Dovepress
Taylor & Francis Group

ORIGINAL RESEARCH

Identification of EGR1 as a Key Diagnostic Biomarker in Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) Through Machine Learning and Immune Analysis

Xuanlin Wu [iD], Tao Pan, Zhihao Fang, Titi Hui, Xiaoxiao Yu, Changxu Liu, Zihao Guo, Chang Liu

Department of General Surgery, Fourth Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, People's Republic of China

Correspondence: Chang Liu, Department of General Surgery, Fourth Affiliated Hospital of Harbin Medical University, 37 Yiyuan Street, Nangang District, Harbin, Heilongjiang, 150001, People's Republic of China, Tel +86-13313699697, Email lc19726666@163.com

Background: Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD), as a common chronic liver condition globally, is experiencing an increasing incidence rate which poses significant health risks. Despite this, the detailed mechanisms underlying the disease's onset and progression remain poorly understood. In this study, we aim to identify effective diagnostic biomarkers for MASLD using microarray data combined with machine learning techniques, which will aid in further understanding the pathogenesis of MASLD.

Methods: We collected six datasets from the Gene Expression Omnibus (GEO) database, using five of them as training sets and one as a validation set. We employed three machine learning methods—LASSO, SVM, and Random Forest (RF)—to identify hub genes associated with MASLD. These genes were further validated using the external dataset GSE164760. Additionally, functional enrichment analysis, immune infiltration analysis, and immune function analysis were conducted. A TF-miRNA-mRNA network was constructed, and single-cell RNA sequencing was used to determine the distribution of key genes within key cell clusters. Finally, the expression of the key genes was further validated using the palmitic acid-induced AML-12 cell line and the MCD mouse model.

Results: In this study, through differential gene expression (DEGs) analysis and machine learning techniques, we successfully identified 10 hub genes. Among these, the key gene EGR1 was validated and screened using an external dataset, with an area under the curve (AUC) of 0.882. Enrichment analyses and immune infiltration assessments revealed multiple pathways involving EGR1 in the pathogenesis and progression of MASLD, showing significant correlations with various immune cells. Furthermore, additional cellular experiments and animal model validations confirmed that the expression trends of EGR1 are highly consistent with our analytical findings.

Conclusion: Our research has confirmed EGR1 as a key gene in MASLD, providing novel insights into the disease's pathogenesis and identifying new therapeutic targets for its treatment.

Keywords: metabolic dysfunction-associated steatotic liver disease, immune infiltration, machine learning, TF-miRNA-mRNA network, EGR1

Introduction

Metabolic dysfunction-associated liver disease (MASLD) is the most prevalent chronic liver condition globally, primarily characterized by excessive fat accumulation within the liver. Unlike non-alcoholic fatty liver disease (NAFLD), MASLD places greater emphasis on metabolic dysfunction.[1] The diagnostic criteria for MASLD include the presence of hepatic steatosis confirmed by imaging or biopsy, along with at least one cardiometabolic risk factor, in the absence of other etiologies that could cause fatty liver changes.[2] MASLD can progress from simple steatosis to more severe forms such as metabolic dysfunction-associated steatohepatitis (MASH), fibrosis, cirrhosis, and MASH-related hepatocellular carcinoma.[3] Approximately 32.4% of the global population is affected by MASLD. Despite its high prevalence and the substantial burden it poses on global health, there remains a significant gap in effective treatment options.[4,5]

1639

The pathogenesis of MASLD is highly complex, as it is a metabolic disorder triggered by multiple factors. The prevailing theory is the "multiple parallel hits hypothesis", which posits that the development of MASLD involves the simultaneous impact of various factors, including insulin resistance, obesity, adipose tissue dysfunction, mitochondrial dysfunction, and endoplasmic reticulum stress.[6,7] Research indicates a close relationship between the gut microbiota and MASLD, with gut microbial metabolites forming a gut-liver axis. Strategies to improve gut microbiota can effectively correct the dysbiosis associated with MASLD and contribute to improved prognosis and insulin resistance recovery.[8] Additionally, studies suggest that interactions between genetic and epigenetic factors also determine individual susceptibility to MASLD.[9]

Furthermore, MASLD is a metabolic disorder closely linked to metabolic syndrome (MetS), type 2 diabetes (T2DM), cardiovascular diseases, and chronic kidney disease.[10–12] Despite this, effective treatments for MASLD are still limited. Currently, the primary management strategies include lifestyle modifications, pharmacological therapy, and, if necessary, surgical interventions.[13] However, these treatments often fail to achieve the desired outcomes. Therefore, understanding the mechanisms driving the onset and progression of MASLD is essential for the development of new therapeutic approaches.

Early growth response 1 (EGR1) is a transcription factor that exerts its biological effects through its characteristic zinc finger domain. It plays a crucial role in processes such as cell growth, proliferation, and development.[14–16] Recent studies suggest that EGR1 may be involved in metabolic diseases, inflammatory responses, and liver diseases. It has been reported that EGR1 may integrate central and peripheral circadian rhythms and regulate lipid metabolic homeostasis in the liver, thereby reducing the formation of lipid droplets.[17] Additionally, EGR1 can mediate the activation of leptin by insulin, contributing to lipid metabolism.[18] In the context of inflammation, EGR1 inhibits the expression of pro-inflammatory genes in macrophages, thereby alleviating inflammatory responses.[19] Furthermore, the transcriptional activation of EGR1 has been shown to promote the repair of acetaminophen-induced liver injury and enhance liver regeneration.[20] Based on these findings, we propose EGR1 as a key diagnostic biomarker and aim to further investigate its role in MASLD.

In this study, we integrated five datasets and conducted differential expression analysis. We used three distinct machine learning algorithms to further filter the differentially expressed genes. The key genes identified were validated using an external dataset and analyzed for immune infiltration using the CIBERSORT algorithm. We also explored the relationships between the key gene and immune cells, conducted enrichment analyses, constructed a TF-miRNA-mRNA network and performed single-cell analysis. Finally, the findings were validated in cellular and animal models.

Materials and Methods
Microarray Datasets Acquisition
Six microarray datasets were initially retrieved from the GEO database (http://www.ncbi.nlm.nih.gov/), including GSE48452 (Control = 14, MASLD = 32), GSE63067 (Control = 7, MASLD = 11), GSE66676 (Control = 34, MASLD = 33), GSE89632 (Control = 24, MASLD = 39), GSE107231 (Control = 5, MASLD = 5), and GSE164760 (Control = 6, MASLD = 74). The flowchart of the current study is illustrated in Figure 1.

Analysis of Datasets
The series matrix expression files for the six datasets were obtained, and the corresponding platform annotation files were employed to transform the probe expression matrices into gene expression matrices. RMA normalization was performed using the "affy" package. In cases where multiple probes corresponded to a single gene, the mean expression value was calculated. The five training sets were merged next, and batch effects were corrected using the "combat" function from the "sva" package.[21] The effectiveness of batch correction was validated by comparing PCA plots before and after correction. The "limma" package[22] within R software was utilized to identify differentially expressed genes (DEGs) between the combined groups of patients with Metabolic dysfunction-associated steatotic liver disease (MASLD) and healthy controls (HC). The thresholds for statistical significance were established at $|Log2FoldChange| \geq 0.5$ and adjusted p-value (adj. P) <0.05. Furthermore, the "ggplot2" and "pheatmap" packages in R were employed to generate volcano plots and heatmaps for these DEGs.
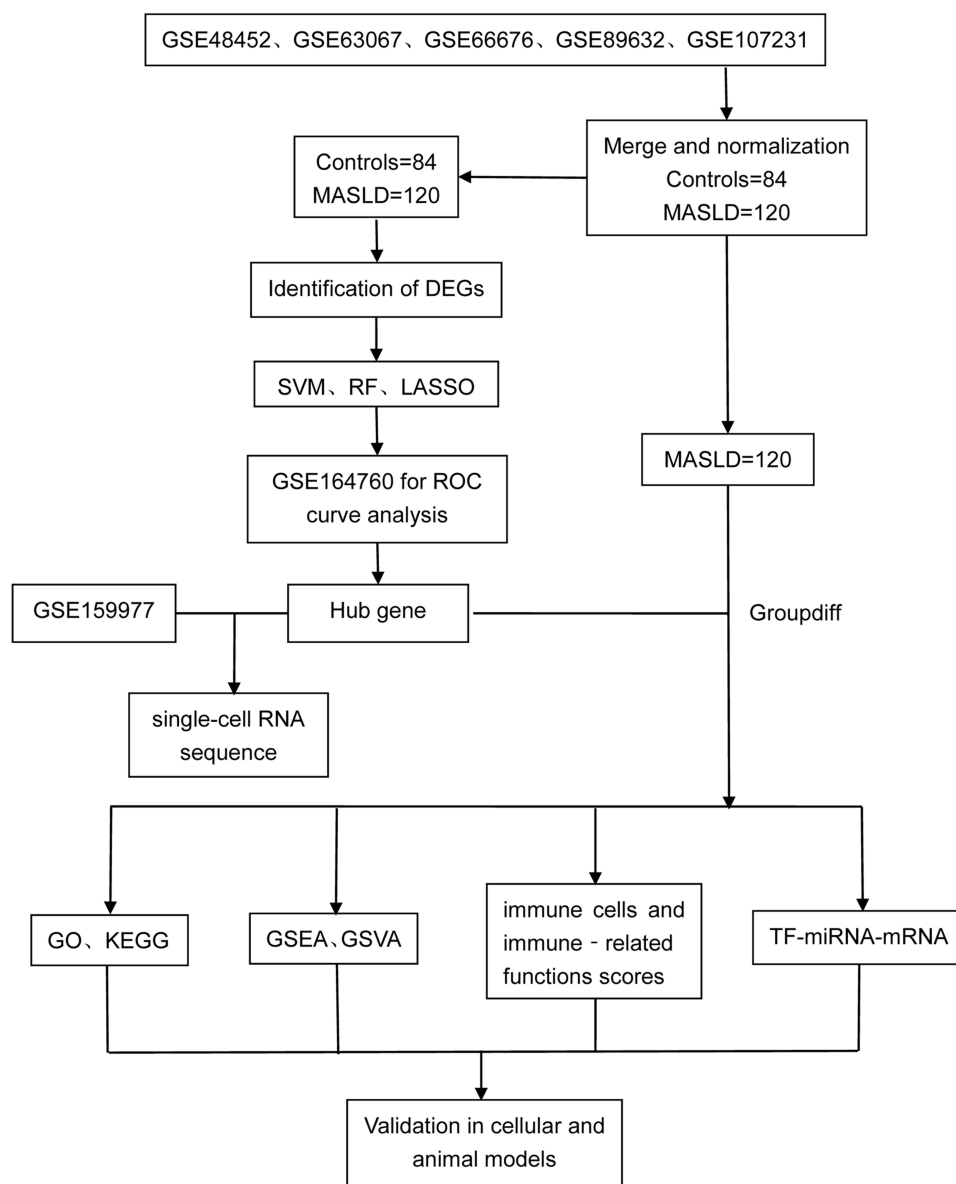
Figure 1 Research Flow Chart.

Machine Learning

The Random Forest (RF) algorithm,[23] Support Vector Machine-Recursive Feature Elimination (SVM-RFE), and Least Absolute Shrinkage and Selection Operator (LASSO) regression[24] were utilized to identify genes with the greatest impact on the prognosis of MASLD. A Venn diagram was employed to determine the intersection of genes identified by these three machine learning techniques, pinpointing hub genes for further research (Details of the parameters and hyperparameters used in the machine learning algorithms are provided in the Supplementary Material).

External Dataset Validation

To further validate intersecting genes, we evaluated them in the external dataset GSE164760 through time-dependent ROC curves. This evaluation included the area under the curve (AUC), specificity, and sensitivity, and also determined the differential expression of these genes between MASLD and control groups. Visualization of the results was performed using the R packages "pROC"[25] and "ggpubr".

Function Enrichment Analysis

To elucidate the potential biological functions and pathways of genes, we employed the "clusterProfiler" R package[26] for functional enrichment analysis. Gene Ontology (GO) is a powerful tool for examining genes' molecular functions, cellular components, and biological processes. The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, a broadly utilized database, enables researchers to explore the biological functions of gene sets and pinpoint the crucial pathways or processes involved in specific biological phenomena.

Gene Set Variation Analysis

The "c2.cp.Kegg.Hs.symbols" and "c5.go.Hs.symbols" files were initially downloaded from the MSigDB database. Subsequently, the R package "GSVA"[27] was utilized to elucidate the enrichment differences between GO and KEGG pathways using a non-parametric, unsupervised gene set variation analysis (GSVA) approach.

Gene Set Enrichment Analysis

We use the "clusterProfiler" package for Gene Set Enrichment Analysis (GSEA), to determine whether KEGG-specific pathways are enriched for a set of related genes.

Immune Function Score and Immune Infiltration Analysis

We utilized the CIBERSORT algorithm to quantify the varying levels of infiltration among 22 immune cell types in tissue samples. In each sample, the cumulative score of all immune cells was normalized to 1. Additionally, we calculated correlations between the immune cells and visualized these relationships through histograms, heatmaps, and various other methods. Based on expression profile information, we employed the "GSVA" package for Single Sample Gene Set Enrichment Analysis (ssGSEA) to examine the differences in immune cell and immune function scores between groups with high and low key gene expression. We assessed the enrichment scores of 16 immune cells and 13 immune functions to further investigate the association between key gene expression and MASLD.

TF-miRNA-mRNA Regulatory Network

Potential transcription factors for target genes were predicted using the Trrust database (https://www.grnpedia.org/trrust), while candidate miRNAs were identified via the Starbase database (https://rnasysu.com/encori). The results were visualized using Cytoscape.

Processing of Single-Cell Sequencing Data

The single-cell RNA sequencing (scRNA-seq) dataset (GSE159977) was retrieved from the publicly accessible GEO database. We processed the scRNA-seq data using the "Seurat" R package.[28] Low-quality cells were excluded based on the following criteria: cells expressing fewer than three genes, containing fewer than 300 genes, or exhibiting mitochondrial gene content exceeding 5%. Data normalization was performed with the "NormalizeData" function using the "LogNormalize" method. To identify the top 2000 highly variable genes, we applied the "FindVariableFeatures" function. Principal component analysis (PCA) was conducted on these highly variable genes using the "RunPCA" function, and the top 15 principal components (PCs) were selected for further analysis. Cell clustering was performed using the "FindNeighbors" and "FindClusters" functions. Subsequently, UMAP analysis was conducted with the "RunUMAP" function, and cells were clustered based on UMAP-1 and UMAP-2 coordinates. To annotate cell clusters, we utilized the "SingleR" R package with the Human Primary Cell Atlas as the reference dataset.

Cell Model

The AML-12 mouse hepatocyte cell line (CL-0602) was purchased from Procell Life Science & Technology Company Limited. The cells were cultured in DMEM/F12 medium (PM150312, Procell, Wuhan, China), supplemented with 10% fetal bovine serum (164210–50, Procell, Wuhan, China) and 1% penicillin/streptomycin (PB180120, Procell, Wuhan, China), and incubated at 37°C in a 5% CO_2 atmosphere. The cells were evenly seeded in a 6-well tissue culture plates at

a density of 1×10^5 cells/well, and treated with 200 μM palmitic acid (KC003, Kunchuang, Xi'an, China) for 48 hours to establish a hepatocyte steatosis model.

Oil Red O Staining

Prepare the Oil Red O staining kit (G1262, Solarbio Science & Technology Co., Ltd., Beijing, China). Wash the prepared cells twice with PBS, then fix them in Oil Red O fixative for 20 minutes. Rinse the cells twice with distilled water and soak in 60% isopropanol for 30 seconds. Add freshly prepared Oil Red O staining solution and incubate for 20 minutes. After rinsing, re-stain the cells with Mayer's hematoxylin solution. Finally, incubate the cells in Oil Red O buffer and observe under a light microscope.

Animal Models

Twelve six-week-old male C57BL/6J mice were used in this experiment, purchased from Liaoning Changsheng Biotechnology Company Limited (Liaoning, China), maintained in standard cages at a controlled ambient temperature of 22°C ± 2°C and subjected to a 12-hour dark/light cycle. There are no restrictions on the availability of water and food. The mice were randomly divided into two groups, with six mice in each group: Control group: standard diet; Treatment group: methionine-choline deficient diet (XSYT-ED-009, 22% fat, 15% protein, 63% carbohydrates). After 8 weeks, blood and liver tissues were collected from the mice. The protocol for these animal experiments received approval from the Professional Committee for Animal Protection of Harbin Medical University, and all procedures adhered to the applicable guidelines and regulations (2023-DWSYLLCZ-04), and performed following Regulations on the Management of Laboratory Animals in Heilongjiang Province and Laboratory Animals General requirements for biosafety in animal experiments.

Histological Procedure

To verify the success of the modeling, serological and histopathological examinations were conducted on the specimens. Fresh liver tissue was fixed in formalin for 24 hours, embedded in paraffin, dehydrated, and sectioned for H&E staining. Additionally, fresh liver tissue was fixed in 4% paraformaldehyde solution, dehydrated, and stained with Oil Red O for frozen sections. To eliminate the potential impact of food on the serum biomarkers, all mice were fasted for 12 hours prior to the experimental treatment, with free access to water. Serum levels of total cholesterol (A111-1-1, JianCheng Bioengineering Institute, Nanjing, China) and triglycerides (A110-1-1, JianCheng Bioengineering Institute, Nanjing, China) were quantified using commercial kits on an automatic blood biochemistry analyzer (VITROS 5600, Abbott, Chicago, USA), according to the manufacturer's instructions. Total cholesterol levels were measured using the COD-PAP method, while triglyceride levels were quantified using the GPO-PAP method. The criteria for diagnosing MASLD in mice include the following: (1) biochemical analysis showing elevated serum levels of total cholesterol (TC) and triglycerides (TG); (2) hepatic histological examination, including H&E staining and Oil Red O staining, to assess liver fat deposition.

Quantitative RT-PCR Analysis

RNA was extracted from cells and mouse liver tissues using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). RNA purity was assessed using a Nanodrop 2000 spectrophotometer, employing UV absorbance measurements. Followed by reverse transcription using PrimeScript reverse transcriptase (Takara, Kusatsu, Japan). Gene expression was analyzed using 2×SYBR Green qPCR (Vazyme, Nanjing, China). The relative expression of hub gene was quantified using the $2^{-\Delta\Delta Ct}$ method, with β-actin serving as the internal reference control. Primer sequences are provided in Table 1.

Western Blot

Total protein was extracted from cell and tissue samples using RIPA buffer, with the addition of a protease inhibitor to prevent protein degradation. After separating proteins by electrophoresis on a 10% SDS-PAGE, they were transferred to a PVDF membrane via electroblotting. The membrane was then blocked with 5% non-fat milk at room temperature for 1 hour. Primary antibodies were incubated with the membrane overnight at 4°C. The primary antibodies included rabbit polyclonal antibody

Table 1 Primers Used for RT-PCR Analysis

Gene	Species	Forward Primer	Reverse Primer
β-actin	Mouse	GTGACGTTGACATCCGTAAAGA	GCCGGACTCATCGTACTCC
EGR-1	Mouse	AACAACCCTATGAGCACCTG	GAGTCGTTTGGCTGGGATAA
TNF-α	Mouse	CAGGCGGTGCCTATGTCTC	CGATCACCCCGAAGTTCAGTAG
IL-6	Mouse	TCCATCCAGTTGCCTTCTTG	TTCCACGATTTCCCAGAGAAC
IL-1β	Mouse	GCCTGTGTTTTCCTCCTTGC	TGCTGCCTAATGTCCCCTTG
CCL2	Mouse	TTAAAAACCTGGATCGGAACCAA	GCATTAGCTTCAGATTTACGGGT

EGR1 (Abmart, PA1373, Shanghai, China) and β-actin (ABclonal, AC006, Wuhan, China). Subsequently, the membrane was incubated with the appropriate secondary antibody (ABclonal, AS014, Wuhan, China) for 1.5 hours at room temperature. Signal detection was performed using a Bio-Rad imaging system (Bio-Rad, CA, USA) with chemiluminescent substrate.

Data Analysis

Biostatistical analysis was carried out using Rstudio (version 4.3.1), while statistical analysis and graphical representation of experimental data were executed using GraphPad Prism version 9.5 software. Group differences for variables with a normal distribution were assessed using Student's t-test. The data in this study are presented as mean ± standard deviation. P-value<0.05 was considered indicative of statistical significance.

Results

Identification of DEGs and Analysis of Immune Infiltration

We downloaded five datasets from the Gene Expression Omnibus (GEO) database: GSE48452, GSE63067, GSE66676, GSE89632, and GSE107231. These datasets were subsequently merged and normalized. Supplementary Figures 1 and 2 illustrates the changes in the data before and after batch correction. Through differential analysis, we identified 42 differentially expressed genes, including 17 upregulated and 25 downregulated genes. The expression patterns of these genes are displayed in a heatmap (Figure 2A) and a volcano plot (Figure 2B).

Complex interactions between various immune cell populations and hepatocytes, stellate cells, and sinusoidal endothelial cells play a critical role during the pathogenesis of MASLD.[29] Therefore, we employed the CIBERSORT algorithm to investigate differences in the immune microenvironment between MASLD patients and healthy controls. Our analysis revealed alterations in the proportions of 22 different immune cell types between the MASLD and control groups (Figure 2C). Figure 2D presents a correlation analysis among these 22 immune cell groups within the MASLD cohort. Figure 2E displays significantly differentially expressed immune cell types between the two groups, notably highlighting an enriched expression of monocytes in the MASLD group. These findings suggest that alterations in the immune microenvironment may play a crucial role in the development of MASLD.

Hub Gene Selection Through Machine Learning Techniques

To further identify Hub differentially expressed genes between the MASLD group and the control group, we employed three machine learning algorithms: LASSO, RandomForest, and SVM-RFE (Figure 3A–F). We used a Venn diagram to intersect the key genes identified by each algorithm, ultimately identifying ten hub genes (Figure 3G), the ten hub genes identified through our analysis are as follows: CYP7A1, PEG10, P4HA1, IGFBP2, IL6, ME1, NR4A2, VIL1, TMEM154, and EGR1.

External Dataset Validation

Significant differential expression of ten characteristic genes was observed between MASLD patients and control samples (Supplementary Figure 3A–J), indicating their potential pivotal roles in MASLD progression. Further, ROC analysis was performed to assess their prognostic value (Supplementary Figures 4A–J), revealing that all genes possessed substantial diagnostic utility.
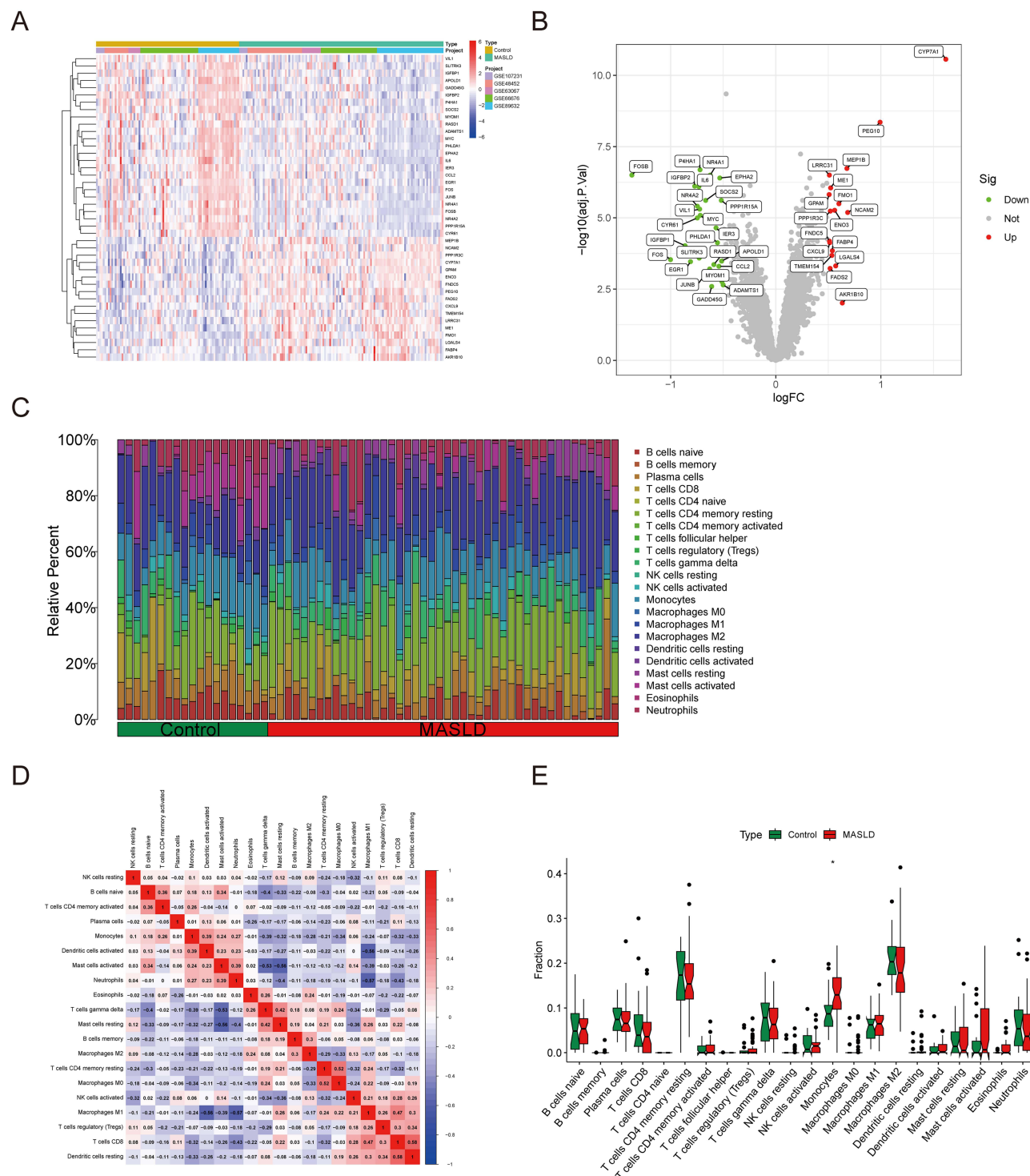
Figure 2 Preprocessing of data and Expression profile of DEGs.(**A**)The heatmap of differential gene expression.(**B**)Volcano plots of DEGs,|log2 fold change (FC)|≥0.5 and p < 0.05 as selection criteria, The red and green dots indicate the up-regulated and down-regulated DEGs, respectively.(**C**)Relative abundance of 22 immune cells between MASLD and control samples.(**D**)The cellular correlation among the 22 immune cells.(**E**)Comparative distribution analysis of 22 immune cell infiltrates between MASLD and control samples.*P < 0.05.

To confirm these key genes' expression patterns and diagnostic capabilities, additional validation was conducted using the external dataset GSE164760, involving gene expression verification and ROC curve analysis. The results demonstrated upregulation in CYP7A1, PEG10, and TMEM154, while EGR1, IGFBP2, NR4A2, and P4HA1 showed downregulation (Figure 4A–G). Further ROC analysis was performed on hub genes using the validation dataset GSE164760, and the area
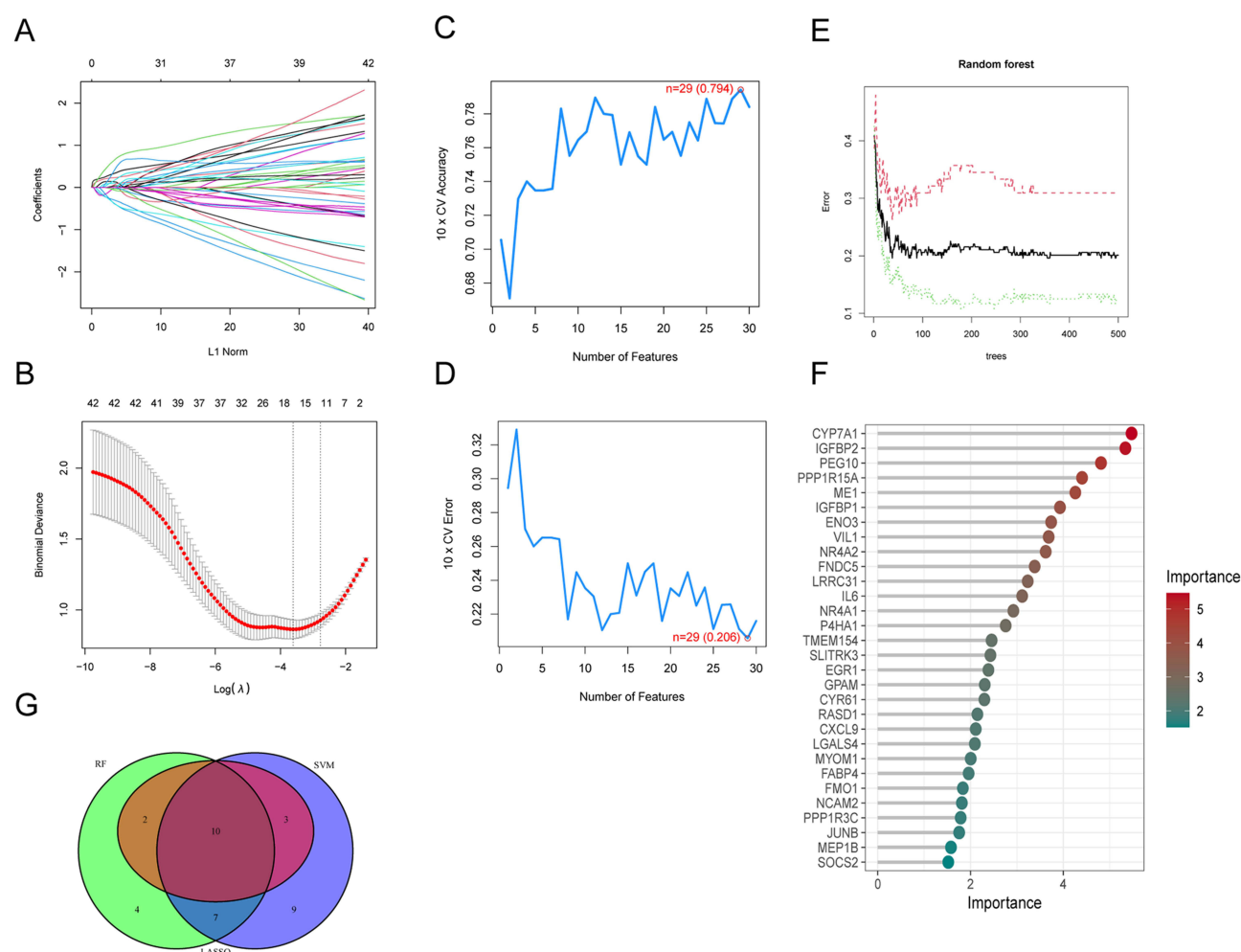
Figure 3 Machine learning to screen hub genes. (**A**)Ten-time cross-verification of adjusted parameter in selection least absolute shrinkage and selection operator (LASSO). Each curve in the picture represents one gene.(**B**)The solid vertical lines represent the partial likelihood deviance SE, the lowest point of the dotted vertical line represent the number of genes (n = 15).(**C**)The accuracy and (**D**) the error of the feature selection for the SVM-RFE algorithm. n=29 has the highest accuracy and the lowest cross-validation error.(**E**)Relationship between the number of random forest trees and error rates.(**F**)Rank genes according to their relative importance.(**G**)The Venn diagram to intersect 3 gene subsets.

under the curve (AUC) values were obtained. The AUC values for IGFBP2 (AUC = 0.869), NR4A2 (AUC = 0.883), TMEM154 (AUC = 0.836), and EGR1 (AUC = 0.882) were all greater than 0.83 (Figure 4H–K), further confirming the strong predictive capabilities of these genes.

Enrichment Analysis

Based on the AUC values from the external validation set, we selected EGR1 as the key gene for further research. Using the R package "limma", we divided patients in the MASLD group into high EGR1 expression and low EGR1 expression groups. Through differential analysis, we identified a total of 278 differentially expressed genes. Figure 5A is a heatmap of the differentially expressed genes, displaying the top 40 genes with expression differences. Figure 5B shows the correlations among these 40 genes, with red indicating positive correlations and blue indicating negative correlations.

Subsequently, we conducted Gene Ontology (GO) (Figure 5C) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Figure 5D) enrichment analyses on these differentially expressed genes. The biological process analysis revealed significant enrichment in regulation of cell-cell adhesion, leukocyte cell-cell adhesion, chemotaxis, taxis, mononuclear cell differentiation and leukocyte migration. The cellular component analysis mainly enriched in external side of plasma membrane, secretory granule membrane, endocytic vesicle, collagen-containing extracellular matrix,
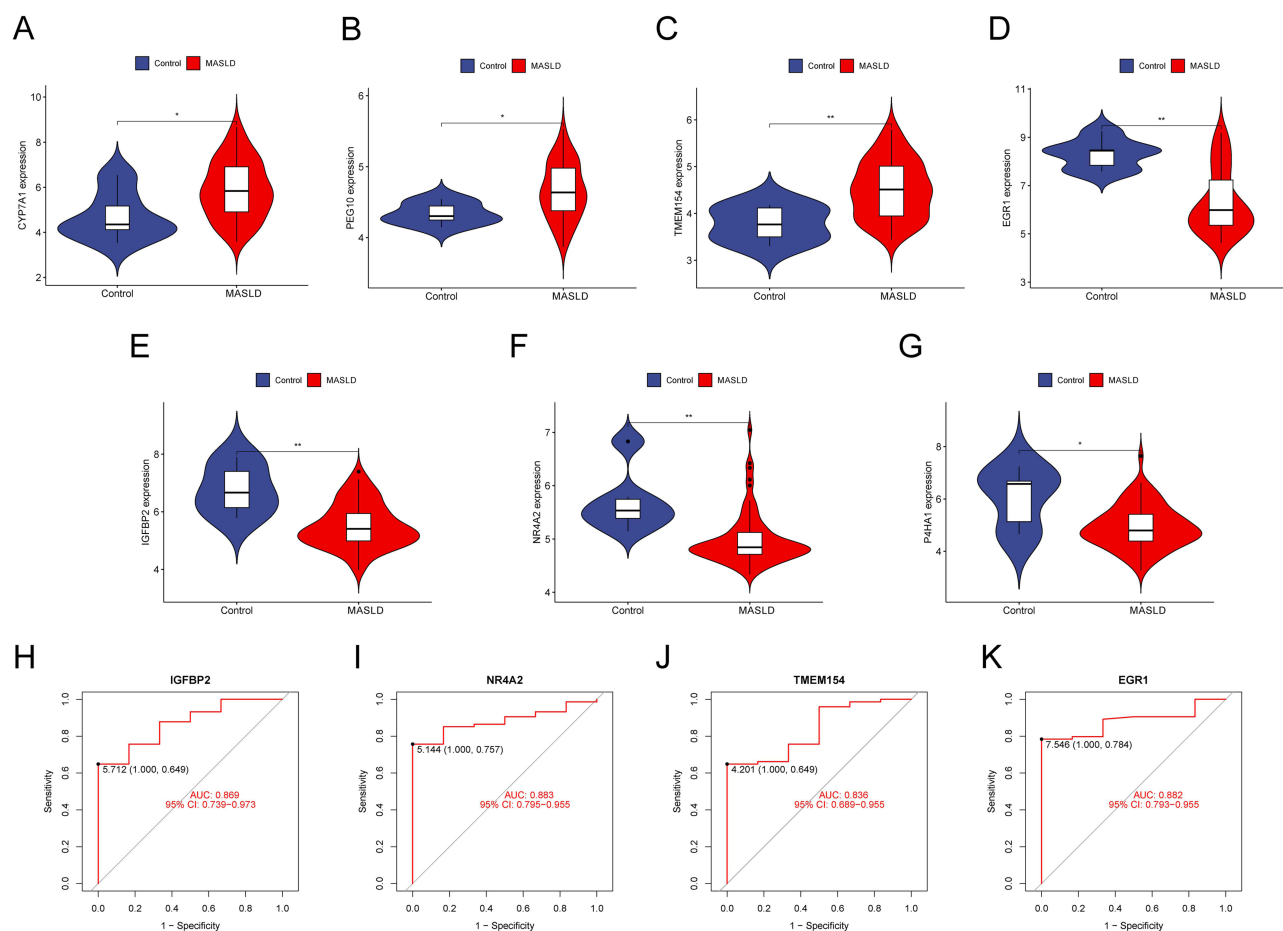
Figure 4 Gene expression in the training set. (**A-G**)Violin plots showed the expression of hub genes in the training set. (**H-K**)ROC analysis of hub genes in the training set. *P < 0.05; **P < 0.01.

endocytic vesicle membrane. The molecular function analysis mainly involved DNA-binding transcription activator activity, RNA polymerase II-specific, DNA-binding transcription activator activity, immune receptor activity, cytokine binding. In the KEGG enrichment analysis, significant enrichment was observed in pathways including osteoclast differentiation, human T-cell leukemia virus 1 infection, cytokine-cytokine receptor interaction, fluid shear stress and atherosclerosis, MAPK signaling pathway.

GSEA and GSVA of Hub Gene

To further investigate the impact of the key gene on the pathogenesis of MASLD, we performed Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA). The GSEA of GO revealed that cell chemotaxis, cellular response to molecule of bacterial origin, granulocyte chemotaxis, leukocyte chemotaxis, positive regulation of miRNA metabolic process are enriched in the high-expression group of EGR1 (Figure 6A); Mitochondrial gene expression, mitochondrial translation, mitochondrial protein containing complex, ribosomal subunit, structural constituent of ribosome were enriched in the low-expression group of EGR1 (Figure 6B). GSEA of KEGG showed that chemokine signaling pathway, cytokine-cytokine receptor interaction, leishmania infection, MAPK signaling pathway, toll-like receptor signaling pathway are enriched in the high expression group of EGR1 (Figure 6C); oxidative phosphorylation, Parkinson's disease, peroxisomes, proteasomes and ribosomes were enriched in the low expression group of EGR1 (Figure 6D).
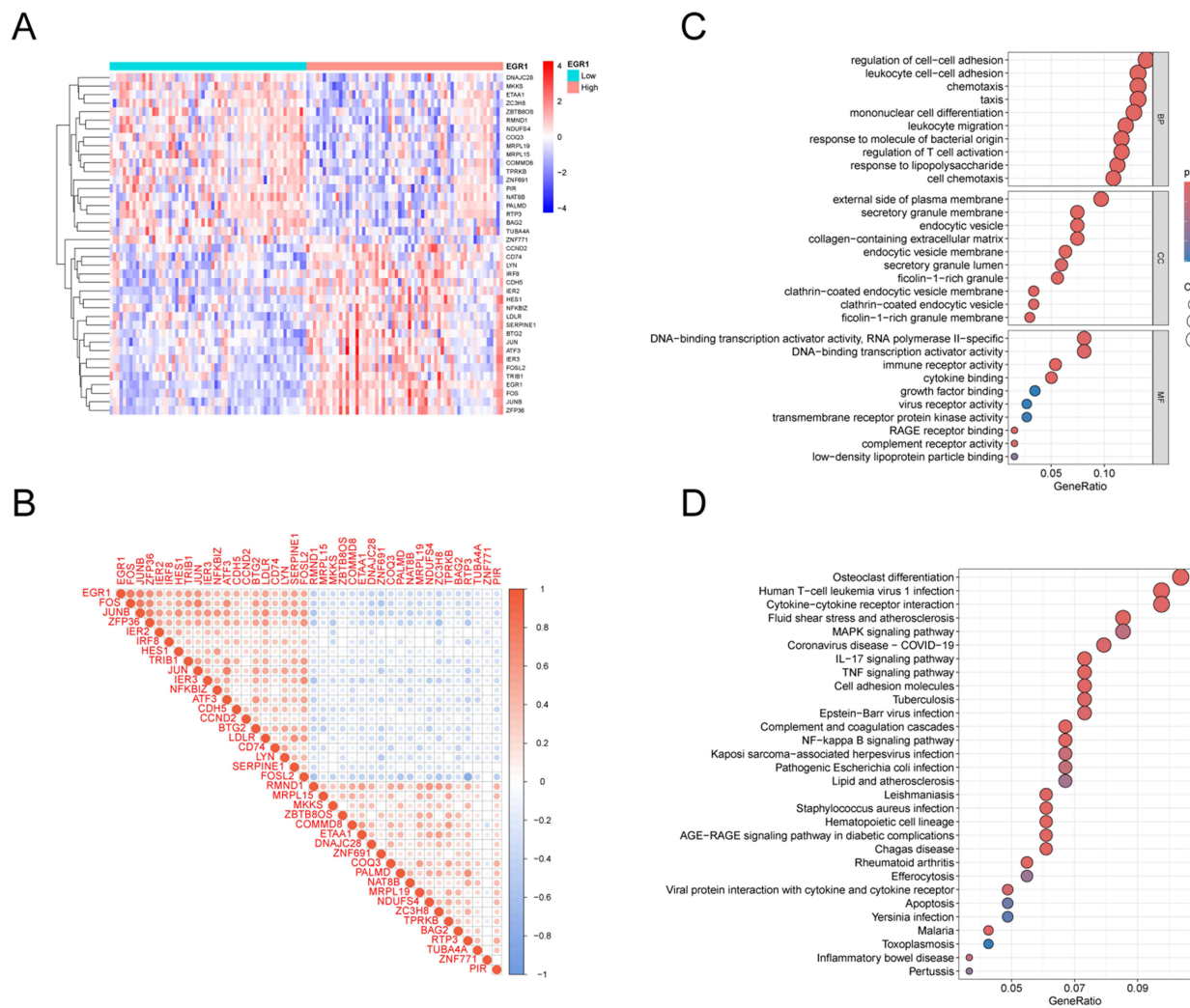
Figure 5 Identification and biological function analysis of core gene high and low expression groups. (**A**)Heat map showing the differentially expressed genes between the two groups. (**B**)Correlation analysis among the top 40 differentially expressed genes. (**C**)GO enrichment analysis showing the functional roles of differentially expressed genes from the perspectives of Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). (**D**)KEGG enrichment analysis reveals the related signaling pathways of differentially expressed genes.

Subsequently, we conducted GSVA to analyze the differences between the high and low expression subgroups of EGR1. In the high expression group of EGR1 (Figure 6E), there is activation of several pathways, including those related to Parkinson's disease, protein export, aminoacyl-tRNA synthesis, ribosome, and oxidative phosphorylation. In the low expression group of EGR1, pathways including prion diseases, MAPK signaling, B-cell receptor signaling, T-cell receptor signaling, and leishmania infection are enriched.

Immune Analysis

Previous research indicates that the immune microenvironment plays a crucial role in the pathogenesis of MASLD. To further investigate the effects of EGR1 on alterations in the immune microenvironment, we analyzed differences in 16 immune cell types and 13 immune functions between groups with high and low expression of hub genes (Figure 7A).

Subsequent analysis explored the associations between the key gene and immune cells. As depicted in Figure 7B, EGR1 exhibits significant positive correlations with neutrophils (r=0.3) and mast cells activated(r=0.25), and notable negative correlations with dendritic cells resting (r=−0.21) and mast cells resting (r=−0.31). These results highlight the integral role of EGR1 in modulating diverse immune cell interactions, emphasizing its potential as an immune-related therapeutic target for MASLD.
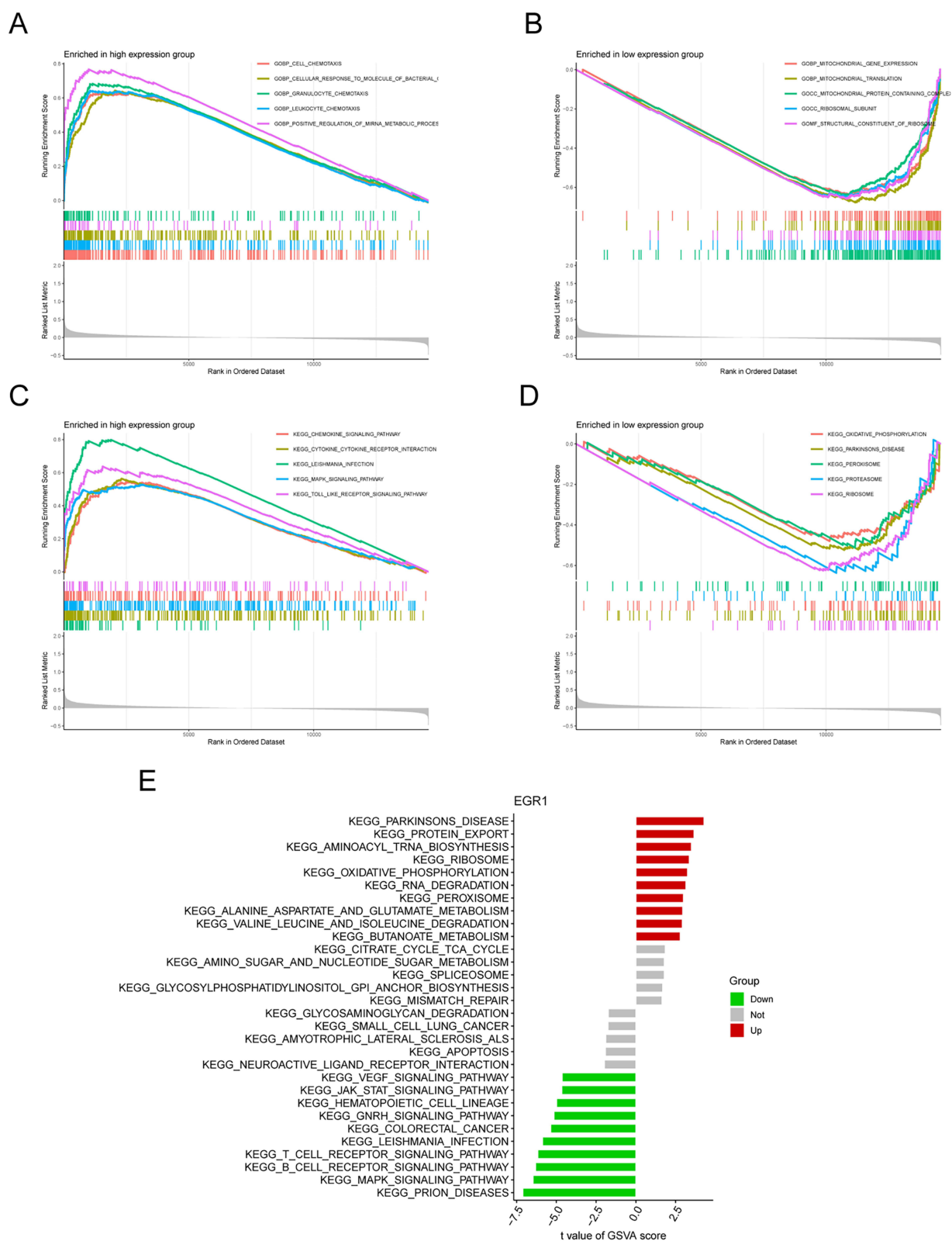
Figure 6 GSEA and GSVA of key gene. (**A-B**)GSEA analysis of core genes using the GO gene sets. (**C-D**)GSEA analysis of key genes using the KEGG gene sets. (**E**) GSVA enrichment method was used to perform enrichment analysis of KEGG pathways for key genes.
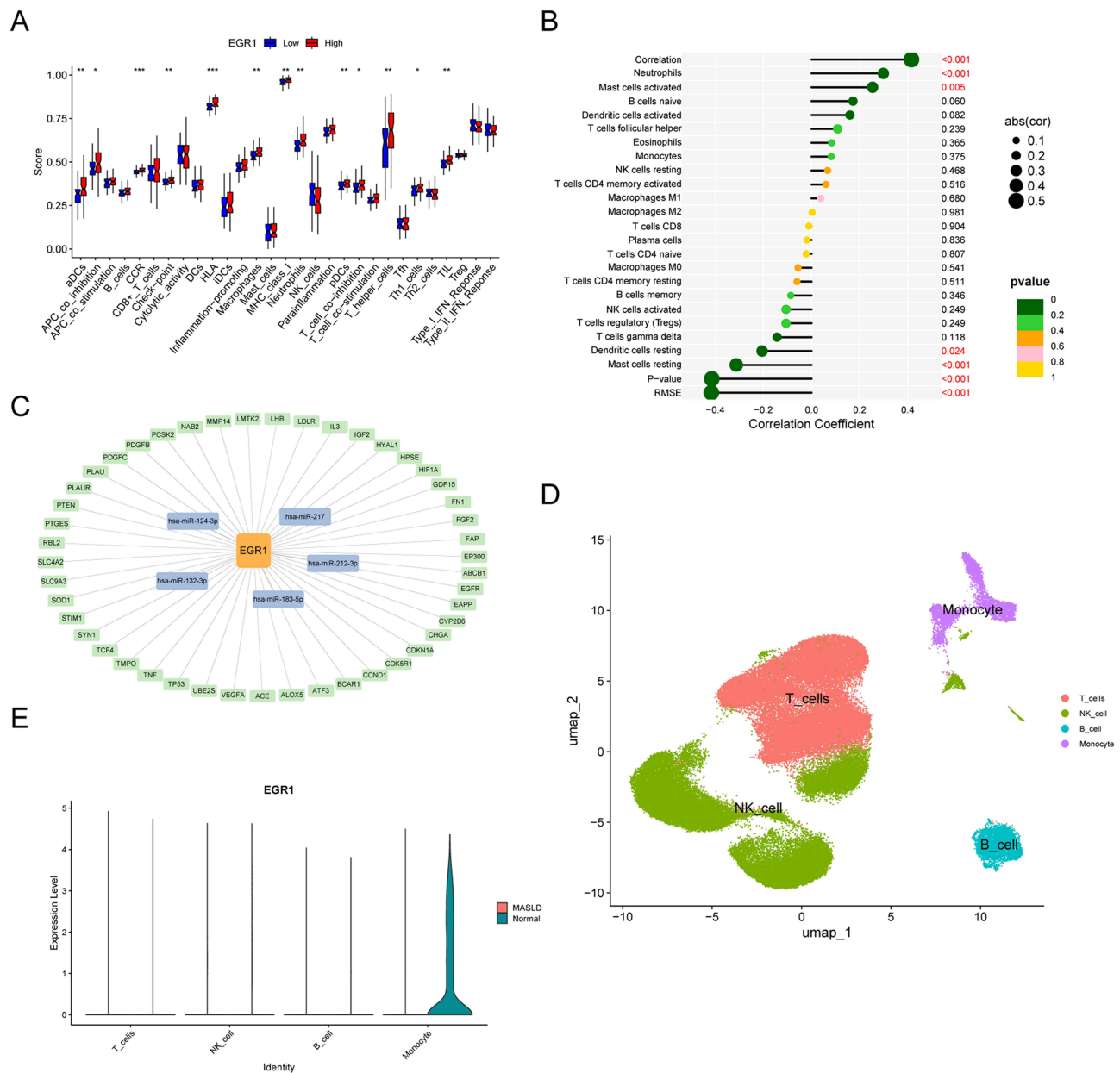
Figure 7 Immune analysis of key genes, TF-miRNA-mRNA regulatory network and single-cell analysis. (**A**)immune-related functions scores. (**B**)The relationship between key genes and 22 types of immune cells. (**C**) TF-miRNA-mRNA regulatory network. (**D**)Analysis of Single-Cell RNA Sequencing Dataset GSE159977. (**E**)Key gene enrichment in immune cells.*P < 0.05; **P < 0.01, ***P < 0.001.

TF-miRNA-mRNA Regulatory Network and Single-Cell Sequencing Data

Potential transcription factors (TFs) were retrieved from the Trrust database, and candidate miRNAs were identified using the Starbase database. We subsequently constructed a visual representation of the interaction network (Figure 7C). The TF-miRNA-mRNA regulatory network models the intricate interplay of gene expression regulation within cells. Researching these networks is essential for elucidating cellular mechanisms of gene regulation, which profoundly impact our understanding of disease processes and the development of innovative therapeutic approaches. In the diagram, green boxes represent transcription factors (TFs), and blue boxes represent microRNAs (miRNAs), which are short non-coding RNA molecules.

Through the analysis of the single-cell dataset GSE159977, we identified the distribution of key gene across four distinct cell clusters (Figure 7D). Specifically, for the key gene EGR1, we observed significant differences in monocytes between the control and MASLD groups (Figure 7E).

Validated by in Vivo and in Vitro Experiments

We established a steatotic liver cell model in mice. Using Oil Red O staining, we observed that lipid droplets were significantly accumulated in the PA group compared to the normal control group (Figure 8A). PCR analysis demonstrated a significant reduction in EGR1 expression levels in the PA group (Figure 8B). At the protein level, EGR1 was also markedly decreased (Figure 8C and D). We further validated the expression levels of key gene in an animal model. Compared to the control group, liver tissue sections from MCD diet-fed mice exhibited prominent inflammation and ballooning degeneration in H&E and Oil Red O staining, with substantial lipid droplet accumulation (Figure 8E). In addition, serum levels of total cholesterol (TC) and triglycerides (TG) in these mice showed significant alterations (Figure 8F and G), confirming the successful establishment of the model.

Subsequently, we measured the expression levels of pro-inflammatory cytokines in liver tissue. Significant changes were observed in CCL2, IL-1β, and TNF-α levels, while IL-6 levels remained unchanged (Figure 8H). The relative mRNA levels of EGR1 in mouse liver tissue showed a marked reduction (Figure 8I), and corresponding protein levels displayed significant differences (Figure 8J and K), consistent with previous findings.

Discussion

Approximately 32.4% of the global population is affected by Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD), and its prevalence has been rising significantly over time, leading to a substantial public health burden.[30] Despite this, effective treatments remain scarce. Therefore, identifying key genes involved in the pathogenesis of MASLD is essential for developing effective therapeutic strategies. In this study, we integrated microarray data analysis with machine learning techniques to identify differentially expressed genes. Validation with external datasets identified EGR1 as a key gene, showing an AUC value significantly higher than that of other genes. Analysis of gene expression patterns in both the training and validation sets showed that EGR1 exhibited marked differences between MASLD patients and healthy controls, highlighting its pivotal role in the onset and progression of MASLD. Additionally, immune infiltration analysis revealed distinct differences in immune cell profiles between MASLD patients and healthy controls.

To improve the accuracy and reliability of identifying biomarkers associated with MASLD development, we employed three machine learning algorithms to further refine 42 differentially expressed genes. This process led to the identification of ten hub genes: CYP7A1, PEG10, P4HA1, IGFBP2, IL6, ME1, NR4A2, VIL1, TMEM154, and EGR1. Notably, CYP7A1, PEG10, and TMEM154 were upregulated in MASLD patients in the training set, while EGR1, IGFBP2, NR4A2, and P4HA1 were downregulated, a pattern corroborated by the validation dataset GSE164760. ROC regression analysis highlighted EGR1 (AUC=0.882) as having strong predictive power. Based on the key gene, we categorized MASLD patients into high-expression and low-expression groups for further analysis, thereby offering new perspectives on MASLD research.

Among the differentially expressed genes in high and low expression groups of core genes, Gene Ontology (GO) enrichment analysis of biological processes revealed significant enrichment in chemotaxis, taxis, mononuclear cell differentiation and leukocyte migration. Chemotaxis and taxis refer to the directional migration of cells in response to external stimuli, a process primarily regulated by chemokines and their receptors. These receptors are classified into four subfamilies: CXC, CC, C, and CX3C. In MASLD, these chemokines are involved not only in the progression from simple steatosis to metabolic dysfunction-associated steatohepatitis (MASH) but also in modulating the ratio of Kupffer cell differentiation into M1 and M2 macrophages in the liver.[31,32] Studies have shown that upon liver inflammation, monocytes are recruited to the site of injury, where they secrete pro-inflammatory cytokines.[33] These monocytes then differentiate into various macrophage subsets depending on the local microenvironmental signals,[34] with an initial predominance of pro-inflammatory M1 macrophages. This phenotype exacerbates the progression of hepatic steatosis to fibrosis. Moreover, therapeutic inhibition of monocyte recruitment has been shown to reduce liver fibrosis in animal models.[35] Additionally, in mouse models of MASLD, there is a significant increase in the proportion of monocyte-
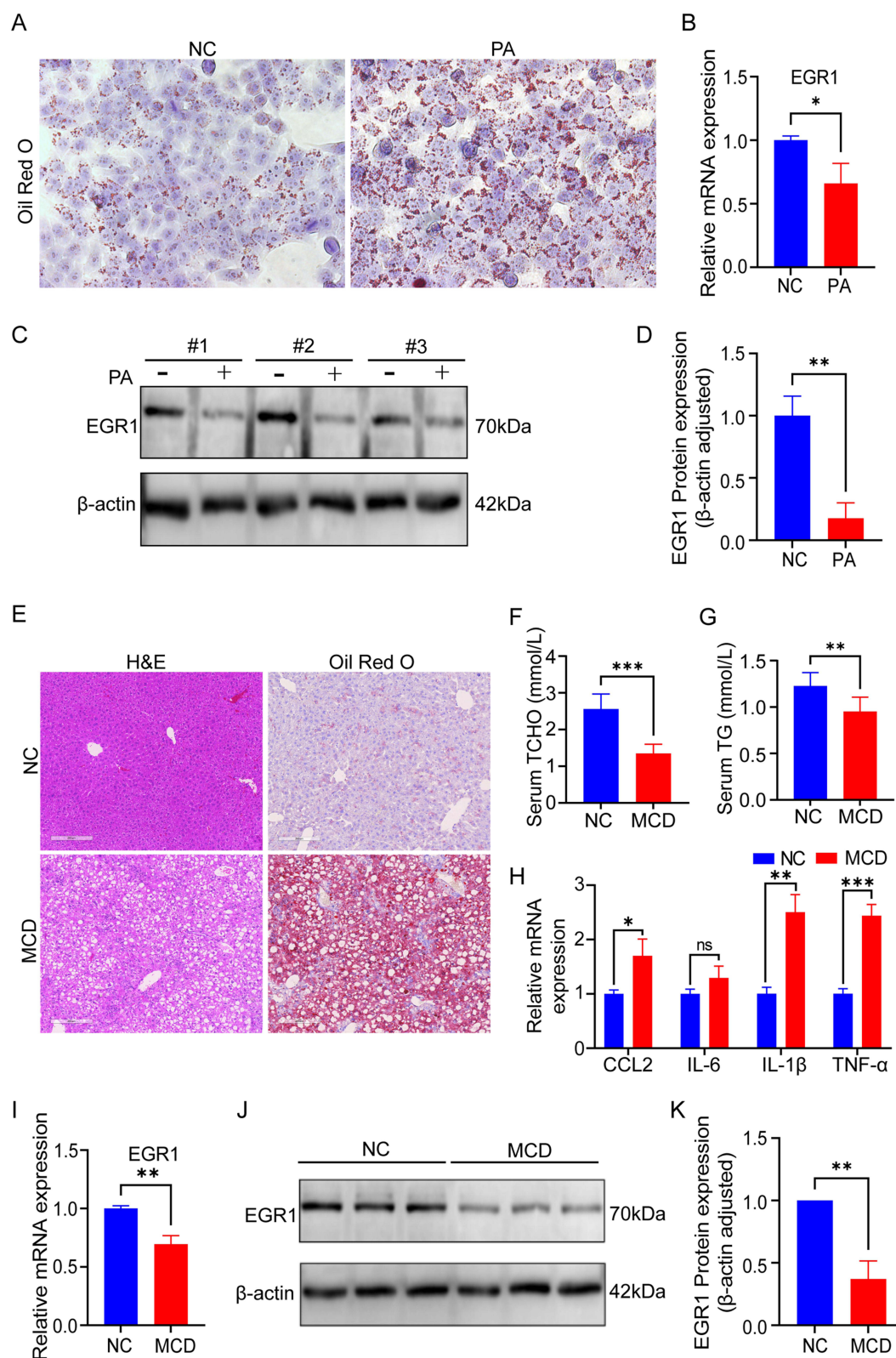
Figure 8 Experimental validation of gene expression. (**A**)Oil Red O staining was used to visualize lipid droplets within the cell model. (**B**)Relative mRNA levels of key gene in cell model. (**C**)Protein levels of EGR1 in AML12 cells. (**D**)Quantification of proteins in each group using Western blotting. (**E**)H&E staining and Oil Red O staining of liver tissues from control and MCD group mice (magnification ×100, scale bar = 200 μm. In the MCD group, hepatocytes are significantly swollen and spherical in shape, with enlarged nuclei that are compressed to one side. Fat vacuoles are visible, indicating fat infiltration. Inflammatory cell infiltration is markedly increased. ORO staining reveals numerous red lipid droplets.). Relative serum TC(**F**) and TG(**G**) levels in control and MCD group mice. (**H**)Relative mRNA levels of pro-inflammatory cytokines in animal model. (**I**)Relative mRNA levels of key gene in animal model. (**J**)Protein levels of EGR1 in animal model. (**K**)Quantification of proteins in each group using Western blotting.*P < 0.05; **P < 0.01, ***P < 0.001.

derived macrophages relative to the total number of liver macrophages, further highlighting the contribution of these cells to disease progression.[36] Leukocyte migration, an essential component of the immune response, ensures that leukocytes are recruited to sites of infection to resolve inflammation and repair tissue damage.[37] Taken together, these processes play a crucial role in the development and progression of MASLD, suggesting that the key gene EGR1 may exert its effects by modulating these pathways, potentially inhibiting the occurrence and progression of MASLD.

We also investigated the differences in immune profiles between high and low core gene expression groups. Increasing evidence suggests that various immune cells play a significant role in the development and progression of MASLD.[38] Immune responses have both advantages and disadvantages, they can provide protection but may also exacerbate tissue damage. Moreover, the same cell type may have different effects at different stages and locations. We observed differences in neutrophils, Th1 helper T cells, and macrophages. Neutrophils are among the first cells to arrive at the site of damage after inflammation begins. These cells produce myeloperoxidase (MPO), a lysosomal enzyme that interacts with hydrogen peroxide (H_2O_2) to generate reactive oxygen species (ROS), contributing to the progression of MASH through oxidative stress and tissue damage.[39] Th1 helper T cells, which are polarized from CD4[+] T helper cells,[40] are pro-inflammatory cells that mediate cell-mediated immunity and secrete cytokines such as interferon-gamma (IFN-γ) and interleukin-2 (IL-2). IFN-γ, in particular, drives the polarization of macrophages into the M1 phenotype and stimulates macrophages to secrete pro-inflammatory factors.[41] Macrophage infiltration, particularly in the portal vein, is considered an early event in MASLD.[42] Due to the varying liver microenvironment, macrophages respond to different metabolites and signals, exhibiting dual functions and classifying into classical activated M1 and alternatively activated M2 types.[43] M1 macrophages primarily produce pro-inflammatory factors such as IL-1, IL-6, and IL-23, which promote inflammation, liver damage, and fibrosis.[44] In contrast, M2 macrophages have opposing effects. The changing ratio of these macrophage types during MASLD progression may be crucial to disease advancement. Therefore, modulating target genes to influence specific immune cells could potentially mitigate the progression of MASLD.

Recent studies have shown that EGR1 plays a crucial regulatory role in various metabolic diseases. For instance, increased hepatic expression of EGR1 has been shown to improve diet-induced fatty liver.[45] Research has found that inhibition of EGR1 expression significantly raises liver lipid levels and promotes the progression of fatty liver to mild fibrosis.[3] Furthermore, in adipocytes exposed to high insulin for prolonged periods, EGR1 and its target gene GGPPS may contribute to the development of insulin resistance by continuously activating the Erk/MAPK signaling pathway.[46] EGR1 is also involved in regulating inflammatory responses. Studies have demonstrated that overexpression of EGR1 inhibits the activation of inflammatory enhancers and significantly reduces the levels of TNF-α and IL-6 released by macrophages.[19] Additionally, EGR1 plays a protective role in liver injury. In a CCL4-treated mouse model, EGR1 knockout mice showed more severe liver necrosis and higher plasma levels of AST and ALT. Further research revealed that EGR1 protects against CCL4-induced liver injury by regulating TNF-α levels.[47] Furthermore, EGR1 has been reported to act as a tumor suppressor in liver cancer. EGR1 induces apoptosis in human liver cancer cell lines HepG2 and Hep3B and enhances cell apoptosis by synthesizing ursodeoxycholic acid derivatives.[48] These findings suggest that EGR1 plays an important regulatory role in various metabolic and liver diseases and may serve as a potential therapeutic target.

In addition, we conducted GSEA and GSVA to perform a comprehensive bioinformatics evaluation of the involvement of key gene in various disease-related signaling pathways, including the MAPK signaling pathway, B cell receptor signaling pathway, and T cell receptor signaling pathway, and Leishmaniasis infection. Leishmaniasis infection, caused by a parasitic organism, can potentially lead to excessive activation of the immune system, resulting in pathological changes such as liver inflammation and fibrosis. Studies have shown a positive correlation between Leishmaniasis infection burden and an increased apoptosis index in hepatocytes, Kupffer cells, and inflammatory infiltrates.[49] In addition, our analysis identified several key biological processes, providing new insights into the pathophysiological mechanisms of MASLD. We also constructed a TF-miRNA-mRNA network, highlighting TFs and miRNAs associated with the key gene. These TFs and miRNAs play crucial roles in the transcription and translation processes of the key gene, suggesting potential regulatory mechanisms. Furthermore, single-cell sequencing data revealed differential expression of target genes within specific cell subpopulations, indicating that these cells may play a significant role in disease progression. Finally, validation using cellular and animal liver models confirmed our findings, with differential expression of the EGR1 gene consistent with our analysis.

However, there are several limitations to our study. First, the dataset used in this research is sourced from the publicly available GEO database, which has a limited sample size. Additionally, the heterogeneity between datasets may affect the generalizability of the model. Second, our research relied solely on a mouse model, with a limited number of samples, and the findings from the bioinformatics analysis could not be validated using clinical samples. Furthermore, additional in vivo and in vitro studies are necessary to comprehensively investigate the gene's role in disease processes.

Conclusion

This study integrated five datasets and performed differential expression analysis. We applied three distinct machine learning algorithms to identify key genes and validated these genes using external datasets. Additionally, we conducted immune infiltration analysis using the CIBERSORT algorithm to explore the relationship between key genes and immune cells. The research also included enrichment analysis, TF-miRNA-mRNA network construction, and single-cell analysis, with validation conducted in both cellular and animal models. This work provides novel biomarkers and targets for the understanding and treatment of MASLD.

Ethics Statement

The animal experiments conducted in this study were approved by the Animal Experimentation Ethics Committee of Harbin Medical University(2023-DWSYLLCZ-04), and performed following Regulations on the Management of Laboratory Animals in Heilongjiang Province and Laboratory Animals General requirements for biosafety in animal experiments. GEO is a public database, and patients contributing to the database have obtained ethical approval. Users can freely download the relevant data for research purposes and publish related articles. Our study is based on open-source data, which has been fully deidentified and does not contain any information that could identify individual subjects. Therefore, the data is considered anonymous and does not involve direct intervention with individuals. As a result, the Ethics Committee of the Fourth Affiliated Hospital of Harbin Medical University has determined that no further review is required.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. Chen L, Tao X, Zeng M, Mi Y, Xu L. Clinical and histological features under different nomenclatures of fa tty liver disease: NAFLD, MAFLD, MASLD and MetALD. *J Hepatol*. 2024;80(2):e64–e66. doi:10.1016/j.jhep.2023.08.021
2. Rinella ME, Lazarus JV, Ratziu V, et al. A multisociety Delphi consensus statement on new fatty liver disease n omenclature. *J Hepatol*. 2023;79 (6):1542–1556. doi:10.1016/j.jhep.2023.06.003
3. European Association for the Study of the L. European Association for the Study of D, European Association for the Study of O. EASL-EASD-EASO clinical practice guidelines on the management of metab olic dysfunction-associated steatotic liver disease (MASLD). *J Hepatol*. 2024;81 (3):492–542. doi:10.1016/j.jhep.2024.04.031
4. Riazi K, Azhari H, Charette JH, et al. The prevalence and incidence of NAFLD worldwide: a systematic review a nd meta-analysis. *Lancet Gastroenterol Hepatol*. 2022;7(9):851–861. doi:10.1016/S2468-1253(22)00165-0
5. Younossi ZM, Blissett D, Blissett R, et al. The economic and clinical burden of nonalcoholic fatty liver disease i n the United States and Europe. *Hepatology*. 2016;64(5):1577–1586. doi:10.1002/hep.28785
6. Buzzetti E, Pinzani M, Tsochatzis EA. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NA FLD). *Metabolism*. 2016;65 (8):1038–1048. doi:10.1016/j.metabol.2015.12.012

7. Tilg H, Moschen AR. Evolution of inflammation in nonalcoholic fatty liver disease: the mul tiple parallel hits hypothesis. *Hepatology*. 2010;52 (5):1836–1846. doi:10.1002/hep.24001

8. Bellanti F, Lo Buglio A, Vendemiale G. Hepatic mitochondria-gut microbiota interactions in metabolism-associa ted fatty liver disease. *Metabolites*. 2023;13(3):322. doi:10.3390/metabo13030322

9. Moretti V, Romeo S, Valenti L. The contribution of genetics and epigenetics to MAFLD susceptibility. *Hepatol Int*. 2024;18(1):1–3. doi:10.1007/s12072-12024-10667-12075

10. Chun HS, Lee M, Lee JS, et al. Metabolic dysfunction associated fatty liver disease identifies subjec ts with cardiovascular risk better than non-alcoholic fatty liver dise ase. *Liver Int*. 2023;43(3):608–625. doi:10.1111/liv.15508

11. Yoon EL, Jun DW. Changing the nomenclature from nonalcoholic fatty liver disease to met abolic dysfunction-associated fatty liver disease is more than a chang e in terminology. *Clin Mol Hepatol*. 2023;29(2):371–373. doi:10.3350/cmh.2023.0086

12. Yang K, Song M. New Insights into the Pathogenesis of Metabolic-Associated Fatty Liver Disease (MAFLD): gut-liver-heart crosstalk. *Nutrients*. 2023;15(18):3970. doi:10.3390/nu15183970

13. Romero-Gómez M, Zelber-Sagi S, Trenell M. Treatment of NAFLD with diet, physical activity and exercise. *J Hepatol*. 2017;67(4):829–846. doi:10.1016/j.jhep.2017.05.016

14. Thiel G, Rössler OG. Glucose homeostasis and pancreatic islet size are regulated by the tra nscription factors elk-1 and Egr-1 and the protein phosphatase calcine urin. *Int J mol Sci*. 2023;24(1):815. doi:10.3390/ijms24010815

15. Liu H-T, Liu S, Liu L, Ma -R-R, Gao P. EGR1-Mediated Transcription of lncRNA-HNF1A-AS1 promotes cell-cycle pr ogression in gastric cancer. *Cancer Res*. 2018;78(20):5877–5890. doi:10.1158/0008-5472.CAN-18-1011

16. Ding Y, Chen X, Liu C, et al. Identification of a small molecule as inducer of ferroptosis and apopt osis through ubiquitination of GPX4 in triple negative breast cancer c ells. *J Hematol Oncol*. 2021;14(1):19. doi:10.1186/s13045-020-01016-8

17. Wu J, Bu D, Wang H, et al. The rhythmic coupling of Egr-1 and Cidea regulates age-related metabol ic dysfunction in the liver of male mice. *Nat Commun*. 2023;14(1):1634. doi:10.1038/s41467-023-36775-8

18. Mohtar O, Ozdemir C, Roy D, Shantaram D, Emili A, Kandror KV. Egr1 mediates the effect of insulin on leptin transcription in adipocy tes. *J Biol Chem*. 2019;294(15):5784–5789. doi:10.1074/jbc.AC119.007855

19. Trizzino M, Zucco A, Deliard S, et al. EGR1 is a gatekeeper of inflammatory enhancers in human macrophages. *Sci Adv*. 2021;7(3):eaaz8836. doi:10.1126/sciadv.aaz8836

20. Wei M, Gu X, Li H, et al. EGR1 is crucial for the chlorogenic acid-provided promotion on liver r egeneration and repair after APAP-induced liver injury. *Cell Biol Toxicol*. 2023;39(6):2685–2707. doi:10.1007/s10565-023-09795-9

21. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variatio n in high-throughput experiments. *Bioinformatics*. 2012;28(6):882–883. doi:10.1093/bioinformatics/bts034

22. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and m icroarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007

23. Alderden J, Pepper GA, Wilson A, et al. Predicting pressure injury in critical care patients: a machine-learni ng model. *Am J Crit Care*. 2018;27 (6):461–468. doi:10.4037/ajcc2018525

24. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate desc ent. *J Stat Softw*. 2010;33(1):1–22. doi:10.18637/jss.v033.i01

25. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC c urves. *BMC Bioinf*. 2011;12 (1):77. doi:10.1186/1471-2105-12-77

26. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among ge ne clusters. *OMICS*. 2012;16 (5):284–287. doi:10.1089/omi.2011.0118

27. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;14:7. doi:10.1186/1471-2105-14-7

28. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell Data. *Cell*. 2019;177(7):1888–1902.e1821. doi:10.1016/j.cell.2019.05.031

29. Huby T, Gautier EL. Immune cell-mediated features of non-alcoholic steatohepatitis. *Nat Rev Immunol*. 2022;22(7):429–443. doi:10.1038/s41577-021-00639-3

30. Eslam M, El-Serag HB, Francque S, et al. Metabolic (dysfunction)-associated fatty liver disease in individuals of normal weight. *Nat Rev Gastroenterol Hepatol*. 2022;19(10):638–651. doi:10.1038/s41575-022-00635-5

31. Nagata N, Chen G, Xu L, Ando H. An Update on the Chemokine System in the Development of NAFLD. *Medicina (Kaunas)*. 2022;58(6):761. doi:10.3390/medicina58060761

32. Ullah A, Ud Din A, Ding W, Shi Z, Pervaz S, Shen B. A narrative review: CXC chemokines influence immune surveillance in ob esity and obesity-related diseases: type 2 diabetes and nonalcoholic f atty liver disease. *Rev Endocr Metab Disord*. 2023;24(4):611–631. doi:10.1007/s11154-023-09800-w

33. Triantafyllou E, Woollard KJ, McPhail MJW, Antoniades CG, Possamai LA. The role of monocytes and macrophages in acute and acute-on-chronic li ver failure. *Front Immunol*. 2018;9:2948. doi:10.3389/fimmu.2018.02948

34. Miura K, Yang L, van Rooijen N, Ohnishi H, Seki E. Hepatic recruitment of macrophages promotes nonalcoholic steatohepatit is through CCR2. *Am J Physiol Gastrointest Liver Physiol*. 2012;302(11):G1310–1321. doi:10.1152/ajpgi.00365.2011

35. Krenkel O, Puengel T, Govaere O, et al. Therapeutic inhibition of inflammatory monocyte recruitment reduces st eatohepatitis and liver fibrosis. *Hepatology*. 2018;67(4):1270–1283. doi:10.1002/hep.29544

36. Tran S, Baba I, Poupel L, et al. Impaired Kupffer cell self-renewal alters the liver response to lipid overload during non-alcoholic steatohepatitis. *Immunity*. 2020;53(3):627–640.e625. doi:10.1016/j.immuni.2020.06.003

37. Nourshargh S, Alon R. Leukocyte migration into inflamed tissues. *Immunity*. 2014;41(5):694–707. doi:10.1016/j.immuni.2014.10.008

38. Nati M, Haddad D, Birkenfeld AL, Koch CA, Chavakis T, Chatzigeorgiou A. The role of immune cells in metabolism-related liver inflammation and development of non-alcoholic steatohepatitis (NASH). *Rev Endocr Metab Disord*. 2016;17(1):29–39. doi:10.1007/s11154-016-9339-2

39. Klebanoff SJ. Myeloperoxidase: friend and foe. *J Leukoc Biol*. 2005;77(5):598–625 doi:10.1189/jlb.1204697.

 1655

40. Schmitt N, Ueno H. Regulation of human helper T cell subset differentiation by cytokines. *Curr Opin Immunol*. 2015;34:130–136 doi:10.1016/j.coi.2015.03.007.

41. Dale DC, Boxer L, Liles WC. The phagocytes: neutrophils and monocytes. *Blood*. 2008;112(4):935–945 doi:10.1182/blood-2007-12-077917.

42. Gadd VL, Skoien R, Powell EE, et al. The portal inflammatory infiltrate and ductular reaction in human nona lcoholic fatty liver disease. *Hepatology*. 2014;59(4):1393–1405 doi:10.1002/hep.26937.

43. Murray PJ, Wynn TA. Protective and pathogenic functions of macrophage subsets. *Nat Rev Immunol*. 2011;11(11):723–737 doi:10.1038/nri3073.

44. Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol*. 2008;8(12):958–969 doi:10.1038/nri2448.

45. Song H, Zheng Z, Wu J, Lai J, Chu Q, Zheng X. White pitaya (Hylocereus undatus) juice attenuates insulin resistance and hepatic steatosis in diet-induced obese mice. *PLoS One*. 2016;11(2):e0149670 doi:10.1371/journal.pone.0149670.

46. Shen N, Yu X, Pan F-Y, Gao X, Xue B, Li C-J. An early response transcription factor, Egr-1, enhances insulin resist ance in type 2 diabetes with chronic hyperinsulinism. *J Biol Chem*. 2011;286(16):14508–14515 doi:10.1074/jbc.M110.190165.

47. Pritchard MT, Cohen JI, Roychowdhury S, Pratt BT, Nagy LE. Early growth response-1 attenuates liver injury and promotes hepatopro tection after carbon tetrachloride exposure in mice. *J Hepatol*. 2010;53(4):655–662 doi:10.1016/j.jhep.2010.04.017.

48. Park SE, Lee SW, Hossain MA, et al. A chenodeoxycholic derivative, HS-1200, induces apoptosis and cell cyc le modulation via Egr-1 gene expression control on human hepatoma cell s. *Cancer Lett*. 2008;270(1):77–86 doi:10.1016/j.canlet.2008.04.038.

49. Verçosa BLA, Muniz-Junqueira MI, Mineiro ALBB, Costa FAL, Melo MN, Vasconcelos AC. Enhanced apoptotic index in hepatocytes, Kupffer cells, and inflammato ry infiltrate showed positive correlation with hepatic lesion intensit y, parasite load, and clinical status in naturally Leishmania-infected dogs. *Microb Pathog*. 2023;181:106194 doi:10.1016/j.micpath.2023.106194.