The role of endoplasmic reticulum stress-related genes in the diagnosis and subtyping of non-alcoholic fatty liver disease

Zihao Guo[1], Xiaoxiao Yu[1], Zhihao Fang[1], Kai Yang[1], Changxu Liu[1], Zhichao Dong[1] and Chang Liu[1]

[1] *Department of General Surgery, The Fourth Hospital of Harbin Medical University, Harbin, Heilongjiang, China*

Abstract. Non-alcoholic fatty liver disease (NAFLD) is the most prevalent liver disease worldwide. Chronic activation of endoplasmic reticulum stress (ERS) in hepatocytes may promote the development of NAFLD, yet endoplasmic reticulum stress-related genes (ERSGs) have not been studied in NAFLD. Our aim is to study the relationship between ERSGs and the immune microenvironment of NAFLD patients and to construct predictive models. We screened 48 endoplasmic reticulum stress-related differentially expressed genes (ERSR-DEGs) using data from two GEO datasets and the GeneCards database. Enrichment analysis revealed that ERSR-DEGs are closely associated with immune-related pathways and functions. The immune infiltration profile of NAFLD was obtained by single sample gene set enrichment analysis (ssGSEA). There were significant differences in immune cell infiltration and immune function between NAFLD group and control group. Using 113 NAFLD samples, we explored two molecular clusters based on ERSR-DEGs. A five-gene SVM model was selected as the best machine learning model, and a nomogram based on five-gene SVM model showed good predictive efficiency. The mRNA expression levels of POR, PPP1R15A, FOS and FAS were significantly different between NAFLD mice and healthy mice. In conclusion, ERS is closely associated with the development of NAFLD. We established a promising and SVM-based predictive model to assess the risk of disease in patients with ERS subtypes and NAFLD.

Key words: Non-alcoholic fatty liver disease — Endoplasmic reticulum stress — Immune infiltration — Machine learning — Prediction model

Abbreviations: DEGs, differentially expressed genes; ERSGs, endoplasmic reticulum stress-related genes; GEO, Gene Expression Omnibus; GO, Gene Ontology; GSVA, gene set variance analysis; HFD, high-fat diet; KEGG, Kyoto Encyclopedia of Genes and Genomes; Kyoto Encyclopedia of Genes and Genomes; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; PCA, principal component analysis; PERK, PKR-like ER kinase; RF, random forest models; ROC, receiver operating characteristic; ssGSEA, single sample gene set enrichment analysis; SVM, support vector machine models; TC, total cholesterol; TG, total triglycerides; UPR, unfolded protein response; XGB, extreme gradient boosting.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is the development of hepatic steatosis without heavy alcohol consumption or causes other than metabolic disorders (drugs or genetic disorders) that constitute a metabolic syndrome. The prevalence of NAFLD has been increasing with the expansion of the world economy and the improvement of living standards, with a global prevalence of approximately 25% (Younossi et al. 2016). Given the increasing understanding of the pathogenesis of NAFLD and its high prevalence, experts have proposed a clinical diagnosis – metabolic dysfunction associated with fatty liver disease (Eslam et al. 2020). The liver pathology of NAFLD ranges from simple hepatic steatosis to steatohepatitis, which increases the risk of developing cirrhosis and cancer (Powell et al. 2021). Despite rapid progress in clinical research into NAFLD treatments as new targets and diagnostics emerge, our understanding of NAFLD remains poor (Friedman et al. 2018). Hence, further illumination of the molecular pathogenesis of NAFLD is necessary for better treatment of NAFLD.

NAFLD is characterized by the synthesis and accumulation of hepatic triglycerides and the lipotoxicity of fatty acids, free cholesterol, and other lipid metabolites, leading to mitochondrial dysfunction with oxidative stress, endoplasmic reticulum stress (ERS), activation of the unfolded protein response, and ultimately chronic liver disease (Yazici and Sezer 2017). Hepatocytes are responsible for lipogenesis, cholesterol biosynthesis, glucose, and xenobiotic metabolism. Hepatocytes contain an abundance of rough endoplasmic reticulum and smooth endoplasmic reticulum to meet the metabolic demands of the individual (Wang and Kaufman 2016). In eukaryotic cells, the endoplasmic reticulum is a key organelle for protein folding and quality control, as well as for regulating intracellular calcium homeostasis and steroid molecule synthesis. The endoplasmic reticulum's homeostasis is crucial to the physiological functions of normal cells (Pagliassotti 2012). When misfolded protein accumulation in the endoplasmic reticulum exceeds a critical threshold, endoplasmic reticulum homeostasis is disrupted, and cells initially trigger an adaptive signaling pathway called the unfolded protein response (UPR). UPR reduces secreted protein load, enhances protein folding (transcription of molecular chaperones and folding enzymes), and attempts to correct this by promoting autophagy and endoplasmic reticulum associated degradation, which increases clearance. UPR is initiated by three endoplasmic reticulum transmembrane proteins: PKR-like ER kinase (PERK), inositol-requiring enzyme 1 (IRE-1), and activating transcription factor-6 (ATF-6). Under normal conditions, IRE1, PERK, and ATF6 remain inactive after binding to GRP78. Once ERS occurs, GRP78 dissociates from the ERS sensor and activates three branches (Doultsinos et al. 2017). IRE1α-XBP1 pathway regulates hepatic lipid metabolism through low-density lipoprotein secretion and lipogenesis (Wang et al. 2012). Activation of the PERK pathway inhibits the translation of IkB, increases NF-κB activity, and promotes the secretion of TNFα, IL-1β, and IL-6 (Meares et al. 2014). Liver biopsies from non-alcoholic steatohepatitis (NASH) patients revealed an increase in ERS marker GRP78 and the pro-apoptotic protein CHOP compared to patients with hepatic steatosis alone, which supports the idea that ERS-induced apoptosis contributes to the transition from steatosis to NASH and beyond (Gonzalez-Rodriguez et al. 2014). This finding suggests the involvement of ERS in the onset and progression of NAFLD. Considering the key role of ERS in the progression of NAFLD, reticulum stress-related genes (ERSGs) may become potential biomarkers and important targets for treating NAFLD.

The present study used the GEO database to analyze the endoplasmic reticulum stress-related differentially expressed genes (ERSR-DEGs) and immune signatures between normal and NAFLD samples. Enrichment analysis of these ERSR-DEGs has the potential to identify biological pathways that play a key role in the development of NAFLD. We divided 113 NAFLD patients into two groups based on the screened 48 ERSR-DEGs, and the differences in immune cells between the two groups were further analyzed. Various machine learning models are also applied to build disease prediction models. Multiple metrics are used to evaluate the effectiveness of the predictive model. The model was also validated in a mouse model of high-fat diet-induced NAFLD, providing a new perspective on the molecular mechanisms underlying NAFLD pathogenesis. The flow chart for this study is shown in Figure S1 (in Supplementary material).

Materials and Methods

Acquisition of datasets and ERS-related genes

We conducted a systematic search of the Gene Expression Omnibus database (GEO) (http://www.ncbi.nlm.nih.gov/geo/) using the terms: "Homo sapiens" and "NAFLD". The GSE89632 (Arendt et al. 2015) dataset (GPL14951 platform) and the GSE164760 (Pinyol et al. 2021) dataset (GPL13667 platform) were selected as training sets, and a total of 30 normal and 113 NAFLD liver tissue samples were included. The GSE151158 (Kriss et al. 2020) dataset (GPL28577 platform) and the GSE66676 (Xanthakos et al. 2015) (GPL6244 platform) were selected as validation sets. The GSE151158 dataset contained 21 normal and 40 NAFLD liver tissue samples. The GSE66676 dataset contained 34 normal and 33 NAFLD liver tissue samples. All NAFLD samples included were diagnosed by liver biopsy, and for healthy samples by imaging and histology in all datasets. Then we used the "limma" and "sva" packages in R software (version 4.2.2) to

perform normalization and batch effect correction on the training set data (Leek et al. 2012).

ERSGs were obtained from the GeneCards database (https://www.genecards.org) and genes with a correlation score >7 were extracted from this study (Safran et al. 2010).

Identification of differentially expressed genes associated with NAFLD and ERS

We used the R package "limma" to identify differentially expressed genes (DEGs) between NAFLD and normal liver tissue by Student's *t*-test (Ritchie et al. 2015). Genes with a *p*-value of < 0.05 and |logFC| > 0.5 were considered DEGs. Finally, the DEGs were intersected with ERSGs, and the intersected genes were the ERSR-DEGs in NAFLD. Moreover, volcano plot is based on differential expression data using the "ggplot2" package, and gene matrix heat plot using the "heatmap" package.

Functional and pathway enrichment analysis

The R packages "clusterProfiler" and "org.Hs.eg.db" were used to elucidate the potential gene function annotation and enrichment pathways associated with ERSR-DEGs (Yu et al. 2012). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to determine statistically significant enrichment using an adjusted *p* value < 0.05 as the cut-off criterion by Student's *t*-test. Histograms and bubble plots were generated with the R package "ggplot2".

Immune infiltration analysis for NAFLD

Assessment of immune infiltration status between normal and fatty liver samples using single sample gene set enrichment analysis (ssGSEA) to calculate the normalized enrichment score (Hanzelmann et al. 2013). We used Mann-Whitney U-test to screened samples using *p* < 0.05. An immune infiltration heat map was produced using the "heatmap" package. Immune cell levels between NAFLD and control were visualized using the "vioplot" package.

Predictive modeling based on various machine learning techniques

We applied the "caret" R package to build machine learning models, including random forest models (RF), support vector machine models (SVM), and extreme gradient boosting (XGB). The residual distributions and feature importance between these machine learning models are also visualised. The aim of RF is to reduce the number of variables required to obtain a prediction in order to reduce the data collection burden and increase efficiency (Rigatti 2017). SVM is

a powerful feature selection algorithm in machine learning classification techniques. The algorithm has good robustness and is now widely used to classify data (Speiser et al. 2019). The XGB algorithm is a representative algorithm based on integral lift, which compensates for the overfitting problem of the gradient lift model. The region under the receiver operating characteristic (ROC) curve is displayed using the "proc" R package. Therefore, the top five most significantly differentially expressed gene genes from the optimal machine learning model were used as the hub genes relevant to NAFLD. ROC curve analysis was performed in the GSE151158 and GSE66676 dataset to validate the diagnostic value of this model. Finally, we also used spearman correlation analysis to explore the relationship between the key predictive genes and clinical indicators. *p* < 0.05 was considered statistically significant.

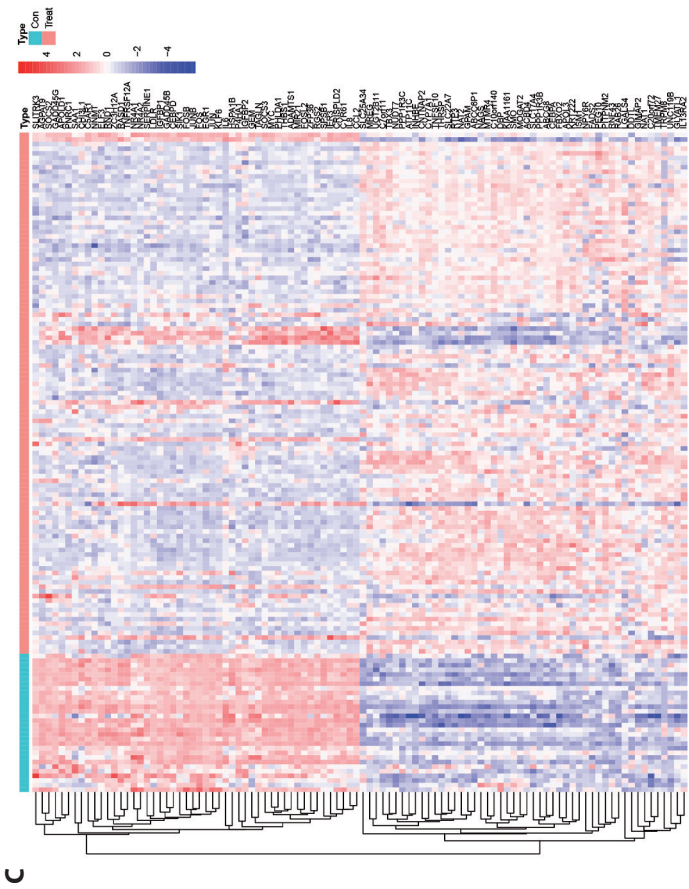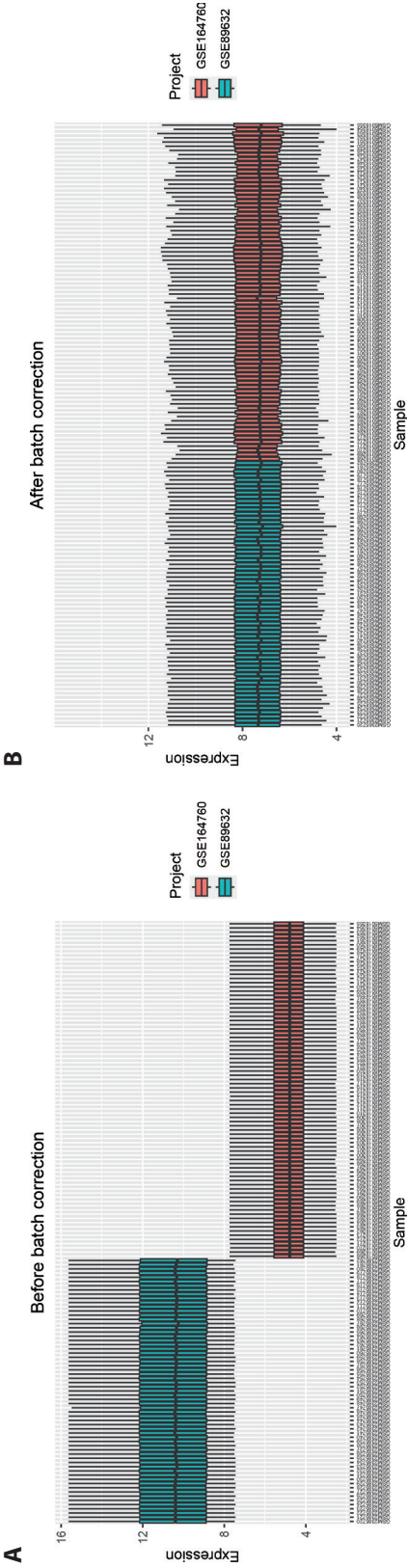### Construction and validation of a nomogram

A nomogram was developed using the "rms" R package to assess the occurrence of NAFLD. Scores for each gene in NAFLD were obtained, summing the scores for these key genes and predicting the prevalence of NAFLD based on the scores. Validation by calibration curves and decision curve analysis (DCA) curves to measure the recognition capability of the nomogram.

Subclusters analysis with 48 endoplasmic reticulum stress-related genes

The "consensusclusterplus" package can be used to identify patterns associated with ERSGs (Wilkerson and Hayes 2010). First, only the disease group sample was retained, and NAFLD patients were divided into subgroups with maximum subgroup classification k = 9, and the optimal subgroup was selected. Subsequently, a principal component analysis (PCA) was performed to quantify the endoplasmic reticulum stress pattern and to further determine whether the above groupings were correct. The R software packages "limma" and "heatmap" were then used to explore the clustered ERSR-DEGs and to study ERSR-DEGs with significant features of molecular subtypes. We performed gene set variance analysis (GSVA) using the marker gene set (C2.cp.kegg.symbols.gmt, C2.go.symbols.gmt) sourced from the MSigDB database. We then analyzed gene enrichment signatures using the R package "profiler". The results can be visualized using the "ggplot2" package.

Animal model and experiment design

Twenty male C57BL/6J mice, body mass 23.45 ± 1.22 g, 6 weeks old, were purchased from the Liaoning Provincial Laboratory Animal Resource Centre (Liaoning, China).
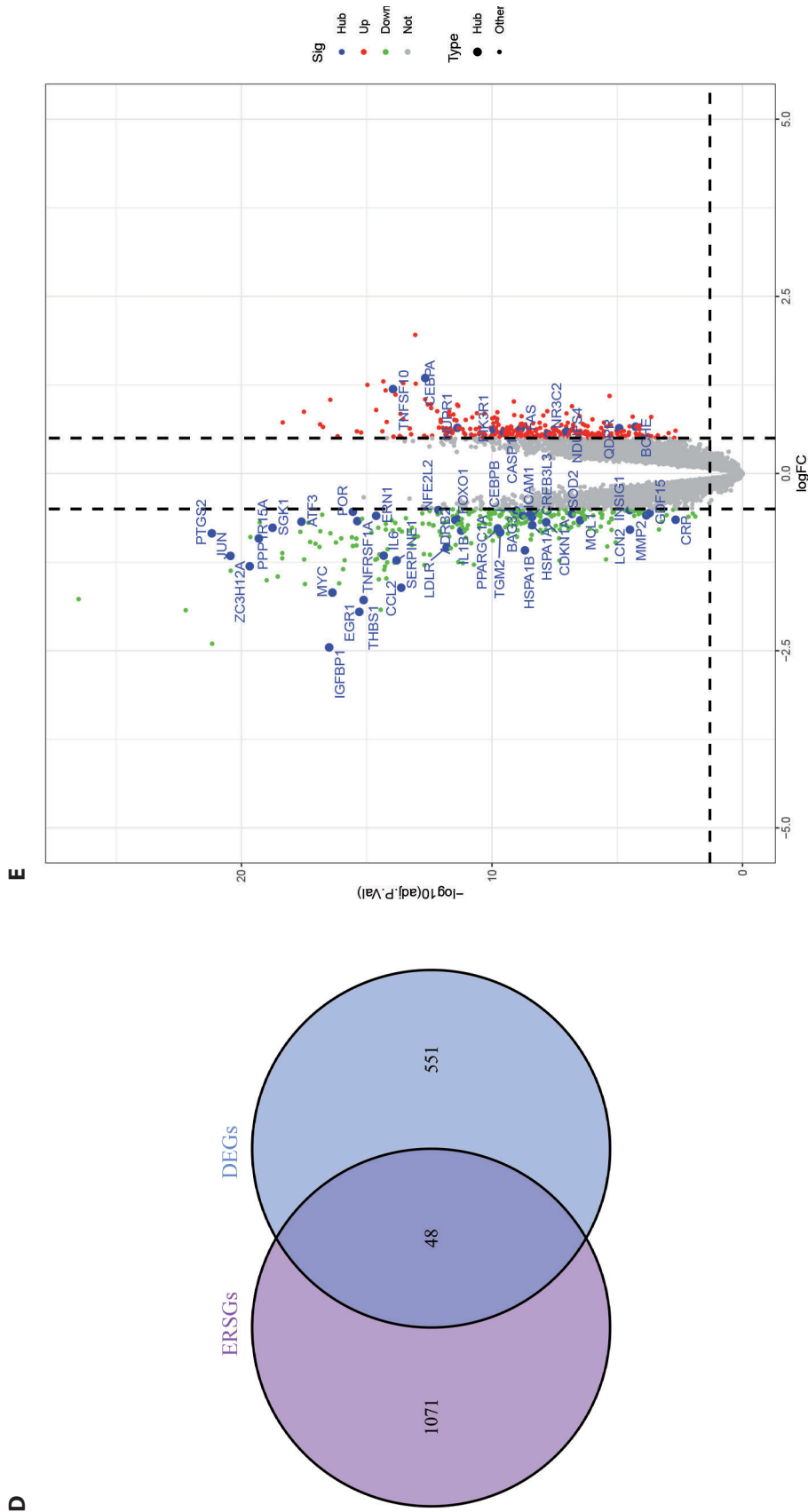
Figure 1. ERSR-DEGs in NAFLD. Boxplot showing the training set batch effects before (**A**) and after removal (**B**). **C**. Heatmap showing DEG expression profiles in NAFLD and non-NAFLD controls. **D**. DEGs and ERSGs were taken to intersect to obtain ERSR-DEGs. Volcano map (**E**) represents ERSR-DEGs. ERSR-DEGs, endoplasmic reticulum stress-related differentially expressed genes; NAFLD, nonalcoholic fatty liver disease.

Before animal testing, all mice were habituated for 7 days at a suitable temperature and a light/dark cycle of 12 hours. Animal management and experimental procedures are in strict compliance with the standard requirements of the Laboratory Animal Centre of the Fourth Hospital of Harbin Medical University (2022-DWSYLLCZ-20). All mice were anesthetized using 2% isoflurane after 12 weeks. After mice were sacrificed by cervical dislocation, liver tissues were collected.

After habituation, 20 mice were randomly divided into the normal control (NC) group ($n = 10$) and NAFLD group ($n = 10$). The normal control group was fed with the standard chow diet (SCD), and the NAFLD group was fed with high-fat diet (HFD) (60% fat) (D12492, xiaoshuyoutai, Beijing, China). Both groups of mice were executed after 12 weeks.

Quantitative real-time PCR

Total RNA was extracted from liver tissue using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), and cDNA was synthetized using PrimeScript reverse transcriptase (Takara, Kusatsu, Japan). Then, 2*SYBR Green qPCR (Vazyme, Nanjing, China) was used to analyze gene expression. $2^{-\Delta\Delta Ct}$ method is used to analyze the expression level of the target gene, and the result is controlled by β-actin. Primer sequences in this study are listed in Table S1 (in Supplementary material).

Triglyceride and cholesterol levels, histology

Mice's liver pathological damage was measured by hematoxylin and eosin (H&E) staining. Oil red O staining was used to detect liver lipid deposition. Total triglycerides (TG) and Total cholesterol (TC) were determined using a TG kit and TC kit (Jiancheng, Nanjing, China), according to the manufacturer's instructions.

Statistical analysis

All statistical analyzes were performed using R version 4.2.2. Differences between two groups were compared using Student's *t*-test or Mann-Whitney U-test according to whether the data conformed to a normal distribution. Correlations between each variable were analyzed using the Spearman correlation test. A value of $p < 0.05$ was considered as statistical significance. All statistical analyses were performed using GraphPad prism8.0.2 software.

Results

ERS-related genes differentially expressed in NAFLD

The GSE89632 and GSE63067 datasets were combined and batch effects were eliminated (Fig. 1A,B). Based on the screening criteria, a total of 598 DEGs were obtained, including 301 up-regulated genes and 297 down-regulated genes. The top 50 DEGs are indicated by a heatmap (Fig. 1C). We used the obtained DEGs to take intersections with ERSGs and obtained a total of 48 ERSR-DEGs (Fig. 1D). The intersecting genes were represented with a volcano plot (Fig. 1E).

Enrichment analysis of differentially expressed ERS-related genes

Enrichment analysis was performed on the 48 ERSR-DEGs obtained above. ERSR-DEGs related to biological processes were significantly enriched in the regulation of positive regulation of inflammatory response, response to unfolded protein, and response to oxidative stress. ERSR-DEGs related to cytological components were significantly enriched in the endoplasmic reticulum lumen, RNA polymerase II transcription regulator complex, and collagen-containing extracellular matrix. ERSR-DEGs related to molecular functions are mainly enriched in DNA-binding transcription activator activity, RNA polymerase II-specific, and low-density lipoprotein particle binding (Fig. S2A in Supplementary material). KEGG analysis revealed that ERSR-DEGs were mainly enriched in TNF signaling pathways, non-alcoholic fatty liver, and fluid shear stress and atherosclerosis (Fig. S2B in Supplementary material).

Immune infiltration analysis

Immune permeability analysis revealed differences in the percentage of infiltration of immune cell types between NAFLD and non-NAFLD controls. As shown in the heatmap and boxplot (Fig. 2A,B), activated CD4 T cells, activated CD8 T cells, activated dendritic cell, CD56 natural killer cells, eosinophil, gamma delta T cells, immature B cell, immature dendritic cell, MDSC, mast cell, neutrophil, plasmacytoid dendritic cell, T follicular helper cell, Type 1 T helper cell, Type 17 T helper cell, Type 2 T helper cell, effector memory CD4 T cells, memory B cells, central memory CD4 T cells, and central memory CD4 T cells showed statistical difference between the NAFLD group and the control group. In terms of immune function (Fig. 2C), cytolytic activity, Type I IFN reponse, APC co-inhibition, APC co-stimulation, CCR, check-point, MHC class I, parainflammation, and T cell co-inhibition was significantly different in the control and NAFLD groups.

Identification of cluster-specific differentially expressed ERS-related genes

Using unsupervised cluster analysis, different subtypes associated with ERS were identified according to the expression levels of 48 ERSR-DEGs. The most stable grouping
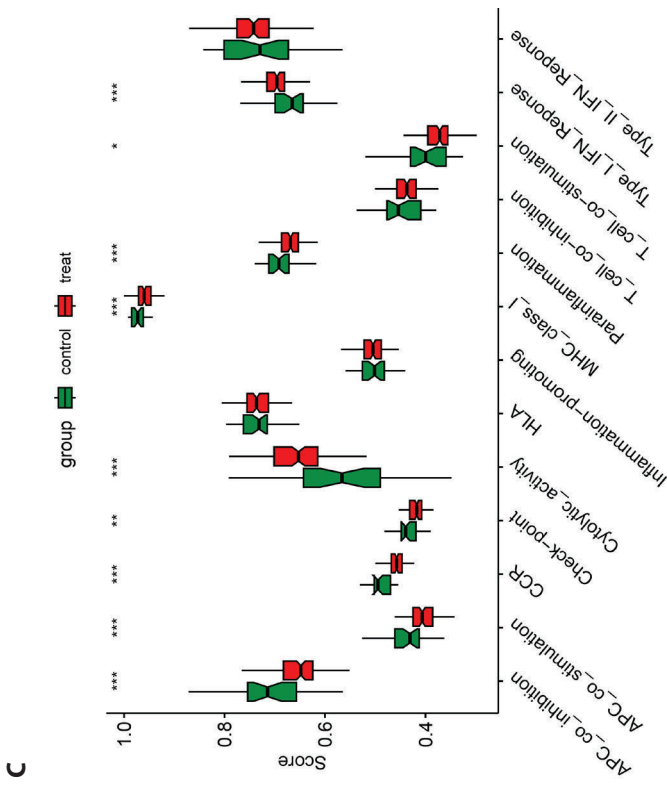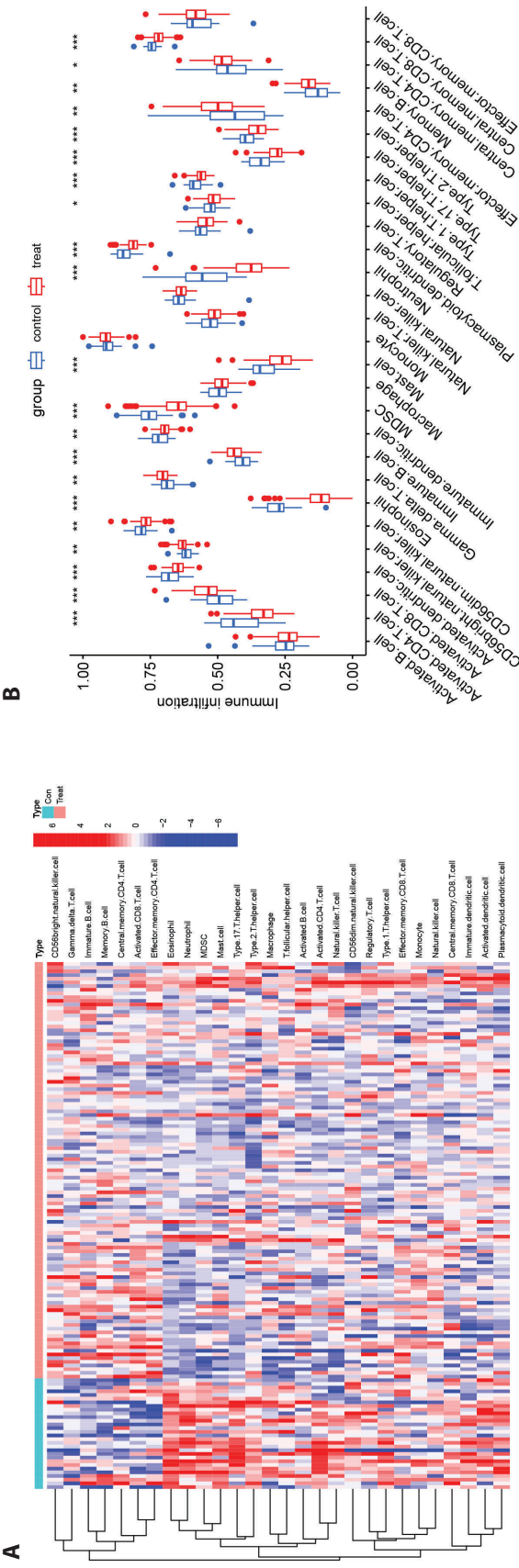
Figure 2. Identification of immune infiltration patterns in NAFLD and normal conditions. **A.** Heatmap showed the expression characteristics of immune cells in NAFLD and non-NAFLD controls. Boxplots showed the differences in immune cells infiltrating (**B**) and immune function (**C**) between NAFLD and non-NAFLD controls. $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$.
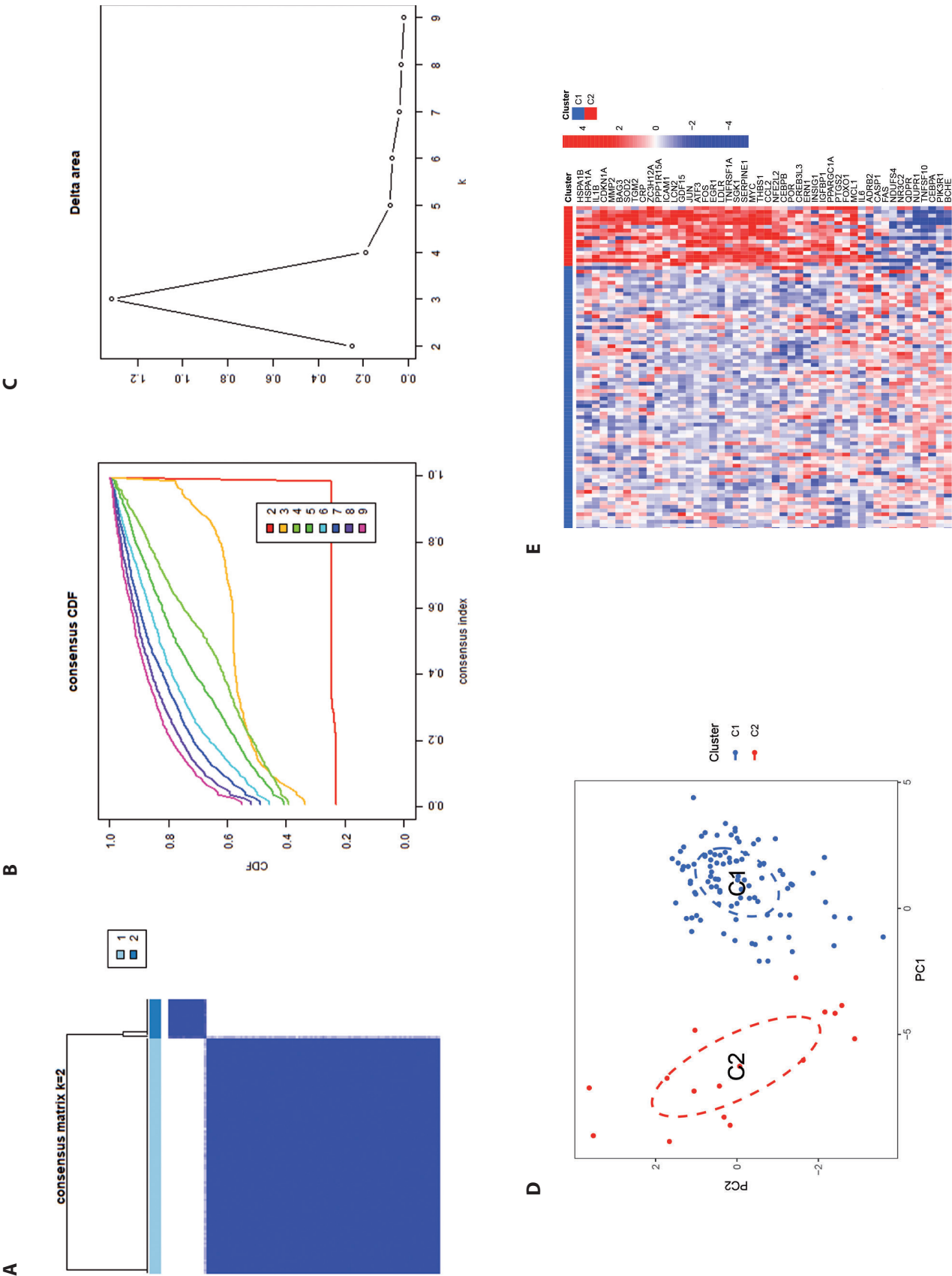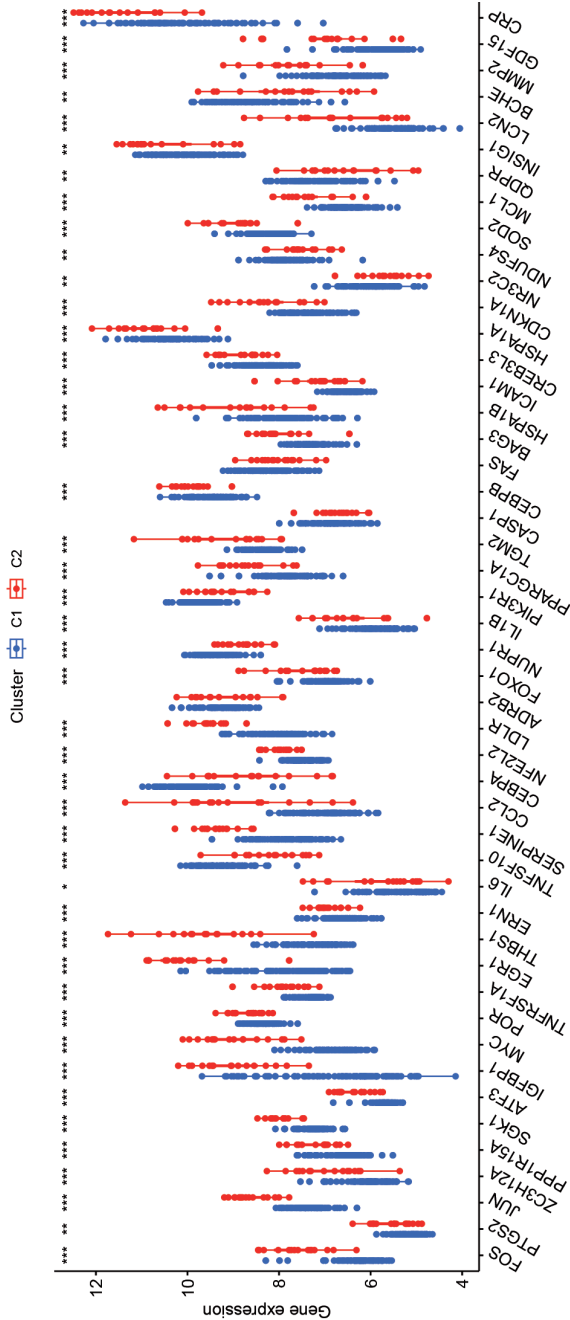
was provided when k = 2 (Fig. 3A). Furthermore, at k = 2, the cumulative distribution function (CDF) curve (Fig. 3B) is the flattest and the delta area curve (Fig. 3C) has an inflection point. Thereafter, the 113 NAFLD samples were divided into two different categories, namely Cluster 1 (*n* = 97) and Cluster 2 (*n* = 16). The samples of the two clusters can be clearly distinguished as shown in the PCA diagram (Fig. 3D). The expression levels of ERSGs were visualized by heatmap and boxplot (Fig. 3E,F). Most of the ERSGs, including FOS, PTGS2, JUN, ZC3H12A, PPP1R15A, and SGK1, were expressed at higher levels in Cluster 2 than in Cluster 1.

Functional annotation and immune infiltration characteristics between endoplasmic reticulum stress clusters

There are differences in the immune microenvironment between Cluster 1 and Cluster 2 (Fig. 4A). Most immune cells including activated B cells, activated CD4 T cells, and activated dendritic cells were significantly elevated in Cluster 2, with the exception of gamma delta T cells. Similarly, Cluster 2 has a higher immunity score (Fig. 4B). GSVA analysis was used to further explore the function between the two clusters in relation to cluster specific DEGs. In the KEGG analysis (Fig. 4C), Cluster 1 was enriched in the metabolism of xenobiotics by cytochrome p450, pentose and glucuronate interconversions, ascorbate, and alternate metabolism; whereas Cluster 2 was upregulated in the intestinal immune network for IgA production, P53 signaling pathway, proximal tubule bicarbonate reclamation. The functional enrichment (Fig. 4D) showed that the transcription factor AP-1 complex negative regulation of interleukin 21 production, and positive regulation of miRNA metabolic process were significantly correlated with Cluster 2; however, thiolester hydrolase activity, monoacylglycerol metabolic process, alpha linolenic acid metabolic process were enriched in Cluster 1.

Construction of predictive model based on multiple machine learning methods

To more accurately identify specific genes with high diagnostic value, we used RF, SVM, and XGB models based on the screened 48 relevant genes. The SVM machine learning model provided the lowest residual (Fig. 5A,B). The important ERSGs in each machine model (Fig. 5C). In addition, we used ROC curves to evaluate the diagnostic performance of the three machine learning algorithms (Fig. 5D–F). The area under the ROC curve of the machine learning SVM model is the largest (SVM, AUC = 0.997; RF, AUC = 0.990; XGB, AUC = 0.990). Overall, in combination with these results, the SVM model proved to have the highest predictive performance. In the end, the five most important characters in
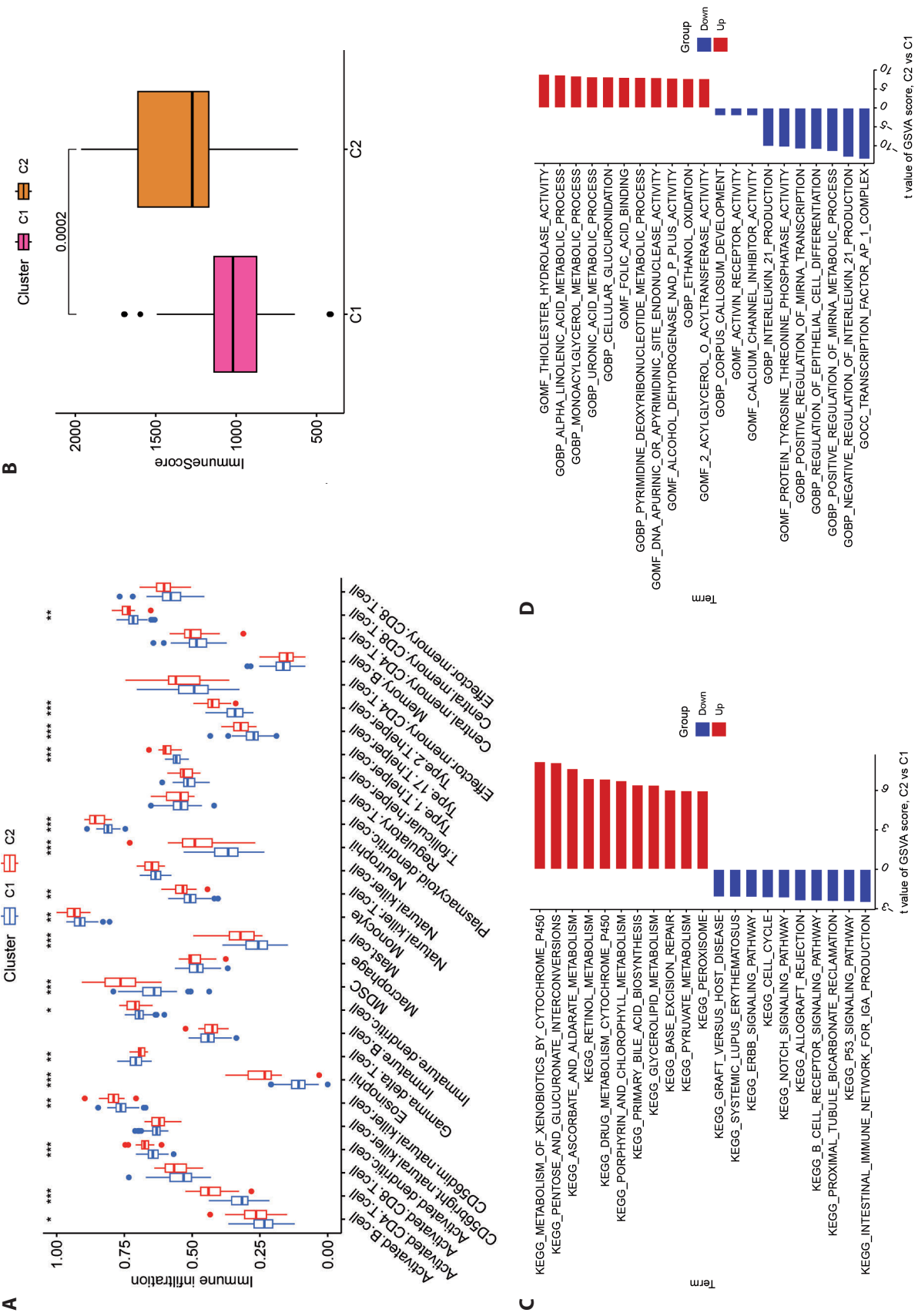

Figure 3. Identification of ERS subtypes in NAFLD. **A.** Two different subtypes of NAFLD. **B.** CDF curves. **C.** Delta area of CDF curves. **D.** PCA analysis showing significant differences in genes between the different models. Heatmap (**E**) and boxplot (**F**) showed differences in the expression of ERSGs in C1 and C2, with significant differences in the expression of 45 genes. * *p* < 0.05, ** *p* < 0.01, *** *p* < 0.001. PCA, principal component analysis. CDF, cumulative distribution function.

Figure 4. Functional enrichment analysis and immune cell infiltration between endoplasmic reticulum stress subtypes. **A.** Correlation matrix of all 28 immune cell subtype compositions. **B.** Boxplot demonstrates the immune score for the two ERS subtypes. The heatmap of GSVA Enrichment analysis was used to visualize biological pathways (**C**) and functions (**D**) between Cluster 1 and Cluster 2. GSVA, gene set variation analysis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

the SVM model (ERN1, POR, PPP1R15A, FOS, and FAS) were selected as hub ERSR-DEGs for further analysis.

To further evaluate the predictive efficiency of the SVM model, we constructed a nomogram based on the five hub ERSR-DEGs described above (Fig. 6A). Based on the calibration curves, the error between the actual and predicted risk of NAFLD was small (Fig. 6B). The study found between 0 and 1 in the decision analysis graph, indicating that the nomogram has a higher clinical value than individual trait genes and has implications for predicting the occurrence of NAFLD (Fig. 6C). The ROC curves showed good results for the 5-gene SVM model with an AUC value of 0.712 in the GSE151158 dataset and 0.710 in the GSE66676 dataset (Fig. 6D,E). After these validations, it was shown that our diagnostic model was effective in distinguishing NAFLD from normal individuals.

To investigate the biological relationship between hub ERSR-DEGs and immune microenvironmental features, the correlation between five ERSR-DEGs and infiltrating immune cells was analysed (Fig. 6F). Correlation analysis showed that five hub ERSR-DEGs were positively correlated with the most immune cells. As an example, POR expression was significantly positively correlated with plasmacytoid dendritic cell, neutrophil natural killer cell, monocyte, MDSC, mast cell, and eosinophil and significantly negatively correlated with immature B cell. It is speculated that hub ERSR-DEGs are a key factor controlling the state of immune infiltration in NAFLD patients.

Correlation of genes with clinical indicators

We validated the correlation of hub ERSR-DEGs with clinical indicators using clinical information from patients in the GSE89632 dataset (Fig. S3 in Supplementary material). We found that FAS was positively correlated with AST ($p = 0.047$, $R = 0.32$), ballooning (intensity) ($p = 0.021$, $R = 0.37$), fibrosis (stage) ($p = 0.021$, $R = 0.37$), and FOS was negatively correlated with HDL ($p = 0.033$, $R = -0.37$).

Changes in expression of the five hub ERSR-DEGs in HFD-induced NAFLD mice

Mice fed HFD developed hepatic steatosis, similar to NAFLD in humans (Van Herck et al. 2017). After 12 weeks of feeding HFD to mice in the NAFLD group, liver pathology showed swollen hepatocytes and balloon-like degeneration in liver pathology sections of mice (Fig. 7A). Oil red O staining showed accumulation of neutral lipids in NAFLD group (Fig. 7B). Mouse liver TG (Fig. 7C) and TC (Fig. 7D) content and adipogenic genes (ACC1 FASN SCD1) expression (Fig. 7E) were significantly elevated in the NAFLD group. The above results indicated that the NAFLD mouse model was successfully established.

We evaluated the five hub ERSR-DEGs. qRT-PCR results (Fig. 7F) showed that the mRNA levels of POR, PPP1R15A, and FOS were significantly down-regulated and FAS was significantly up-regulated in the HFD group compared to the NC group. In contrast, there were no significant changes in other genes. In short, these results suggest that ERSGs are well differentiated for NAFLD, validating the analysis of the microarray data.

Discussion

As obesity, diabetes, and metabolic syndrome continue to rise, epidemiological projections indicate that the prevalence of NAFLD will continue to rise through 2030; liver-related disease mortality will double (Estes et al. 2018). Numerous NAFLD-causing factors are also strongly associated with hepatocellular carcinoma. It is challenging to detect NAFLD-associated hepatocellular carcinoma through exploration of the pathogenesis, so relevant biomarkers, initial treatment, and prevention of NAFLD are imminent (Ioannou 2021).

The important role of ERS in NAFLD has been found in fatty liver mouse models and many clinical samples (Lebeaupin et al. 2018). The endoplasmic reticulum is the main site of intracellular protein synthesis and modification. When protein precursors are synthesized over the quality control capacity of the endoplasmic reticulum, an unfolded protein response is triggered, which in turn induces ERS, disrupting hepatic lipid metabolism and apoptosis of hepatocytes and hepatic stellate cells. This study provides the first insight into the biological importance of ERSGs and their relationship to immune infiltration in NAFLD. Furthermore, NAFLD subtypes were predicted using genetic markers linked to ERS.

We identified 48 ERSR-DEGs using GEO and GeneCards databases to conduct the first comprehensive analysis of the expression profile of ERS regulators in liver tissue from normal and non-alcoholic fatty liver patients. We further demonstrated the level of immune cell infiltration in NAFLD patients through the ssGSEA algorithm. activated CD4 T cells, activated CD8 T cells, activated dendritic cell, CD56 natural killer cells, eosinophil, gamma delta T cells, immature B cell, immature dendritic cell, MDSC, mast cell, neutrophil, plasmacytoid dendritic cell, T follicular helper cell, Type 1 T helper cell, Type 17 T helper cell, Type 2 T helper cell, effector memory CD4 T cells, memory B cells, central memory CD4 T cells, and central memory CD4 T cells showed statistical difference between the NAFLD group and the control group, as determined by the immune cell infiltration analysis. This finding – consistent with previous studies (Wen et al. 2021; Zhang et al. 2022) – suggests that immunity plays an essential role in NAFLD. Furthermore, based on 48 ERSR-DEGs, we identified two ERS-related clusters using unsupervised cluster analysis. Cluster 2 exhibited relatively high
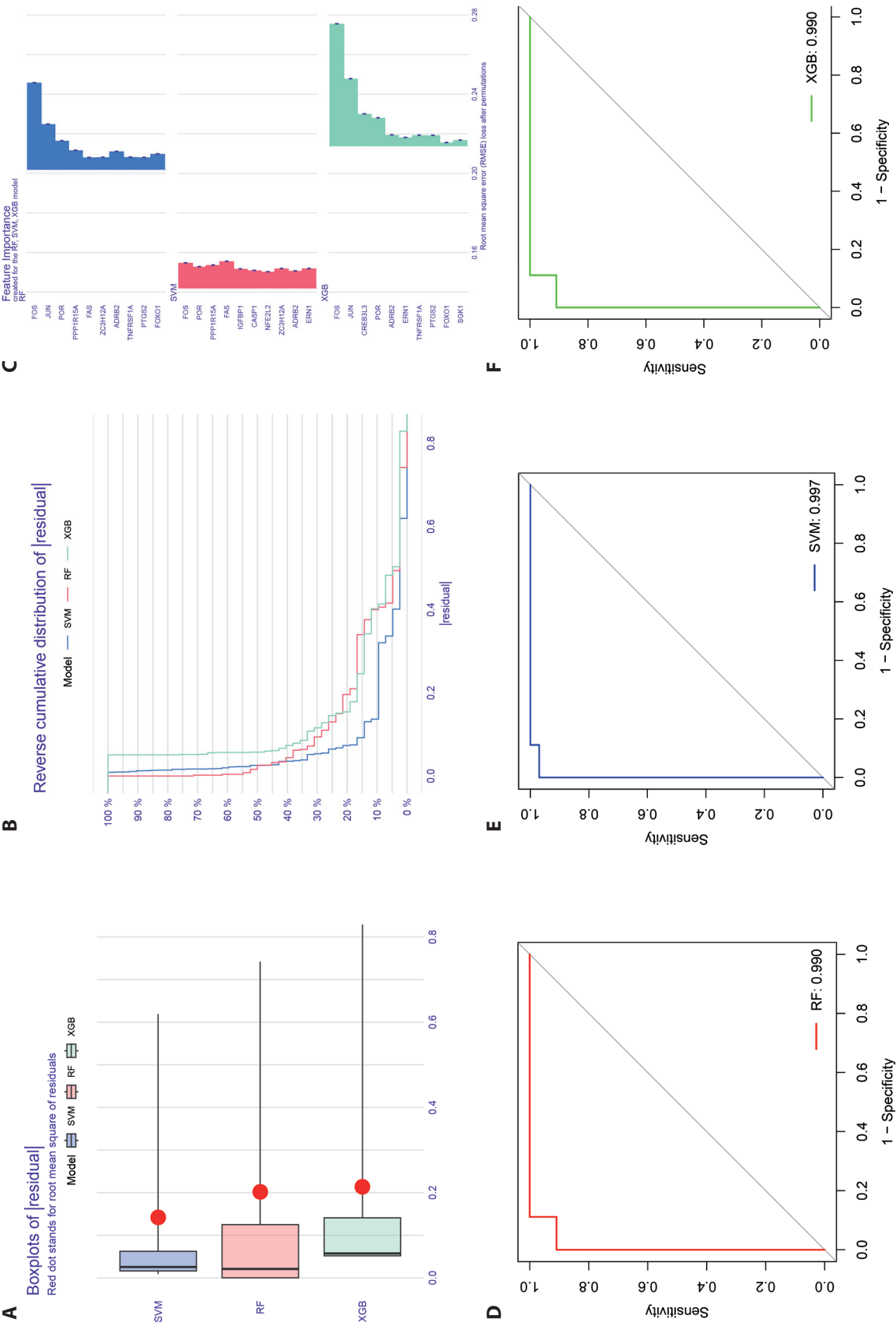
Figure 5. Construction and evaluation of RF, SVM, and XGB machine models. **A.** Reverse cumulative distribution of residuals of each machine learning model. **B.** Boxplots showed the residuals of each model. **C.** The important features in RF, SVM, and XGB machine models. ROC analysis of RF (**D**), SVM (**E**), and XGB (**F**) machine learning models based on 5-fold cross-validation in the testing cohort. x-axis: FPR, false positive rate; y-axis: TPR, true positive rate; AUC, the area under the curve; RF, random forest; SVM, support vector machine; XGB, extreme gradient boosting; ROC, receiver operating characteristic; CI, confidence interval.

levels of immune infiltration and immune scores. In KEGG analysis, C2 was enriched in the intestinal immune network for IgA production, P53 signaling pathway, proximal tubule bicarbonate reclamation; in GO analysis, C2 was enriched in the transcription factor AP-1 complex, negative regulation of interleukin 21 production, and positive regulation of miRNA metabolic process. Therefore, it is reasonable to hypothesize that Cluster 2 may be able to block the progression of NAFLD by producing a variety of immune cells, such as T cells and B cells, and ultimately exhibit a better prognosis for NAFLD. To construct reliable prediction models, we compared the prediction properties of three machine algorithms (RF, SVM, and XGB) and finally selected the prediction model of SVM. The SVM model provided the smallest residuals and had the highest test cohort prediction efficiency (AUC = 0.997), indicating that SVM-based machine learning performs satisfactorily in predicting NAFLD.

We used the GSE151158 and GSE66676 datasets for validation, and the results showed that the area under the ROC curve was 0.712 and 0.710, respectively. The SVM model can accurately predict NAFLD. This prediction provides a new way of diagnosing NAFLD. The five ERSGs (ERN1, POR, PP-P1R15A, FOS, and FAS) with the most significant differences in the SVM model were included in the risk study; a nomogram was constructed to calculate total scores to analyze the probability of risk in NAFLD patients. Correlations between immune cell infiltration and hub ERSR-DEGs were then estimated to elucidate ERSR-DEGs in the immunological context of NAFLD. AST and HDL are routine blood tests for NAFLD patients, and ballooning and fibrosis are important indicators to assess the severity of NAFLD. Therefore, we correlated the five genes in the SVM model with them. The results showed that FAS was positively correlated with AST, ballooning, and fibrosis, while FOS was negatively correlated with HDL. After the successful induction of NAFLD in mice, mRNA levels of ERSGs were analyzed using qRT-PCR. The results demonstrated that compared with the control group, the levels of POR, PPP1R15A, FOS, and FAS in the liver tissue of HFD-fed mice were significantly changed, supporting the microarray analysis.

PPP1R15A/GADD34 – a cell cycle protein – is up-regulated under ERS conditions and reduces endoplasmic reticulum stress-related apoptosis by selectively inhibiting the phosphorylation of the ERS protein eIF2α, an important molecule for restoring protein synthesis during ERS (Walter and Ron 2011; Moreno et al. 2012). Overexpression of GADD34 phosphatase in the mice liver leads to dephosphorylation of eIF2α, which has been shown to lead to a weakening of the adipose-forming nuclear receptor PPARγ, lipogenic enzymes, and metabolic transcriptional regulators C/EBP-α and C/ EBP-β. This resulted in glucose tolerance, increased insulin sensitivity and a reduced incidence of hepatic steatosis in mice fed a high-fat diet compared to mice

fed a normal diet (Oyadomari et al. 2008). Van Herck et al. (2017) demonstrated that GADD34-deficient mice on a normal diet became obese and developed fatty liver, followed by cirrhosis, hepatocellular carcinoma, and insulin resistance. And liver steatosis can be exacerbated by aging and a high-fat diet. Unlike the results of Oyadomari et al. (2008), eIF2α phosphorylation did not increase significantly in GADD34-deficient mice due to HFD and aging. Obesity and steatosis in GADD34-deficient mice are caused by up-regulation of insulin-Akt signaling. Therefore, the mechanism of the effect of GADD34 on fatty liver is still controversial, which deserves further investigation.

FAS/FASL pathway is one of the major apoptosis pathways, and FASL is a ligand for FAS. FASL expression increases in response to lipotoxicity and binds to FAS, causing FAS to trimerize and form an apoptosis-inducing complex, which initiates apoptotic signaling and cell death (Zhang et al. 2012). FAS-mediated apoptosis requires endoplasmic reticulum-mediated calcium release, which is dependent on phospholipase C-gamma1 (PLC-gamma1) activation and Ca2+ release from inositol 1,4,5-trisphosphate receptor (IP3R) channels (Wozniak et al. 2006). FAS is significantly expressed in liver samples from NASH patients and can mediate hepatocyte apoptosis (Wozniak et al. 2006). Inhibition of FAS receptor signaling and inhibition of downstream effectors (including caspase3 and NF-κB) by injection of small interfering RNA (siRNA) into mouse tail vein reduces apoptosis and inflammation progression in mouse hepatocytes (Savari et al. 2019). FAS not only has an effect on apoptosis and inflammation in fatty liver cells, but also on fat metabolism. In normal diet mice with FAS overexpression, FAS impaired mitochondrial function and fatty acid oxidation by BH3 interaction domain death agonist, thereby promoting hepatic lipid accumulation and insulin resistance (Item et al. 2017).

Cytochrome P450 reductase (POR) is a microsomal electron transporter found in the endoplasmic reticulum. It is an obligate electron donor of cytochrome P450 (P450), which includes P450 required for drug metabolism, cholesterol, steroids, and bile acid synthesis (Henderson et al. 2003). POR deficient mice showed increased liver triglyceride content and decreased non-fasting plasma triglyceride and cholesterol levels (Weng et al. 2005). Porter et al. (2011) successfully replicated hepatic lipid accumulation in POR-deficient mice through RNA (siRNA) to inhibit POR in a cellular model. Bile acid deficiency and loss of farnesol X receptor stimulation led to excessive triglyceride synthesis in hepatocytes.

The FOS proto-oncogene (also known as c-FOS) is an immediate early expression gene and a major member of the FOS gene family. It can be dimerized with the JUN family to form an activating protein 1 (AP-1) complex. c-FOS, as a transcription factor, is involved in cell growth, proliferation, differentiation and death (Eferl and Wagner 2003). In our
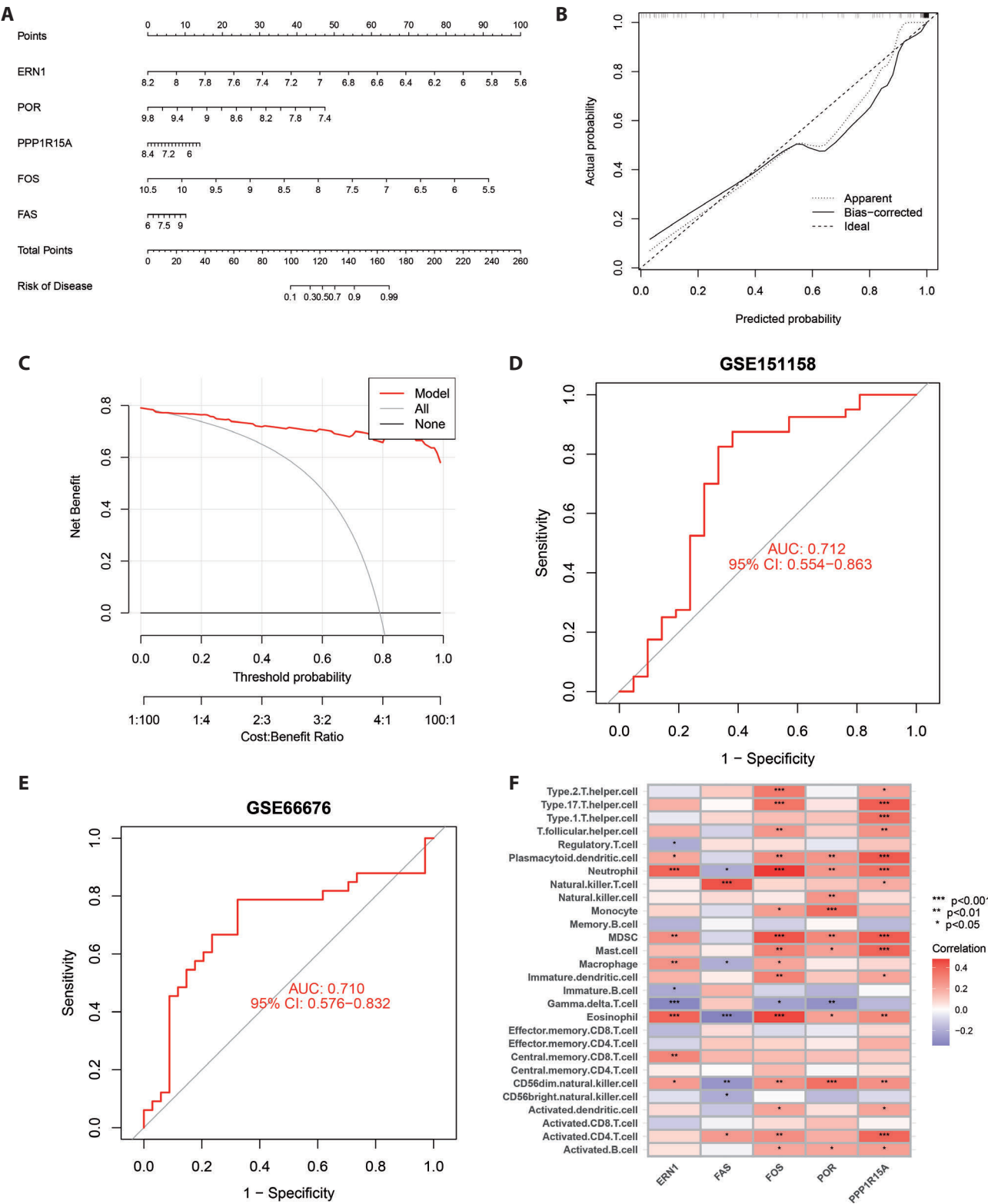
Figure 6. Validation of the 5-gene-based SVM model. **A.** Construction of a nomogram for predicting the risk of NAFLD based on the 5-gene-based SVM model. To assess the predictive efficiency of the nomogram model, calibration curves (**B**) and DCA (**C**) were constructed. ROC analysis of the 5-gene-based SVM model in GSE151158 datasets (**D**) and GSE66676 (**E**) datasets. **F.** Correlation between immune infiltration and five hub ERSR-DEGs. DCA, decision curve analysis.

study, the downregulation of c-FOS expression in NAFLD group was verified by microarray differential gene analysis and animal experiments. This is consistent with previous research findings (Chen et al. 2020; Jiang et al. 2021). After cluster analysis, we found that c-FOS expression was significantly higher in group C2 than in group C1. We speculated that the expression of c-FOS may also be different in different NAFLD subtypes. Although the contribution of c-FOS to the etiology of NAFLD seems to be little or unknown, there have been several bioinformatics articles on fatty liver, which have screened c-FOS as a hub gene (Ye et al. 2020; Ma et al. 2021). Therefore, it is necessary to further clarify their functions in the future. c-FOS is involved in the onset and development of HCC. Hepatocyte-specific expression of c-FOS leads to

hepatic inflammation, hepatocyte proliferation, DNA damage response, and precancerous transformation, but specific deletion of c-Fos reverses this phenotype. Mechanistically, c-FOS decreases the expression and activity of the nuclear receptor LXRα, leading to increased hepatic cholesterol and accumulation of bile acids (Bakiri et al. 2017).

Despite the encouraging results, the present study has several limitations. First, our data came from public databases and lacked raw sequencing data, leading to some bias in the results. Second, we could not collect clinical samples, so the results of the bioinformatics analysis could not be validated on liver tissue from NAFLD patients. Finally, more in-depth studies are required to understand the potential mechanisms of ERSGs in NAFLD.
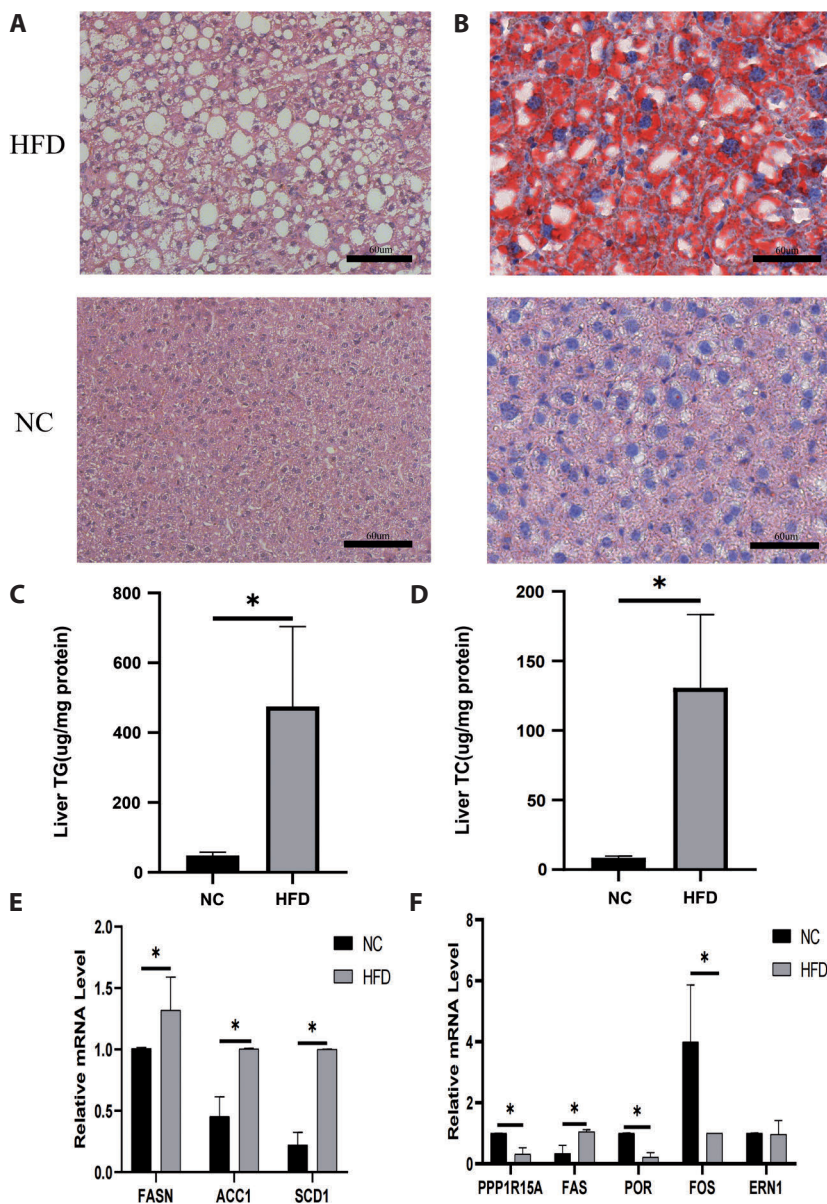

Figure 7. Hub genes expression in HFD-fed mice. Hematoxylin and eosin staining (**A**) and Oil red O staining (**B**) (original magnification × 400). Intrahepatic TG (**C**) and TC (**D**) content of the HFD group were higher than those in the NC group. The relative mRNA expression levels of hepatic lipogenesis-related genes (**E**) and ERSGs (**F**) in the NC group and HFD group were detected by qRT-PCR (n = 10 mice *per* group). * $p < 0.05$. NC, normal control; TG, total triglyceride; TC, total cholesterol; HFD, high-fat diet.

Conclusions

In conclusion, our study demonstrated the association between ERSGs and immune system infiltration in NAFLD patients through different bioinformatics techniques and explained the considerable immune heterogeneity in NAFLD patients with different ERS groups. POR, PPP1R15A, FOS, and FAS may predict the occurrence risk of NAFLD patients and play important roles in the pathogenesis of immune infiltration in NAFLD. Our findings elucidate the involvement of ERSGs in the progress of NAFLD and may provide new insight into the disease typing and diagnosis.

References

Arendt BM, Comelli EM, Ma DW, Lou W, Teterina A, Kim T, Fung SK, Wong DKH, McGilvray I, Fischer SE, Allard JP (2015): Altered hepatic gene expression in nonalcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. Hepatology **61,** 1565-1578
https://doi.org/10.1002/hep.27695

Bakiri L, Hamacher R, Grana O, Guio-Carrion A, Campos-Olivas R, Martinez L, Dienes DP, Thomsen MK, Hasenfuss SC, Wagner EF (2017): Liver carcinogenesis by FOS-dependent inflammation and cholesterol dysregulation. J. Exp. Med. **214,** 1387-1409
https://doi.org/10.1084/jem.20160935

Chen F, Zhou Y, Wu Z, Li Y, Zhou W, Wang Y (2020): Integrated analysis of key genes and pathways involved in nonalcoholic steatohepatitis improvement after roux-en-Y gastric bypass surgery. Front. Endocrinol. (Lausanne) **11,** 611213
https://doi.org/10.3389/fendo.2020.611213

Doultsinos D, Avril T, Lhomond S, Dejeans N, Guedat P, Chevet E (2017): Control of the unfolded protein response in health and disease. SLAS Discov. **22,** 787-800
https://doi.org/10.1177/2472555217701685

Eferl R, Wagner EF (2003): AP-1: a double-edged sword in tumorigenesis. Nat. Rev. Cancer **3,** 859-868
https://doi.org/10.1038/nrc1209

Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, Zelber-Sagi S, Wai-Sun Wong V, Dufour JF, Schattenberg JM, et al. (2020): A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. J. Hepatol. **73,** 202-209
https://doi.org/10.1016/j.jhep.2020.03.039

Estes C, Anstee QM, Arias-Loste MT, Bantel H, Bellentani S, Caballeria J, Colombo M, Craxi A, Crespo J, Day CP, Eguchi Y, et al. (2018): Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030. J. Hepatol. **69,** 896-904
https://doi.org/10.1016/j.jhep.2018.05.036

Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ (2018): Mechanisms of NAFLD development and therapeutic strategies. Nat. Med. **24,** 908-922
https://doi.org/10.1038/s41591-018-0104-9

Gonzalez-Rodriguez A, Mayoral R, Agra N, Valdecantos MP, Pardo V, Miquilena-Colina ME, Vargas-Castrillón J, Lo Iacono O, Corazzari M, Fimia GM, et al. (2014): Impaired autophagic flux is associated with increased endoplasmic reticulum stress during the development of NAFLD. Cell Death Dis. **5,** e1179
https://doi.org/10.1038/cddis.2014.162

Hanzelmann S, Castelo R, Guinney J (2013): GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics **14,** 7
https://doi.org/10.1186/1471-2105-14-7

Henderson CJ, Otto DM, Carrie D, Magnuson MA, McLaren AW, Rosewell I, Wolf CR (2003): Inactivation of the hepatic cytochrome P450 system by conditional deletion of hepatic cytochrome P450 reductase. J. Biol. Chem. **278,** 13480-13486
https://doi.org/10.1074/jbc.M212087200

Ioannou GN (2021): Epidemiology and risk-stratification of NAFLD-associated HCC. J. Hepatol. **75,** 1476-1484
https://doi.org/10.1016/j.jhep.2021.08.012

Item F, Wueest S, Lemos V, Stein S, Lucchini FC, Denzler R, Fisser MC, Challa TD, Pirinen E, Kim Y, et al. (2017): Fas cell surface death receptor controls hepatic lipid metabolism by regulating mitochondrial function. Nat. Commun. **8,** 480
https://doi.org/10.1038/s41467-017-00566-9

Jiang ZY, Zhou Y, Zhou L, Li SW, Wang BM (2021): Identification of key genes and immune infiltrate in nonalcoholic steatohepatitis: a bioinformatic analysis. Biomed. Res. Int. **2021,** 7561645
https://doi.org/10.1155/2021/7561645

Kriss M, Golden-Mason L, Kaplan J, Mirshahi F, Setiawan VW, Sanyal AJ, Rosen HR (2020): Increased hepatic and circulating chemokine and osteopontin expression occurs early in human NAFLD development. PLoS One **15,** e0236353
https://doi.org/10.1371/journal.pone.0236353

Lebeaupin C, Vallee D, Hazari Y, Hetz C, Chevet E, Bailly-Maitre B (2018): Endoplasmic reticulum stress signalling and the pathogenesis of non-alcoholic fatty liver disease. J. Hepatol. **69,** 927-947
https://doi.org/10.1016/j.jhep.2018.06.008

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

Ma M, Xie W, Li X (2021): Identification of autophagy-related genes in the progression from non-alcoholic fatty liver to non-alcoholic steatohepatitis. Int. J. Gen. Med. **14,** 3163-3176
https://doi.org/10.2147/IJGM.S317785

Meares GP, Liu Y, Rajbhandari R, Qin H, Nozell SE, Mobley JA, Corbett JA, Benveniste EN (2014): PERK-dependent activation of JAK1 and STAT3 contributes to endoplasmic reticulum stress-induced inflammation. Mol. Cell. Biol. **34,** 3911-3925
https://doi.org/10.1128/MCB.00980-14

Moreno JA, Radford H, Peretti D, Steinert JR, Verity N, Martin MG, Halliday M, Morgan J, Dinsdale D, Ortori CA, et al. (2012): Sustained translational repression by eIF2alpha-P mediates prion neurodegeneration. Nature **485,** 507-511
https://doi.org/10.1038/nature11058

Oyadomari S, Harding HP, Zhang Y, Oyadomari M, Ron D (2008): Dephosphorylation of translation initiation factor 2alpha enhances glucose tolerance and attenuates hepatosteatosis in mice. Cell Metab. **7,** 520-532
https://doi.org/10.1016/j.cmet.2008.04.011

Pagliassotti MJ (2012): Endoplasmic reticulum stress in nonalcoholic fatty liver disease. Annu. Rev. Nutr. **32,** 17-33
https://doi.org/10.1146/annurev-nutr-071811-150644

Pinyol R, Torrecilla S, Wang H, Montironi C, Pique-Gili M, Torres-Martin M, Wei-Qiang L, Willoughby CE, Ramadori P, Andreu-Oller C, et al. (2021): Molecular characterisation of hepatocellular carcinoma in patients with non-alcoholic steatohepatitis. J. Hepatol. **75,** 865-878
https://doi.org/10.1016/j.jhep.2021.04.049

Porter TD, Banerjee S, Stolarczyk EI, Zou L (2011): Suppression of cytochrome P450 reductase (POR) expression in hepatoma cells replicates the hepatic lipidosis observed in hepatic POR-null mice. Drug Metab. Dispos. **39,** 966-973
https://doi.org/10.1124/dmd.111.038562

Powell EE, Wong VW, Rinella M (2021): Non-alcoholic fatty liver disease. Lancet **397,** 2212-2224
https://doi.org/10.1016/S0140-6736(20)32511-3

Rigatti SJ (2017): Random Forest. J. Insur. Med. **47,** 31-39
https://doi.org/10.17849/insm-47-01-31-39.1

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. **43,** e47
https://doi.org/10.1093/nar/gkv007

Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, et al. (2010): GeneCards Version 3: the human gene integrator. Database (Oxford) **2010,** baq020
https://doi.org/10.1093/database/baq020

Savari F, Badavi M, Rezaie A, Gharib-Naseri MK, Mard SA (2019): Evaluation of the therapeutic potential effect of Fas receptor gene knockdown in experimental model of non-alcoholic steatohepatitis. Free Radic. Res. **53,** 486-496
https://doi.org/10.1080/10715762.2019.1608982

Speiser JL, Miller ME, Tooze J, Ip E (2019): A comparison of random forest variable selection methods for classification prediction modeling. Expert. Syst. Appl. **134,** 93-101
https://doi.org/10.1016/j.eswa.2019.05.028

Van Herck MA, Vonghia L, Francque SM (2017): Animal models of nonalcoholic fatty liver disease-a starter's guide. Nutrients **9**
https://doi.org/10.3390/nu9101072

Walter P, Ron D (2011): The unfolded protein response: from stress pathway to homeostatic regulation. Science **334,** 1081-1086
https://doi.org/10.1126/science.1209038

Wang M, Kaufman RJ (2016): Protein misfolding in the endoplasmic reticulum as a conduit to human disease. Nature **529,** 326-335
https://doi.org/10.1038/nature17041

Wang S, Chen Z, Lam V, Han J, Hassler J, Finck BN, Davidson NO, Kaufman RJ (2012): IRE1alpha-XBP1s induces PDI expression to increase MTP activity for hepatic VLDL assembly and lipid homeostasis. Cell. Metab. **16,** 473-486
https://doi.org/10.1016/j.cmet.2012.09.003

Wen W, Wu P, Zhang Y, Chen Z, Sun J, Chen H (2021): Comprehensive analysis of NAFLD and the therapeutic target identified. Front. Cell. Dev. Biol. **9,** 704704
https://doi.org/10.3389/fcell.2021.704704

Weng Y, DiRusso CC, Reilly AA, Black PN, Ding X (2005): Hepatic gene expression changes in mouse models with liver-specific deletion or global suppression of the NADPH-cytochrome P450 reductase gene. Mechanistic implications for the regulation of microsomal cytochrome P450 and the fatty liver phenotype. J. Biol. Chem. **280,** 31686-31698
https://doi.org/10.1074/jbc.M504447200

Wilkerson MD, Hayes DN (2010): ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics **26,** 1572-1573
https://doi.org/10.1093/bioinformatics/btq170

Wozniak AL, Wang X, Stieren ES, Scarbrough SG, Elferink CJ, Boehning D (2006): Requirement of biphasic calcium release from the endoplasmic reticulum for Fas-mediated apoptosis. J. Cell. Biol. **175,** 709-714
https://doi.org/10.1083/jcb.200608035

Xanthakos SA, Jenkins TM, Kleiner DE, Boyce TW, Mourya R, Karns R, Brandt ML, Harmon CM, Helmrath MA, Michalsky MP, et al. (2015): High prevalence of nonalcoholic fatty liver disease in adolescents undergoing bariatric surgery. Gastroenterology **149,** 623-634
https://doi.org/10.1053/j.gastro.2015.05.039

Yazici D, Sezer H (2017): Insulin resistance, obesity and lipotoxicity. Adv. Exp. Med. Biol. **960,** 277-304
https://doi.org/10.1007/978-3-319-48382-5_12

Ye J, Lin Y, Wang Q, Li Y, Zhao Y, Chen L, Wu Q, Xu C, Zhou C, Sun Y, et al. (2020): Integrated multichip analysis identifies potential key genes in the pathogenesis of nonalcoholic steatohepatitis. Front. Endocrinol. (Lausanne), **11,** 601745
https://doi.org/10.3389/fendo.2020.601745

Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M (2016): Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology **64,** 73-84
https://doi.org/10.1002/hep.28431

Yu G, Wang LG, Han Y, He QY (2012): clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS **16,** 284-287
https://doi.org/10.1089/omi.2011.0118

Zhang F, Zhang Z, Li Y, Sun Y, Zhou X, Chen X, Sun S (2022): Integrated bioinformatics analysis identifies robust biomarkers and its correlation with immune microenvironment in nonalcoholic fatty liver disease. Front. Genet. **13,** 942153

https://doi.org/10.3389/fgene.2022.942153

Zhang L, Kaizuka Y, Hanagata N (2012): Imaging of Fas-FasL membrane microdomains during apoptosis in a reconstituted cell-cell junction. Biochem. Biophys. Res. Commun. **422,** 298-304
https://doi.org/10.1016/j.bbrc.2012.04.152

The role of endoplasmic reticulum stress-related genes in the diagnosis and subtyping of non-alcoholic fatty liver disease

Zihao Guo[1], Xiaoxiao Yu[1], Zhihao Fang[1], Kai Yang[1], Changxu Liu[1], Zhichao Dong[1] and Chang Liu[1]

[1] *Department of General Surgery, The Fourth Hospital of Harbin Medical University, Harbin, Heilongjiang, China*
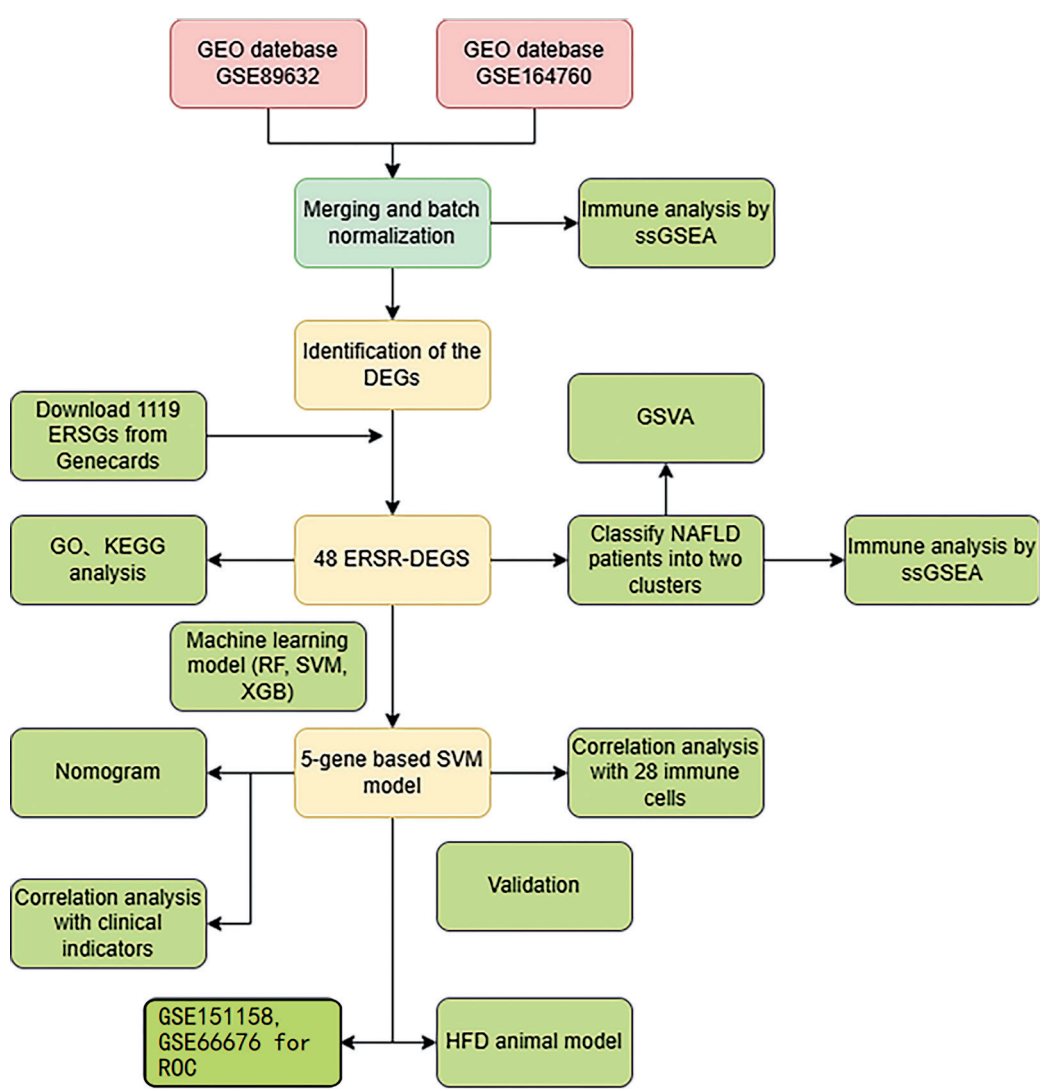
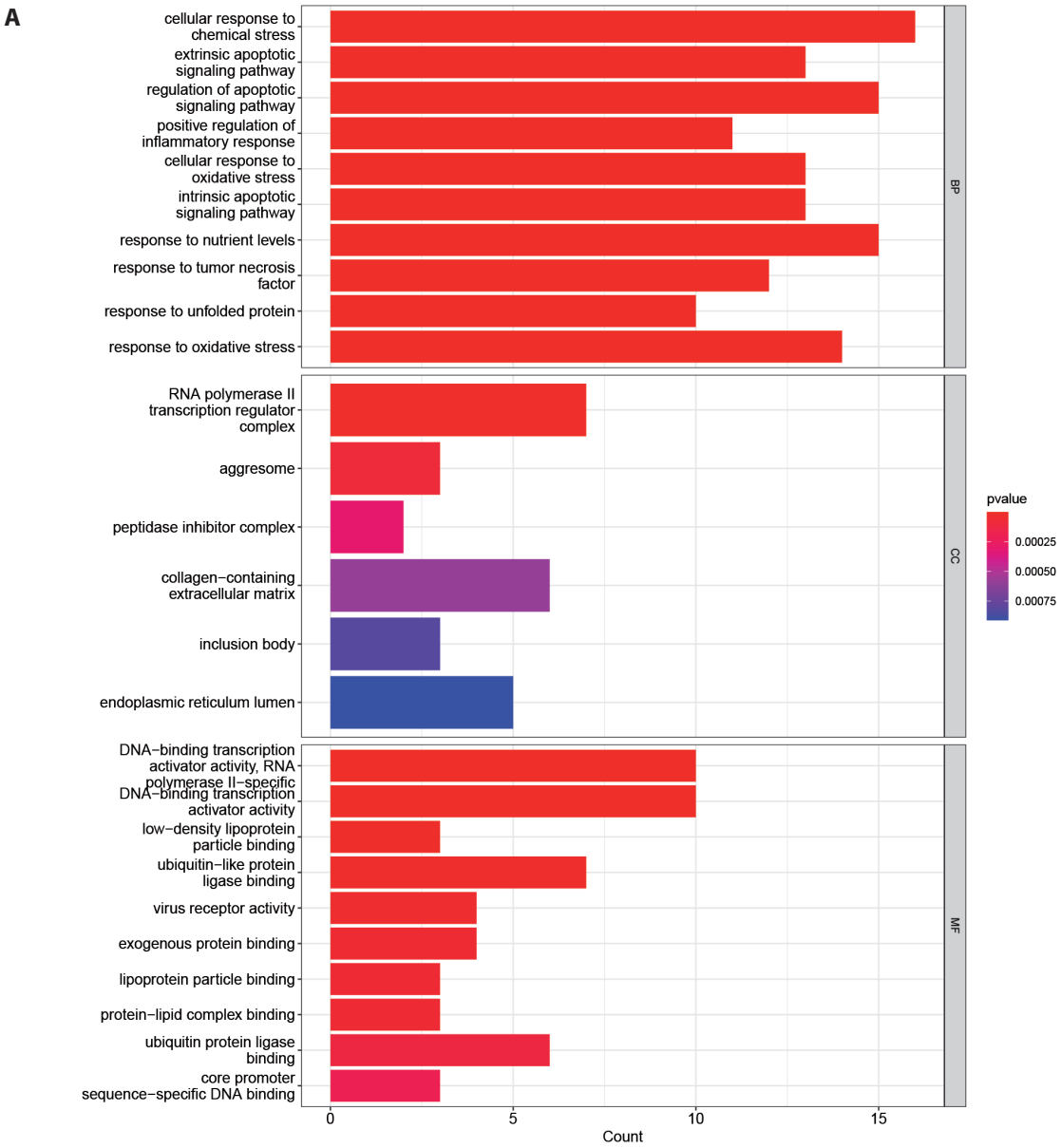**Supplementary Figures**

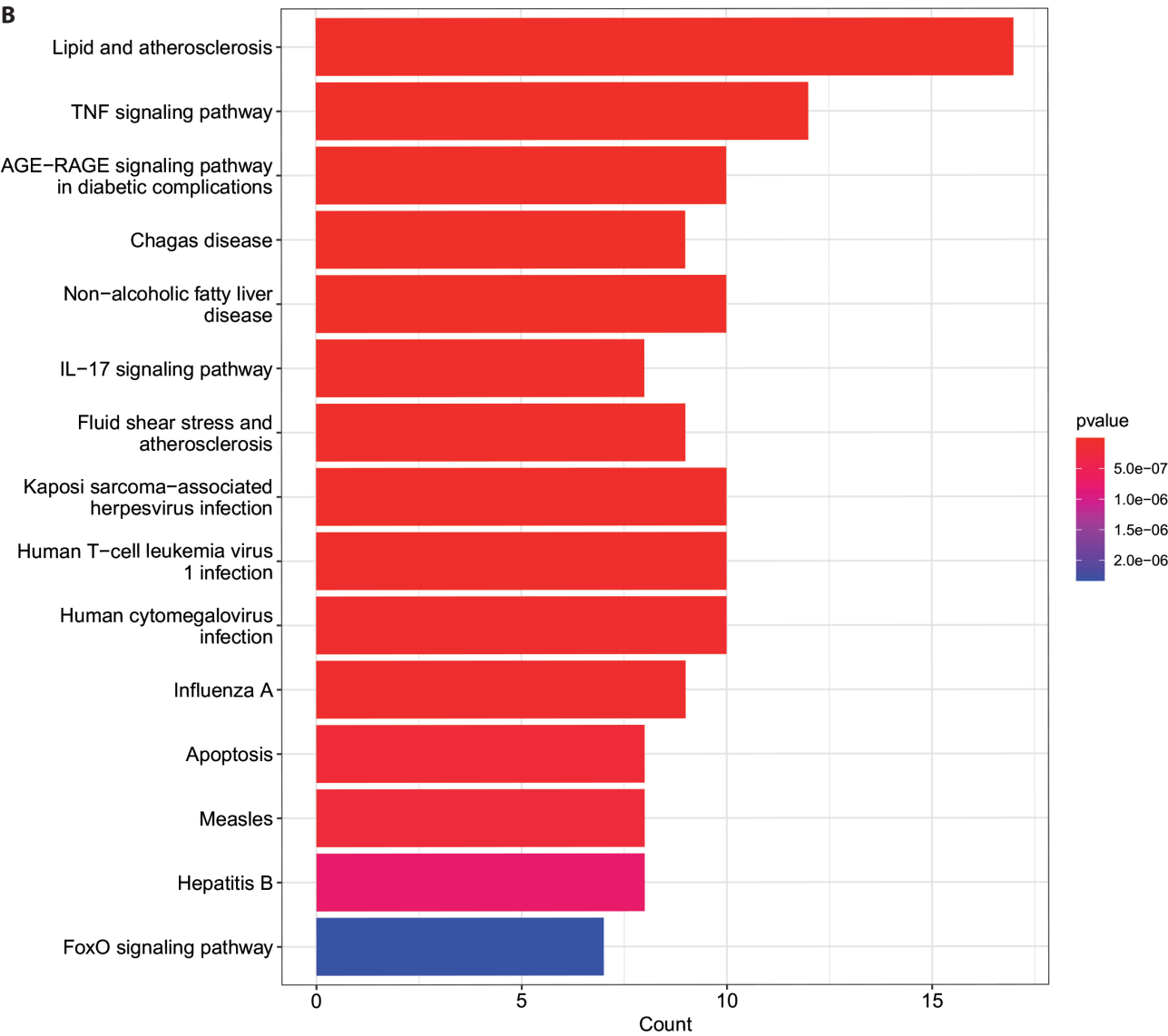**Figure S1**. Flow chart.

A

Figure S2. Functional enrichment analysis of the 48 ERSR-DEGs. **A.** Histogram of GO analysis for ERSR-DEGs, including biological processes, cellular components, and molecular functions. **B.** Bubble diagram of KEGG analysis for ERSR-DEGs. GO, Gene Ontology; BP, Biological process; CC, Cellular component; MF, Molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes.
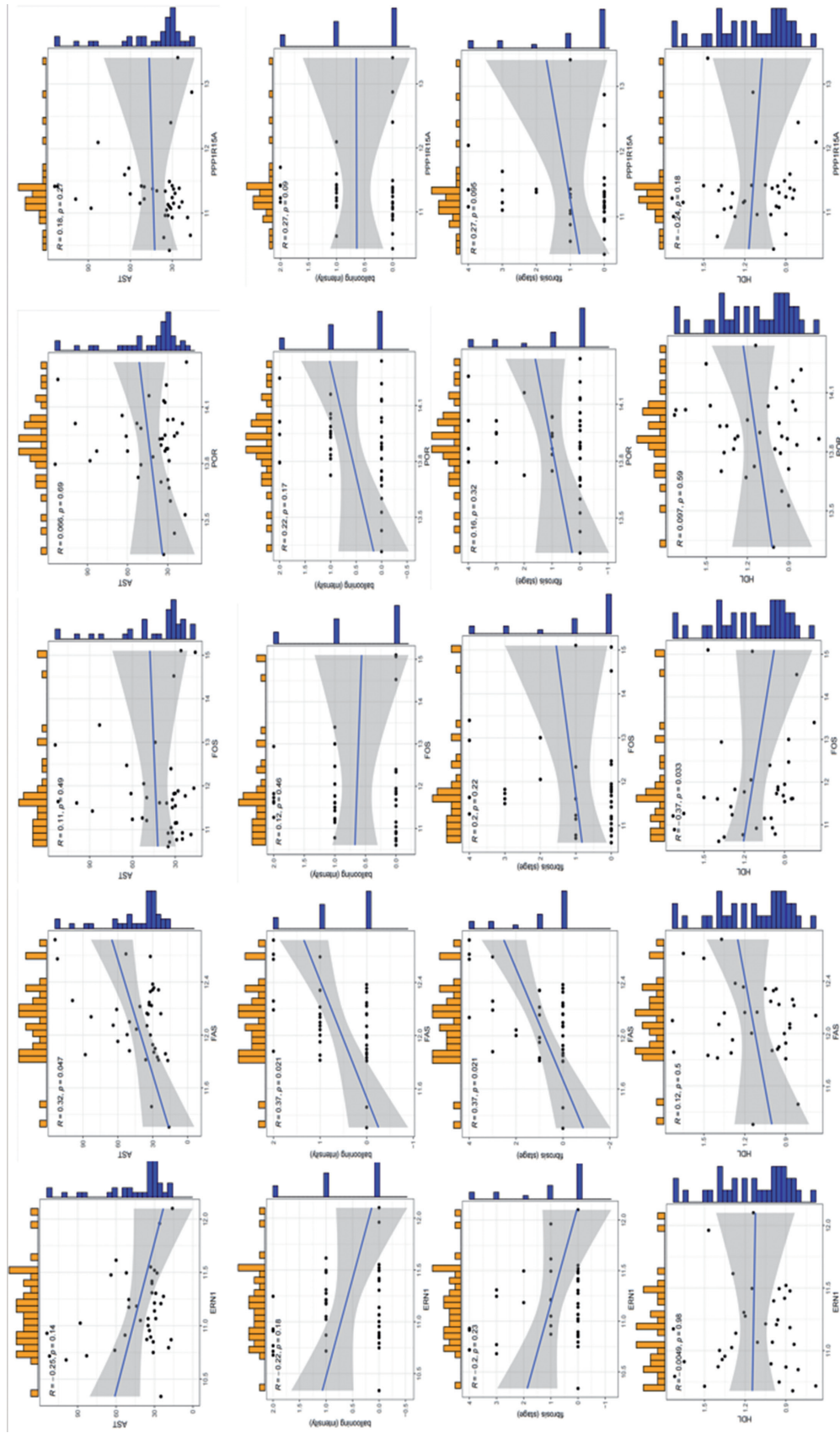
Figure S3. Correlation of genes with clinical indicators. The correlation of five hub genes with AST (**A**), ballooning (intensity) (**B**), fibrosis (stage) (**C**) and HDL (**D**).

Table S1. Primer information in study

Gene	Forward primer	Reverse primer
ERN1	ATCTCAGGATGTGGAGGAGAA	CAGGAAGGTGCTCAGGATAATG
POR	AGATCGACAAGTCCCTGGTA	CTCCTGCAGCCAATCATAGAA
PPP1R15A	AGAGGACACAGAGGAAGAAGA	TTCAGGAAGGCACTTGTATGG
FOS	GGTGAAGACCGTGTCAGGAG	AGTTGATCTGTCTCCGCTTGG
FAS	TGTCAGCCTGGTGAACGAAA	CTTGGTATTCTGGGTCCGGG
FASN	CACAGTGCTCAAAGGACATGCC	CACCAGGTGTAGTGCCTTCCTC
ACC1	GCTGCTCGGATCACTAGTGAA	TTCTGCTATCAGTCTGTCCAG
SCD1	CACACCTTCCCCTTCGACTA	TGACTCCCGTCTCCAGTTCT
β-actin	GTCGTACCACAGGCATTGTGATGG	GCAATGCCTGGGTACATGGTGG