



OPEN

Exploring shared pathways and the shared biomarker ERRFI1 in Obstructive sleep apnoea and atherosclerosis using integrated bioinformatics analysis

Bowen Chen¹, Liping Dong¹, Jihua Zhang², Ying Hao¹, Weiwei Chi¹ & Dongmei Song^{1,2}✉

Obstructive sleep apnea (OSA) is an upper airway disorder occurring during sleep and is associated with atherosclerosis (AS). AS is a cardiovascular disease caused by environmental and genetic factors, with a high global mortality rate. This study investigated common pathways and potential biomarkers of OSA and AS. Microarray data were downloaded from the Gene Expression Omnibus (GEO) database and used to screen for differentially expressed genes (DEGs) in the OSA and AS datasets. A weighted gene co-expression network analysis (WGCNA) was used to identify the co-expression modules of OSA and AS. The least absolute shrinkage and selection operators (LASSO) were used to determine critical biomarkers. Immune cell infiltration analysis was used to investigate the correlation between immune cell infiltration and common biomarkers of OSA and AS. Results revealed that differentially expressed genes may be involved in inflammatory processes, chemokine signaling pathways, and molecular changes in cell adhesion. ERBB receptor feedback inhibitor 1 (ERRFI1) was the best-shared biomarker for OSA and AS. Immune infiltration analysis showed that ERRFI1 expression was correlated with immune cell changes. Changes in immune pathways, inflammatory processes, and cell adhesion molecules may underlie the pathogenesis of both diseases, and ERRFI1 may be a potential diagnostic marker for patients with OSA and AS.

OSA is a major public health problem affecting 5–20% of the general population, with a worldwide prevalence of approximately 100 million individuals aged 30–69 years¹. OSA is a condition of apnea and hypoventilation caused by the collapse of the upper airway during sleep, accompanied by snoring, disturbed sleep architecture, decreased oxygen saturation, and daytime sleepiness². Long-term OSA can lead to cardiovascular diseases, diabetes, depression, motor vehicle accidents, and workplace accidents, resulting in a considerable burden on the global healthcare system and economic burden³.

Atherosclerosis is a chronic inflammation of blood vessels triggered by a complex interaction of risk factors and arterial wall cells. Previous studies have shown that OSA can lead to atherosclerosis (AS). In a series of prospective studies, the incidence of coronary artery disease was found to be higher in patients with OSA (16.2%) than in those without OSA (5.4%)⁴. And the prevalence of atherosclerosis increases significantly with the severity of OSA, up to 42% in patients with moderate to severe OSA⁵. Recent evidence also suggests that individuals at high risk for OSA are more likely to develop coronary plaque⁶. OSA can cause atherosclerosis and can induce AS through the activation of inflammatory pathways including hypoxia-inducible factor, nuclear factor kappa-light-chain-enhancer of activated B cells pathway, Toll-like receptor 4 (TLR4), adhesion molecules, and tumor necrosis factor⁷. It also contributes to the development of AS by causing abnormal platelet aggregation⁸. Oxidative stress⁹ and abnormal glucolipid metabolism mediate the development of AS¹⁰. Previous studies have shown that a high-cholesterol diet (HCD) is the traditional induction modality for inducing AS in mice¹¹. Notably, OSA and a high-cholesterol diet (HCD) may cause AS by distinct mechanisms¹². The main pathological process of OSA is chronic intermittent hypoxia (CIH)¹³. Previous studies have reported that CIH is a weaker inducer of

¹Clinical Biobank, The First Hospital of Hebei Medical University, Shijiazhuang, China. ²Department of Otolaryngology, The First Hospital of Hebei Medical University, Shijiazhuang, China. ✉email: songdongmei@hebm.edu.cn

AS8 and there are important differences between CIH- and HCD-induced activation of inflammatory pathways. Histologically, CIH-induced atherosclerotic plaques did not have a necrotic core, whereas HCD-induced plaques have a typical necrotic core and a fibrous cap. Furthermore, HCD leads to the formation of macrophage foam cells, but CIH does not. These findings reflect the inherent differences in pathology and underlying mechanisms between the two types of atherosclerosis in terms of the nature of the pathology and underlying mechanisms. OSA and AS are closely related and mechanistically different, so it is important to study the mechanisms of the two diseases for early intervention and treatment. However, there is a paucity of research into the similarities and differences between the two diseases and further exploration is needed.

In this study, we collected data from the GEO database, then screened for differential genes, performed enrichment analysis, WGCNA analysis, Gene Set Enrichment Analysis (GSEA) analysis¹⁴, and LASSO¹⁵ regression models to screen for shared pathways and key biomarkers for OSA and AS, and analyzed the expression levels and diagnostic value of the genes, validated the dataset for validation, and examined monocytes from blood samples collected from patients. By these similar methods, Zhang et al.¹⁶ found that the MAPK signaling pathway may be associated with both pathogenesis of ankylosing spondylitis and ulcerative colitis and that poly(A) specific ribonuclease subunit PAN3 Gene (PAN3) may be a potential diagnostic marker for patients with ulcerative colitis complicated by ankylosing spondylitis. Similarly, Gao¹⁷ used these similar approaches to find that immune responses may be associated with both epilepsy and subarachnoid hemorrhage pathogenesis and that Matrix metalloproteinase-9 (MMP9) and Complement C3a Receptor 1 (C3aR1) may be potential diagnostic markers for subarachnoid hemorrhage complicated by epilepsy. Using these methods, we aim to explore possible shared pathways of action between OSA and AS and identify new possible diagnostic and therapeutic strategies for patients with AS secondary to OSA (Fig. 1).

Results

Differential genetic screening. A total of 566 DEGs were identified in patients with OSA and normal controls, including 369 upregulated and 197 downregulated genes, as shown in the volcano plot (Fig. 2A). A total of 2189 DEGs (see Supplementary Table S1 online) of patients with AS and normal controls were screened, identifying 1256 up-regulated and 933 downregulated genes, as shown in the volcano plot (Fig. 2B). The DEGs were analyzed using the Kyoto Encyclopedia of Gene and Genome (KEGG, <https://www.genome.jp/kegg/>) pathway enrichment analysis. KEGG analysis showed that both OSA and AS were enriched in the cell adhesion molecule pathways (Fig. 2C,D, Supplementary Table S2 online). The two GEGs were taken to intersect to obtain differentially expressed genes common to AS and OSA, with a total of 93 genes (Fig. 2E, Supplementary Table S3 online). The co-expressed differentially expressed genes were subjected to KEGG and Gene Ontology (GO) analysis (see Supplementary Table S4 online), and GO analysis (Fig. 3A) showed that a series of important processes were involved, myeloid leukocyte migration, leukocyte chemotaxis, and response to chemokine, etc., and in terms of KEGG (Fig. 3B), several important pathways are involved, such as Cytokine-cytokine receptor interaction, Chemokine signaling pathway, and these results strongly suggest that inflammatory processes are associated with both OSA and AS are related.

Construction and module analysis of weighted gene co-expression network analysis (WGCNA). WGCNA was used to identify differentially expressed co-expressed clusters of genes between OSA and AS, and to calculate the correlation between the combined modules and disease characterization. Two datasets, GSE135917 and GSE100927, were used for the WGCNA analysis (see Supplementary Table S5 online). Outliers were checked using sample clustering, and no samples were removed from either GSE135917 or GSE100927 (see Supplementary Figs. S1, S2 online). According to the approximate scale-free topology criterion, $\beta=6$ was chosen to determine the soft threshold in the OSA model, and $\beta=5$ was chosen to determine the soft threshold in the AS model (see Supplementary Figs. S3, S4 online). Cluster dendrogram of the co-expression of AS and OSA (Fig. 4A,C). After merging similar gene modules, four modules (Fig. 4B) were identified in the OSA model and five modules (Fig. 4D) were identified in the AS model. Among the OSA modules, the gray module had the strongest positive correlation with OSA ($R=0.39$), and the turquoise module had the strongest negative correlation with the occurrence of OSA ($R=-0.61$). The turquoise module positively correlated with the occurrence of AS ($R=0.72$) (see Supplementary Table S6 online).

Identifying shared genes and shared pathways. In total, 109 genes (Fig. 5A, see Supplementary Table S7 online) overlapped between the strongest positive and negative modules of OSA and AS, which may be associated with the pathogenesis of OSA and AS. Enrichment analysis was performed for these 109 genes. KEGG analysis indicated that these genes may be involved in chemokine signaling pathways and cell adhesion molecule pathways (Fig. 5B, see Supplementary Table S8 online). GO analysis showed that these genes were present in the extracellular matrix (Fig. 5C, see Supplementary Table S9 online). GSEA was then performed on OSA and AS samples and immune responses were found to be involved in common pathogenic processes (Fig. 5D,E). Therefore, we propose that the onset of OSA and AS is driven by a combination of immune responses and changes in cell adhesion molecules.

Potential shared diagnostic genes selection by least absolute shrinkage and selection operator. After taking the intersection of the Venn diagrams for DEGs and modular hub genes (Fig. 6A), 34 shared genes were identified (see Supplementary Table S10 online). To narrow the range of potentially shared diagnostic gene biomarkers among the 34 DEGs, we used the LASSO model (Fig. 6B,C,E,F, see Supplementary Table S11 online). Five shared key genes (Fig. 6D, see Supplementary Table S12 online) were mined using LASSO regression: ERRFI1, FNDC1, HLA-DRB1, HSD11B1, and PERP.

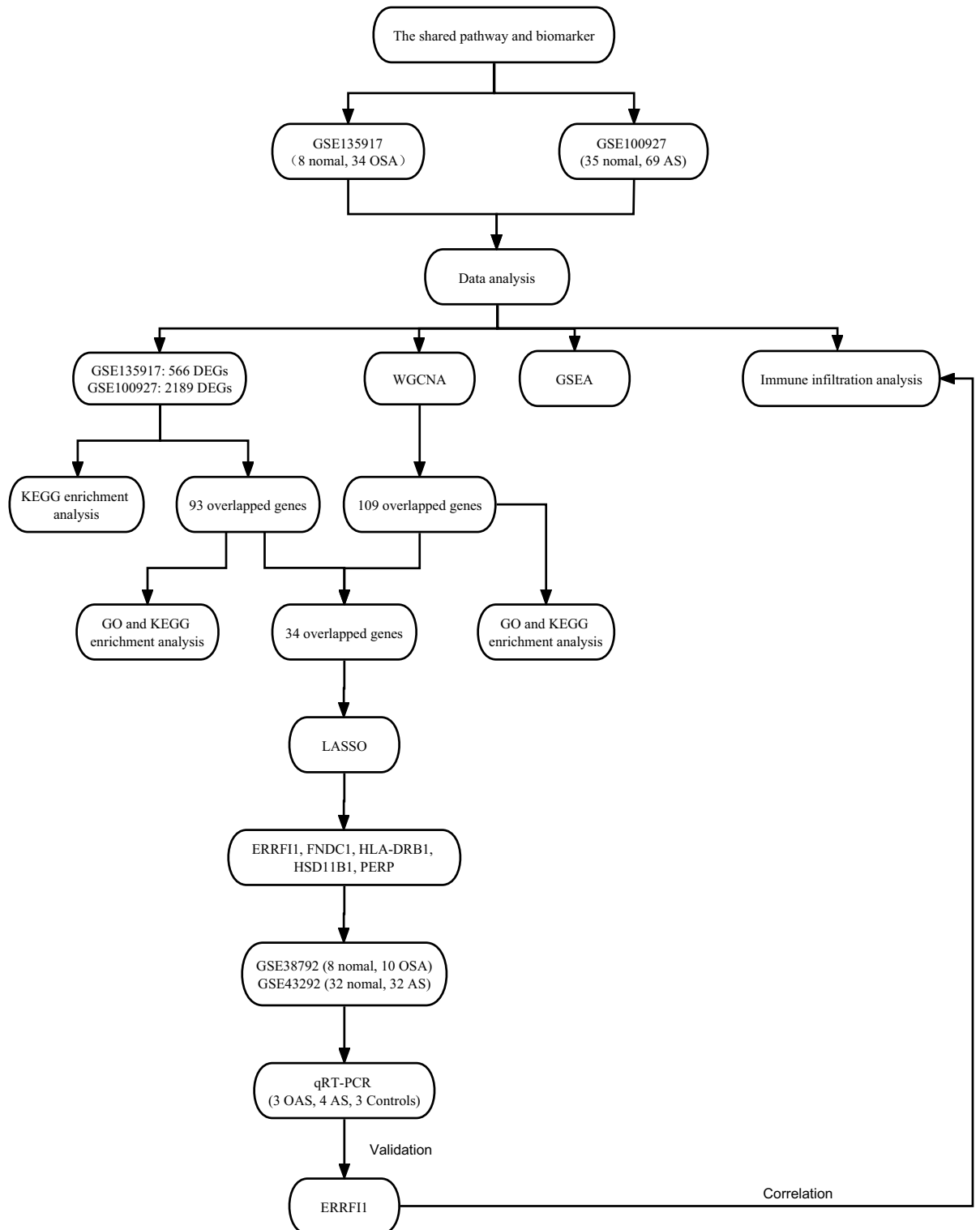


Figure 1. Schematic presentation of the analysis process.

Candidate biomarker expression levels and diagnostic value. Expression of the five key genes identified in the GSE135917 and GSE100927 datasets. Interestingly, differential expression analysis showed that ERRFI1 and FNDC1 were significantly under-expressed in OSA, and HLA-DRB1, HSD11B1, and PERP were significantly overexpressed in OSA (Fig. 7A–E). However, the FNDC1 and HLA-DRB1 expression was significantly elevated, and the ERRFI1, HSD11B1, and PERP expression were significantly decreased in AS (Fig. 7F–J). We also determined the diagnostic capabilities of these five shared markers based on receiver operating characteristic (ROC) analysis. In GSE135917, we obtained an area under the curve (AUC) > 0.6 for all five genes

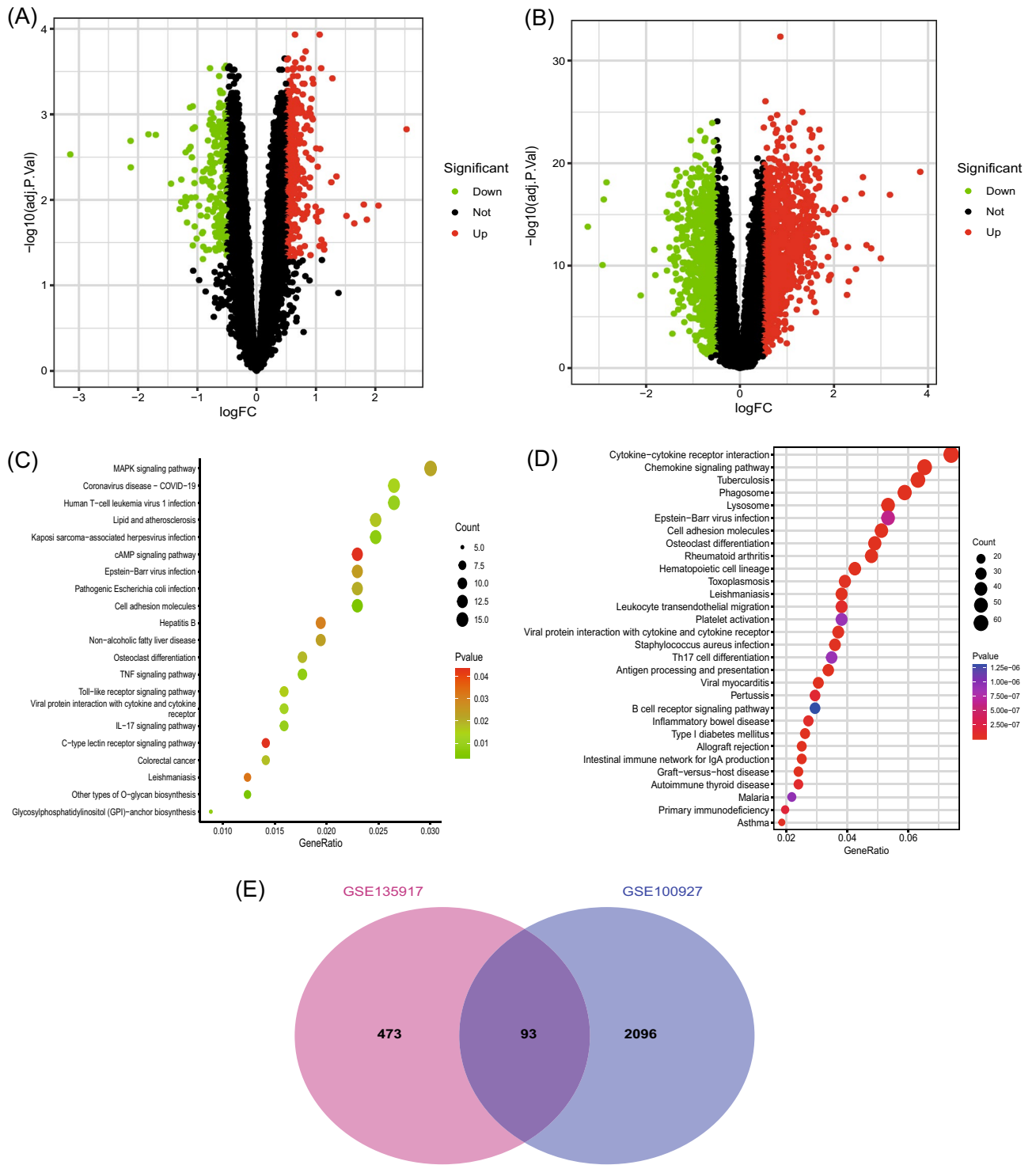


Figure 2. Differential genetic screening. (A) The volcano plots of DEGs in GSE135917 ($n = 566$, $p < 0.05$). (B) The volcano plots of DEGs in GSE100927 ($n = 2189$, $p < 0.05$). (C) The KEGG analysis of DEGs in GSE135917. (D) The KEGG analysis of DEGs in GSE100927. (E) Venn diagram of DEGs from GSE100927 and GSE135917.

(Fig. 7K–O). In GSE100927, the AUC of the five genes obtained was > 0.6 (Fig. 7P–T), indicating relatively satisfactory diagnostic efficiency.

Validation of external datasets and qRT-PCR results. We validated the expression of ERFF1, FNDC1, HLA-DRB1, HSD11B1, and PERP in OSA using the external validation dataset, GSE38792. Differential expression analysis showed significant differences ($p < 0.05$) in ERFF1 and HLA-DRB1, similar to the

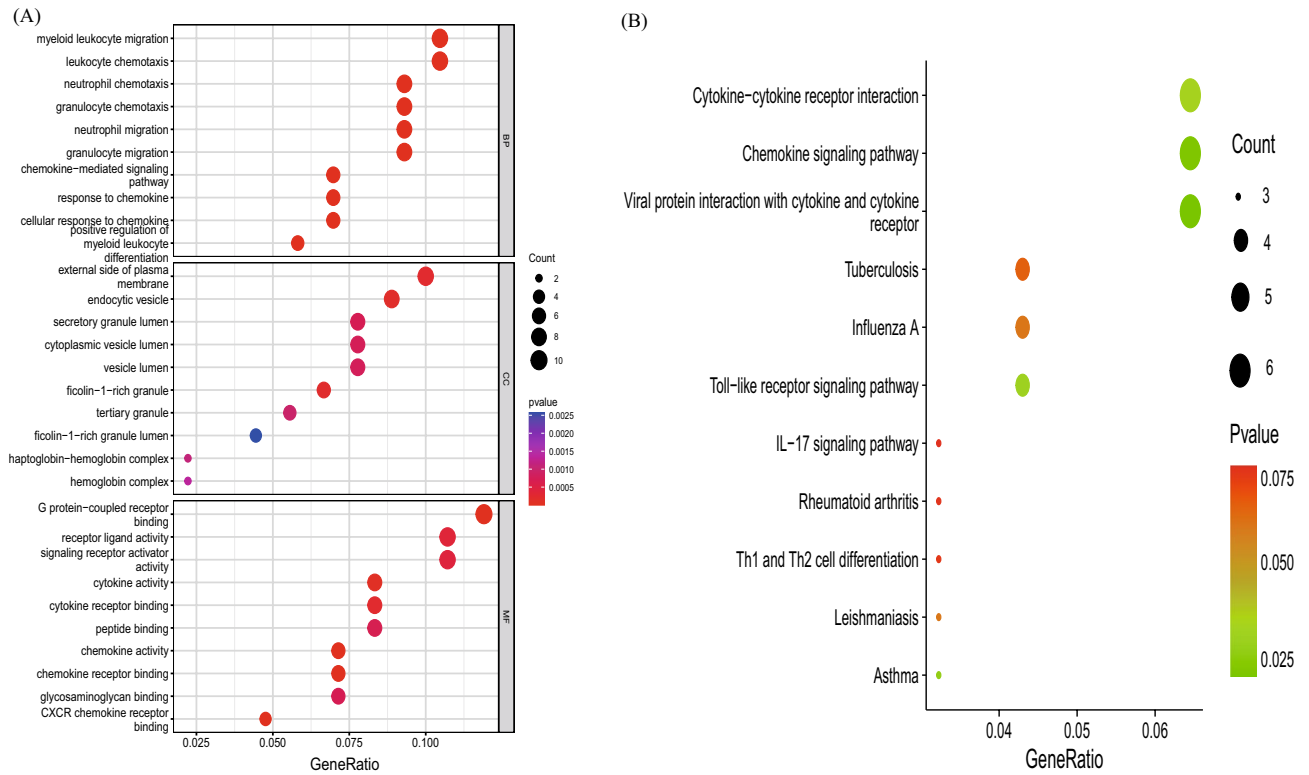


Figure 3. (A) The GO analysis of 93 genes. (B) The KEGG analysis of 93 genes.

above results (Fig. 8A–E); however, the remaining three genes were not statistically significant ($p > 0.05$) in the validation set. In another independent AS validation set, GSE43292, the expression levels of ERFFI1 and PERP differed significantly across samples, similar to the above results (Fig. 8F–J); however, the remaining three genes were not statistically significant ($p > 0.05$) in the validation set. Interestingly, there were significant differences ($p < 0.05$) in ERFFI1 expression in the OSA and AS groups compared to controls. For this differential gene, fresh whole blood samples were collected from ten patients, PBMC was extracted, and qRT-PCR analysis was performed to further validate the differential expression of ERFFI1 in the patient samples. The results showed reduced ERFFI1 expression in patients with OSA compared to normal subjects, and similarly reduced ERFFI1 expression in patients with AS ($p < 0.05$) (Fig. 8K,L), suggesting that ERFFI1 has the potential to be a shared diagnostic marker for both diseases.

Immune cell infiltration and its correlation with candidate biomarkers. This study investigated the differences in immune cell infiltration in different samples, taking into account the important role of the immune response in the development of OSA and AS. In GSE135917, 28 immune cells were identified and shown in the heat and violin plots (Fig. 9A,B), and in GSE100927 28 immune cells were identified and shown in the heat and violin plots (Fig. 9D,E). Additionally, the screened candidate biomarkers were found to be closely related to immune cells (Fig. 9C,F). ERFFI1 expression was significantly associated ($p < 0.05$) with T follicular helper cells, natural killer T cells, Memory B cells, macrophages, immature B cells, eosinophils, CD56dim natural killer cell, CD56brght natural killer cells, activated dendritic cells, and Activated CD8 T cells in OSA samples. In the AS samples, ERFFI1 expression was significantly associated ($p < 0.05$) with type 17 T helper cells, type 1 T helper cells, T follicular helper cells, regulatory T cells, natural killer T cells, natural killer cells, monocytes, MDSC, macrophages, immature dendritic cells, immature B cells, gamma delta T cells, effector memory CD4 T cells, central memory CD4 T cells, CD56dim natural killer cell, CD56brght natural killer cells, activated dendritic cells, activated CD8 T cells, activated CD4 T cells, and activated B cells.

Discussion

OSA has a high prevalence globally and is one of the main causes of atherosclerosis¹⁸. Previous studies have shown a close association between OSA and AS¹⁹ and have also identified mechanistic differences¹². We consider that OSA and AS share a common mechanism of response to systemic injury; however, there is little research on this mechanism. We, therefore, conducted a systematic biological approach in order to explore possible shared pathways of action in OSA and AS, as well as new possible potential diagnostic and therapeutic targets in OSA and AS, and found that OSA and AS may be driven by a combination of immune responses and molecular changes in cell adhesion, with ERFFI1 as a key shared gene steadily declining in OSA and AS patients with a well diagnostic role.

We obtained common differential genes by taking the intersection of differential genes from the OSA and AS datasets. Enrichment analysis of these common differential genes revealed the involvement of immune,

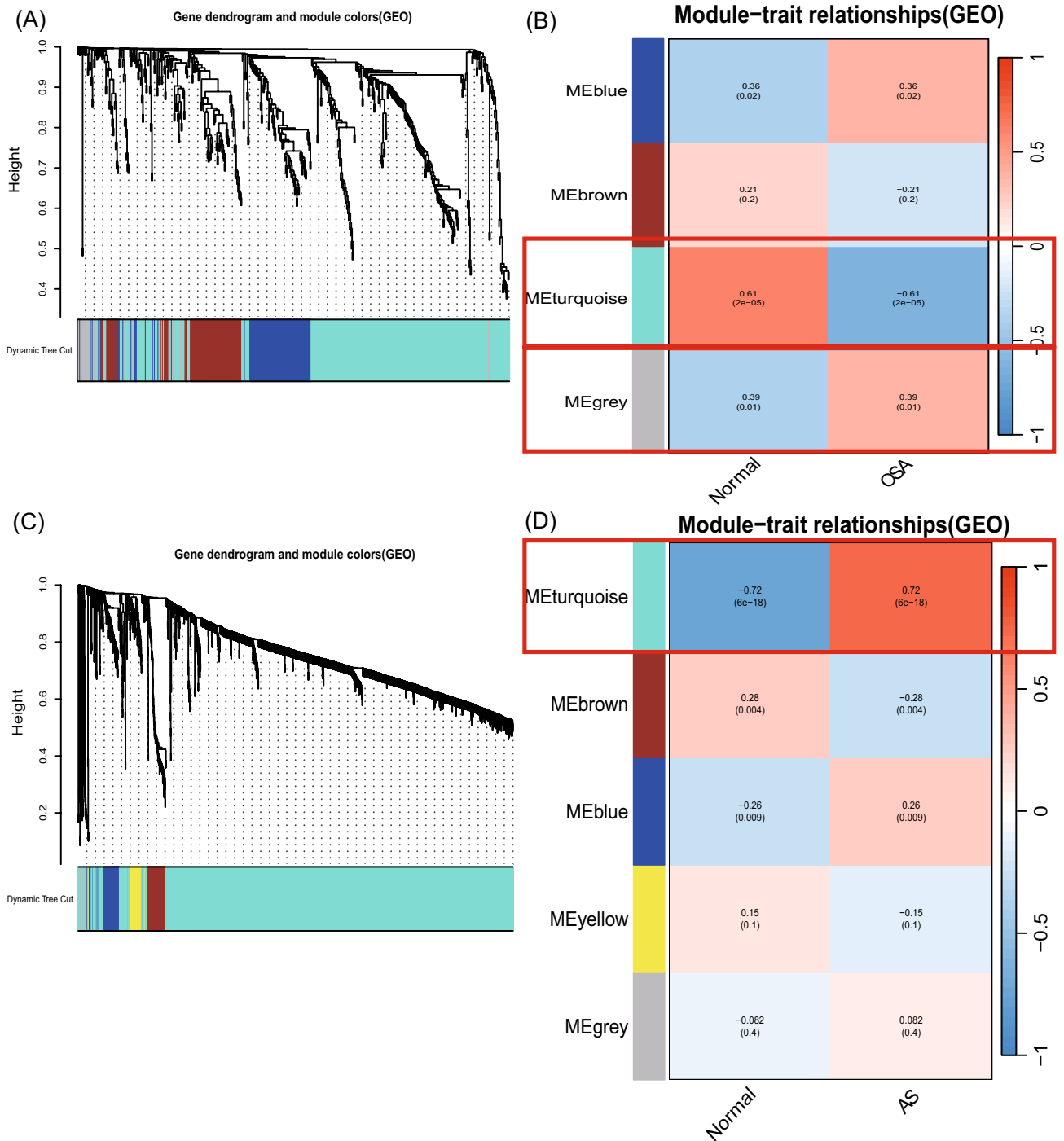


Figure 4. Construction and module analysis of weighted gene co-expression network (WGCNA). (A) The cluster dendrogram of co-expression in OSA. (B) Correlation between modules and clinical traits in OSA. (C) The cluster dendrogram of co-expression in AS. (D) Correlation between modules and clinical traits in AS.

inflammatory processes, and cell adhesion molecular pathways. Previous studies have shown the involvement of immune responses in the pathogenesis of both OSA and AS, with the presence of proinflammatory factors²⁰ and immune cell aggregation in atherosclerotic plaques²¹. Immune cells such as B lymphocytes, NK cells, and CD8 + / CD56 + cells are involved in the inflammatory process, leading to the development of cardiovascular complications in OSA^{22,23}. Cell adhesion molecules are molecules that cause adhesion between cells, between cells and the matrix, or between cell-matrix cells. They are involved in cell recognition, activation, signaling, cell proliferation, and differentiation and are the molecular basis of a series of important physiological and pathological processes, such as immune response, inflammatory response, and tumor metastasis. Intermittent hypoxia increases cell permeability by generating reactive oxygen species that cause VE-cadherin to be excluded²⁴. Potential role of

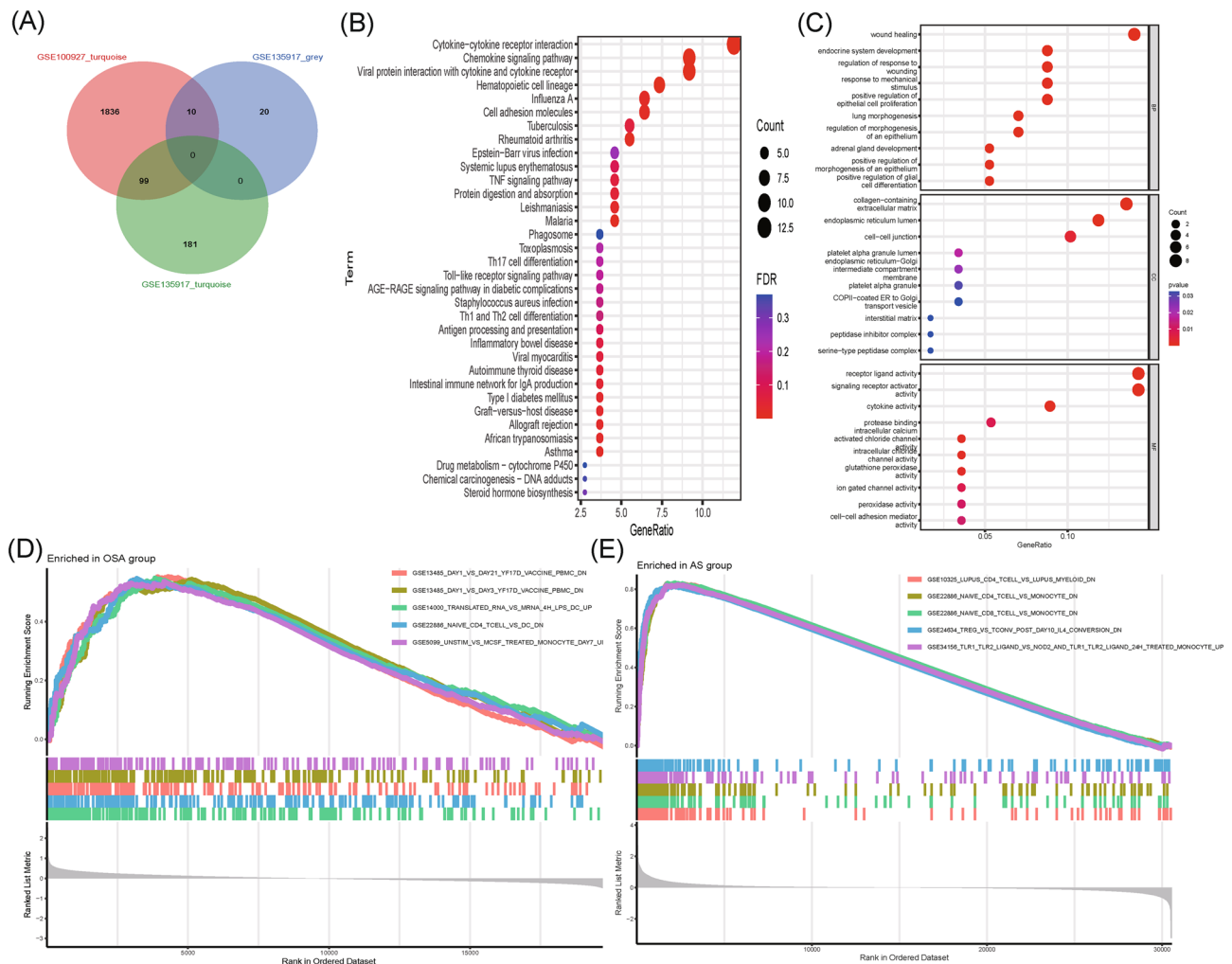


Figure 5. Identifying shared genes and shared pathways. **(A)** Venn diagram showing the overlap of 109 genes in the OSA and AS modules. **(B)** Bubble map of KEGG analysis of 109 shared genes between OSA and AS from the WGCNA screen. **(C)** Bubble map of GO analysis of 109 shared genes between OSA and AS screened by WGCNA. **(D)** Results of the Single-gene GSEA analyses of DEGs in OSA. **(E)** Results of the Single-gene GSEA analyses of DEGs in AS.

monocyte adhesion, chemotaxis, and macrophage polarization in the development of OSA-induced cardiovascular disease²⁵. OSA also acts on tumors through integrin-based adhesion²⁶. Similarly, cell adhesion factors play an important role in the disease progression. Cell adhesion molecules have been found to potentially contribute to atherosclerotic plaque rupture during AS progression of AS²⁷. Therefore, we consider that OSA and AS share common pathogenesis, involving immune responses, inflammatory processes, and cell adhesion factors.

A total of 109 genes overlapped in the strongest positive and negative modules of the WGCNA for OSA and AS. Thus, these genes may be involved in the pathogenesis of OSA and AS. We identified 34 shared genes based on the intersection of DEGs and modular hub genes using Venn diagrams. To narrow down the range of potentially shared diagnostic genes among the 34 DEGs, we used the LASSO model. Using LASSO regression, five shared key genes were mined: *ERRFI1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, and *PERP*. We then validated these five genes with a validation dataset, collected blood samples from patients with OSA and AS, and found that only the *ERRFI1* gene was stably decreased in OSA and AS. This is the first time that *ERRFI1* has been identified and reported as a potential diagnostic marker for these two diseases, which has never been addressed in previous studies. *ERRFI1* (also known as *Mig6*, *RALT*, or gene 33) is an adapter protein with complex functions in cell biology and human diseases. Interestingly, *ERRFI1* was found to play different roles in different cells and tissues²⁸. In human umbilical vein endothelial cells, *ERRFI1* exerts antiapoptotic and antiangiogenic effects by inhibiting *EGFR* signaling²⁹. Knockdown of *ERRFI1* in human lung microvascular endothelial cells promotes apoptosis³⁰, which induces endothelial proliferation of vascular smooth muscle cells and increased cell migration and proliferation of vascular smooth muscle cells³¹. *ERRFI1* can also act on *EGFR* expressed in intimal smooth muscle cells of human atherosclerotic plaques and plays a key role in the development of atherosclerosis. It has also been shown that smooth muscle cells in atherosclerosis show increased *EGFR* downstream signaling and *EGFR* phosphorylation³¹. Furthermore, *ERRFI1* reduces the production of inflammatory mediators³² and

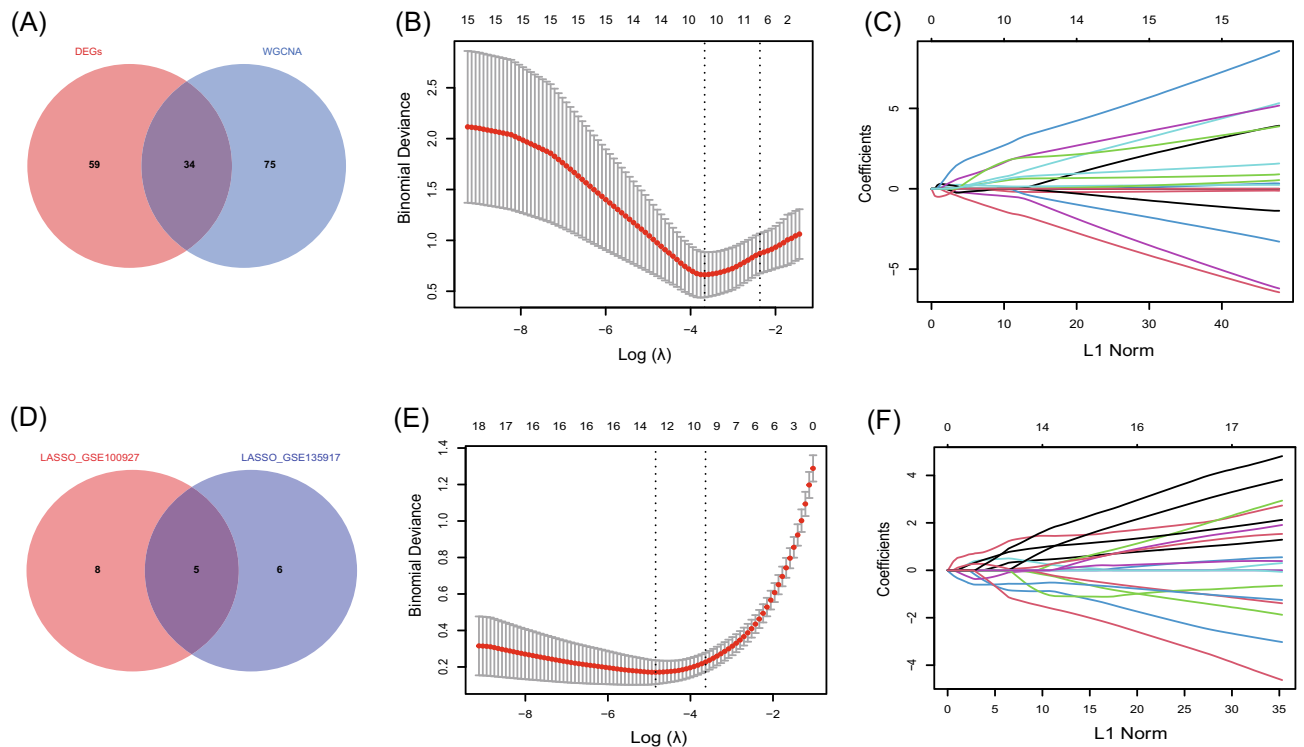


Figure 6. Potential shared diagnostic genes selection by least absolute shrinkage and selection operator. **(A)** The Venn diagram shows that 34 key genes are overlapping between the genes screened by WGCNA and those screened by DEGs. **(B)** LASSO regression models were used to identify potential shared diagnostic genes. Ten-time cross-validation was performed to select the optimal tuning parameter $\log(\lambda)$ in the GSE135917 database. **(C)** LASSO coefficient profiles of shared genes in the GSE135917 dataset. **(D)** Venn diagram showing the key shared biomarkers. **(E)** Tenfold cross-validation to select the optimal tuning parameter $\log(\lambda)$ in the GSE100927 database. **(F)** LASSO coefficient profiles of shared genes in the GSE100927 dataset.

regulates excessive inflammatory responses by regulating the activation of EGFR in endotoxemia. This inflammatory factor expression is significantly increased ($p < 0.05$) when ERRFI1 is knocked down³³. Therefore, we postulated that decreased ERRFI1 in OSA and AS could cause apoptosis of vascular endothelial cells, induce vascular smooth muscle cell proliferation, and increase inflammation. These findings suggest a reduced protective factor against the pathogenesis of both diseases. This study provides new ideas for future research on the protective mechanisms of both OSA and AS.

Considering the important role of the immune response in the development of OSA and AS, we investigated the differences in immune cell infiltration in different samples. Our results showed that both OSA and AS were closely associated with immune cell infiltration, consistent with the findings of previous studies²³. OSA has a significantly effect on the number of circulating inflammatory cells, lymphocytes, natural killer (NK and NKT-like cells)³⁴. Furthermore, aggregation of pro-inflammatory factors and immune cells was also present in atherosclerotic plaques^{20,21}. Interestingly, Activated CD8 T cells and CD56 bright natural killer cells showed the same trends in OSA and AS. Previous studies have shown that T lymphocyte abnormalities play an important role in endothelial cell dysfunction³⁵. The study Zhang et al. showed elevated levels of CD8 + T lymphocytes in patients with OSA compared to controls¹⁶ and that activated CD8 + T lymphocytes can have a killing effect on vascular endothelial cells, thus aiding the progression of secondary atherosclerosis. Atherosclerosis is a chronic inflammatory disease with immune infiltration³⁶, and Patients with atherosclerosis have significantly elevated CD8 T cells³⁷, mainly in the fibrous cap area. Second, NKs cells play an important role in atherosclerosis^{38,39}. A study of 48 patients with OSA found that both NK and NKT cells were increased in patients with OSA compared with controls³⁴. CD56bright natural killer cells cause cell lysis, vascular endothelial cell damage, and atherosclerosis by releasing proinflammatory factors, antibody-dependent cytotoxicity, perforins, and granzymes^{40,41}. In the correlations between ERRFI1 and 28 immune cells, we found a significant negative correlation between ERRFI1 and cellular immunity, suggesting a decrease in ERRFI1 and an increase in cellular immunity, which may be associated with inflammation and impaired vascular endothelial function in both diseases.

Our study has certain limitations. First, we only studied peripheral blood lymphocytes from patients; therefore, we could not explore the systemic changes in ERRFI1 in OSA and AS. Second, the small patient sample raises the possibility of bias. We plan to expand the collection of blood and tissue samples in future research in order to validate these results.

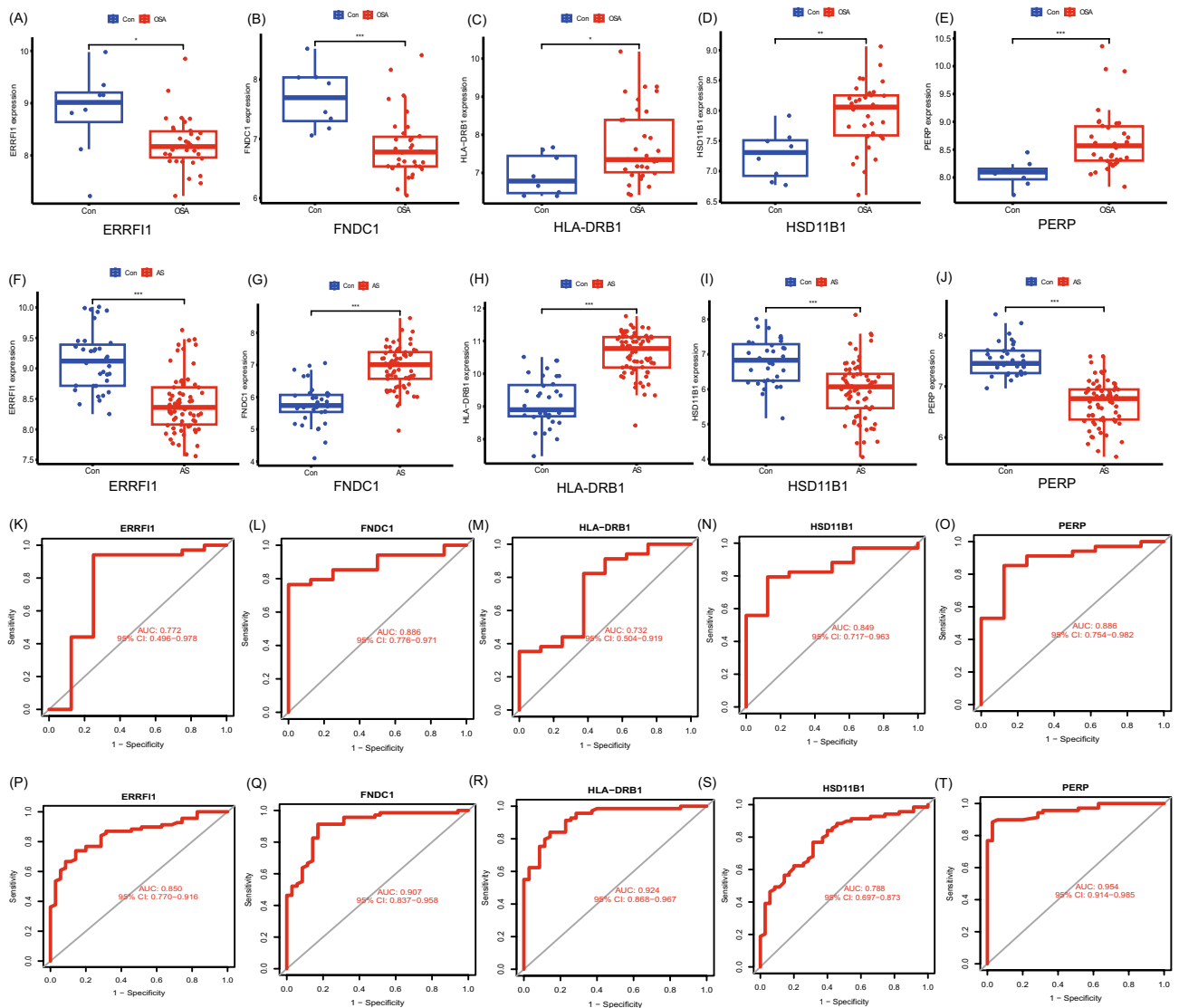


Figure 7. Candidate gene expression levels and diagnostic value. (A–E) Expression of *ERRF1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, and *PERP* in GSE135917. (F–J) Expression of *ERRF1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, and *PERP* in GSE100927. (K–O) ROC curves of *ERRF1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, and *PERP* genes in GSE135917. (P–T) ROC curves of *ERRF1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, *PERP* genes in GSE100927. Con: control; * represents $p < 0.05$; ** represents $p < 0.01$; *** represents $P < 0.001$.

Conclusion

Despite these limitations, this is the first study to use bioinformatics to explore the common mechanisms of action of OSA and AS, and our results suggest that immune pathways, inflammatory processes, and cell adhesion molecules are implicated in the pathogenesis of OSA and AS. Furthermore, we identified for the first time that *ERRF1* may be a potential diagnostic marker for OSA and AS, declining in both diseases. This suggests that a common underlying mechanism involving *ERRF1* may exist. Immune cell infiltration results showed that both OSA and AS were closely associated with immune cell infiltration, and *ERRF1* was closely associated with cellular immunity. This study offers a new perspective for exploring the common pathogenesis of OSA and AS and the intrinsic link between the two diseases and screening for a stable and reliable shared disease marker that can provide potential diagnostic and therapeutic targets for both diseases. We plan to further investigate the mechanisms of cell adhesion molecular pathways and *ERRF1* expression in OSA and AS based on both in vivo and in vitro experiments.

Methods

Datasets and data pre-processing. Figure 1 illustrates the workflow chart of data preparation, processing, analysis, and validation. Gene expression data for OSA and atherosclerosis were downloaded from the GEO database⁴². Further information is available online (<https://www.ncbi.nlm.nih.gov/geo/> datasets). The dataset screening criterion was that each dataset should have no fewer than 12 samples (Control sample + sample size

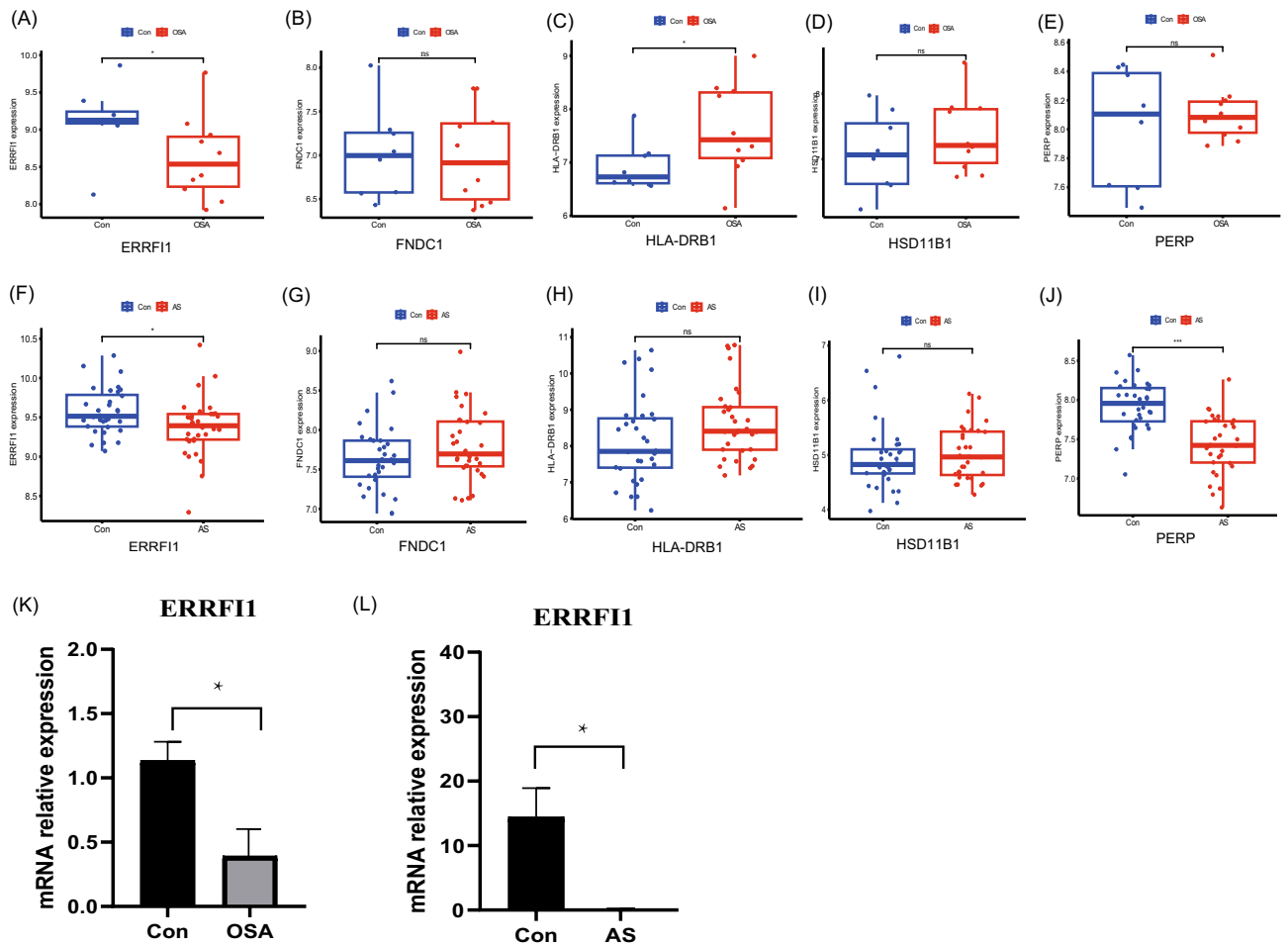


Figure 8. Validation of external datasets and qRT-PCR results. (A–E) Expression of *ERRF1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, and *PERP* in GSE38792. (F–J) Expression of *ERRF1*, *FNDC1*, *HLA-DRB1*, *HSD11B1*, and *PERP* in GSE43292. (K, L) qRT-PCR analysis of mRNA expression levels of *ERRF1* in PBMC from patients and healthy controls. * represents $p < 0.05$, *** represents $p < 0.001$, ns represents no statistical difference.

of disease patients) to ensure the accuracy of the WGCNA. The GSE135917 dataset, based on the GPL6244 platform, was then downloaded and contained samples from eight control groups and 34 patients with OSA. To validate the diagnostic effect of the gene, the GSE38792 dataset based on GPL6244 was downloaded. The data set contained samples from 8 control subjects and 10 patients with OSA. Adipose tissue samples were collected from patients in both datasets.

To investigate atherosclerosis, the GSE100927 dataset based on the GPL17077 platform, which included 35 controls and 69 atherosclerotic samples, was downloaded. To validate the diagnostic value of the gene, the GSE43292 dataset based on GPL6244, which contained samples from 32 controls and 32 atherosclerotic patients, was downloaded. Samples from both datasets were collected from arterial blood vessels. All data annotation and extraction were performed using the R software ((R.4.2.1, <https://www.r-project.org/>).

Differential genetic screening. Screening of the DEGs datasets for GSE135917 and GSE10092 was conducted. Limma R package to screen DEGs. An adjusted p -value < 0.05 and $|\text{LogFC}| > 0.5$ were selected as the cut-off standard. R software was used to plot the differential gene clustering of the volcanoes.

Construction and module analysis of weighted gene co-expression network analysis (WGCNA). (Weighted Gene Co-expression Network Analysis) WGCNA⁴³ is a bioinformatics analysis method used to describe gene association patterns among different samples, The WGCNA software package is used to perform WGCNA analysis. Co-expression networks corresponding to the clinical features of DEGs in OSA and AS were constructed using the WGCNA R package. First, hierarchical cluster analysis was performed using the cluster function in R to exclude outlier samples. Then, according to the criteria for scale-free networks, the "pickSoftThreshold" function in the WGCNA package was used to select a suitable soft-power threshold β (with values ranging from 1 to 20) for automatic network construction. The results were clustered using a topological overlap matrix analysis, which contained module assignments labeled by color and module eigengenes. Pearson's correlation analysis was used to calculate the correlation between module feature (ME) and clinical

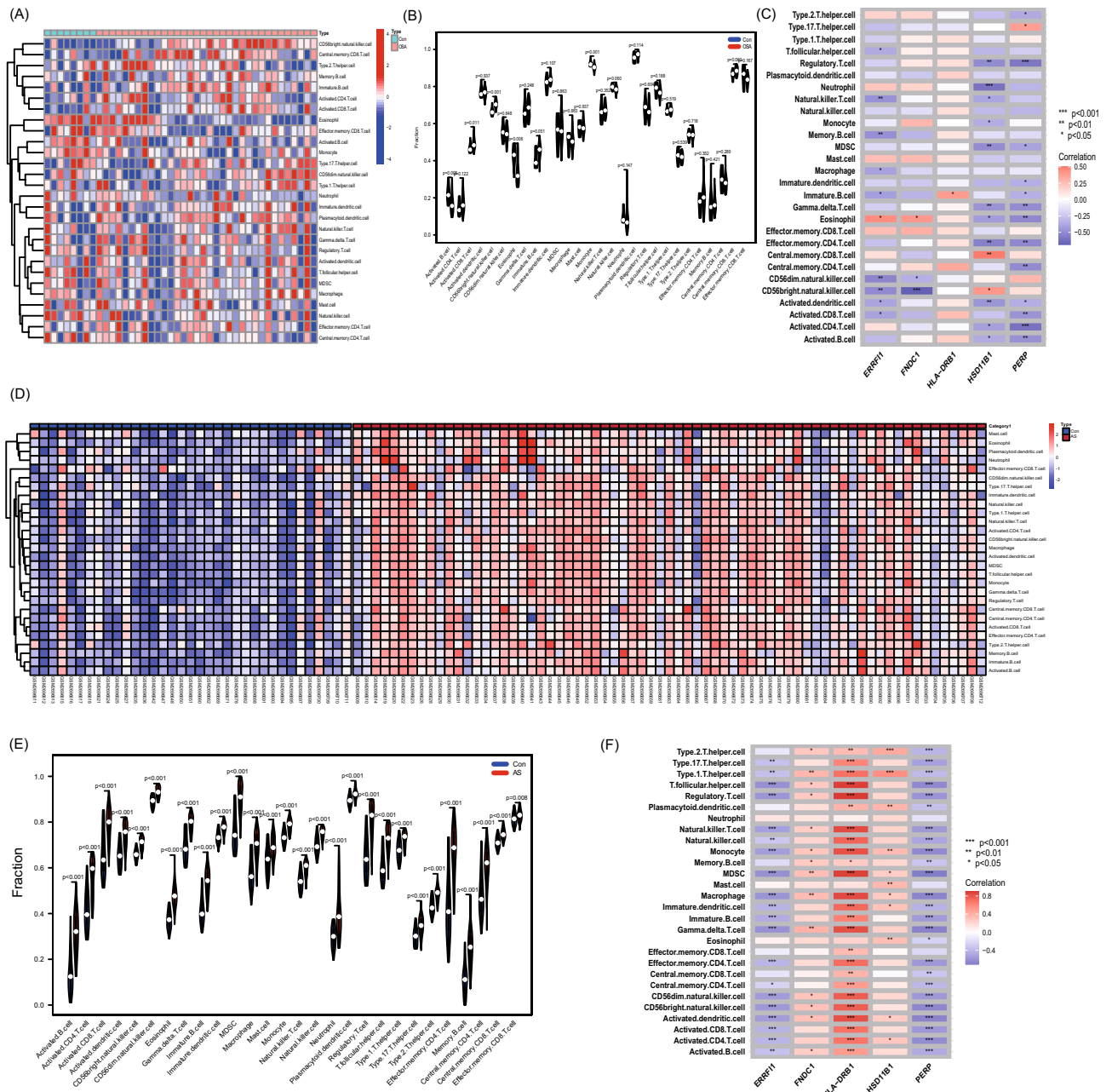


Figure 9. Analysis of immune infiltration associated with OSA and AS. Heatmap (A) and violin plot (B) show the distribution of 28 immune cells in the GSE135917 sample. (C) Association of key shared genes with immune cell infiltration in GSE135917 samples. Heatmap (D) and violin plot (E) show the distribution of 28 immune cells in the GSE100927 sample. (F) Association of key shared genes with immune cell infiltration in GSE100927 samples.

characteristics. $|\text{correlation coefficient } (R)| > 0.3$ and p -values < 0.05 were considered significant for interactions with clinical features.

Enrichment analysis. GO enrichment analysis is a bioinformatics method commonly used for comprehensive information retrieval from large-scale genetic data. The KEGG pathway enrichment analysis has been widely used to understand biological mechanisms and functions⁴⁴. The GO plot program package was used to visualize the GO and KEGG pathways. Finally, important signaling pathways were explored using the cluster profile and GSVA packages. The gene sets and expression matrices of OSA and AS were analyzed using GSEA to explore their possible regulatory pathways.

Identification of shared genes and Feature selection by the least absolute shrinkage and selection operator. Combinatorial analysis of genes from the WGCNA and DEG screens was performed by

plotting Venn diagrams. Overlapping genes were considered core-shared genes. Lasso is an important method for regression that uses an ℓ_1 penalty to achieve a sparse solution. We used the "glmnet" package in R to perform LASSO regression to filter the best predictors of OSA and AS in the intersection of the above DEGs and WCGNA.

Candidate biomarker expression levels and diagnostic value. The expression levels of key shared genes ($p < 0.05$) were detected using the R software ggplot2 package box plots (GSE135917, GSE38792, GSE100927, and GSE43292). The AUC of ROC was used to determine the diagnostic value of potential biomarkers in the dataset (GSE135917, and GSE100927) using the pROC R package.

Extraction of peripheral blood mononuclear cells (PBMC). Purple anticoagulation tubes were used to collect 10 ml of whole blood samples from three patients with OSA, four patients with AS, and three normal controls. Human peripheral blood and phosphate buffer were mixed in a homogeneous ratio of 1:1 (10 ml:10 ml). Ten milliliters of lymphocyte separation liquid (SIGMA, 10,771) was applied in each 50 ml centrifuge tube, and 20 ml of blood was gently added to the phosphate buffer on the upper surface of the lymphocyte separation liquid. The tubes were centrifuged in a centrifuge for 30 min (2000 rpm, 20 °C). The middle and upper interface cloud layers were carefully collected and transferred to new 50 mL tubes. Finally, the PBMC was washed with phosphate buffer, centrifuged, and the supernatant was discarded. This study was approved by the Ethics Committee of First Hospital of Hebei Medical University and all participants provided their written informed consent to participate in the study. And the study was conducted in accordance with the relevant guidelines and the regulations.

Quantitative Real-Time PCR (qRT-PCR). Total RNA was extracted from the PBMC using the Eastep® Super Total RNA Extraction Kit (Promega Shanghai Ltd.). Pipette 1.5 μ L of total RNA through the Nanodrop 1000 for measurement. RNA-to-cDNA transcription was performed using GoScript™ Reverse Transcription Mix, Random Primer Protocol, and Oligo (dT) Protocol from Promega. qRT-PCR was performed using a Roche LightCycler 480 qRT-PCR amplifier following the manufacturer's instructions. Relative mRNA levels were normalized to the level of glyceraldehyde-3-phosphate dehydrogenase, using the $2^{-\Delta\Delta C_t}$ method. (*ERRF1*-Fwd: 5'-GAAGACCTACTGGAGCAGTCG-3'; *ERRF1*-Rev: 5'-GACTTTTGAGATGGACCATTCTG-3').

ssGSEA. ssGSEA analysis using the "GSVA" R package was used to analyze the infiltration of 28 immune cells in lesions and normal samples. The correlation between core genes and the abundance of infiltrating immune cells was analyzed using Spearman's correlation coefficients and a p -value < 0.05 was considered statistically significant.

Data availability

The datasets GSE135917, GSE38792, GSE100927 and GSE43292 for this study can be found in the GEO datasets (<https://www.ncbi.nlm.nih.gov/geo/>).

Received: 13 April 2023; Accepted: 6 September 2023

Published online: 12 September 2023

References

1. Benjafield, A. V. *et al.* Estimation of the global prevalence and burden of obstructive sleep apnoea: A literature-based analysis. *Lancet Respir. Med.* **7**, 687–698 (2019).
2. Rundo, J. V. Obstructive sleep apnea basics. *Cleve Clin. J. Med.* **86**, 2–9 (2019).
3. Lyons, M. M., Bhatt, N. Y., Pack, A. I. & Magalang, U. J. Global burden of sleep-disordered breathing and its implications. *Respirology* **25**, 690–702 (2020).
4. Peker, Y., Carlson, J. & Hedner, J. Increased incidence of coronary artery disease in sleep apnoea: A long-term follow-up. *Eur. Respir. J.* **28**, 596–602 (2006).
5. Marin, J. M., Carrizo, S. J., Vicente, E. & Agusti, A. G. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: An observational study. *Lancet* **365**, 1046–1053 (2005).
6. Walia, H. K. *et al.* Atherosclerotic plaque in individuals without known cardiovascular disease but with established obstructive sleep apnea and at high risk of obstructive sleep apnea. *Am. J. Prev. Cardiol.* **14**, 100497 (2023).
7. Luo, B. *et al.* Intermittent Hypoxia and Atherosclerosis: From molecular mechanisms to the therapeutic treatment. *Oxid. Med. Cell Longev.* **2022**, 1438470 (2022).
8. Oga, T. *et al.* Effects of obstructive sleep apnea with intermittent hypoxia on platelet aggregability. *J. Atheroscler. Thromb.* **16**, 862–869 (2009).
9. Liu, Q. B. *et al.* The induction of reactive oxygen species and loss of mitochondrial Omi/HtrA2 is associated with S-nitrosoglutathione-induced apoptosis in human endothelial cells. *Toxicol. Appl. Pharmacol.* **244**, 374–384 (2010).
10. Sharma, S. K. *et al.* Prevalence and risk factors of syndrome Z in urban Indians. *Sleep Med.* **11**, 562–568 (2010).
11. Luo, S. *et al.* Continuous hypobaric hypoxia may promote atherosclerosis progression in apolipoprotein E-deficient mice. *Int. J. Med. Sci.* **20**, 849–857 (2023).
12. Song, D., Fang, G., Greenberg, H. & Liu, S. F. Chronic intermittent hypoxia exposure-induced atherosclerosis: A brief review. *Immunol. Res.* **63**, 121–130 (2015).
13. Wang, F. *et al.* Effects of chronic intermittent hypoxia and chronic sleep fragmentation on gut microbiome, serum metabolome, liver and adipose tissue morphology. *Front. Endocrinol.* **13**, 820939 (2022).
14. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U S A* **102**, 15545–15550 (2005).
15. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).

16. Zhang, M. *et al.* Exploration of the shared pathways and common biomarker PAN3 in ankylosing spondylitis and ulcerative colitis using integrated bioinformatics analysis. *Front. Immunol.* **14**, 1089622 (2023).
17. Gao, H., Li, J., Li, Q. & Lin, Y. Identification of hub genes significantly linked to subarachnoid hemorrhage and epilepsy via bioinformatics analysis. *Front. Neurol.* **14**, 1061860 (2023).
18. Chiu, H. Y. *et al.* Obstructive sleep apnea in young Asian adults with sleep-related complaints. *Sci. Rep.* **12**, 20582 (2022).
19. Tobaldini, E. *et al.* Sleep, sleep deprivation, autonomic nervous system and cardiovascular diseases. *Neurosci. Biobehav. Rev.* **74**, 321–329 (2017).
20. Wolf, D. & Ley, K. Immunity and inflammation in atherosclerosis. *Circ. Res.* **124**, 315–327 (2019).
21. Roy, P., Orecchioni, M. & Ley, K. How the immune system shapes atherosclerosis: Roles of innate and adaptive immunity. *Nat. Rev. Immunol.* **22**, 251–265 (2022).
22. Dyugovskaya, L., Lavie, P., Hirsh, M. & Lavie, L. Activated CD8+ T-lymphocytes in obstructive sleep apnoea. *Eur. Respir. J.* **25**, 820–828 (2005).
23. Domagała-Kulawik, J., Osińska, I., Piechuta, A., Bielicki, P. & Skirecki, T. T, B, and NKT cells in systemic inflammation in obstructive sleep apnoea. *Mediators Inflamm.* **2015**, 161579 (2015).
24. Harki, O. *et al.* VE-cadherin cleavage in sleep apnoea: New insights into intermittent hypoxia-related endothelial permeability. *Eur. Respir. J.* **58**, 2004518 (2021).
25. Zhou, J., Bai, W., Liu, Q., Cui, J. & Zhang, W. Intermittent hypoxia enhances THP-1 monocyte adhesion and chemotaxis and promotes M1 macrophage polarization via RAGE. *Biomed Res Int* **2018**, 1650456 (2018).
26. Santamaria-Martos, F. *et al.* Biomarkers of carcinogenesis and tumour growth in patients with cutaneous melanoma and obstructive sleep apnoea. *Eur. Respir. J.* **51**, 1701885 (2018).
27. Yin, M. *et al.* Cell adhesion molecule-mediated therapeutic strategies in atherosclerosis: From a biological basis and molecular mechanism to drug delivery nanosystems. *Biochem Pharmacol* **186**, 114471 (2021).
28. Xu, D. & Li, C. Gene 33/Mig6/ERRF11, an adapter protein with complex functions in cell biology and human diseases. *Cells* **10**, 1574 (2021).
29. Miao, Y., Cui, L., Chen, Z. & Zhang, L. Gene expression profiling of DMU-212-induced apoptosis and anti-angiogenesis in vascular endothelial cells. *Pharm. Biol.* **54**, 660–666 (2016).
30. Jin, N. *et al.* Mig-6 is required for appropriate lung development and to ensure normal adult lung homeostasis. *Development* **136**, 3347–3356 (2009).
31. Lee, J. H. *et al.* Mig-6 gene knockout induces neointimal hyperplasia in the vascular smooth muscle cell. *Dis. Markers* **2014**, 549054 (2014).
32. Chen, W. *et al.* Mig6 reduces inflammatory mediators production by regulating the activation of EGFR in LPS-induced endotoxemia. *J. Cell Physiol.* **233**, 6975–6983 (2018).
33. Ma, P., Zhang, C., Huo, P., Li, Y. & Yang, H. A novel role of the miR-152-3p/ERRF11/STAT3 pathway modulates the apoptosis and inflammatory response after acute kidney injury. *J. Biochem. Mol. Toxicol.* **34**, e22540 (2020).
34. Ludwig, K., Huppertz, T., Radsak, M. & Gouveris, H. Cellular immune dysfunction in obstructive sleep apnea. *Front. Surg.* **9**, 890377 (2022).
35. Incalza, M. A. *et al.* Oxidative stress and reactive oxygen species in endothelial dysfunction associated with cardiovascular and metabolic diseases. *Vascul. Pharmacol.* **100**, 1–19 (2018).
36. Kobiyama, K. & Ley, K. Atherosclerosis. *Circ. Res.* **123**, 1118–1120 (2018).
37. Kolbus, D. *et al.* CD8+ T cell activation predominate early immune responses to hypercholesterolemia in Apoe^{-/-} mice. *BMC Immunol.* **11**, 58 (2010).
38. Olson, N. C. *et al.* Associations of innate and adaptive immune cell subsets with incident type 2 diabetes risk: The MESA study. *J. Clin. Endocrinol. Metab.* **105**, e848–857 (2020).
39. Olson, N. C. *et al.* Innate and adaptive immune cell subsets as risk factors for coronary heart disease in two population-based cohorts. *Atherosclerosis* **300**, 47–53 (2020).
40. Abel, A. M., Yang, C., Thakar, M. S. & Malarkannan, S. Natural killer cells: Development, maturation, and clinical utilization. *Front. Immunol.* **9**, 1869 (2018).
41. Palano, M. T. *et al.* When a friend becomes your enemy: Natural killer cells in atherosclerosis and atherosclerosis-associated risk factors. *Front. Immunol.* **12**, 798155 (2021).
42. Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
43. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
44. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, D587–D592 (2023).

Acknowledgements

We acknowledge GEO database for providing their platforms and contributors for uploading their meaningful datasets. We gratefully acknowledge all the individuals whose participation and contributions made this work possible.

Author contributions

B.W.C. contributed to conception and design. B.W.C. and L.P.D. analyzed the data. W.W.C., Y.H. and J.H.Z. validated the method and data. B.W.C. wrote this manuscript. D.M.S. edited the manuscript and provided constructive comments. All authors read and approved the final manuscript.

Funding

Medical Science Research Projects of Hebei Province grant number 20221413, The Provincial level Science and Technology Program of Hebei Province grant number 21377734D and Training Programme Foundation for the Talents of Clinical Medicine sponsored by the Government grant number LS202003.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-42184-0>.

Correspondence and requests for materials should be addressed to D.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023