7). PBC is a chronic inflammation of the liver bile ducts that is mediated by autoimmune reactions. Patients with PBC usually present with symptoms such as fatigue, pruritus, and jaundice. Laboratory tests typically show elevated serum alkaline phosphatase, cholestatic liver enzymes, and anti-mitochondrial antibodies. In a few cases, some patients with PBC may develop cirrhosis (8–11). Observational studies have shown that the proportion of patients with PBC also suffering from SLE can reach 3.7% (12–15). Although the prevalence is not very high, there are certain common clinical features between patients with PBC and SLE. One example is the existence of anti-nuclear antibodies (ANA) in their immune profile (13, 14, 16–18). In addition, individuals with SLE-PBC have a higher occurrence of blood, muscle, and pulmonary involvement than those with SLE. Moreover, the combination of PBC also has a negative impact on the survival rate of SLE (19). Thus, it is essential for clinicians to have an exhaustive understanding of this overlap. However, whether PBC and SLE have a causal relationship or their coexistence is coincidental is uncertain. Confounding factors and reverse causation influence the results of observational studies (20,

21). Therefore, a more robust study design is needed to assess the causal relationship between PBC and SLE.

Utilizing genetic variations as instrumental variables (IVs), Mendelian randomization (MR) provides insight into the causal connection between exposure factors and clinical outcomes (22, 23). This approach is valuable in assessing and identifying potential causal associations, especially when confounding factors can bias causal relationships in observational studies, and randomized controlled trials are challenging to implement (24, 25). Through MR analysis, this study seeks to establish the bidirectional causal relationship between PBC and SLE. There is no substantial evidence to establish the mechanistic link between the two diseases. Therefore, this study also utilizes transcriptomic data from PBC and SLE to explore potential interacting genes and shared diagnostic genes between the two diseases to gain insight into the underlying pathological processes that may connect them.

Materials and methods

Bidirectional Mendelian randomization

Study design and data sources

This investigation employed bidirectional MR analysis to examine the causal connection between PBC and SLE. Genetic association estimates for PBC were obtained from a recent genome-wide association study (GWAS) involving 2764 European PBC patients and 10475 controls (http://gwas.mrcieu.ac.uk/) (26). To increase the reliability of the results, we also used the SNPs of 8021 European PBC patients extracted from a GWAS meta-analysis for

validation (27). SLE summary statistics were collected from a large GWAS of the European population comprising 5,201 cases and 9,066 controls (http://gwas.mrcieu.ac.uk/) (28). In addition, a multivariate MR (MVMR) analysis was employed to exclude the influence of confounders, including body mass index (BMI), smoking and alcohol consumption. The dataset for BMI (ukb-b-19953) was obtained from the MRC-IEU consortium, smoking (ieu-b-25) and alcohol consumption (ieu-b-73) were obtained from the GSCAN consortium (https://gwas.mrcieu.ac.uk/datasets/). The study flow diagram is outlined in Figure 1.

SNPs selection

During the initial selection phase, we employed strict criteria to choose preliminary candidate single nucleotide polymorphisms (SNPs) for our study. We focused on autosomal biallelic SNPs with a P-value below 5×10^{-8}. To ensure the independence of these genetic variants, we conducted a clumping analysis, specifically examining linkage disequilibrium with an R^2 value below 0.001 within a 10000 kb window. This analysis helped us identify SNPs that were not closely linked to each other. To further validate our chosen SNPs, we manually searched for them in the PhenoScanner V2, a comprehensive database (http://www.phenoscanner.medschl.cam.ac.uk/). This step aimed to confirm that none of the selected SNPs had potential confounding effects on each other. By ensuring their independence and lack of confounding, we increased the reliability of our findings. Additionally, we sought to ensure a robust association between IVs and exposure factors. To achieve this, we excluded IVs with F values below 10, as calculated using the formula (R2/(1-R2) *(N-2)] (29). This step allowed us to prioritize IVs that exhibited strong associations with the relevant exposure factors, thereby enhancing the validity and accuracy of our study.

Statistical analysis

In this study, we employed a diverse set of five methods to assess the bidirectional causal effects between PBC and SLE. These methods included inverse-variance weighted (IVW), weighted median, simple mode, weighted mode, and MR-Egger. The IVW method served as our primary approach, as it offers reliable estimates in the absence of pleiotropy and heterogeneity. For MVMR analysis, the multivariate IVW method was employed. To address potential issues like pleiotropy, we conducted an MR-Egger regression analysis. We also evaluated heterogeneity using Cochran's Q statistics. Additionally, we performed sensitivity analyses using leave-one-out techniques and funnel plots. Given the potential presence of horizontal pleiotropy, we utilized the MR-PRESSO method to filter out SNPs that could introduce bias and yield more reliable causal inference results. Moreover, we employed the IVW Radial method, which enabled us to generate radial plots and identify any outlier SNPs.

To determine the significance of the causal relationship, we considered multiple factors. A causal relationship was deemed significant if the P value obtained from the IVW method was less than 0.05. Furthermore, the P value from the MR-Egger intercept test needed to be greater than 0.05, indicating the absence of directional pleiotropy. Additionally, there should be consistency in the direction of estimates among the IVW, MR-Egger, and weighted median methods. The "TwoSampleMR," "MRPRESSO," "MVMR," and "MendelianRandomization," package for R version 4.3.0 was used to perform all statistical analyses.

Transcriptomic analyses

Data collection and download

We utilized gene expression data from two datasets to investigate the association between PBC and SLE. The first dataset, GSE119600, sourced from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), comprised microarray-based measurements using RNA from whole blood samples. It included 370 samples, of which we selected 90 individuals diagnosed with PBC and 47 healthy controls for data analysis. The second dataset, GSE65391, also obtained from the GEO database, consisted of 996 samples, including 924 SLE cases and 72 healthy controls. Those datasets provided valuable insights into gene expression patterns related to PBC and SLE.

Differential expression analysis

In the initial analysis, we utilized R (4.3.0) software to process and normalize the expression matrix. By employing the "DEseq2" R package, we identified differentially expressed genes (DEGs) from GSE119600 and GSE65391 datasets, considering an adjusted P value < 0.05. To visually represent the differential gene patterns, we utilized R software to create a heatmap to cluster the DEGs. The Gene Ontology (GO) analysis was also employed using the "ClusterProfiler" R package to delve deeper into the biological mechanisms of DEGs.

FIGURE 1
Schematic diagram of the study design.
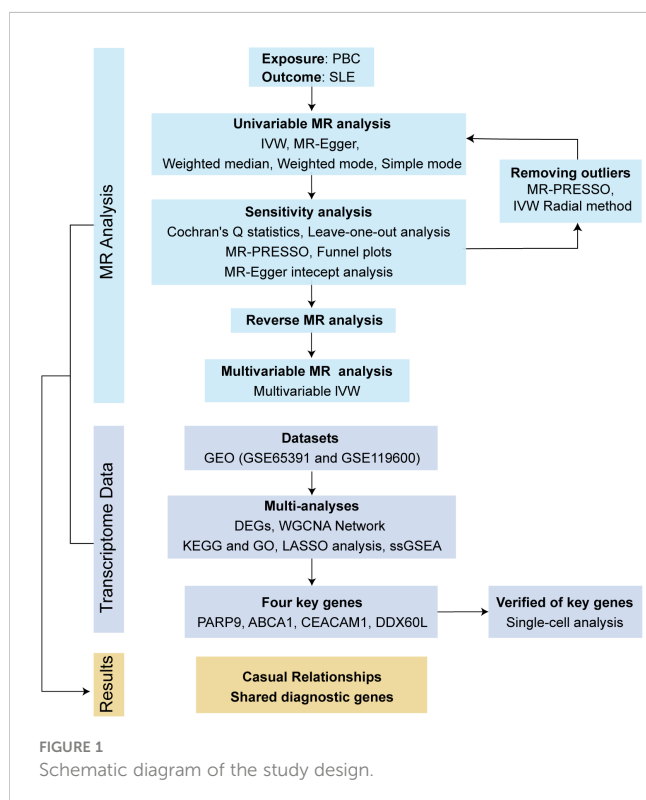
WGCNA network building and module identification

Utilizing WGCNA, a powerful bioinformatic analysis approach, assists in unraveling the gene correlation patterns across various samples. To construct the co-expression network, we utilized the WGCNA R package and considered genes with an adjusted P value < 0.05. The procedure began with hierarchical clustering to identify outliers utilizing the "Hculst" function. Next, we achieved a scale-free network by selecting the optimal soft thresholding power β through the "pick Soft Threshold" function. The matrix portraying the similarities in gene expression was subsequently transformed into an adjacency matrix by utilizing the "adjacency" function, incorporating the selected soft-thresholding parameter β. To enhance data quality and reduce spurious associations, we converted the adjacency matrix into a topological overlap matrix (TOM). Subsequently, we identified modules using hierarchical clustering and the dynamic tree-cut method. In order to determine the correlation between modules and patient clinical characteristics, Pearson correlation analysis was carried out, with results deemed statistically significant at P < 0.05.

Detection of shared genes and pathway enrichment

Through the utilization of Venn diagrams, we conducted an integrated assessment of the genes derived from both WGCNA and DEGs. The resultant common genes were deemed central shared genes and retained for further investigation into functional enrichment. Using R packages like "clusterProfiler," "enrich plot," and "ggplot2", we conducted GO analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. The resulting bar chart displays the top 10 pathways, with p-values below 0.01, highlighting their significance.

Identification of shared diagnostic genes using LASSO regression analysis

Utilizing the "glmnet" R package, we employed LASSO regression analysis, a popular technique that employs an ℓ1 penalty to achieve a sparse solution. This analysis aimed to identify the most influential predictors of SLE and PBC among the DEGs and the intersecting genes identified by WGCNA. LASSO regression allowed us to pinpoint the critical factors associated with these diseases, aiding in understanding their underlying mechanisms.

Evaluation of shared diagnostic gene expression levels

Using the "GSVA" R package, we performed single-sample gene set enrichment analysis (ssGSEA) to score the shared core genes separately in patient and normal samples. This approach allows us to quantify the degree to which a group of genes shows coherent overexpression or underexpression in an individual sample. To further examine the expression levels of these crucial genes, we employed boxplots from the "gglplot2" package in R, considering a significance threshold of P < 0.05. This visualization technique concisely represents how pivotal genes are expressed across diverse conditions, enhancing our understanding of their functional roles.

Single-cell data processing and clustering

The single-cell sequencing analysis started with downloading the raw data from the GSE174188 dataset on GEO, which was then processed using the Seurat software package (version 5.0.0). As per the established protocols, quality control of the scRNA-seq data was stringently performed. Subsequently, we focused on each sample's top 2000 highly variable genes (HVGs), identified based on variance stabilizing transformation (vst). These HVGs were then normalized for further analysis. We scaled these genes using the ScaleData function, and dimensionality reduction was performed using the RunPCA function, where we specifically chose 30 dimensions (dim = 30). Following this, cell clustering was conducted using the FindNeighbors and FindClusters functions, identifying 17 distinct cell clusters. These clusters were visualized using the RunUMAP function, allowing for a clear representation of the cellular heterogeneity and the distinct cell populations within the dataset.

Result

Causal effects of PBC on SLE and sensitivity analyses

To evaluate the causal effects of PBC on SLE, we analyzed 1,124,241 PBC-related SNPs from the GWAS study involving 2,764 European PBC patients and 10475 controls. To remove the influence of confounders, we excluded SNPs strongly correlated with SLE (rs10488631, rs2304256, rs9303277) using the PhenoScanner V2 website, resulting in a preliminary selection of 21 SNPs as IVs for PBC (Supplementary Table S1).

Our primary analyses observed a significant causal association between PBC and an SLE (IVW, OR: 1.319, 95% CI: 1.168 - 1.489, P < 0.001). The other four MR methods revealed similar results (Supplementary Table S2). MR-Egger regression analyses showed no pleiotropy (Intercept, 0.034, P = 0.603). However, Cochran's Q test showed the exist of heterogeneity (IVW, Q =131.319, P < 0.001; MR-Egger, Q = 129.418, P < 0.001) (Supplementary Table S3). To address this concern, we utilized the MR-PRESSO (rs35464393, rs3771317) and the IVW Radial method (rs11065987, rs12924729, rs3745516) to identify outlier SNPs (Supplementary Figure S1).

After excluding five outlier SNPs, the final analysis still demonstrated a significant and consistent causal association between PBC and an increased risk of SLE. The OR estimates obtained from five MR methods were as follows: 1.347 (95% CI: 1.276 - 1.422) for the IVW method, 1.560 (95% CI: 1.293 - 1.882) for MR Egger, 1.363 (95% CI: 1.258 - 1.477) for weighted median, 1.393 (95% CI: 1.234 - 1.573) for weighted mode, and 1.397 (95% CI: 1.223 - 1.597) for simple mode, all indicating statistically significant associations between PBC and SLE (P < 0.001 across five MR methods) (Figures 2, 3A, B; Supplementary Table S2).

FIGURE 2
Forest plot of bidirectional mendelian randomization.
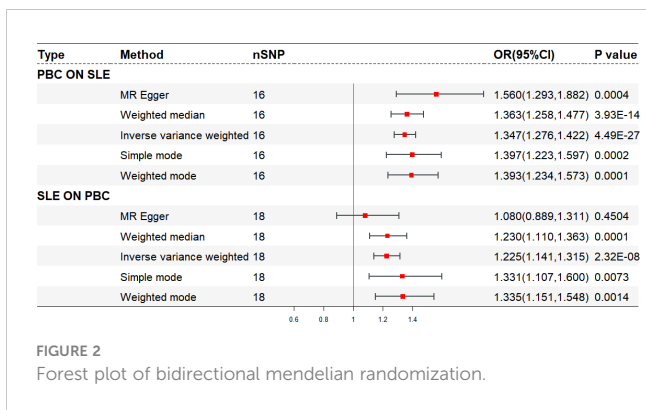
Moreover, no evidence of heterogeneity (IVW, Q = 14.418, P = 0.494; MR-Egger, Q = 11.865, P = 0.617) or horizontal pleiotropy (Intercept, -0.043, P = 0.132) was observed (Supplementary Table S3). A leave-one-out analysis revealed that no single SNP significantly influenced the causal relationship between PBC and SLE (Figure 3C). Moreover, a relatively symmetrical distribution of variant effects for SLE was observed in the Funnel plot (Figure 3D).

Causal effects of SLE on PBC and sensitivity analyses

In reverse analysis, 27 SNPs were screened as IVs for SLE to evaluate causality with PBC (Supplementary Table S4). Initial analysis indicated a significant causal relationship between SLE and PBC (IVW, OR: 1.255, 95% CI: 1.116 - 1.411, P < 0.001), details are shown in Supplementary Table S5. Sensitivity analyses revealed no evidence of pleiotropy (Intercept, 0.072, P = 0.098). While the Cochran's Q test identified potential heterogeneity in the initial analysis (IVW, Q = 168.677, P < 0.001; MR-Egger, Q = 150.810, P <

0.001) (Supplementary Table S6). Therefore, we employed the MR-PRESSO (rs35000415, rs35251378, rs389884, rs597808) and the IVW Radial method (rs2573219, rs353608, rs4274624, rs6671847, rs9852014) to identify outlier SNPs (Supplementary Figure S2).
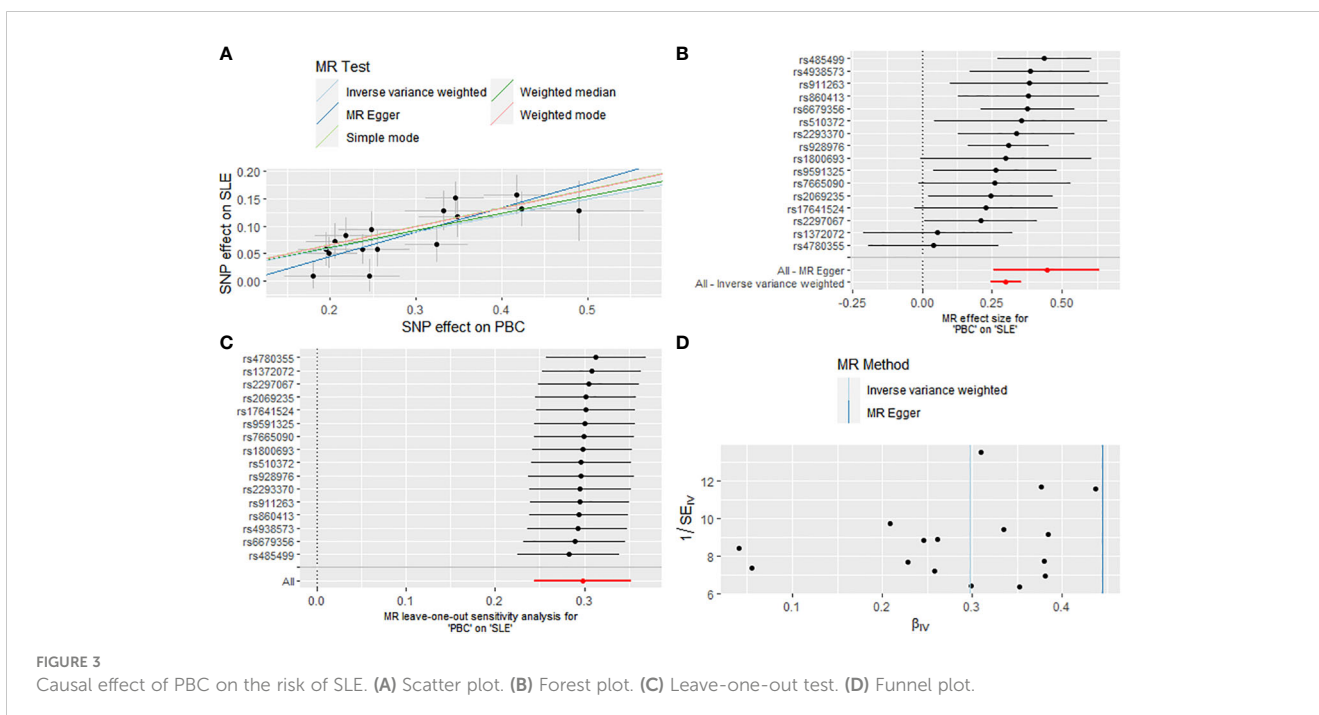
The final analysis confirmed the continued presence of a causal connection between SLE and the increased risk of PBC (OR: 1.225, 95% CI: 1.141 - 1.315, P < 0.001 for IVW; OR: 1.230, 95% CI: 1.110 - 1.363, P < 0.001 for weighted median; OR: 1.335, 95% CI: 1.151 - 1.548, P = 0.001 for weighted mode; OR: 1.331, 95% CI: 1.107 - 1.600, P = 0.007 for simple mode; OR: 1.080, 95% CI: 0.889 - 1.311, P = 0.450 for MR Egger) (Figures 2, 4A, B; Supplementary Table S5), with no evidence of heterogeneity (IVW, Q = 13.503, P = 0.702; MR-Egger, Q = 11.643, P = 0.768) and pleiotropy (Intercept, 0.040, P = 0.191) (Supplementary Table S6). Similarly, we utilized a leave-one-out analysis and funnel plots to supplement the reliability of the results (Figures 4C, D).

Validation of bidirectional MR analysis

To verify the robustness of the results of the bidirectional MR analysis described above, we also used the SNPs of 8021 European PBC patients extracted from a GWAS meta-analysis for analysis. The results of the validation cohort were consistent with previous studies and supported a bidirectional causal relationship between PBC and SLE, as shown in Supplementary Tables S7-12, Supplementary Figures S3, 4.

Multivariable MR analysis

Within the MVMR, significant direct causal associations were observed between PBC and an increased risk of SLE (OR = 1.34126,

FIGURE 3
Causal effect of PBC on the risk of SLE. **(A)** Scatter plot. **(B)** Forest plot. **(C)** Leave-one-out test. **(D)** Funnel plot.

FIGURE 4
Causal effect of SLE on the risk of PBC. **(A)** Scatter plot. **(B)** Forest plot. **(C)** Leave-one-out test. **(D)** Funnel plot.
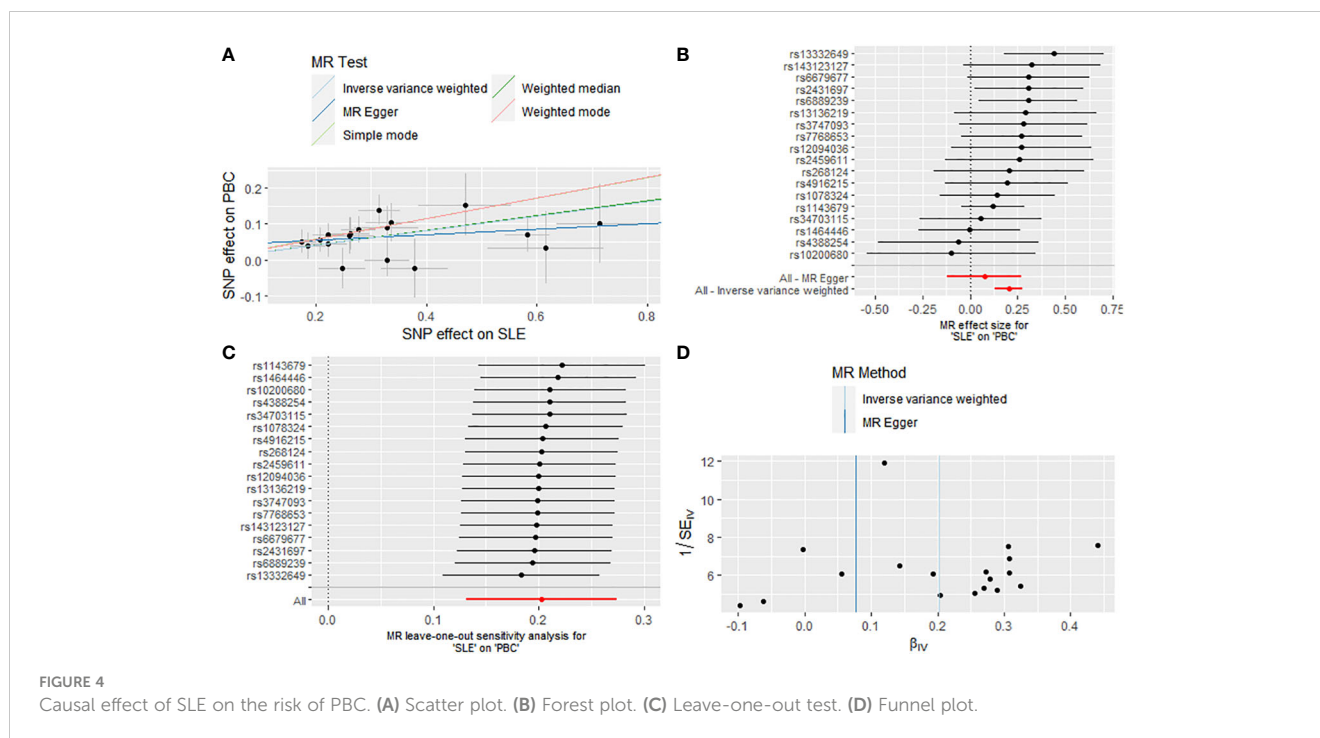
95% CI: 1.23731 - 1.45393, P < 0.001). Conversely, SLE was also directly associated with an increased risk of PBC (OR = 1.19932, 95% CI: 1.12629 - 1.27708, P < 0.001). BMI, smoking and alcohol consumption did not confound the results (Supplementary Table S13).

Identification of differentially expressed genes

In the SLE dataset (GSE65391), we discovered 328 DEGs, with 217 upregulated and 111 downregulated DEGs. Similarly, in the PBC dataset (GSE119600), we identified 7096 DEGs, comprising 3628 upregulated and 3468 downregulated DEGs. Heatmaps (Figures 5A, B) showcased the top 30 DEGs for both diseases. Notably, we found 107 DEGs overlapping between SLE and PBC (Figure 5C). Conducting GO enrichment analysis on these 107 genes revealed their significant involvement in immune-related processes, such as immune effector activity, leukocyte activation, cell activation, immune response, and interspecies interaction between organisms (Figure 5D).

Correlation analysis between modules and clinical traits with WGCNA

Sample clustering was performed to identify aberrant samples. No abnormal samples were detected in GSE65391, whereas 16 outlier samples were removed from GSE119600 (Figures 6A, B). To achieve a scale-free network, we evaluated the scale-free fit index and average connectivity. The soft threshold for GSE65391 was determined as $\beta=8$, while for GSE119600, β was set to 12. The co-

expression network analysis identified ten modules for SLE samples and seven for PBC samples (Figures 6C, D).

To explore disease progression-related genes, we assessed the relationship between modules and clinical phenotypes, enabling us to uncover potential genetic associations. In the GSE65391 dataset for SLE, the pink module exhibited the most potent positive correlation (r = 0.37, P < 0.001), whereas the turquoise module showed the most significant negative correlation (r = -0.27, P < 0.001). Similarly, for PBC in the GSE119600 dataset, the blue module displayed the most potent positive correlation (r = 0.35, P < 0.001), while the turquoise module had the most significant negative correlation (r = -0.41, P < 0.001). By overlapping the hub modules of PBC and SLE using a Venn diagram, we identified 112 shared genes (Figure 6E).

GO enrichment analysis of these 112 genes revealed their involvement in pathways such as cell activation, immune effector process, immune response, killing of cells of other organisms, and leukocyte activation (Figure 6F), similar to the pathways enriched by the DEGs.

Identification of potential shared genes

A set of eight genes (DDX60L, CEACAM1, PARP9, IL1RN, ABCA1, TMEM140, TNFSF10, IFI16) were identified as common genes that intersected between the genes identified through WGCNA and DEGs. These genes hold the potential as crucial links between the two diseases. To explore shared regulatory pathways, we performed GO and KEGG enrichment analyses on these eight genes. GO analysis revealed their association with defense response to virus, response to virus, regulation of innate immune responses, and interleukin-1 production (Figure 7A). Additionally, KEGG analysis indicated their
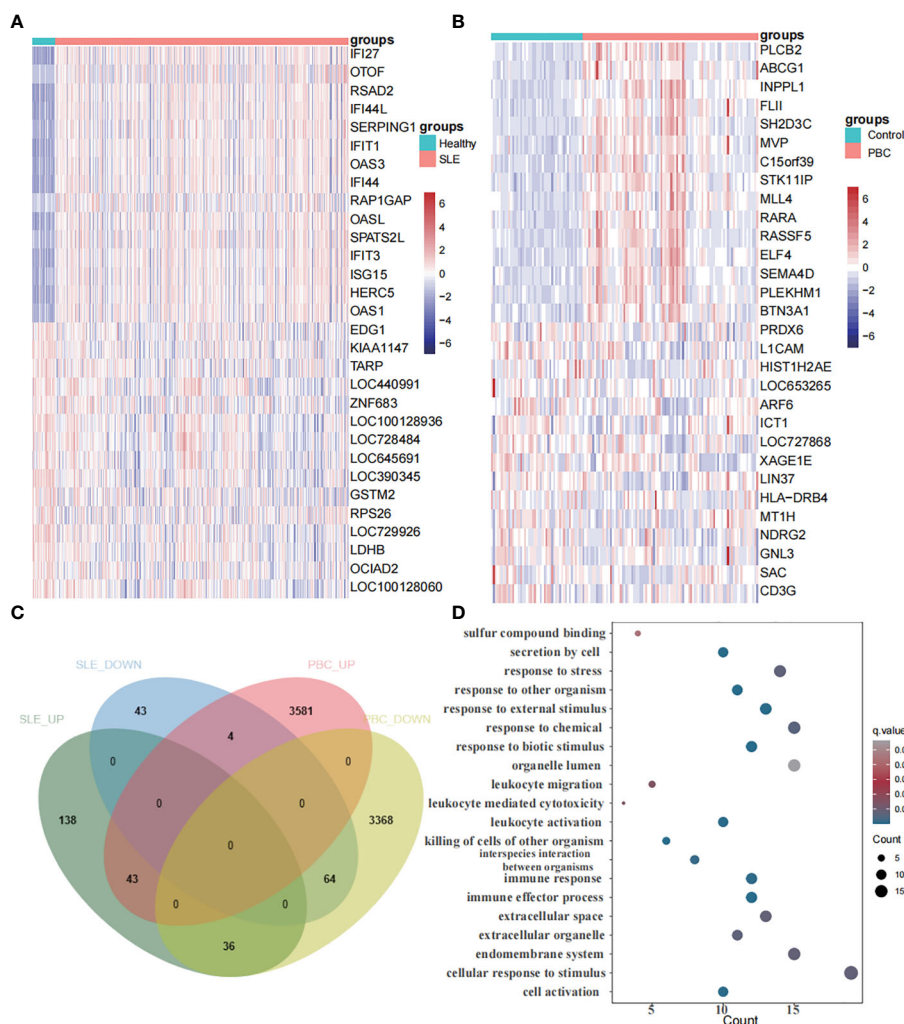
FIGURE 5
Identification of differentially expressed genes. **(A)** A heatmap of the top 30 DEGs in GSE65391. **(B)** A heatmap of the top 30 DEGs in GSE119600.
(C) Venn diagram shows that 107 genes overlap in the SLE and PBC. **(D)** Gene ontology enrichment analysis of these 107 genes.

involvement in pathways such as lipid and atherosclerosis, cytokine-cytokine receptor interaction, and natural killer cell-mediated cytotoxicity (Figure 7B).

Identification of shared diagnostic genes via LASSO

We then utilized the LASSO regression method to uncover potential shared diagnostic genes. In GSE65391, the LASSO examination pinpointed 4 of the 8 core intersecting genes with the optimal $\lambda = 0.00403$ (Supplementary Figures S5A, B). Similarly, in GSE119600, the LASSO analysis identified 7 of the 8 core intersecting genes with the most suitable $\lambda = 0.00342$ (Supplementary Figures S5C, D). As a result, we discovered 4 common genes (PARP9, ABCA1, CEACAM1, DDX60L) as the top shared diagnostic biomarkers for both PBC and SLE (Supplementary Figure S5E).

Evaluation of shared diagnostic gene expression levels

In Figures 8A, C, the expression levels of the 4 core genes in the two datasets are displayed. Remarkably, the genes PARP9, ABCA1, CEACAM1, and DDX60L were significantly upregulated in both SLE and PBC (P < 0.05). Furthermore, when scoring these four genes as a gene set for the two datasets, significant differences were observed between the two groups, affirming the potential of these four genes as shared diagnostic biomarkers for both diseases (Figures 8B, D).

Identification of shared diagnostic genes in PBMCs-specific cell populations of SLE patients

The single-cell transcriptomic analysis of PBMCs from SLE patients and healthy individuals revealed 14 distinct cell
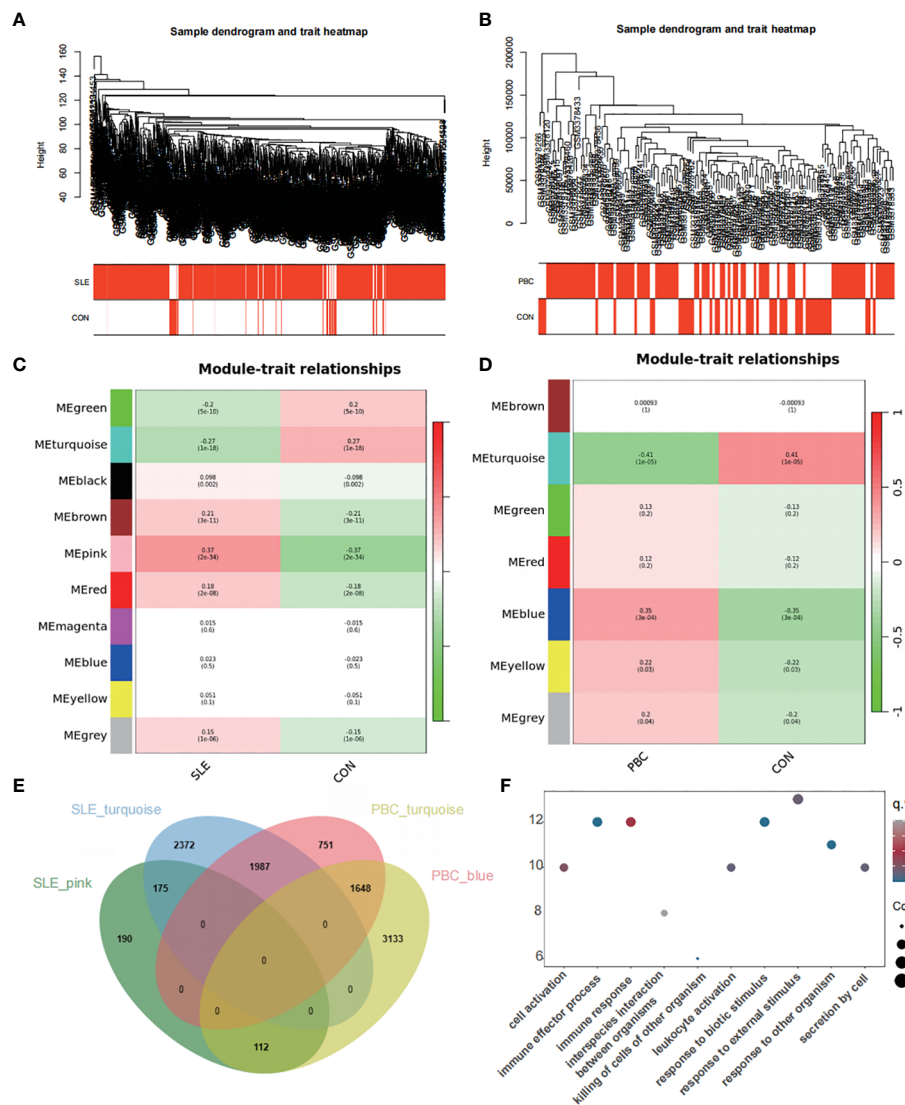
FIGURE 6
Co-expression network analysis for differentially expressed genes. (A) Sample dendrogram and trait heatmap in GSE65391 (B) Sample dendrogram and trait heatmap in GSE119600. (C) Heatmap of the module-trait relationships in GSE65391. (D) Heatmap of the module-trait relationships in GSE119600. (E) Venn diagram identifies 112 shared genes by overlapping the hub modules of PBC and SLE. (F) Gene ontology enrichment analysis of these 112 genes.

FIGURE 7
Functional enrichment analyses of the shared genes. (A) GO analysis of the shared genes. (B) KEGG pathway enrichment analysis of the shared genes.

FIGURE 8
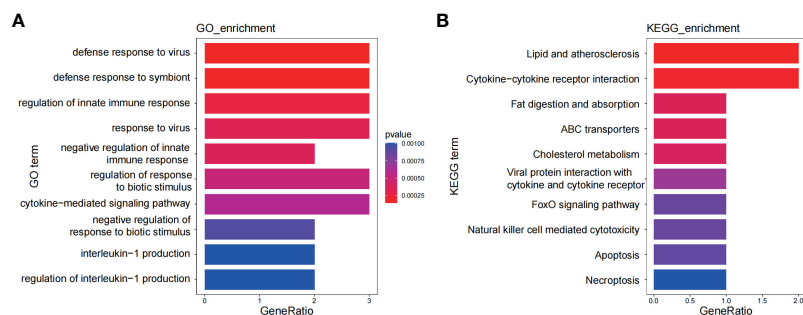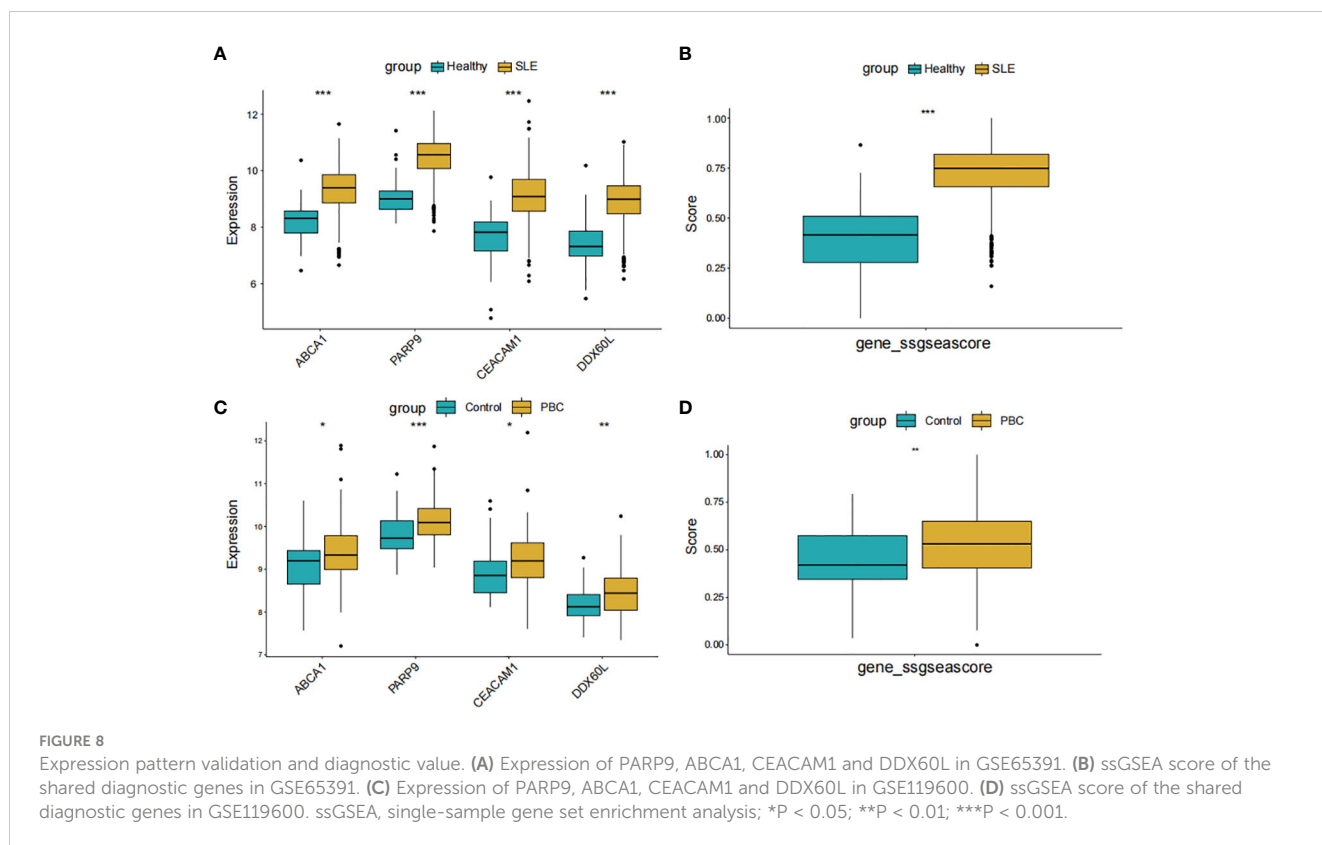Expression pattern validation and diagnostic value. **(A)** Expression of PARP9, ABCA1, CEACAM1 and DDX60L in GSE65391. **(B)** ssGSEA score of the shared diagnostic genes in GSE65391. **(C)** Expression of PARP9, ABCA1, CEACAM1 and DDX60L in GSE119600. **(D)** ssGSEA score of the shared diagnostic genes in GSE119600. ssGSEA, single-sample gene set enrichment analysis; *P < 0.05; **P < 0.01; ***P < 0.001.
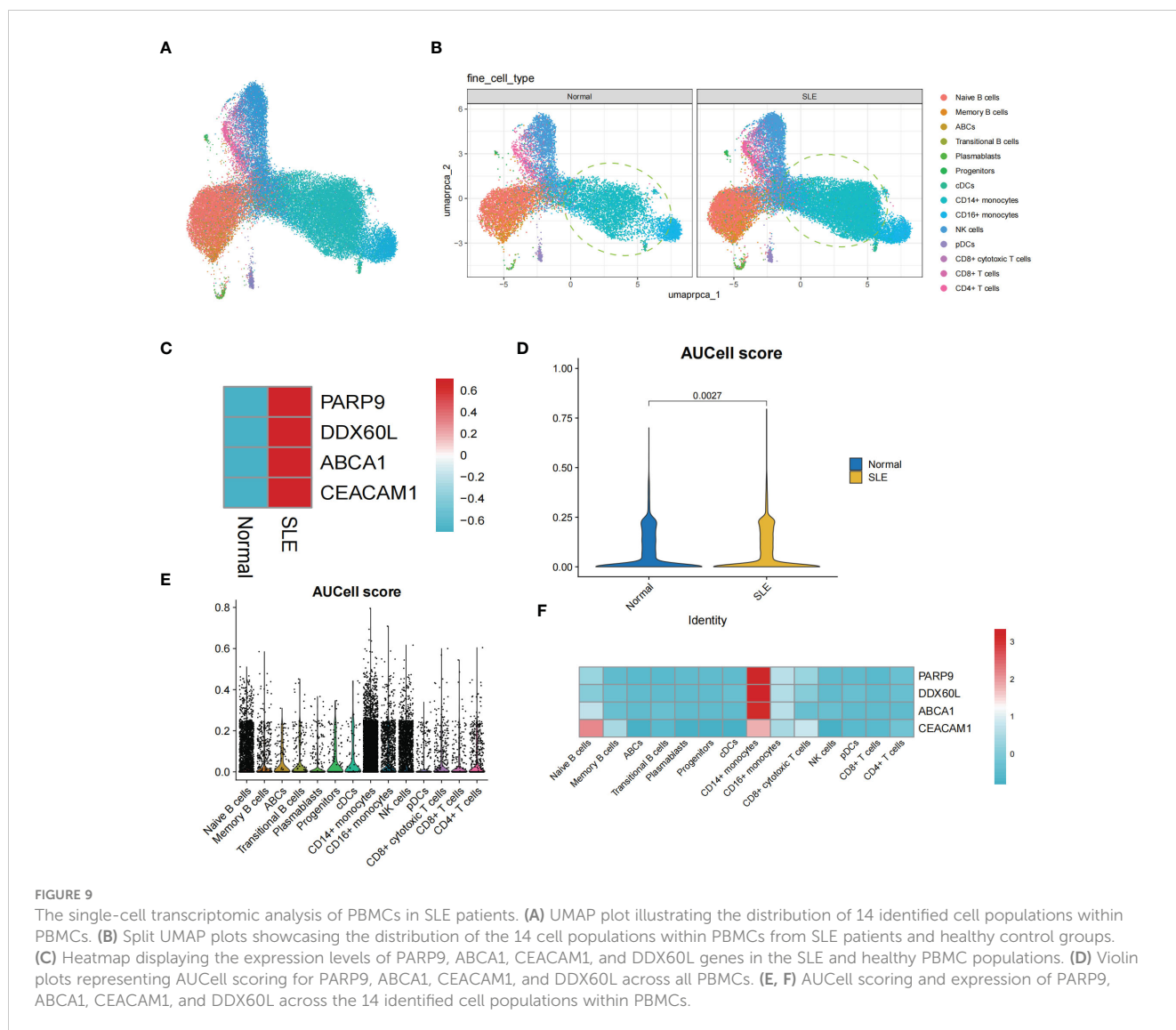
populations within the PBMCs (Figure 9A). Notably, we observed an increased presence of CD14+ monocytes in the PBMCs of SLE patients compared to healthy individuals (Figure 9B). Upon conducting an expression analysis of the identified Shared Diagnostic Genes (PARP9, ABCA1, CEACAM1, DDX60L) in the PBMCs of both SLE patients and healthy individuals, it was found that these genes exhibited higher expression levels in the PBMCs of SLE patients (Figure 9C). This finding underscores the potential feasibility of these four genes as shared diagnostic biomarkers. Further, after performing aucell scoring using these four genes on the PBMCs, it was evident that their expression was significantly higher in SLE patients compared to healthy individuals (P < 0.05, Figure 9D). Most compellingly, a more detailed analysis revealed that the expression of PARP9, ABCA1, CEACAM1, DDX60L was notably elevated in the CD14+ monocyte population compared to other cell groups within the PBMCs (Figures 9E, F).

Discussion

This study conducted a bidirectional MR analysis to investigate the causal relationship between PBC and SLE and integrated transcriptomic data from PBC and SLE to explore shared diagnostic genes between the two diseases. The results revealed a bidirectional causal relationship between PBC and SLE. Specifically, it was observed that PBC increased the risk of SLE (IVW, OR: 1.347, 95% CI: 1.276-1.422, p<0.001), and conversely, SLE could also increase the risk of PBC (IVW, OR: 1.225, 95% CI: 1.141-1.315, p<0.001). In addition, transcriptomic analyses identified PARP9, ABCA1, CEACAM1, and DDX60L as promising diagnostic biomarkers for PBC and SLE. These genes are highly expressed in CD14+ monocytes in PBMCs of SLE patients and may be associated with innate immune responses and immune activation.

The relationship between PBC and SLE is receiving increasing attention, and several retrospective and descriptive studies have summarized the coexistence of PBC and SLE. Overall, the current studies suggest that the prevalence of SLE is higher in patients with PBC than in the control population (12, 14). The two diseases have some common clinical features, and their coexistence may complicate organ involvement and worsen clinical prognosis (16–19). However, the exact causal relationship between the two diseases is uncertain. This study utilized MR analysis to provide evidence for a bidirectional causal relationship between PBC and SLE from a genetic perspective. At the same time, we validated the results of the MR analysis using another set of datasets, further confirming the reliability of the results. A retrospective study has observed a higher prevalence of SLE in patients with PBC than in the control population and a higher prevalence of SLE in relatives of patients with PBC, suggesting a possible genetic link between the two diseases (14). In addition, previous GWAS studies have also identified the IRF5-TNPO3 gene haplotype locus, which is shared by PBC and SLE (30, 31). Recently, a GWAS study has revealed that CD58 is a shared genetic susceptibility locus between SLE and PBC, and rs10924104 is associated with the regulating of CD58

FIGURE 9
The single-cell transcriptomic analysis of PBMCs in SLE patients. **(A)** UMAP plot illustrating the distribution of 14 identified cell populations within PBMCs. **(B)** Split UMAP plots showcasing the distribution of the 14 cell populations within PBMCs from SLE patients and healthy control groups. **(C)** Heatmap displaying the expression levels of PARP9, ABCA1, CEACAM1, and DDX60L genes in the SLE and healthy PBMC populations. **(D)** Violin plots representing AUCell scoring for PARP9, ABCA1, CEACAM1, and DDX60L across all PBMCs. **(E, F)** AUCell scoring and expression of PARP9, ABCA1, CEACAM1, and DDX60L across the 14 identified cell populations within PBMCs.

expression and the intensity of autoimmune disease susceptibility (32). These studies provide indirect support for our findings of a genetic association between PBC and SLE.

This study integrated PBC and SLE transcriptome data and identified PARP9, ABCA1, CEACAM1 and DDX60L as potential diagnostic biomarkers. Single-cell analysis detected increased expression of these four genes in PBMCs from SLE patients compared to healthy individuals, mainly in CD14+ monocyte subsets. Notably, CD14+ monocytes were also validated as the dominant cell population in PBMCs from SLE patients, suggesting that CD14+ monocytes in PBMCs may be a key cell population for diagnosing and understanding the common pathogenesis of SLE and PBC. Previous studies have reported that in European and Asian SLE populations, expression levels of common SLE GWAS SNPs are significantly increased in CD14+ monocytes, with significant enrichment for DNase I hypersensitive sites ((DHSs). DHSs are important sites for the regulation of gene transcription and are generally poorly methylated. During the pathogenesis of SLE, CDHs in CD14+ monocytes are mainly related to the regulation of the innate immune responses of and

the activation of immune effects (33). Type I interferon (IFN-1) plays a crucial role in the development of SLE (34). Studies have observed that there are highly expressed gene transcripts that induce IFN-1 production in monocytes of SLE patients, and these IFN-1 regulatory genes have generally low methylation levels, including the PARP9 (35, 36). In PBC, it has been reported that the number of CD14+ monocytes is increased in PBC and that this is associated with α-SMA-positive myofibroblasts in liver tissue (37). However, the expression of these four diagnostic genes in PBC CD14+ monocytes cannot be further investigated due to the current lack of single-cell transcriptomic data, and we will continue to investigate this area.

PARP9 is a protein-coding gene primarily involved in DNA damage repair, transcriptional regulation and immune responses, including IFN-mediated antiviral defenses (38). PARP9 has been reported to recognize and bind viral RNA, activate the PI3K/AKT pathway, and significantly contribute to the initiation and amplification of IFN-1 generation (39). IFN-1 is an important mediator in the pathogenesis of SLE, and viral infections are considered to be a causative factor in the pathogenesis (35).

Therefore, it cannot be excluded that the role of PARP9 in CD14+ monocytes from SLE patients is related to PARP9-mediated production of type I IFN in the initial immune cells. DDX60L belongs to the DExD/H-box RNA helicase family and is involved in RNA metabolism, innate immune responses, and antiviral defense (40). There are few studies on DDX60L and one study have reported that it may be an interferon-stimulated gene product and may enhance IFN-1 production (41). So far, there has been limited research on the role of PARP9 and DDX60L in PBC. ABCA1 is mainly involved in cholesterol and lipid metabolism (42). but it has been reported that ABCA1 is also implicated in phagocytosing apoptotic cells (43–45). In SLE peripheral blood, monocytes and their differentiated macrophages have high capacity to phagocytose apoptotic cells (46). Whether the increased expression of ABCA1 in CD14+ monocytes of SLE patients is related to this physiological process requires further investigation. Interestingly, the pathogenesis of PBC is associated with phagocytosis by apoptotic cholangiocytes (47). The expression of ABCA1 is significantly increased in liver tissue (48), and whether this is directly related to the phagocytosis of apoptotic cholangiocytes remains unclear. CEACAM1 is a signaling receptor. In monocytes, it regulates monocyte survival and differentiation through the ERK/MEK and PI3K/Akt pathways (49). In addition, CEACAM1 can also promote the initiation of T-cell responses (50). CEACAM1 expression is upregulated in bile duct epithelial cells from PBC patients (51), but the exact mechanism of action needs to be verified.

Nevertheless, it is vital to acknowledge the constraints of our study. Firstly, our study only involved individuals of European descent to better control for confounding introduced by heterogeneity. Thus, the applicability of the findings may have some limitations. Secondly, genetic susceptibility and epigenetic modifications contribute to the development of autoimmune diseases, and our study did not specifically address the role of epigenetics in gene expression. Therefore, our conclusions may require further validation and refinement in future studies. Lastly, although our screening process found some shared diagnostic genes for these two diseases, more mechanistic studies are required to verify these results and prove their biological correlation. In the future, we will simultaneously collect clinical patients for further study.

Conclusion

Our study confirmed the bidirectional causal relationship between PBC and SLE, and identified PARP9, ABCA1, CEACAM1, and DDX60L genes as the most potential shared diagnostic genes between the two diseases, providing insights for exploring the underlying mechanisms of those diseases.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The human participants included in this article were gathered from multiple previous studies. This study utilized comprehensive GWAS datasets rather than individual-level data, making ethical approval irrelevant.

Author contributions

TT: Writing – original draft. AT: Writing – original draft. LL: Writing – original draft. JY: Writing – original draft. LW: Writing – original draft. LZ: Writing – review & editing. JC: Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1270401/full#supplementary-material

References

1. Durcan L, O'Dwyer T, Petri M. Management strategies and future directions for systemic lupus erythematosus in adults. *Lancet (London England)*. (2019) 393:2332–43. doi: 10.1016/S0140-6736(19)30237-5

2. Kaul A, Gordon C, Crow MK, Touma Z, Urowitz MB, van Vollenhoven R, et al. Systemic lupus erythematosus. *Nat Rev Dis Primers*. (2016) 2:16039. doi: 10.1038/nrdp.2016.39

3. Xu L, Zhao J, Sun Q, Xu X, Wang L, Liu T, et al. Loss-of-function variants in SAT1 cause X-linked childhood-onset systemic lupus erythematosus. *Ann Rheumatic Dis*. (2022) 81:1712–21. doi: 10.1136/ard-2022-222795

4. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, Kelly JA, Comeau ME, Marion MC, et al. Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat Commun*. (2017) 8:16021. doi: 10.1038/ncomms16021

5. Nachbar F, Korting HC, Hoffmann RM, Kollmann M, Meurer M. Unusual coexistence of systemic lupus erythematosus and primary biliary cirrhosis. *Dermatology*. (1994) 188:313–7. doi: 10.1159/000247174

6. Beisel C, Weiler-Normann C, Teufel A, Lohse AW. Association of autoimmune hepatitis and systemic lupus erythematodes: a case series and review of the literature. *World J Gastroenterol*. (2014) 20:12662–7. doi: 10.3748/wjg.v20.i35.12662

7. Adiga A, Nugent K. Lupus hepatitis and autoimmune hepatitis (Lupoid hepatitis). *Am J Med Sci*. (2017) 353:329–35. doi: 10.1016/j.amjms.2016.10.014

8. Corpechot C, Chretien Y, Chazouilleres O, Poupon R. Demographic, lifestyle, medical and familial factors associated with primary biliary cirrhosis. *J Hepatol*. (2010) 53:162–9. doi: 10.1016/j.jhep.2010.02.019

9. Carey EJ, Ali AH, Lindor KD. Primary biliary cirrhosis. *Lancet (London England)*. (2015) 386:1565–75. doi: 10.1016/S0140-6736(15)00154-3

10. Lleo A, Wang GQ, Gershwin ME, Hirschfield GM. Primary biliary cholangitis. *Lancet (London England)*. (2020) 396:1915–26. doi: 10.1016/s0140-6736(20)31607-x

11. Faisal MS, Gonzalez HC, Gordon SC. Primary biliary cholangitis: epidemiology, diagnosis, and presentation. *Clinics Liver Dis*. (2024) 28:63–77. doi: 10.1016/j.cld.2023.06.005

12. Wang L, Zhang FC, Chen H, Zhang X, Xu D, Li YZ, et al. Connective tissue diseases in primary biliary cirrhosis: a population-based cohort study. *World J Gastroenterol*. (2013) 19:5131–7. doi: 10.3748/wjg.v19.i31.5131

13. Floreani A, Franceschet I, Cazzagon N, Spinazze A, Buja A, Furlan P, et al. Extrahepatic autoimmune conditions associated with primary biliary cirrhosis. *Clin Rev Allergy Immunol*. (2015) 48:192–7. doi: 10.1007/s12016-014-8427-x

14. Gershwin ME, Selmi C, Worman HJ, Gold EB, Watnik M, Utts J, et al. Risk factors and comorbidities in primary biliary cirrhosis: a controlled interview-based study of 1032 patients. *Hepatol (Baltimore Md)*. (2005) 42:1194–202. doi: 10.1002/hep.20907

15. Wang CR, Tsai HW. Autoimmune liver diseases in systemic rheumatic diseases. *World J Gastroenterol*. (2022) 28:2527–45. doi: 10.3748/wjg.v28.i23.2527

16. Fan X, Men R, Ni P, Lu C, Si T, Ma Y, et al. Concomitant systemic lupus erythematosus might have a negative impact on the biochemical responses to treatment in patients with primary biliary cholangitis. *Clin Rheumatol*. (2020) 39:795–803. doi: 10.1007/s10067-019-04853-2

17. Efe C, Purnak T, Ozaslan E, Ozbalkan Z. The importance of autoimmune hepatitis and primary biliary cirrhosis in patients with systemic lupus erythematosus. *Lupus*. (2011) 20:112. doi: 10.1177/0961203310389098

18. Ahmad A, Heijke R, Eriksson P, Wirestam L, Kechagias S, Dahle C, et al. Autoantibodies associated with primary biliary cholangitis are common among patients with systemic lupus erythematosus even in the absence of elevated liver enzymes. *Clin Exp Immunol*. (2021) 203:22–31. doi: 10.1111/cei.13512

19. Cheng C, Wang Z, Wang L, Zhao J, Wang Q, Tian X, et al. Clinical characteristics and prognosis of concomitant systemic lupus erythematosus and primary biliary cholangitis. *Clin Rheumatol*. (2021) 40:1819–26. doi: 10.1007/s10067-020-05457-x

20. Smith GD, Timpson N, Ebrahim S. Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. *Ann Med*. (2008) 40:524–41. doi: 10.1080/07853890802010709

21. Smith GD. Mendelian randomization for strengthening causal inference in observational studies: application to gene x environment interactions. *Perspect Psychol Sci*. (2010) 5:527–45. doi: 10.1177/1745691610383505

22. Wang C, Wu W, Yang H, Ye Z, Zhao Y, Liu J, et al. Mendelian randomization analyses for PCOS: evidence, opportunities, and challenges. *Trends Genet*. (2022) 38:468–82. doi: 10.1016/j.tig.2022.01.005

23. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. (2014) 23:R89–98. doi: 10.1093/hmg/ddu328

24. Sekula P, Del Greco MF, Pattaro C, Kottgen A. Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol*. (2016) 27:3253–65. doi: 10.1681/ASN.2016010098

25. Richmond RC, Davey Smith G. Mendelian randomization: concepts and scope. *Cold Spring Harb Perspect Med*. (2022) 12. doi: 10.1101/cshperspect.a040501

26. Cordell HJ, Han Y, Mells GF, Li Y, Hirschfield GM, Greene CS, et al. International genome-wide meta-analysis identifies new primary biliary cirrhosis risk loci and targetable pathogenic pathways. *Nat Commun*. (2015) 6:8019. doi: 10.1038/ncomms9019

27. Cordell HJ, Fryett JJ, Ueno K, Darlay R, Aiba Y, Hitomi Y, et al. An international genome-wide meta-analysis of primary biliary cholangitis: Novel risk loci and candidate drugs. *J Hepatol*. (2021) 75:572–81. doi: 10.1016/j.jhep.2021.04.055

28. Bentham J, Morris DL, Graham DSC, Pinder CL, Tombleson P, Behrens TW, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet*. (2015) 47:1457–64. doi: 10.1038/ng.3434

29. Boef AG, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int J Epidemiol*. (2015) 44:496–511. doi: 10.1093/ije/dyv071

30. Kottyan LC, Zoller EE, Bene J, Lu X, Kelly JA, Rupert AM, et al. The IRF5-TNPO3 association with systemic lupus erythematosus has two components that other autoimmune disorders variably share. *Hum Mol Genet*. (2015) 24:582–96. doi: 10.1093/hmg/ddu455

31. Liu X, Invernizzi P, Lu Y, Kosoy R, Lu Y, Bianchi I, et al. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet*. (2010) 42:658–60. doi: 10.1038/ng.627

32. Hitomi Y, Ueno K, Aiba Y, Nishida N, Kawai Y, Kawashima M, et al. rs10924104 in the expression enhancer motif of CD58 confers susceptibility to human autoimmune diseases. *Hum Genet*. (2024) 143:19–33. doi: 10.1007/s00439-023-02617-2

33. Liu L, Yin X, Wen L, Yang C, Sheng Y, Lin Y, et al. Several critical cell types, tissues, and pathways are implicated in genome-wide association studies for systemic lupus erythematosus. *G3 (Bethesda Md)*. (2016) 6:1503–11. doi: 10.1534/g3.116.027326

34. Psarras A, Wittmann M, Vital EM. Emerging concepts of type I interferons in SLE pathogenesis and therapy. *Nat Rev Rheumatol*. (2022) 18:575–90. doi: 10.1038/s41584-022-00826-z

35. Crow MK. Pathogenesis of systemic lupus erythematosus: risks, mechanisms and therapeutic targets. *Ann Rheumatic Dis*. (2023) 82:999–1014. doi: 10.1136/ard-2022-223741

36. Ulff-Møller CJ, Asmar F, Liu Y, Svendsen AJ, Busato F, Grønbaek K, et al. Twin DNA methylation profiling reveals flare-dependent interferon signature and B cell promoter hypermethylation in systemic lupus erythematosus. *Arthritis Rheumatol (Hoboken NJ)*. (2018) 70:878–90. doi: 10.1002/art.40422

37. Leicester KL, Olynyk JK, Brunt EM, Britton RS, Bacon BR. Differential findings for CD14-positive hepatic monocytes/macrophages in primary biliary cirrhosis, chronic hepatitis C and nonalcoholic steatohepatitis. *Liver Int*. (2006) 26:559–65. doi: 10.1111/j.1478-3231.2006.01255.x

38. Zhang Y, Mao D, Roswit WT, Jin X, Patel AC, Patel DA, et al. PARP9-DTX3L ubiquitin ligase targets host histone H2BJ and viral 3C protease to enhance interferon signaling and control viral infection. *Nat Immunol*. (2015) 16:1215–27. doi: 10.1038/ni.3279

39. Xing J, Zhang A, Du Y, Fang M, Minze LJ, Liu YJ, et al. Identification of poly (ADP-ribose) polymerase 9 (PARP9) as a noncanonical sensor for RNA virus in dendritic cells. *Nat Commun*. (2021) 12:2681. doi: 10.1038/s41467-021-23003-4

40. Ullah R, Li J, Fang P, Xiao S, Fang L. DEAD/H-box helicases:Anti-viral and pro-viral roles during infections. *Virus Res*. (2022) 309:198658. doi: 10.1016/j.virusres.2021.198658

41. Grunvogel O, Esser-Nobis K, Reustle A, Schult P, Muller B, Metz P, et al. DDX60L is an interferon-stimulated gene product restricting hepatitis C virus replication in cell culture. *J Virol*. (2015) 89:10548–68. doi: 10.1128/JVI.01297-15

42. Yvan-Charvet L, Wang N, Tall AR. Role of HDL, ABCA1, and ABCG1 transporters in cholesterol efflux and immune responses. *Arterioscler Thromb Vasc Biol*. (2010) 30:139–43. doi: 10.1161/ATVBAHA.108.179283

43. Babashamsi MM, Koukhaloo SZ, Halalkhor S, Salimi A, Babashamsi M. ABCA1 and metabolic syndrome; a review of the ABCA1 role in HDL-VLDL production, insulin-glucose homeostasis, inflammation and obesity. *Diabetes Metab Syndrome*. (2019) 13:1529–34. doi: 10.1016/j.dsx.2019.03.004

44. Chen W, Li L, Wang J, Zhang R, Zhang T, Wu Y, et al. The ABCA1-efferocytosis axis: A new strategy to protect against atherosclerosis. *Clin Chim Acta*. (2021) 518:1–8. doi: 10.1016/j.cca.2021.02.025

45. AlFadhli S, Ghanem AA, Nizam R. Genome-wide differential expression reveals candidate genes involved in the pathogenesis of lupus and lupus nephritis. *Int J Rheumatic Dis*. (2016) 19:55–64. doi: 10.1111/1756-185x.12745

46. Uribe-Querol E, Rosales C. Phagocytosis: our current understanding of a universal biological process. *Front Immunol*. (2020) 11:1066. doi: 10.3389/fimmu.2020.01066

47. Trivedi PJ, Cullen S. Etiopathogenesis of primary biliary cirrhosis: an overview of recent developments. *Hepatol Int*. (2013) 7:28–47. doi: 10.1007/s12072-012-9362-7

48. Takeyama Y, Uehara Y, Anan A, Morihara D, Yokoyama K, Takata K, et al. Increased hepatic ABCA1 transporter is associated with hypercholesterolemia in a cholestatic rat model and primary biliary cholangitis patients. *Med Mol Morphol*. (2017) 50:227–37. doi: 10.1007/s00795-017-0166-7

49. Yu Q, Chow EM, Wong H, Gu J, Mandelboim O, Gray-Owen SD, et al. CEACAM1 (CD66a) promotes human monocyte survival via a phosphatidylinositol 3-kinase- and AKT-dependent pathway. *J Biol Chem*. (2006) 281:39179–93. doi: 10.1074/jbc.M608864200

50. Kammerer R, Stober D, Singer BB, Obrink B, Reimann J. Carcinoembryonic antigen-related cell adhesion molecule 1 on murine dendritic cells is a potent regulator of T cell stimulation. *J Immunol (Baltimore Md: 1950)*. (2001) 166:6537–44. doi: 10.4049/jimmunol.166.11.6537

51. Baba N, Kobashi H, Yamamoto K, Terada R, Suzuki T, Hakoda T, et al. Gene expression profiling in biliary epithelial cells of primary biliary cirrhosis using laser capture microdissection and cDNA microarray. *Trans Res*. (2006) 148:103–13. doi: 10.1016/j.trsl.2006.04.007