Electronic Journal of Biotechnology 83 (2026) 100719

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

journal homepage: www.elsevier.com/locate/ejbt

Research article

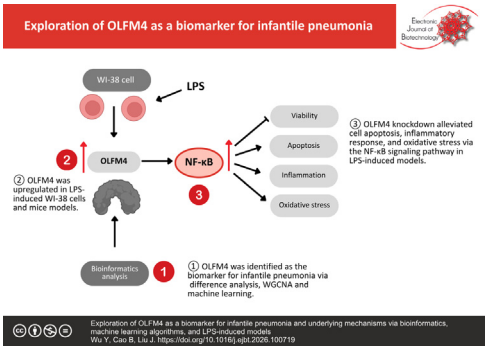
Exploration of OLFM4 as a biomarker for infantile pneumonia and underlying mechanisms via bioinformatics, machine learning algorithms, and LPS-induced models ☆

Ying Wu [1], Baocen Cao [1], Jie Liu *

Department of Pediatrics, Huangshi Central Hospital, Huangshi, China

GRAPHICAL ABSTRACT

Exploration of OLFM4 as a biomarker for infantile pneumonia and underlying mechanisms via bioinformatics, machine learning algorithms, and LPS-induced models.



ARTICLE INFO

Article history:
Received 3 January 2026
Accepted 28 April 2026
Available online 2 July 2026

Keywords:
Bioinformatics
Biomarker
Infantile pneumonia
Inflammation
Lipopolysaccharide (LPS)
NF-κB signaling pathway

ABSTRACT

Background: Infantile pneumonia is a common health concern worldwide, with elevated morbidity and mortality rates among affected children. This study aims to identify key genes associated with infantile pneumonia using bioinformatics and unravel the underlying mechanisms.
Results: OLFM4 was the only biomarker identified for infantile pneumonia. Besides, OLFM4 expression was promoted in the serum of infantile pneumonia patients, and LPS-stimulated cells and mouse models. OLFM4 knockdown repressed cell apoptosis, levels of TNF-α, IL-6, IL-1β, MPO, MDA, ROS, and activation of the NF-κB pathway, and facilitated the SOD level in LPS-induced models. OLFM4 knockdown alleviated the lung injury of the LPS-induced mouse model.
Conclusions: OLFM4 knockdown alleviated cell apoptosis, inflammatory response, and oxidative stress via the NF-κB signaling pathway in LPS-induced WI-38 cells and mouse model. The findings suggest that OLFM4 may pave the way for the treatment of infantile pneumonia.

☆ Audio abstract available in Supplementary material.

* Corresponding author.
E-mail address: 13581289523@163.com (J. Liu).

[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.ejbt.2026.100719

Machine learning (ML)
OLFM4
Oxidative stress
Reactive oxygen species
Weighted gene co-expression network analysis (WGCNA)

How to cite: Wu Y, Cao B, Liu J. Exploration of OLFM4 as a biomarker for infantile pneumonia and underlying mechanisms via bioinformatics, machine learning algorithms, and LPS-induced models. Electron J Biotechnol 2026;83. https://doi.org/10.1016/j.ejbt.2026.100719.

1. Introduction

Pediatric pneumonia is a prevalent respiratory illness in newborns and young children, potentially leading to heart failure and other consequences in severe instances [1,2]. Pneumonia is the primary cause of mortality in children under five years old [3]. Bacterial, viral, prokaryotic microorganisms, and fungal infections are the primary causes of pneumonia [4,5]. Currently, the predominant treatment modalities for infantile pneumonia are antiviral and antibacterial therapies [6,7]. Nevertheless, the escalating use of antibiotics has led to heightened pathogen resistance, complicating treatment of the condition [8,9]. Consequently, comprehensive investigation into the pathophysiology of infantile pneumonia and the identification of more efficient therapeutic targets are crucial for its treatment [10].

In recent years, bioinformatics has expanded its scope, facilitating the integration of various biological data types, including multi-omics, pathological, imaging, and clinical data [11–15]. The integration enables the modeling of intricate interactions within biological systems and provides a thorough understanding of the molecular landscape of epidemic diseases, cancer, and chronic diseases [16–18]. Employing systems biology methodology, machine learning, and network-based tools, researchers acquire insights into the complex relationships among genes, proteins, and pathways associated with human diseases [11,12,19]. This comprehensive perspective enhances the understanding of the biological processes underlying illnesses and highlights essential biological mechanisms and signaling pathways that are amenable to therapeutic intervention [12]. The incidence rate of infantile pneumonia varies among children, and the severity of infantile pneumonia ranges from mild to severe [2,4]. However, the underlying causes for these variations are not precisely known. Bioinformatics may help to elucidate the intricate molecular landscape associated with infantile pneumonia.

This study examined the mRNA dataset pertaining to infantile pneumonia within the GEO database. Weighted gene co-expression network analysis (WGCNA) and machine learning (ML) were utilized to screen possible targets for infantile pneumonia. A literature review revealed olfactomedin 4 (OLFM4) as the sole gene involved in the inflammatory processes of many diseases and in the exacerbation of lung damage. OLFM4 is a conserved glycoprotein found in neutrophils, lung tissue, cardiomyocytes, and renal epithelia [20–23]. OLFM4 was independently linked to heightened mortality in both pediatric and adult patients with sepsis-induced acute respiratory distress syndrome (ARDS) [24]. Numerous studies have indicated the function of OLFM4 in cardiomyocyte sepsis, lipopolysaccharide (LPS)-induced acute respiratory infections in rat models, and various malignancies [25–29]. In addition, OLFM4 has been established as correlated with disease severity in pediatric patients suffering from viral lower respiratory tract infections [30]. Nevertheless, there is a paucity of studies examining the role of OLFM4 and its underlying mechanisms in infantile pneumonia. This work examined the regulatory roles and related pathways of OLFM4 in LPC-induced pneumonia in WI-38 cells and mouse model. The results indicated that OLFM4 could serve as an effective target for alleviating infantile pneumonia.

2. Materials and methods

2.1. Screening of differentially expressed genes (DEGs)

Screening of DEGs of the GSE103119 dataset was performed using the online sangerbox platform (https://sangerbox.com/) and "limma" package. The raw counts were normalized to minimize potential bias. DEGs were filtered according to the criterion of adjusted p-value < 0.05 and |log2FC| > 1. The heatmap and volcano plot were visualized using sangerbox.

2.2. WGCNA

The "WGCNA" package was used to create the gene co-expression matrix. The expression profile data of 6000 genes across 172 samples following pre-processing before WGCNA include filtering, normalization, and log2 transformation. A total of 172 samples were first clustered, and the optimal soft threshold (β) was determined to be 7. Parameters minModuleSize = 25 and MergeCutHeight = 0.25 were used for the construction of a topological overlap matrix. Hierarchical clustering was performed to identify gene modules and calculate the correlation between module genes and infantile pneumonia.

2.3. Screening of candidate genes by machine learning

Three machine learning techniques, including the lasso model, SVM-RFE, and RF, were utilized to narrow down the candidate genes. The LASSO algorithm was executed with the "glmnet 4.1.8" R package. This methodology employed five-fold cross-validation to assess the model's efficacy and ultimately identified the candidate genes based on the optimal λ. The "e1071 1.7.16" R package was utilized to execute the SVM-RFE algorithm, facilitating the ranking of candidate feature genes based on their significance and employing five-fold cross-validation for error evaluation. The "randomForest 4.7.1.2" R program was utilized to execute the RF algorithm, consolidating the outcomes for classification, regression, and feature selection. Ultimately, Venn diagrams were constructed to illustrate the intersection of the three algorithms.

2.4. Evaluation of receiver operating characteristic (ROC)

The diagnostic accuracy of the OLFM4 for infantile pneumonia was evaluated using the ROC analysis, emphasizing the associated area under the curve (AUC) values.

2.5. Human participants

Three milliliters of blood were obtained from 35 patients with infantile pneumonia and 35 healthy individuals (undergoing physical examinations). Prior to participation in the study, informed consent was obtained from the participants or their guardians. The study received approval from the Ethical Committee of Huangshi Central Hospital. The serum from the blood samples was collected and preserved for OLFM4 mRNA expression analysis.

2.6. Cell culture and treatment

Human embryonic lung (WI-38) cells were obtained from Procell (Wuhan, China) and grown in DMEM (SUNNCELL, Wuhan, China) supplemented with 10% fetal bovine serum (SUNNCELL) at 37°C in an incubator with 95% air and 5% CO2. LPS was purchased from Yeasen (Shanghai, China). Following treatment of WI-38 cells with LPS (0, 0.5, 1, and 2 μg/ml) for 12 h, the mimic infantile pneumonia cell damage model was established for subsequent investigations. For LPS + Pyrrolidinedithiocarbamate (PDTC) treatment, WI-38 cells were preincubated with PDTC (10 μM, Sigma-Aldrich) before addition of LPS [31].

2.7. Cell transfection

The GenePharma (Shanghai, China) designed the siRNA targeting OLFM4 (si-OLFM4) and negative controls of siRNA (si-NC). The transient transfection of si-RNA was performed with GP-transfect-Mate (GenePharma). The pcDNA-based OLFM4 overexpression (OE-OLFM4) plasmids and control vector (OE-NC) were obtained from GenePharma and transfected into WI-38 cells using Lipofectamine 3000 (Invitrogen).

2.8. Cell counting kit- 8 (CCK-8) test

The assessment of cell viability was conducted using the CCK-8 test kit (Beyotime, Shanghai, China) at a concentration of 5×10^3 cells per well. The cells were cultured in 96-well microplates for 24 h, followed by treatment with LPS (1 μg/mL), LPS (1 μg/mL) + si-NC, LPS (1 μg/mL) + si-OLFM4, or left untreated (Control) for 12 h. Thereafter, the cells in each well were treated with 10 μL of CCK-8 solution and incubated for 4 h. The cell viability was ultimately assessed using a microplate reader.

2.9. Measurement of malondialdehyde (MDA) and superoxide dismutase (SOD)

The WI-38 cells were cultured in 12-well plates (1×10^5 cells per well) until reaching 70% confluence. Subsequently, the cells received the following treatments, including LPS (1 μg/mL), LPS (1 μg/mL) + si-NC, LPS (1 μg/mL) + si-OLFM4, or without (Control) for 12 h. Next, the cells and cell supernatant were separated via centrifugation. The levels of MDA in the supernatant were analyzed using a lipid peroxidation MDA assay kit (Beyotime). The collected cells were mixed with SOD working solution for cell lysis. The mixture was centrifuged to obtain the supernatant for SOD detection with SOD Assay Kit with WST-8 (Beyotime).

2.10. Enzyme-Linked Immunosorbent Assay (ELISA)

The levels of tumor necrosis factor-alpha (TNF-α), IL-6, and IL-1β were analyzed with ELISA assay kits (MEIMIAN, Jiangsu, China). Briefly, 1×10^5 WI-38 cells were seeded in the 24-well microplate for 24 h. Then, the cells were subsequently subjected to treatment with LPS (1 μg/mL), LPS (1 μg/mL) + si-NC, LPS (1 μg/mL) + si-OLFM4, or without (Control) for 12 h. Next, the supernatant samples were collected. The blood samples from the mice were centrifuged to obtain serum in the upper layer. The contents of pro-inflammatory cytokines in the supernatants and serum were quantified using the ELISA reagents.

2.11. Flow cytometry

Cellular apoptosis and reactive oxygen species (ROS) were detected using flow cytometry. In short, 2×10^5 WI-38 cells were cultured in 6-well microplates to 70% confluence. Then, the cells were treated with LPS (1 μg/mL), LPS (1 μg/mL) + si-NC, LPS (1 μg/mL) + si-OLFM4, or without (Control) for 12 h. Cells were collected and rinsed with precooling PBS. The Annexin V-FITC/PI Apoptosis Detection Kit (Yeasen) in this research was employed to estimate cell apoptosis. Cell suspension (2×10^5 cells) was mixed with Annexin V-FITC and PI and incubated at room temperature for 15 min. After that, the flow cytometry device (Thermo Fisher Scientific, MA, USA) was utilized to quantify the apoptosis ratio.

Relative ROS levels were detected using ROS Assay Kit (Bjbalb, Beijing, China). Cell suspension (2×10^5 cells) was mixed with DCFH-DA and cultivated at 37°C for 20 min, with rotation every 4 min. After washing the free DCFH-DA, cells were measured with the flow cytometry system.

2.12. Real-time quantitative PCR (RT-qPCR)

Total RNA was isolated from serum or cells and purified utilizing TRIzol reagent (Tiangen, Beijing, China). Complementary DNA was produced with the BeyoRT™ III First Strand cDNA Synthesis Kit (Beyotime). Following amplification, the relative expression level of OLFM4 was determined using the $2^{-\Delta\Delta Ct}$ method, with GAPDH as the internal standard. The sequences of the specific primers are displayed in Table 1.

2.13. Western blot

The precooling RIPA lysis buffer (Beyotime) was used to lyse treated WI-38 cells or lung tissues. The protein extracts (30 μg) were separated using 12% SDS-PAGE (Vazyme) and shifted to the polyvinylidene fluoride membrane (Beyotime). The membranes in this experiment were incubated with primary antibodies of anti-OLFM4 (#14369, 1:1000, Cell Signaling Technology (CST), MA, USA), p-P65 (#3033, 1:1000, CST), anti-P65(#8242, 1:1000, CST), anti-p-IκBα (#2859, 1:1000, CST), anti-IκBα (#4812, 1:1000, CST), and anti-GAPDH (ab8245, 1:4000, Abcam, Cambridge, England), followed by being probed with secondary antibodies. Finally, the results were observed and quantified.

2.14. Animal study

Male C57BL/6 mice and OLFM4 knockdown (Ad-sh-NC or Ad-sh-OLFM4) mice (aged 8 weeks and weighing 22–28 g) were purchased from Cyagen Biosciences and housed under specific pathogen-free conditions with controlled environmental parameters (22–25°C, 45–60% humidity, 12 h/12 h light–dark cycle). These mice were categorized into four groups, including Sham (intraperitoneally received saline), LPS (intraperitoneally received LPS 2 mg/kg), Ad-sh-NC + LPS, and Ad-sh-OLFM4 + LPS groups (with five mice in each group). All mice received corresponding treatment for 24 h. After the mice were euthanized, blood was collected for subsequent ELISA and lung tissues were gathered for RT-qPCR, western blot, lung wet/dry ratio measurement, and pathological analysis. The animal experiment was approved by the Animal Research Committee of Huangshi Central Hospital.

2.15. Tissue staining

For pathological analysis, the lung tissues were fixed, dehydrated, and embedded for pathological investigation. Five-micron serial slices were produced for Masson's trichrome staining (Beyotime) as well as hematoxylin and eosin (HE, Servicebio, Wuhan, China) staining. The dyed slices were observed and photographed. The lung damage ratings were evaluated in accordance with a prior study [32].

Table 1
Primer sequences used for qPCR.

Name		Primers for PCR (5′-3′)
OLFM4	Forward	TTGTCAGAGGTCTAGGGGCA
	Reverse	CCCTCGAGGAGCTCCAGATA
GAPDH	Forward	AATGGGCAGCCGTTAGGAAA
	Reverse	GCGCCCAATACGACCAAATC

2.16. Lung wet/dry (WD) ratio

To evaluate pulmonary edema, the left lung was immediately weighed for wet weight, then dried and weighed to obtain dry weight, and the WD ratio was determined.

2.17. Measurement of myeloperoxidase (MPO) activity

Measurement of MPO activity was performed with the MPO Activity Assay Kit (BC5710, Solarbio, Beijing, China). 0.1 g of lung tissue was mixed with the reagent of the kit and homogenized. After centrifugation, the supernatant was used for subsequent MPO activity measurement according to the instructions. Absorption at 460 nm was recorded via a spectrophotometer.

2.18. Statistical analysis

All measurements are presented as mean ± SD and analyzed using GraphPad Prism 9.0 software. After the Shapiro-Wilk normality test and Brown-Forsythe test, comparisons between groups were performed using one-way ANOVA with Tukey's test, Student's *t*-test, or Mann-Whitney *U* test. $p < 0.05$ was considered to be statistically significant.

3. Results

3.1. Identification of DEGs

Transcriptomic analysis of the GSE103119 dataset, containing 20 healthy controls and 152 infantile pneumonia patients, was performed, and identified 984 DEGs, comprising 576 upregulated and 408 downregulated genes. These results were visually displayed using a volcano plot (Fig. 1A). Additionally, the expression patterns of the top 10 upregulated and downregulated DEGs were further presented using a heatmap (Fig. 1B).

3.2. WGCNA analysis

In this study, WGCNA was utilized to identify gene modules potentially related to infantile pneumonia. This analysis constructed a scale-free network, which facilitated the identification of an optimal soft-thresholding power set at 7, as depicted in Fig. 2A. This specific thresholding power, when applied alongside hierarchical clustering employing average linkage techniques, classified these genes into seven distinct clusters (Fig. 2B, C). Core gene clusters with particular traits of infantile pneumonia samples were recognized with a significance level of $p < 0.05$ and |correlation coefficient| > 0.3. Additionally, a significant correlation was observed between the module membership (MM) and gene significance (GS) within the red module (Fig. 2D). Within the red module, 22 genes important for module function were identified and used for further analysis. Through Venn diagram analysis, co-expressed genes between the DEGs and the WGCNA-derived feature module genes were identified (Fig. 2E).

3.3. Identification of characteristic genes via machine learning algorithms

To choose potential genes for infantile pneumonia, three machine learning algorithms were utilized in further analysis. In the lasso model, the presence of infantile pneumonia was linked to the expression of four genes, including OLFM4, CEACAM6, COL17A1, and CAMP (Fig. 3A, B). SVM identified 11 genes, including COL17A1, OLFM4, CEACAM6, MPO, CAMP, LTF, DEFA1, DEFA3, DEFA1B, ELANE, and CEACAM8 (Fig. 3C,D). In the random forest algorithm, there were ten signature genes that had relative relevance scores that were higher than one. These genes were OLFM4, COL17A1, CAMP, CEACAM6, DEFA3, LTF, MPO, OLR1, ELANE, and AZU1 (Fig. 3E,F). Through the intersection of genes resulting from the three distinct machine learning models, four genes (OLFM4, CEACAM6, COL17A1, and CAMP) were identified (Fig. 3G). Through literature review, OLFM4 was the only gene that participated in the inflammatory processes of multiple diseases and contributed to promoting lung injury. Furthermore, compared to the other three genes, the mRNA level of OLFM4 showed the greatest increase following 2 μg/mL LPS treatment (Fig. S1). However, studies on the function of OLFM4 in infantile pneumonia have not yet been reported. OLFM4 was selected as the core gene in the following experiments.

3.4. The expression of OLFM4 was upregulated in pediatric pneumonia

To explore the function of OLFM4 in the process of pediatric pneumonia, the expression of OLFM4 was first investigated. The mRNA levels of OLFM4 were significantly higher in infantile pneumonia patients compared to the healthy controls in the GSE103119 dataset (Fig. 4A). The results of the hospital revealed that serum OLFM4 exhibited markedly higher levels of expression in the 55 infantile pneumonia patients, as illustrated in Fig. 4B. ROC analysis showed the diagnostic value of OLFM4 for infantile pneumonia, with AUC = 0.9501 (Fig. 4C). In order to determine the role of OLFM4 in human lung fibroblast cells, WI-38 cells were stimulated with LPS at concentrations of 0, 0.5, 1, and 2 μg/ml for 12 h. RT-qPCR and WB analyses revealed a notable increase in OLFM4 expression at both transcript and protein levels in WI-38 cells in a dose-dependent manner (Fig. 4D, E). Subsequent experiments utilized 1 μg/ml LPS. These results indicated that OLFM4 was upregulated in both infantile pneumonia patients and LPS-stimulated WI-38 cells.

3.5. Knockdown of OLFM4 alleviated LPS-stimulated WI-38 cell inflammation damage

Next, the role of OLFM4 was explored in LPS-induced WI-38 cells. The downregulation of OLFM4 in WI-38 cells transfected with si-OLFM4 or si-NC was verified via western blot (Fig. 5A). The cell vitality of WI-38 cells after being treated with LPS was repressed compared with the Control group, while OLFM4 knockdown could rescue this effect (Fig. 5B). The apoptosis was expedited after WI-38 cells were treated with LPS, while si-OLFM4 could weaken this action (Fig. 5C). The inflammation factor (IL-6, IL-1β, and TNF-α) productions were increased by LPS, which was weakened by OLFM4 knockdown (Fig. 5D,F). Downregulated OLFM4 suppressed LPS-stimulated facilitation of MDA in WI-38 cells (Fig. 5G). LPS repressed the levels of SOD, which was reversed by si-OLFM4 (Fig. 5H). The ROS production in LPS-stimulated WI-38 cells was weakened by OLFM4 knockdown (Fig. 5I). Taken together, OLFM4 knockdown hindered LPS-stimulated WI-38 inflammation injury.

3.6. Knockdown of OLFM4 inhibited NF-κB activation

The NF-κB signaling pathway plays an important role in human diseases, especially inflammation-related diseases, such as infantile pneumonia. The association between OLFM4 and the NF-κB signaling pathway was explored. The results revealed that, relative to the Control, LPS-stimulated WI-38 cells transfected with si-NC or with-

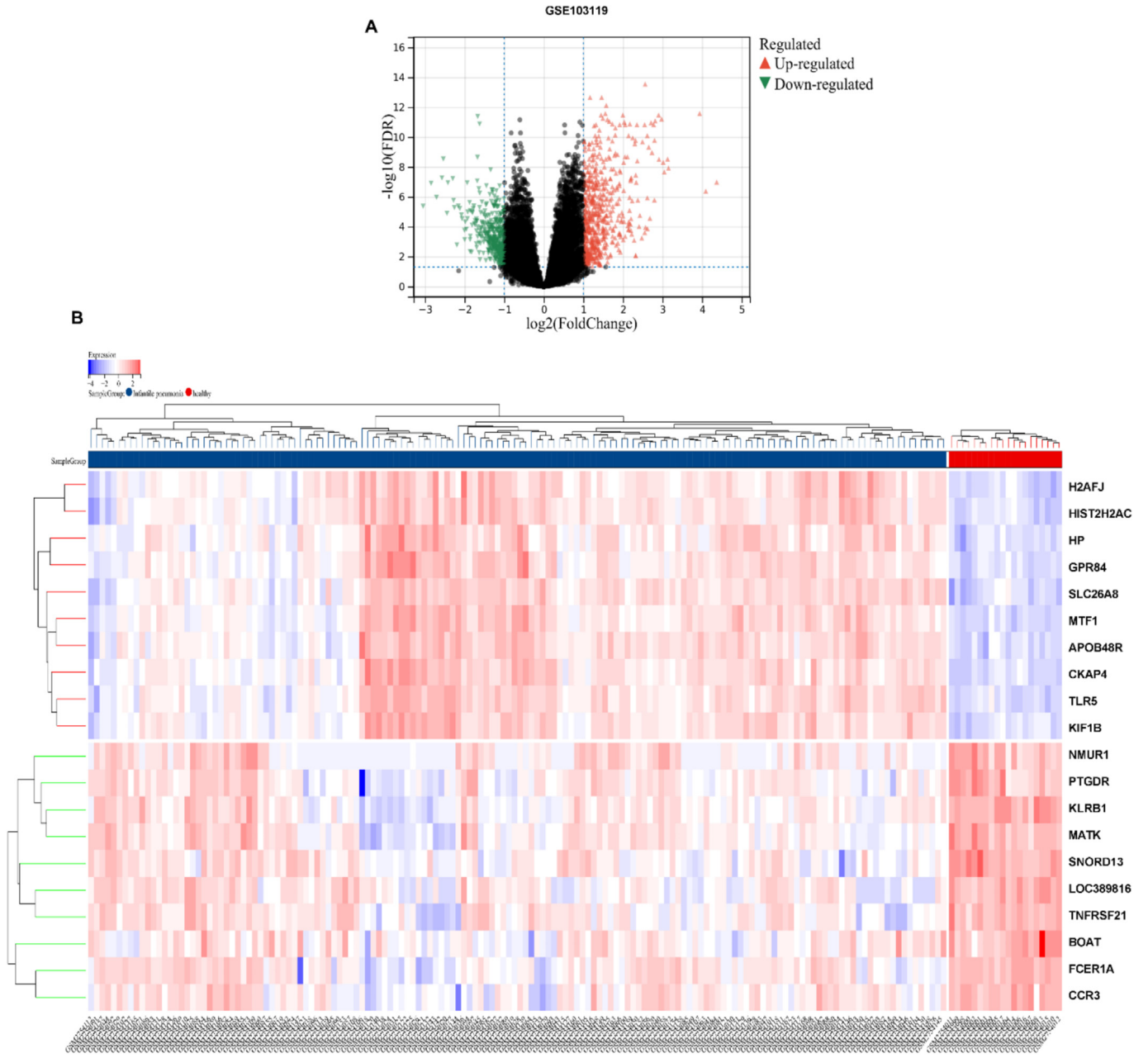


Fig. 1. Identification of differentially expressed genes (DEGs) in GSE103119 dataset. (A) Comparative analysis was conducted between 20 healthy controls and 152 infantile pneumonia patients. Up-regulated genes are depicted in red, while down-regulated genes are depicted in green. (B) A heatmap visualizes the differential expression pattern of the top 10 up-regulated and down-regulated DEGs between healthy samples (in red) and infantile pneumonia samples (in blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

out groups exhibited significantly elevated levels of p-p65 and p-IκBα, while p65 and IκBα levels did not show a significant difference (Fig. 6A, B). However, the elevated levels of p-p65 and p-IκBα were reversed by OLFM4 knockdown. These findings indicated that OLFM4 regulated the LPS-stimulated inflammatory response in WI-38 cells via mechanisms dependent on the NF-κB pathway.

3.7. OLFM4 overexpression aggravated LPS-stimulated inflammation damage via NF-κB pathway

Furthermore, complementary overexpression experiments were performed with NF-κB pathway inhibitor to explore the function of OLFM4. The overexpression of OLFM4 in WI-38 cells transfected with OE-OLFM4 or OE-NC was verified via western blot (Fig. S2A). The WI-38 cells were treated with LPS (1 μg/mL), LPS + OE-NC, LPS + OE-OLFM4, LPS + OE-OLFM4 + PDTC (10 μM), or without (Control). The cell vitality of LPS-treated WI-38 cells was further repressed by OLFM4 overexpression, while inhibition of NF-κB (PDTC) rescued this effect (Fig. S2B). In parallel, OLFM4 overexpression aggravated apoptosis, inflammation response (TNF-α and IL-1β), and ROS production in LPS-treated cells, which were alleviated by PDTC treatment (Fig. S2C-F). Taken together, OLFM4 overexpression aggravated LPS-stimulated WI-38 inflammation injury via NF-κB pathway.

3.8. Knockdown of OLFM4 alleviated the lung injury of LPS-induced mice

Finally, knockdown of OLFM4 *in vivo* was analyzed through a mouse model. Four groups of mice were used for *in vivo* study,

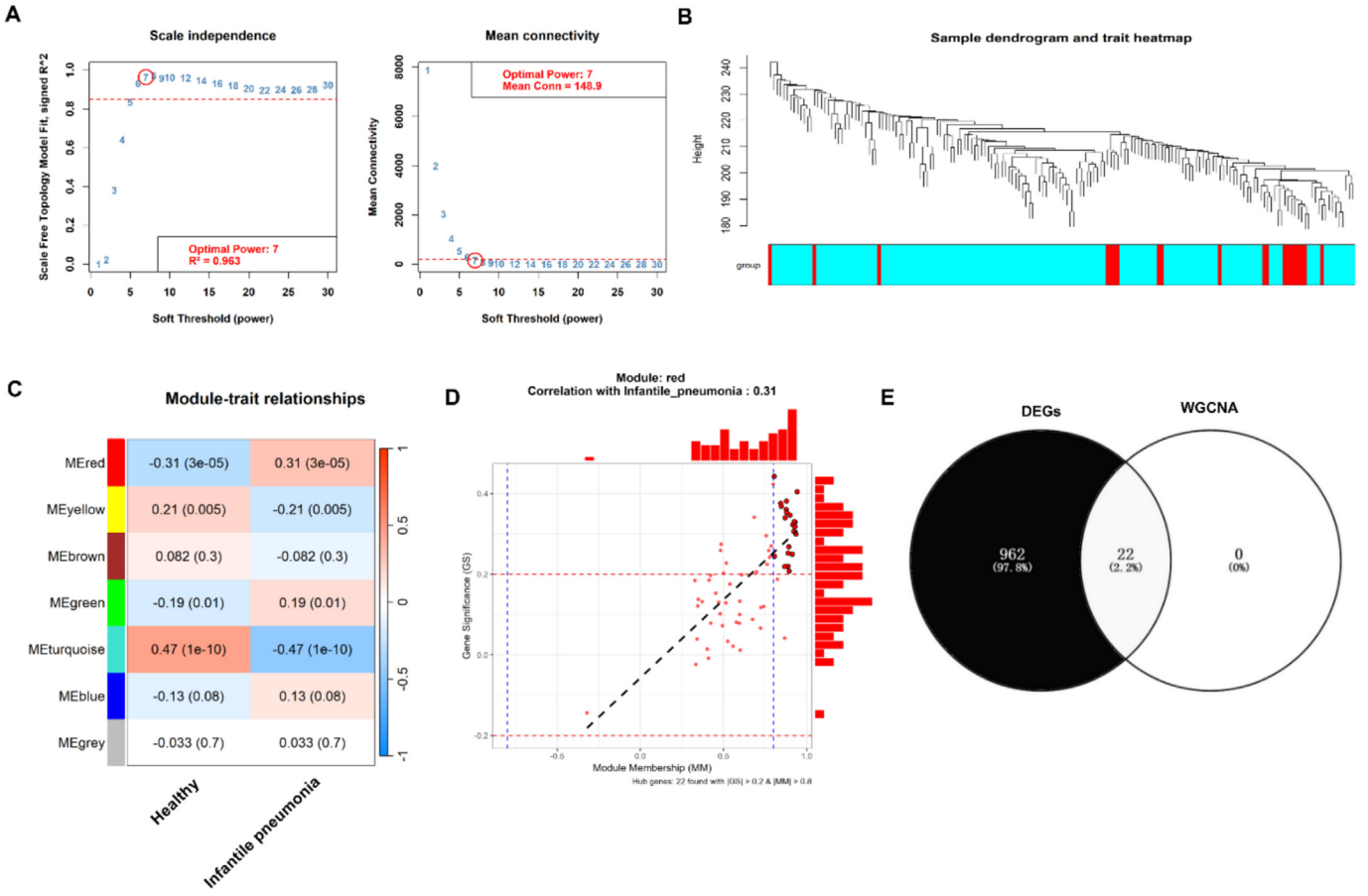


Fig. 2. The identification of pneumonia-related hub genes. (A) Determination of the soft threshold power through analysis, with the optimal value found to be 7. (B) Sample clustering analysis of all samples involved in WGCNA. (C) Extraction of seven gene modules through WGCNA. (D) In the red module, a scatter plot illustrates the relationship between module membership and gene significance. (E) A Venn diagram shows the intersection of DEGs with genes from the red module in WGCNA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

including Sham, LPS, LPS + Ad-sh-NC and LPS + Ad-sh-OLFM4. As shown in Fig. 7A,B, the lung tissue in the Sham group had a normal cell structure; besides, alveoli and bronchi were intact, and there was no inflammatory cell infiltration. In the LPS and LPS + Ad-sh-NC groups, the alveolar walls were significantly thickened, the alveolar structure was lost, and a large number of inflammatory cells infiltrated when compared with the Sham group, which was weakened by OLFM4 knockdown. Compared with the Sham group, lung WD ratio and MPO levels were remarkably increased in the LPS group, which was rescued by OLFM4 knockdown (Fig. 7C,D). The downregulation of OLFM4 in Ad-sh-OLFM4 mice was verified via RT-qPCR and western blot (Fig. 7E). IL-6, TNF-α, and IL-1β levels in the LPS group were increased when compared with the Sham group, while OLFM4 knockdown abrogated these effects (Fig. 7F, G,H). The p-p65 and p-IκBα expression was increased in the LPS group when compared with the Sham group, which was abolished by OLFM4 knockdown (Fig. 7I,J). Overall, OLFM4 knockdown remitted lung injury *in vivo*.

4. Discussion

In-depth research on pathophysiology reveals that children with infantile pneumonia exhibit not only severe pulmonary disease but also damage to many organs throughout the body [4]. Notwithstanding recent progress, the molecular mechanisms driving infantile pneumonia remain a pressing need. Meanwhile, as various informatics technologies advance, machine learning algorithms like WGCNA have matured and are extensively utilized for predicting disease markers and treatment targets. This study examined transcriptome data from the GEO database using several informatics tools and discovered four hub genes associated with infantile pneumonia: OLFM4, CEACAM6, COL17A1, and CAMP. CEACAM6 is involved in interstitial lung disease in early systemic sclerosis [33]. CAMP encodes antimicrobial peptide LL-37, which plays a significant role in the pathogenesis of various diseases, including psoriasis, systemic lupus erythematosus, diabetes mellitus, rheumatoid arthritis, cardiovascular pathologies, and multiple types of cancer [34]. COL17A1 encodes BPAG2, a protein involved in maintenance of skin structure [35]. OLFM4 was the sole gene linked to the inflammatory response of lung and was confirmed in the study.

LPS, a gram-negative bacterial endotoxin, significantly contributes to the pro-inflammatory responses linked to pneumonia [36,37]. LPS has been demonstrated to promote inflammation factor production, oxidative stress, and cellular apoptosis rate in WI-38 cells [38,39]. Consistently, this study demonstrated that LPS-induced WI-38 cells exhibited decreased cell viability and SOD levels, as well as elevated cell apoptosis, inflammation factors, MDA, and ROS levels. OLFM4 belongs to the olfactomedin family and demonstrates dysregulated expression in numerous inflammatory disorders, playing a role in their control [20,26]. OLFM4 is significantly expressed in polymorphonuclear myeloid-derived suppressor cells (PMN-MDSCs) and facilitates the transition from colitis to colorectal cancer by the NF-κB/PTGS2 pathway, whereas

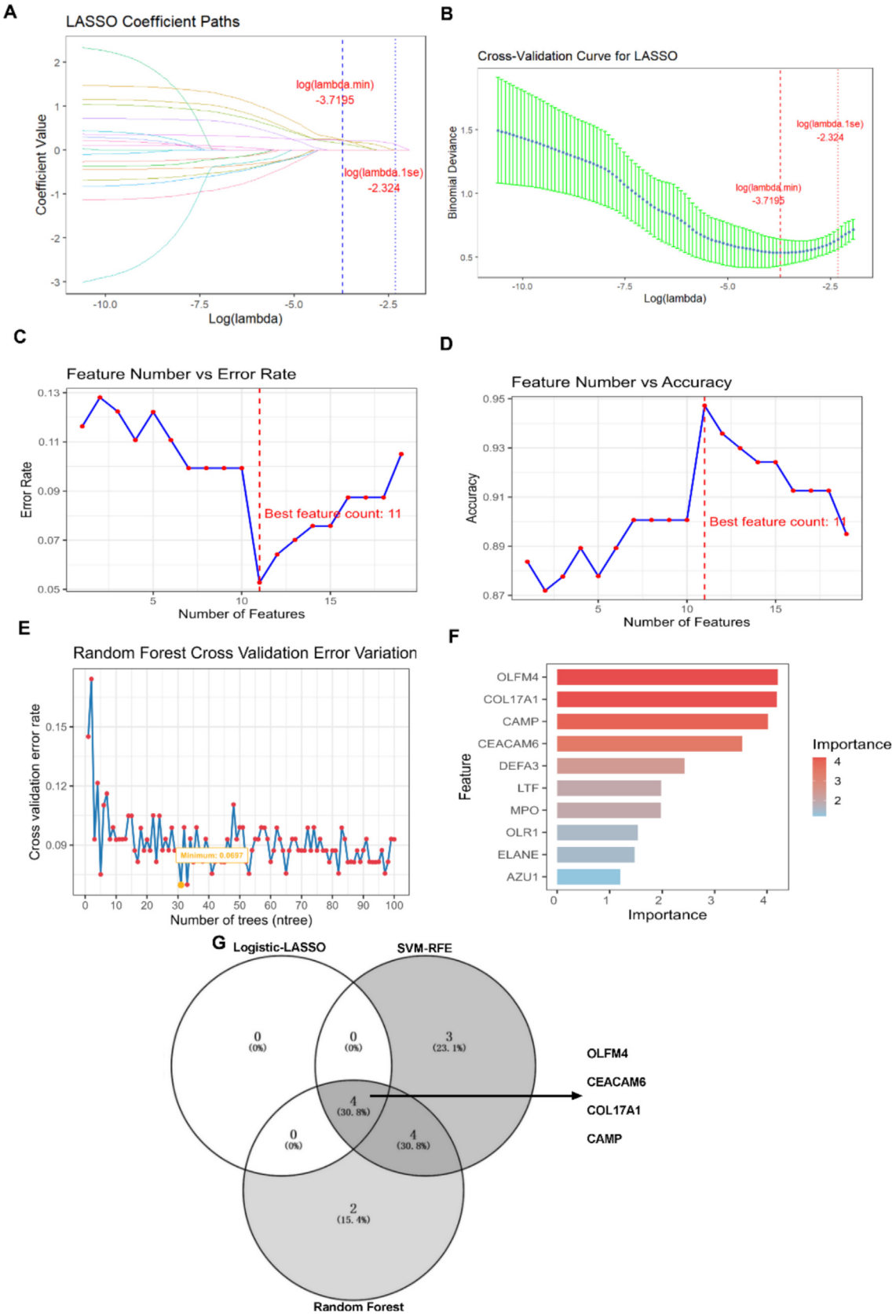


Fig. 3. Machine learning for identifying candidate genes in infantile pneumonia. (A) Path diagram of the LASSO coefficients for the hub genes in infantile pneumonia. (B) LASSO regression cross-validation curve. Optimal λ values were determined using 10-fold cross-validation. (C) Accuracy plot of the SVM-RFE algorithm. (D) Error plot of the SVM-RFE algorithm. (E) Correlation between the number of random forest trees and model errors. (F) RF importance score results (10 genes received importance score). (G) Venn diagram showing the intersection of the results from three machine learning algorithms. LASSO: Least Absolute Shrinkage and Selection Operator; SVM-RFE: Support Vector Machine-Recursive Feature Elimination; RF: Random Forest.

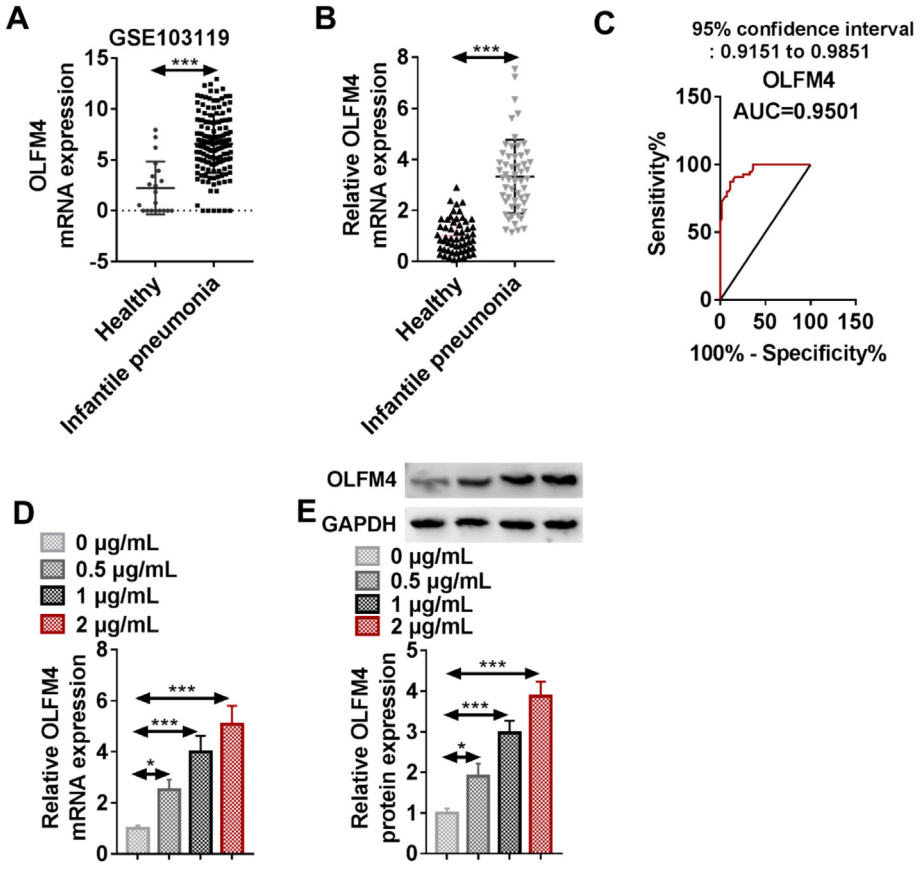


Fig. 4. OLFM4 levels are increased in LPS-stimulated WI-38 cells and in the serum of patients with pediatric pneumonia. (A) OLFM4 expression between 20 healthy controls and 152 infantile pneumonia patients from GSE103119 dataset. (B) OLFM4 expression in the serum of healthy individuals (n = 55) and patients with infantile pneumonia (n = 55) was detected using RT-qPCR. (C) Receiver operating characteristic curves for evaluating the diagnostic value of OLFM4. (D and E) WI-38 cells were induced with different doses of LPS (0, 0.5, 1, and 2 μg/ml) for 12 h, and then the mRNA and protein levels of OLFM4 were examined using RT-qPCR and western blot. $^{*}p < 0.05$, $^{***}p < 0.001$.

the absence of OLFM4 retards this progression [29]. In sepsis-induced ARDS, OLFM4 has been demonstrated to influence the NF-kB signaling pathway, thereby enhancing lung epithelial cell functionality [22]. The overexpression of OLFM4 significantly inhibits LPS-induced pro-inflammatory responses in lung epithelial cells and reduces NF-κB activation by decreasing ROS generation and HIF1α expression [22]. The current work revealed the overexpression of OLFM4 in patients with infantile pneumonia, LPS-induced WI-38 cells, and a murine model. ROC analysis demonstrated significant potential of serum OLFM4 for diagnosing infantile pneumonia. Knockdown of OLFM4 resulted in enhanced cell survival, decreased apoptosis rate, diminished production of inflammatory factors (IL-6, IL-1β, and TNF-α), lower levels of MDA and ROS, increased SOD levels, and inhibited activation of the NF-κB pathway in LPS-induced WI-38 cells. Moreover, OLFM4 knockdown ameliorated lung injury, inflammation, and oxidative stress, while inhibiting the activation of the NF-κB signaling pathway *in vivo*, suggesting that OLFM4 silencing relieved the severity of infantile pneumonia through the NF-κB pathway. In parallel, OLFM4 knockdown protects cardiomyocytes from sepsis via the NF-κB pathway [25]. Although OLFM4 knockdown has alleviated progression of several diseases, including sepsis, gastric cancer, respiratory diseases, and cervical cancer [20,25,27,28], few studies have focused on OLFM4 inhibitors. Qinbaohong Zhike oral liquid, a marketed traditional Chinese medicine product for respiratory diseases, has alleviated lung tissue damage and inflammatory reaction via inhibiting OLFM4 expression in LPS-treated immature rats [26]. However, the study does not explore the active components of Qinbaohong Zhike oral liquid that block OLFM4 gene expression. Therefore, considerable progress remains before OLFM4 may be regarded as a viable target for pneumonia treatment. The data indicated that OLFM4 could serve as a potential diagnostic and therapeutic target for infantile pneumonia.

This study has several limitations that warrant consideration. First, the study was performed using a single human lung fibroblast cell line (WI-38), which, despite its prevalent application in pulmonary inflammation research, fails to fully recapitulate the intricacies of lung tissue responses *in vivo*. The lung microenvironment comprises diverse cell types, including epithelial cells, immune cells, and endothelial cells, whose interactions are crucial in mediating the pathophysiology of pneumonia. To mitigate this restriction, future studies should integrate more physiologically relevant models, such as human bronchial epithelial cells (BEAS-2B), A549 cells, and advanced 3D lung organoid cultures. Second, the precise regulation mechanism of OLFM4 in infantile pneumonia development remains inadequately explored. Third, considering the fact that infantile pneumonia can be caused by viral, bacterial, and fungal pathogens, the present study was limited to explore OLFM4 in the LPS-induced (bacterial) model.

5. Conclusions

In summary, the present study identified OLFM4 as a biomarker associated with infantile pneumonia via bioinformatics and

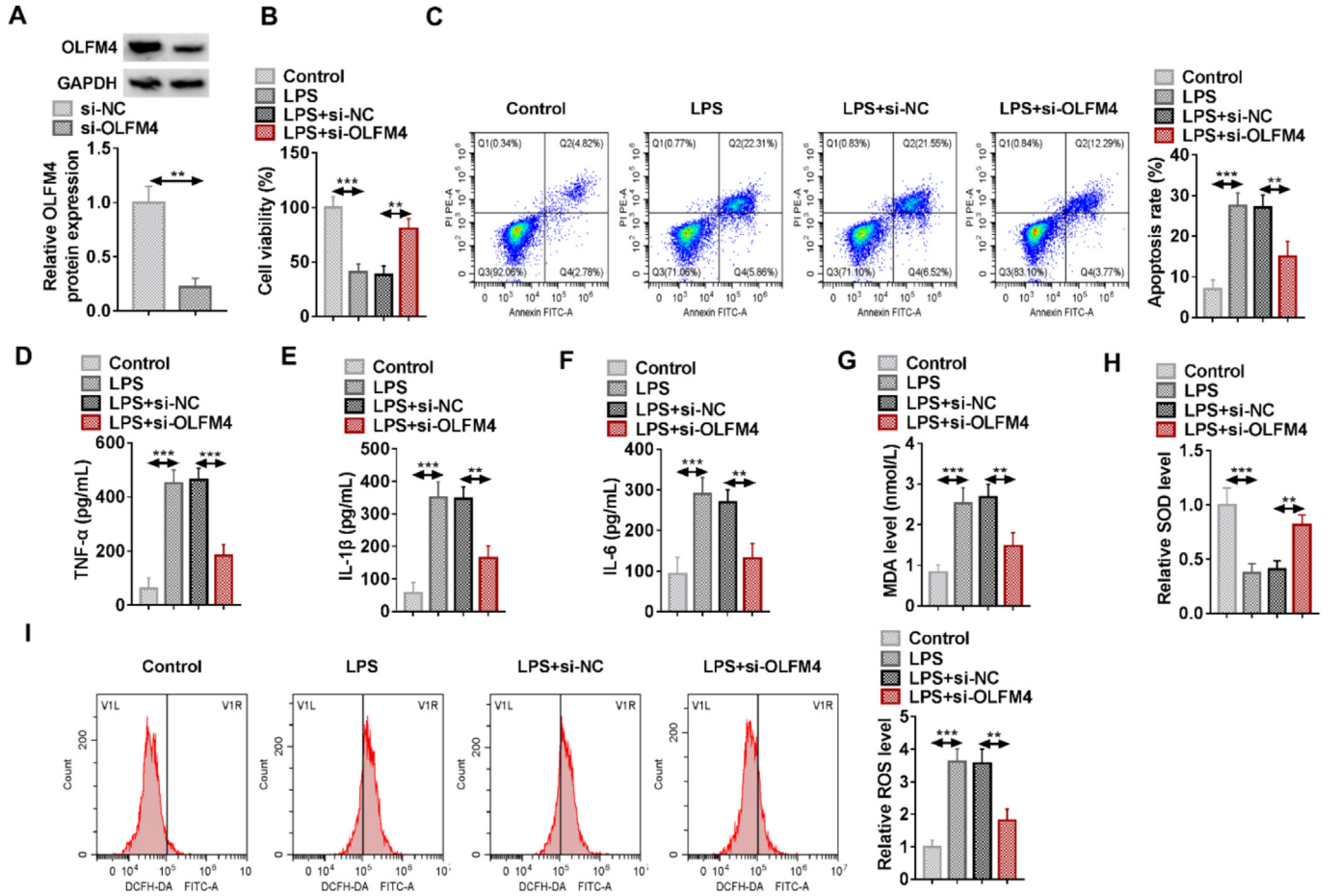


Fig. 5. OLFM4 knockdown alleviated LPS-stimulated WI-38 cell inflammation damage. The WI-38 cells were treated with LPS (1 μg/mL), LPS (1 μg/mL) + si-NC, LPS (1 μg/mL) + si-OLFM4 or without (Control). (A) OLFM4 expression of WI-38 cells transfected with si-OLFM4 or si-NC was verified via western blot. (B) CCK8 was employed to test cell viability. (C) Measurement of cell apoptosis rate by flow cytometry. (D, E, F) The corresponding ELISA assay kits were used to analyze the IL-6, TNF-α and IL-1β concentrations. (G, H) The MDA and SOD levels were measured by using the test kits. (I) Measurement of relative ROS level with flow cytometry. $^{**}p < 0.01$, $^{***}p < 0.001$.

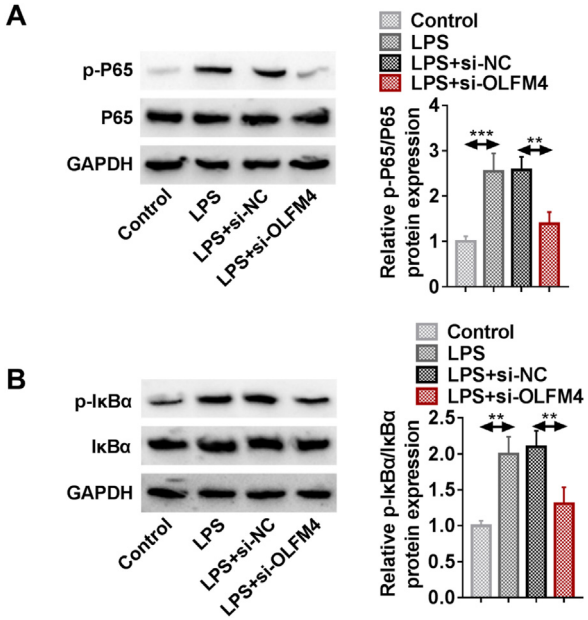


Fig. 6. OLFM4 knockdown affected NF-κB activation. The WI-38 cells were treated with LPS (1 μg/mL), LPS (1 μg/mL) + si-NC, LPS (1 μg/mL) + si-OLFM4 or without (Control). (A, B) Western blot and quantification of p-p65, p-IκBα, p65, and IκBα levels. $^{**}p < 0.01$, $^{***}p < 0.001$.

machine learning. A verification study showed that OLFM4 knockdown alleviated cellular apoptosis, inflammation response, and oxidative stress through the regulation of the NF-κB pathway in LPS-induced WI-38 cells or a mouse model. The results of this study might provide a new direction for clinical treatment of infantile pneumonia.

CRediT authorship contribution statement

Ying Wu: Writing – original draft, Validation, Supervision, Resources, Project administration, Formal analysis. **Baocen Cao:** Software, Methodology, Data curation. **Jie Liu:** Writing – review & editing, Visualization, Investigation, Funding acquisition, Conceptualization.

Ethical approval

The study received approval from the Ethical Committee of Huangshi Central Hospital. The animal experiment was approved by the Animal Research Committee of Huangshi Central Hospital.

Financial support

This study was supported by the Natural Science Foundation of Hubei Province (ZRMS2018001394).

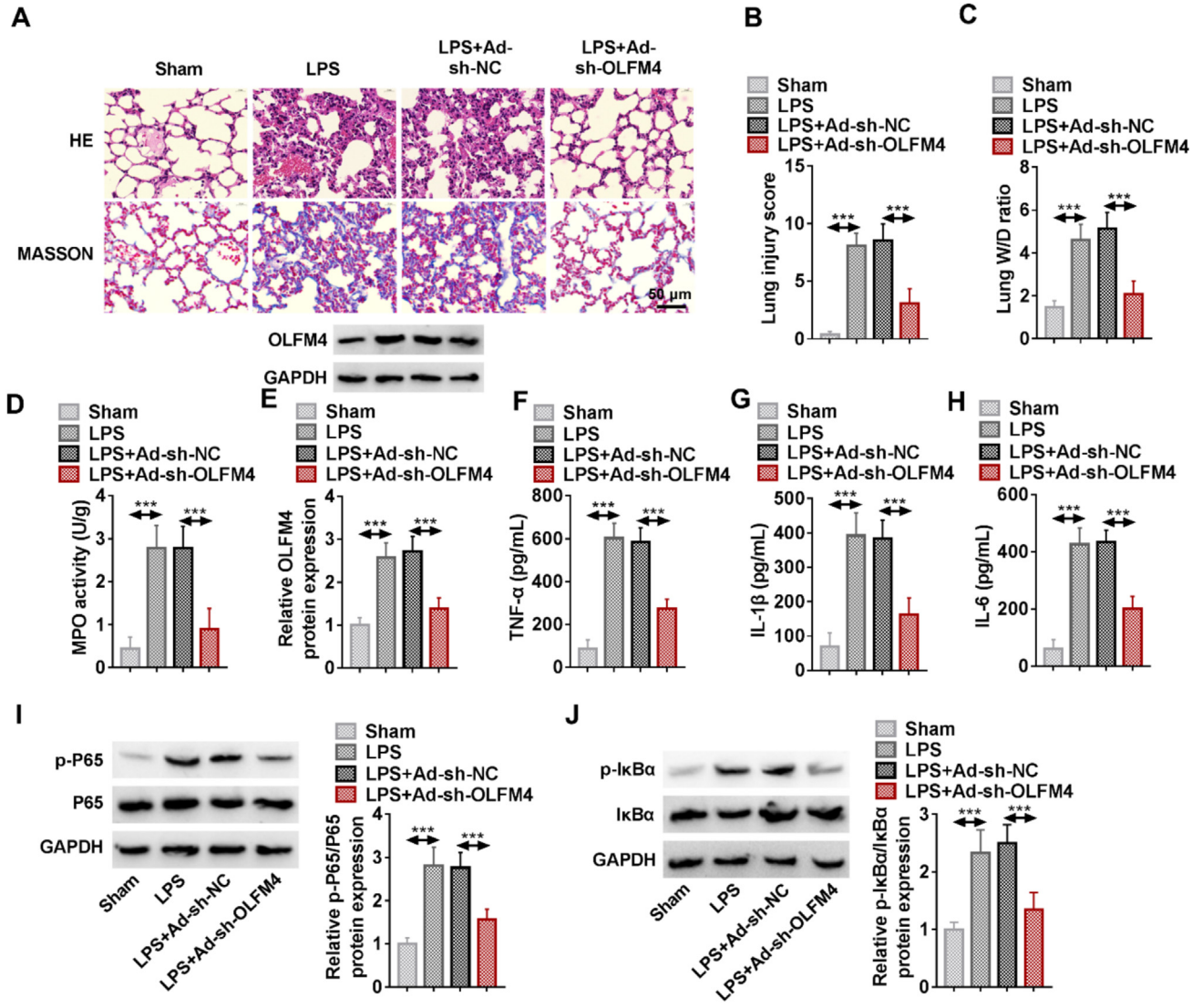


Fig. 7. OLFM4 knockdown ameliorated the lung injury induced by LPS in the model. The C57BL/6 mice were divided into four groups, including Sham (intraperitoneally received saline), LPS (intraperitoneally received LPS 2 mg/kg), Ad-sh-NC + LPS, and Ad-sh-OLFM4 + LPS groups, with five mice in each group. The mice received the treatment for 24 h. (A) The pathological feature of lung tissues was visualized using hematoxylin-eosin staining and Masson's trichrome staining. (B) Calculation of lung injury score. (C and D) Measurement of lung wet/dry (WD) ratio and MPO activity. (E) Western blot and quantification of lung OLFM4 expression. (F, G, H) TNF-α, IL-1β, and IL-6 concentrations in the blood of mice were measured with ELISA reagents. (I, J) Western blot and quantification of p-p65, p-IκBα, p65, and IκBα levels from the lung of the mice. $^{***}p < 0.001$.

Data availability

Data are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors have no conflict of interest to declare.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejbt.2026.100719.

References

[1] Zhang J, Wang CH, Liu XJ, et al. Efficacy and safety analysis of dopamine combined with creatine phosphate sodium in the treatment of infantile pneumonia combined with heart failure. J Biol Regul Homeost Agents 2020;34(6):2103–8. https://doi.org/10.23812/20-300-L. PMid: 33203205.
[2] Ojuawo O, Ojuawo A, Aladesanmi A, et al. Childhood pneumonia diagnostics: A narrative review. Expert Rev Respir Med 2022;16(7):775–85. https://doi.org/10.1080/17476348.2022.2099842. PMid: 35815336.
[3] Yadav KK, Awasthi S. Childhood pneumonia: What's unchanged, and what's new? Indian J Pediatr 2023;90(7):693–9. https://doi.org/10.1007/s12098-023-04628-3. PMid: 37204597.
[4] Meyer Sauteur PM. Childhood community-acquired pneumonia. Eur J Pediatr 2024;183(3):1129–36. https://doi.org/10.1007/s00431-023-05366-6. PMid: 38112800.
[5] Pratt MTG, Abdalla T, Richmond PC, et al. Prevalence of respiratory viruses in community-acquired pneumonia in children: A systematic review and meta-analysis. Lancet Child Adolesc Health 2022;6(8):555–70. https://doi.org/10.1016/S2352-4642(22)00092-X. PMid: 35636455.
[6] Li Q, Zhou Q, Florez ID, et al. Short-course vs long-course antibiotic therapy for children with nonsevere community-acquired pneumonia: A systematic review and meta-analysis. JAMA Pediatr 2022;176(12):1199–207. https://doi.org/10.1001/jamapediatrics.2022.4123. PMid: 36374480.
[7] Chee E, Huang K, Haggie S, et al. Systematic review of clinical practice guidelines on the management of community acquired pneumonia in children. Paediatr Respir Rev 2022;42:59–68. https://doi.org/10.1016/j.prrv.2022.01.006. PMid: 35210170.

[8] Lyon E, Olarte L. Community-acquired bacterial pneumonia in children: An update on antibiotic duration and immunization strategies. Curr Opin Pediatr 2024;36(2):144–9. https://doi.org/10.1097/MOP.0000000000001325. PMid: 38169463.
[9] Wang YS, Zhou YL, Bai GN, et al. Expert consensus on the diagnosis and treatment of macrolide-resistant *Mycoplasma pneumoniae* pneumonia in children. World J Pediatr 2024;20(9):901–14. https://doi.org/10.1007/s12519-024-00831-0. PMid: 39143259.
[10] Chen B, Lin N, Hu Y. Network pharmacology analysis of children pneumonia: Exploration of inflammation-related signaling pathways and potential therapeutic intervention targets. Lett Drug Des Discovery 2025;22(11):100220. https://doi.org/10.1016/j.lddd.2025.100220.
[11] Reel PS, Reel S, Pearson E, et al. Using machine learning approaches for multi-omics data analysis: A review. Biotechnol Adv 2021;49:107739. https://doi.org/10.1016/j.biotechadv.2021.107739. PMid: 33794304.
[12] Morabito A, De Simone G, Pastorelli R, et al. Algorithms and tools for data-driven omics integration to achieve multilayer biological insights: A narrative review. J Transl Med 2025;23(1):425. https://doi.org/10.1186/s12967-025-06446-x. PMid: 40211300.
[13] Sanches PHG, de Melo NC, Porcari AM, et al. Integrating molecular perspectives: Strategies for comprehensive multi-omics integrative data analysis and machine learning applications in transcriptomics, proteomics, and metabolomics. Biology 2024;13(11):848. https://doi.org/10.3390/biology13110848. PMid: 39596803.
[14] Cheng C, Li B, Li J, et al. Multi-stain deep learning prediction model of treatment response in lupus nephritis based on renal histopathology. Kidney Int 2025;107(4):714–27. https://doi.org/10.1016/j.kint.2024.12.007. PMid: 39733792.
[15] Wang L, Liu J, Zhang D, et al. Serum proteomics reveals differential proteins between individuals with vestibular migraine and healthy subjects. Curr Proteomics 2024;21(6):751–9. https://doi.org/10.2174/0115701646331735241231070050.
[16] Smelik M, Zhao Y, Li X, et al. An interactive atlas of genomic, proteomic, and metabolomic biomarkers promotes the potential of proteins to predict complex diseases. Sci Rep 2024;14(1):12710. https://doi.org/10.1038/s41598-024-63399-9. PMid: 38830935.
[17] Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779. PMid: 25826379.
[18] Ahsan MA, Liu Y, Feng C, et al. Bioinformatics resources facilitate understanding and harnessing clinical research of SARS-CoV-2. Brief Bioinform 2021;22(2):714–25. https://doi.org/10.1093/bib/bbaa416. PMid: 33432321.
[19] Xia Y, Zhang Z, Liu Q. Screening of immune-related secretory proteins linking primary Sjögren's syndrome with pulmonary arterial hypertension based on comprehensive bioinformatics analysis and machine learning. Curr Proteomics 2025;22(4):100032. https://doi.org/10.1016/j.curpro.2025.100032.
[20] O'Sullivan R, Alder MN, Dixon CG, et al. Olfactomedin-4 elevation in pediatric acute respiratory distress syndrome. Am J Physiol-Lung Cellular Mol Physiol 2025;329(1):L172–82. https://doi.org/10.1152/ajplung.00040.2025. PMid: 40499529.
[21] Kangelaris KN, Clemens R, Fang X, et al. A neutrophil subset defined by intracellular olfactomedin 4 is associated with mortality in sepsis. Am J Physiol-Lung Cellular Mol Physiol 2021;320(5):L892–902. https://doi.org/10.1152/ajplung.00090.2020. PMid: 33355521.
[22] Gong F, Li R, Zheng X, et al. OLFM4 Regulates lung epithelial cell function in sepsis-associated ARDS/ALI via LDHA-mediated NF-κB signaling. J Inflamm Res 2021;14:7035–51. https://doi.org/10.2147/JIR.S335915. PMid: 34955649.
[23] Hasson DC, Krallman K, VanDenHeuvel K, et al. Olfactomedin 4 as a novel loop of Henle-specific acute kidney injury biomarker. Physiol Rep 2022;10(18):e15453. https://doi.org/10.14814/phy2.15453. PMid: 36117416.
[24] Kangelaris KN, Prakash A, Liu KD, et al. Increased expression of neutrophil-related genes in patients with early sepsis-induced ARDS. Am J Physiol-Lung Cellular Mol Physiol 2015;308:L1102–13. https://doi.org/10.1152/ajplung.00380.2014. PMid: 25795726.
[25] Chen H, Liu S, Fang G. Knockdown of OLFM4 protects cardiomyocytes from sepsis by inhibiting apoptosis and inflammatory responses. Allergol Immunopathol 2024;52(5):15–20. https://doi.org/10.15586/aei.v52i5.1145. PMid: 39278846.
[26] Zhang F, Li Y, Xi Y, et al. Qinbaohong Zhike oral liquid attenuates LPS-induced acute lung injury in immature rats by inhibiting OLFM4. Oxid Med Cell Longev 2022;2022:7272371. https://doi.org/10.1155/2022/7272371. PMid: 36035204.
[27] Wen F, Han Y, Zhang H, et al. Epstein-Barr virus infection upregulates extracellular OLFM4 to activate YAP signaling during gastric cancer progression. Nat Commun 2024;15(1):10543. https://doi.org/10.1038/s41467-024-54850-6. PMid: 39627192.
[28] Li J, Liu C, Li D, et al. OLFM4 inhibits epithelial-mesenchymal transition and metastatic potential of cervical cancer cells. Oncol Res 2019;27(7):763–71. https://doi.org/10.3727/096504018X15399955297355. PMid: 30764901.
[29] Chen Z, Zhang X, Xing Z, et al. OLFM4 deficiency delays the progression of colitis to colorectal cancer by abrogating PMN-MDSCs recruitment. Oncogene 2022;41(22):3131–50. https://doi.org/10.1038/s41388-022-02324-8. PMid: 35487976.
[30] Brand HK, Ahout IM, de Ridder D, et al. Olfactomedin 4 serves as a marker for disease severity in pediatric Respiratory Syncytial Virus (RSV) infection. PLoS One 2015;10(7):e0131927. https://doi.org/10.1371/journal.pone.0131927. PMid: 26162090.
[31] Zhu-Ge D, Yang YP, Jiang ZJ. Knockdown CRNDE alleviates LPS-induced inflammation injury *via* FOXM1 in WI-38 cells. Biomed Pharmacother 2018;103:1678–87. https://doi.org/10.1016/j.biopha.2018.04.192. PMid: 29864958.
[32] Smith KM, Mrozek JD, Simonton SC, et al. Prolonged partial liquid ventilation using conventional and high-frequency ventilatory techniques: Gas exchange and lung pathology in an animal model of respiratory distress syndrome. Crit Care Med 1997;25(11):1888–97. https://doi.org/10.1097/00003246-199711000-00030. PMid: 9366775.
[33] Yokoyama K, Mitoma H, Kawano S, et al. CEACAM 1, 3, 5 and 6 -positive classical monocytes correlate with interstitial lung disease in early systemic sclerosis. Front Immunol 2022;13:1016914. https://doi.org/10.3389/fimmu.2022.1016914. PMid: 36341379.
[34] Voronko OE, Khotina VA, Kashirskikh DA, et al. Antimicrobial peptides of the cathelicidin family: Focus on LL-37 and its modifications. Int J Mol Sci 2025;26(16):8103. https://doi.org/10.3390/ijms26168103. PMid: 40869425.
[35] Sun Y, Qian Q, Xu L, et al. Mechanical stretch-induced interlayer coordination between MMP2 and COL17A1 exacerbates regenerative exhaustion in skin. Adv Sci 2025;12(41):e11474. https://doi.org/10.1002/advs.202511474. PMid: 40960906.
[36] Brooks D, Barr LC, Wiscombe S, et al. Human lipopolysaccharide models provide mechanistic and therapeutic insights into systemic and pulmonary inflammation. Eur Respir J 2020;56(1):1901298. https://doi.org/10.1183/13993003.01298-2019. PMid: 32299854.
[37] Nova Z, Skovierova H, Calkovska A. Alveolar-capillary membrane-related pulmonary cells as a target in endotoxin-induced acute lung injury. Int J Mol Sci 2019;20(49):831. https://doi.org/10.3390/ijms20040831. PMid: 30769918.
[38] Bai D, Han A, Cong S. The effect of down-regulation of CCL5 on lipopolysaccharide-induced WI-38 fibroblast injury: A potential role for infantile pneumonia. Iran J Basic Med Sci 2018;21(5):449–54. https://doi.org/10.22038/ijbms.2018.27165.6640. PMid: 29922423.
[39] Xiao H, Ma W, Zha L, et al. Curcumin alleviates LPS-induced WI-38 cell inflammation injury by regulating PTGS2 expression. Hereditas 2025;162(1):81. https://doi.org/10.1186/s41065-025-00441-4. PMid: 40380246.