doi: 10.4149/gpb_2024017

Bioinformatics analysis of potential ferroptosis and non-alcoholic fatty liver disease biomarkers

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

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

Abstract. Ferroptosis plays a crucial role in the development of non-alcoholic fatty liver disease (NAFLD). In this study, we aimed to use a comprehensive bioinformatics approach and experimental validation to identify and verify potential ferroptosis-related genes in NAFLD. We downloaded the microarray datasets for screening differentially expressed genes (DEGs) and identified the intersection of these datasets with ferroptosis-related DEGs from the Ferroptosis database. Subsequently, ferroptosis-related DEGs were obtained using SVM analysis; the LASSO algorithm was then used to identify six marker genes. Furthermore, the CIBERSORT algorithm was used to estimate the proportion of different types of immune cells. Subsequently, we constructed drug regulatory networks and ceRNA regulatory networks. We identified six genes as marker genes for NAFLD, demonstrating their robust diagnostic abilities. Subsequent functional enrichment analysis results revealed that these marker genes were associated with multiple diseases and play a key role in NAFLD *via* the regulation of immune response and amino acid metabolism, among other pathways. The expression of hepatic EGR1, IL-6, SOCS1, and NR4A1 was significantly downregulated in the NAFLD model. Our findings provide new insights and molecular clues for understanding and treating NAFLD. Further studies are needed to assess the diagnostic potential of these markers for NAFLD.

Key words: Bioinformatics analysis — Non-alcoholic fatty liver disease — Ferroptosis — Biomarkers

Introduction

Non-alcoholic fatty liver disease (NAFLD) is usually diagnosed through exclusion criteria, including alcohol use, viral infections, and other liver-damaging factors. It manifests as an accumulation of excessive fat and is often associated with metabolic dysfunctions, including insulin resistance and hyperlipidemia (Targher et al. 2021). The majority of hepatic lipid metabolism abnormalities stem from imbalances in lipid synthesis and metabolism, resulting in the excessive accumulation of lipids and subsequent dysfunction of hepatocytes. NAFLD comprises two distinct conditions: simple hepatic steatosis and non-alcoholic steatohepatitis (NASH). Left untreated, these conditions may advance to cirrhosis and liver cancer, underscoring NAFLD's significance as a critical public health issue. This prevalent chronic liver disorder affects a quarter of the global population, with its incidence escalating annually (Younossi et al. 2018). According to a mathematical model, the epidemic of obesity and diabetes is expected to increase the incidence of NAFLD, and the death rates from associated illnesses will double (Estes et al. 2018). Two concepts, namely "second hit" and "multiple hits," elucidate the pathophysiology of NAFLD.

* These authors contributed equally to this work.

Electronic Supplementary material. The online version of this article (doi: 10.4149/gpb_2024017) contains Supplementary material.

Correspondence to: Chang Liu, Department of General Surgery, Fourth Affiliated Hospital of Harbin Medical University, No.37 Yiyuan Street, Nangang District, Harbin, Heilongjiang Province, 150001, China
E-mail: changliu72@163.com

Current comprehension indicates that the onset of NAFLD is influenced by genetic and epigenetic factors, mitochondrial dysfunction, endoplasmic reticulum stress, and insulin resistance (Chen et al. 2020). Nevertheless, the absence of sensitive and specific biomarkers for early clinical detection and treatment of NAFLD presents a significant challenge. Consequently, investigating the mechanisms underlying NAFLD is imperative for identifying prospective therapeutic targets and devising innovative treatment approaches.

Ferroptosis is a recently discovered form of programmed cell death triggered by iron-mediated lipid peroxidation (Li J et al. 2020). Hepatocyte lipid accumulation-induced programmed cell death can lead to liver tissue damage and inflammation (Manne et al. 2018). Ferroptosis could significantly activate the inflammatory response in NASH. Iron overload is prevalent among NAFLD patients, with iron-induced lipid peroxidation playing a substantial role in NAFLD pathogenesis (Gao et al. 2021). Malondialdehyde and 4-hydroxynonenal, the secondary products of lipid peroxidation, have been demonstrated in multiple studies to serve as markers of oxidative stress in individuals with NAFLD (He et al. 2023). Antioxidants that inhibit lipid peroxidation effectively reduce serum aminotransferases in patients with NAFLD (Violi and Cangemi, 2010). Furthermore, regulators of iron metabolism show a substantial elevation in NAFLD mice induced by the methione-choline deficient (MCD) diet. Administration of ferroptosis inhibitors (Ferrostatin-1 and Liproxstain-1) effectively attenuated liver injury, inflammation, and fibrosis in MCD diet-fed mice (Li X et al. 2020). Hepatic ferroptosis was shown to occur before apoptosis in hepatocytes (Tsurusaki et al. 2019; Zhang et al. 2021), suggesting that it may be a promising therapeutic target for treating or preventing NAFLD. Hence, we employed bioinformatics analysis to investigate and validate the effectiveness of genes linked to ferroptosis as biomarkers for NAFLD, along with their roles in the hepatic immune system.

Material and Methods

Data sources

In this study, the microarray datasets of NAFLD and normal samples were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). The GSE89632 and GSE63067 datasets, which contained 31 healthy control liver samples and 50 NAFLD liver samples, were obtained from the GPL14951 and GPL570 platforms (Arendt et al. 2015; Frades et al. 2015), ferroptosis-related genes (FRGs), which include 369 driver genes, 348 suppressor genes, and 11 marker genes, were discovered in the FerrDb database (http://www.zhounan.org/ferrdb) and the detailed genes were shown in Supplementary material (Table S1) (Zhou and Bao 2020). The drug-gene interaction database (DGIdb) (https://dgidb.org/) was employed for drug-targeting marker gene prediction (Wagner et al. 2016).

Data processing and analysis of expression differences

Initially, the probes undergo annotation using the dataset's annotation file. In cases where multiple queries correspond to the same gene, the average expression value is adopted as the gene expression value following ID transformation. GSE89632 and GSE63067 were amalgamated as the training set for analysis in the main body of this study, and batch correction of the dataset was performed using the "SVA" package in R software (Leek et al. 2012). PCA plots were used to illustrate the batch effect of the normalized dataset. In addition, the limma package was used for differential analysis to identify differentially expressed genes (DEGs) (Table S2) based on the screening criteria $|\log FC| > 1$ and $p.adj < 0.05$ (Colaprico et al. 2016), and intersected with FRGs to obtain differentially expressed ferroptosis genes (DE-FRGs) for subsequent analysis.

Functional analyses for the DE-FRGs

Gene ontology (GO) functional enrichment analysis comprises three components: biological process (BP), cellular component (CC), and molecular function (MF). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis is widely employed to elucidate biological mechanisms and pathways. The "clusterProfiler" R package (version 3.16.1) was utilized to conduct GO and KEGG pathway enrichment analyses (Yu et al. 2012). Results with significant differential enrichment were filtered out using p-value < 0.05 and p-adj < 0.01 as a criterion.

Identification of optimal diagnostic gene biomarkers for NAFLD

The least absolute shrinkage and selection operator (LASSO) with 10-fold cross-validation and support vector machine-recursive feature elimination (SVM-RFE) were employed to discover supplementary gene markers for diagnosing NAFLD. The LASSO regression algorithm, a renowned linear prediction technique, predicts outcomes by utilizing regression coefficients (Motamedi et al. 2022). SVM-RFE is a machine learning method rooted in support vector machines, which identifies optimal core genes through iterative elimination of feature vectors generated by the support vector machine (Zhao et al. 2022). The SVM algorithm projects input data into a higher-dimensional feature space by mapping a kernel function, thus facilitating classification compared with the original feature space (Uddin et al.

2019). Our objective was to ascertain the optimal diagnostic gene markers for NAFLD by amalgamating the biomarkers acquired from both algorithms. Additionally, we conducted receiver operating characteristic (ROC) curve analysis on the training set utilizing the "pROC" package (Chai et al. 2023). The area under the curve (AUC) was calculated to validate the diagnostic effectiveness of the primary gene marker. Furthermore, our objective was to build a logistic regression model leveraging the predictive capabilities of the GLM package in R, incorporating six marker genes. This model was utilized to predict sample types within the training set and evaluate diagnostic performance using the ROC curve.

Single-gene Gene Set Enrichment Analysis (GSEA) enrichment analysis

The R package GSEA tool was utilized to calculate correlations between the six marker genes and all other genes in the training set, enabling further exploration of the pathways correlated with the marker genes (Subramanian et al. 2005). The KEGG signaling pathway set was used as a preset set to find its abundance in the gene set. Each marker gene's specific enrichment findings were merged into Table S3.

Single-gene Gene Set Variation Analysis (GSVA) enrichment analysis

GSVA represents an unsupervised and nonparametric approach to gene set enrichment analysis that estimates the score attributed to a particular pathway or signature based on transcriptomic data (Hänzelmann et al. 2013). In the current study, we conducted GSVA analysis for each marker gene utilizing the KEGG pathway set as the background gene set. We downloaded the "c2.cp.kegg.symbols" file from GSVA's MSigDB database (Qin et al. 2023). Additionally, the R package "limma" was employed to assess the differences in GSVA scores among samples categorized by high and low expression levels of the marker genes. For cases where $t > 0$, activation of the pathway was assumed in the high-expression group and conversely, if $t < 0$, initiation of the pathway was presumed in the low-expression group.

Immune infiltration analysis

As previous studies have done, the CIBERSORT algorithm calculates the proportion of different immune cell types based on the expression levels of immune cell-related genes (Yu et al. 2023; Zhao et al. 2023). The output of the 22 infiltrated immune cells was integrated to generate a matrix of immune cell fractions for analysis (Liu et al. 2024). The CIBERSORT program package was employed to assess the proportion of 22 immune cell types in liver samples from the training set in this study. Each sample's total proportion of the 22 immune cell types equaled 1, with $p < 0.05$ indicating a significant correlation (Table S4). Bar graphs illustrate the distribution of the 22 immune cell types across different samples, and the vioplot was utilized to visualize differences between NAFLD and normal immune cell groups. A heat map depicting the correlation between marker genes and the 22 immune cell types was generated using the "corrplot" R package (Hu 2020).

Construction of ceRNA network

Targetscan, MiRanda, and MiRDB databases were used to predict mRNA-miRNA couples based on six marker genes, and only miRNAs predicted by all three databases concurrently were kept (John et al. 2004; McGeary et al. 2019; Chen and Wang 2020). A ceRNA network of mRNA-miRNA-lncRNA was created by searching the anticipated miRNAs in the spongeScan database and screening the miRNA-lncRNA pairings (Furió-Tarí et al. 2016). Finally, we used Cytoscape software to visualize the ceRNA network (Doncheva et al. 2019).

Establishment of the mouse NAFLD model

A total of twenty C57BL/6 mice (males, six weeks old) were purchased from Liaoning Changsheng biotechnology Co. Ltd (Benxi, China). The mice were housed in a temperature-controlled chamber (22 ± 2°C) with a 12-hour light/dark cycle. Twenty male C57BL/6 mice, aged six weeks, were randomly assigned to either a high-fat diet (HFD) group or a control (CON) group, each consisting of 10 mice. Mice in the CON group were fed a standard chow diet for approximately 20 weeks, with 12% of calories from fat, 29% from protein, and 59% from carbohydrates. Mice in the HFD group were fed a diet containing 60% fat (d12492, medicience, Jiangsu, China) for about 20 weeks before euthanasia and sampling. After 20 weeks, all mice were anesthetized with 2% isoflurane, followed by euthanasia *via* cervical dislocation, with efforts made to minimize pain. Tissue samples were collected, fixed in 10% formalin for 24 hours, dehydrated, and embedded in paraffin. Tissue sections, 4 μm thick, were prepared, and hematoxylin and eosin (H&E) staining was performed. All methods were carried out in accordance with the Animal Care Guidelines of Harbin Medical University (2022-DWSYLLCZ-20), and all methods are reported in accordance with ARRIVE guidelines for the reporting of animal experiments.

Quantitative RT-PCR analysis

Total RNA was extracted from homogenized tissues using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). Subsequently, 1 μg of total RNA was reverse transcribed using

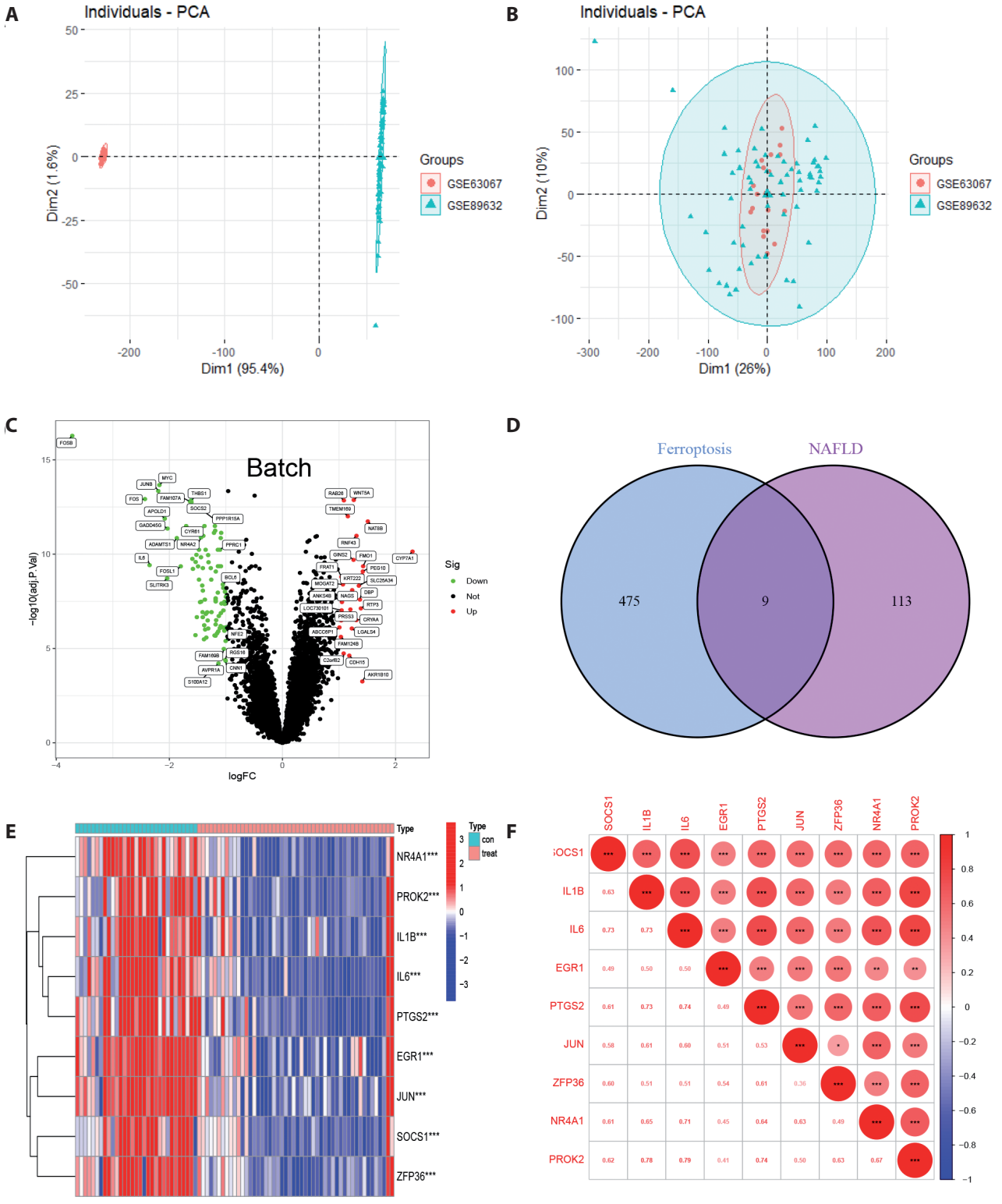

Figure 1. PCA plots illustrating batch effects between the data sets before (**A**) and after (**B**) normalization. **C.** After batch correction, we identified the DEGs in the training set, and in the volcano plot, up-regulated genes are shown in red and down-regulated genes are shown in green. **D.** DEGs were intersected with ferroptosis-related genes to obtain 9 ferroptosis-related differential genes (DE-FRGs). **E.** Heat map showing the expression pattern of NAFLD DE-FRGs in the sample. **F.** The correlation of 9 NAFLD DE-FRGs. They are all positively correlated with each other. (See online version for color figure.)

PrimeScript reverse transcriptase (Takara, Kusatsu, Japan), and 2*SYBR Green qPCR (Vazyme, Nanjing, China) was then performed for gene expression analysis. The reaction system was 10 μl, with 2 * SYBR Green qPCR 5 μl and primers (each) 0.2 μl. The PCR conditions were 95°C for 30 s, followed by 45 cycles of 95°C for 10 s, 60°C for 30 s. The primer sequences are listed in Table S5.

Statistical analysis

The experimental data were analyzed using GraphPad Prism 9.0. Results are expressed as mean ± standard deviation (SD). Student's *t*-test was employed to compare differences between the two groups, with significance set at $p < 0.05$.

Results

Identification of potential DE-FRGs in NAFLD

The GSE89632 and GSE63067 datasets were merged, normalized, and subjected to batch correction. In Figure 1A, principal component analysis (PCA) results before batch correction are presented for the two datasets, each represented by different colors. The plot indicates complete separation between the two datasets without overlap. Figure 1B displays PCA results after batch correction, revealing the successful elimination of the batch effect, thus enabling the combination of the two datasets for further analysis. According to the screening conditions of $|\log FC| > 1$ and $p.adj < 0.05$, 122 DEGs were significantly differentially expressed in NAFLD tissues compared to normal tissues, which were 28 up-regulated genes and 94 down-regulated genes. Among them, nine genes exhibited differential expression in NAFLD tissues compared to normal tissues and were identified as ferroptosis-related genes, all of which were downregulated (Fig. 1D). The expression patterns of DEGs and DE-FRGs in the samples were illustrated in the volcano plot and the clustering heatmap, respectively (Fig. 1C,E). Figure 1F displays the interactions between these genes. Strong positive correlations were observed between IL-6 and PTGS2, NR4A1, PROK2, IL-1B, and SOCS1. Furthermore, PTGS2 and PROK2 exhibited a significant positive correlation with IL-1B. Although the associations between EGR1, JUN, ZFP36, and other DE-FRGs were notable, the correlations were not particularly high.

Functional analyses for the DE-FRGs

GO and KEGG enrichment studies clarified the biological processes and pathways connected to NAFLD DE-FRGs. The results showed that BP enrichment analysis indicated that DE-FRGS was significantly associated with "response to lipopolysaccharide," "response to molecule of bacterial origin," and "fat cell differentiation." CC enrichment was associated with "cytoplasmic ribonucleoprotein granule," "ribonucleoprotein granule." Moreover, MF was associated with "DNA-binding transcription activator activity, RNA polymerase II-specific" and "DNA-binding transcription activator activity" (Fig. 2A,B). In particular, KEGG analysis showed significant enrichment in several immune-related pathways (IL-17 signaling pathway, C-type lectin receptor signaling pathway, TNF signaling pathway) (Fig. 2C,D).

Six DE-FRGs were identified as diagnostic genes for NAFLD

To evaluate the diagnostic potential of DE-FRGs and distinguish between individuals with NAFLD and healthy controls, the machine learning algorithms LASSO and SVM-RFE were employed to identify significant DE-FRGs. Following core gene filtration, 10-fold cross-validation of LASSO identified six genetic markers, as shown in Figures 3A and B. Subsequently, after identifying nine DE-FRGs, SVM-RFE determined the optimal combination of feature genes for NAFLD diagnosis (Fig. 3C). The intersection of marker genes from LASSO and SVM-RFE models led to the selection of six marker genes (EGR1, IL-6, JUN, NR4A1, SOCS1, and ZFP36) as diagnostic genes for NAFLD for further investigation (Fig. 3D).

Logistic regression models were developed using the R package GLM based on the aforementioned six characteristic genes. The ROC curves demonstrated that these models, utilizing the six marker genes, effectively discriminated between standard and NAFLD samples (AUC = 0.946; Fig. 3E). Additionally, ROC curves were constructed for these six marker genes to evaluate their capability in distinguishing between samples with NAFLD and those without it. As illustrated in Figure 3F, the AUCs of all six genes exceeded 0.8, indicating their high diagnostic accuracy. These findings suggest that the logistic regression model surpasses individual marker genes in terms of accuracy and specificity when differentiating NAFLD samples from control samples.

Marker genes were intimately connected to several NAFLD-related pathways

We conducted a single-gene GSEA-KEGG pathway analysis to further investigate the potential role of marker genes in distinguishing NAFLD samples from control samples. Figure S1 displays the top six pathways enriched for each marker gene. After comprehensive research, we found that these genes are involved in ribosomes, cell cycle, immune response (T-cell receptor signaling pathway, extracellular matrix-receptor interaction, and B-cell receptor signaling pathway), amino acid synthesis and metabolism (calcium, leucine, and isoleucine degradation, tyrosine metabolism, and amino acid

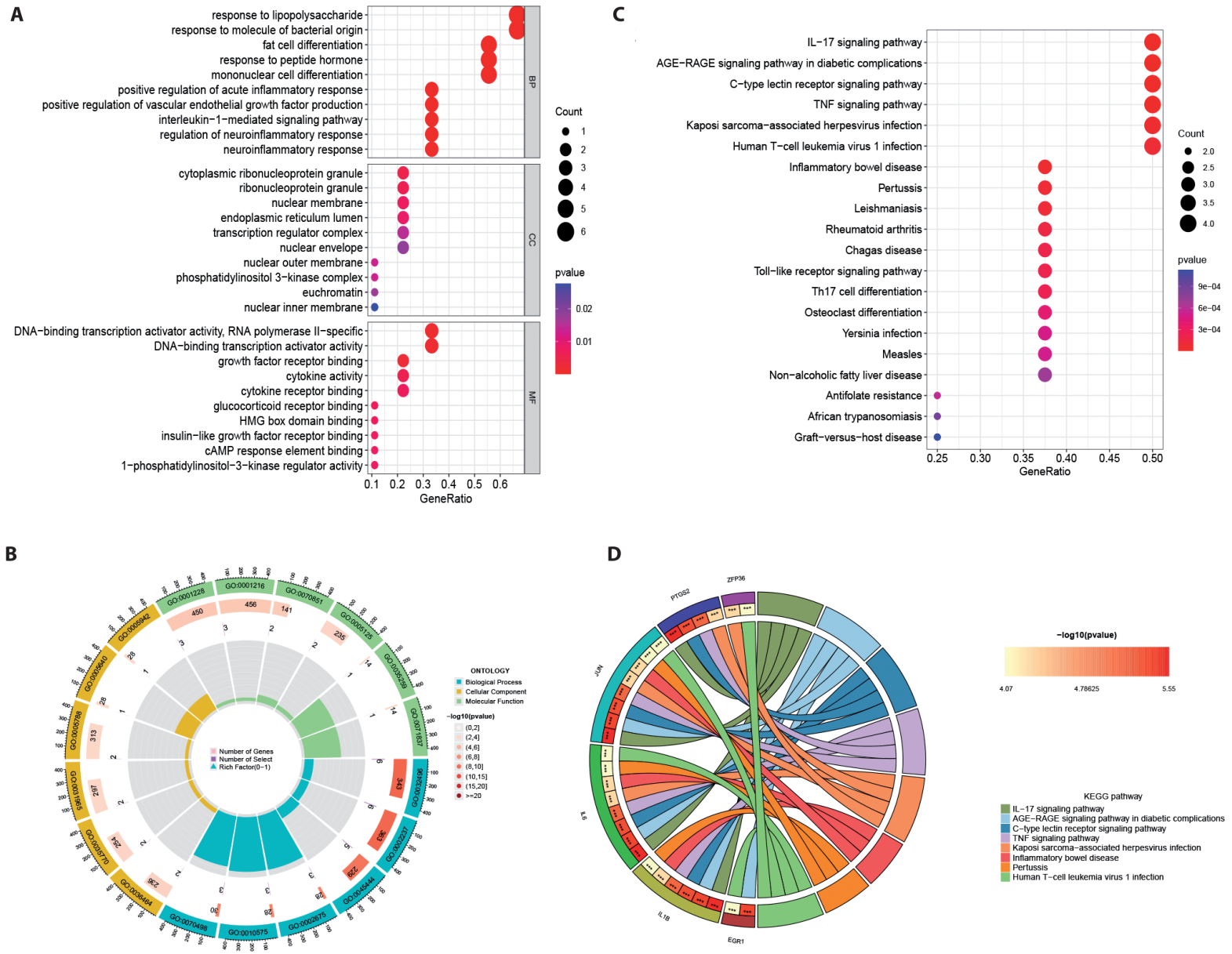

Figure 2. DE-FRG functional analyses. Circles (**A**) and barplots (**B**) were used to display the GO findings. Bubble (**C**) and circle (**D**) graphs were used to display the KEGG findings. (See online version for color figure.)

biosynthesis), and various disease pathways (type 2 diabetes, hypertrophic cardiomyopathy, dilated cardiomyopathy, tumors, and renal cell carcinoma) were abundant. Additionally, we observed that these marker genes were enriched in the "P53 signaling pathway," "MAPK signaling pathway," "NOD-like receptor signaling pathway," "chemokine signaling pathway," "JAK-STAT signaling pathway," "Toll-like receptor signaling pathway," and "Wnt signaling pathways.

We used GSVA to compare the activation pathways across groups with high and low expression levels based on the degree of expression of each marker gene. As shown in Figure S2, the low expression of EGR1 may be involved in the pathogenesis of NAFLD through immune response (extracellular matrix receptor interaction), "hematopoietic cell signaling," and "cytokine interaction." In contrast, its high expression is associated only with the "protein export" pathway. Upregulation of IL-6 activates pathways related to the cell cycle, amino acid metabolism, and DNA production, whereas its downregulation activates "hypertrophic cardiomyopathy," "right ventricular arrhythmogenic," and "myocardial contraction." In addition to EGR1 and IL-6, JUN, NR4A1, and SOCS1 were associated with the NOD-like receptor signaling pathway in the pathogenesis of NAFLD. Furthermore, low expression of IL-6, JUN, NR4A1, SOCS1, and ZFP36 was all directly related to the "JAK-STAT signaling pathway."

Immune landscape analysis

The immunological microenvironment and NAFLD are inextricably linked (Van Herck et al. 2019). Thus, using the CIBERSORT method, we investigated the variations in the immune microenvironment between NAFLD and normal samples. The results showed that the top three classes of immune cells in NAFLD tissues were M2 macrophages, γδ T cells, and resting-memory CD4 T cells, which accounted for approximately 50% of the immune cell infiltration in NAFLD (Fig. 4A). We observed that the infiltration of neutrophils, plasma cells, activated dendritic cells, activated mast cells, and naive B cells was considerably reduced in NAFLD tissues compared to normal tissues. Conversely, the infiltration of M1 and M2 macrophages, resting dendritic cells, and resting mast cells was significantly increased in NAFLD tissue (Fig. 4B).

Finally, we aimed to further study the relationship between our discovered marker genes and immune cells. To this end, we investigated the correlation between the expression of the six marker genes and the presence of 22 resistant cell types, as shown in Figure 4C. Regarding the expression of IL-6, NR4A1, SOCS1, and ZFP36, activated mast cells and resting mast cells had positive and negative correlations, respectively. Additionally, we found that the expression of IL-6 was positively and negatively associated with activated dendritic cells and M2 macrophages, respectively. Furthermore, SOCS1 was positively correlated with activated dendritic cells and monocytes but negatively correlated with γδ T cells. ZFP36 was negatively correlated with resting dendritic cells but positively correlated with monocytes and activated Mast cells. Finally, we observed that NR4A1 was positively correlated with monocytes.

Prediction of marker gene-targeted drugs

We utilized the DGIdb database to identify additional potential therapeutic targets for the marker genes and utilized Cytoscape tools to visualize the associations (Fig. S3). We assessed 53 medications targeting marker genes: 2 for SOCS1, 20 for IL-6, 1 for EGR1, 20 for JUN, and 10 for NR4A1. However, no drugs targeting ZFP36 were found in the DGIdb database. Among the 20 drugs targeting IL-6, six were IL-6 inhibitors. The relationship between marker genes and medications targeting them is presented in Table S6.

A ceRNA network based on marker genes

After that, we used the Targetscan, MiRanda, MiRDB, and spongeScan databases to construct a ceRNA network based on the six marker genes. In total, 290 edges and 248 nodes were present in the network containing six marker genes, 119 miRNAs, and 123 lncRNAs (Fig. S4). Our analysis revealed that 51 lncRNAs might competitively bind hsa-miR-766-3p, hsa-miR-149-3p, hsa-miR-561-3p, hsa-miR-324a-5p, and 11 other miRNA-regulated SOCS1. Moreover, hsa-miR-766-3p and hsa-miR-149-3p were found to be regulated by 39 lncRNAs. For JUN, we identified 30 lncRNAs that regulated JUN expression through competitive binding with hsa-miR-524-5p, hsa-miR-1972, hsa-miR-542-3p, hsa-miR-758-3p, and hsa-miR-940. Among them, lncRNA LINC00917, AC079586.1, and RP11-157B13.7 can target both hsamiR-542-3p and hsa-miR-758-3p. We also found that 20 lncRNAs can regulate JUN expression through the modulation of hsa-miR-515-5p, hsa-miR-760, hsa-miR 302a-5p, hsa-miR-149-5p, which in turn affects IL-6 expression. Among these, hsa-miR-515-5p expression was regulated by 15 lncRNAs. In the ceRNA network of NR4A1, eight and nine lncRNAs were identified to bind to hsa-miR-3425p and hsa-miR-665, respectively, to regulate the gene. In total, five lncRNAs were found to compete with hsa-miR-3787-3p for binding, leading to the regulation of EGR1 expression. Detailed information on the ceRNA network is provided in Table S7.

Expression of the marker gene in the NAFLD mouse model

To validate the reliability of the identified marker genes, NAFLD mouse models were induced by feeding C57BL/6 mice with an HFD for an additional 20 weeks. H&E stain-

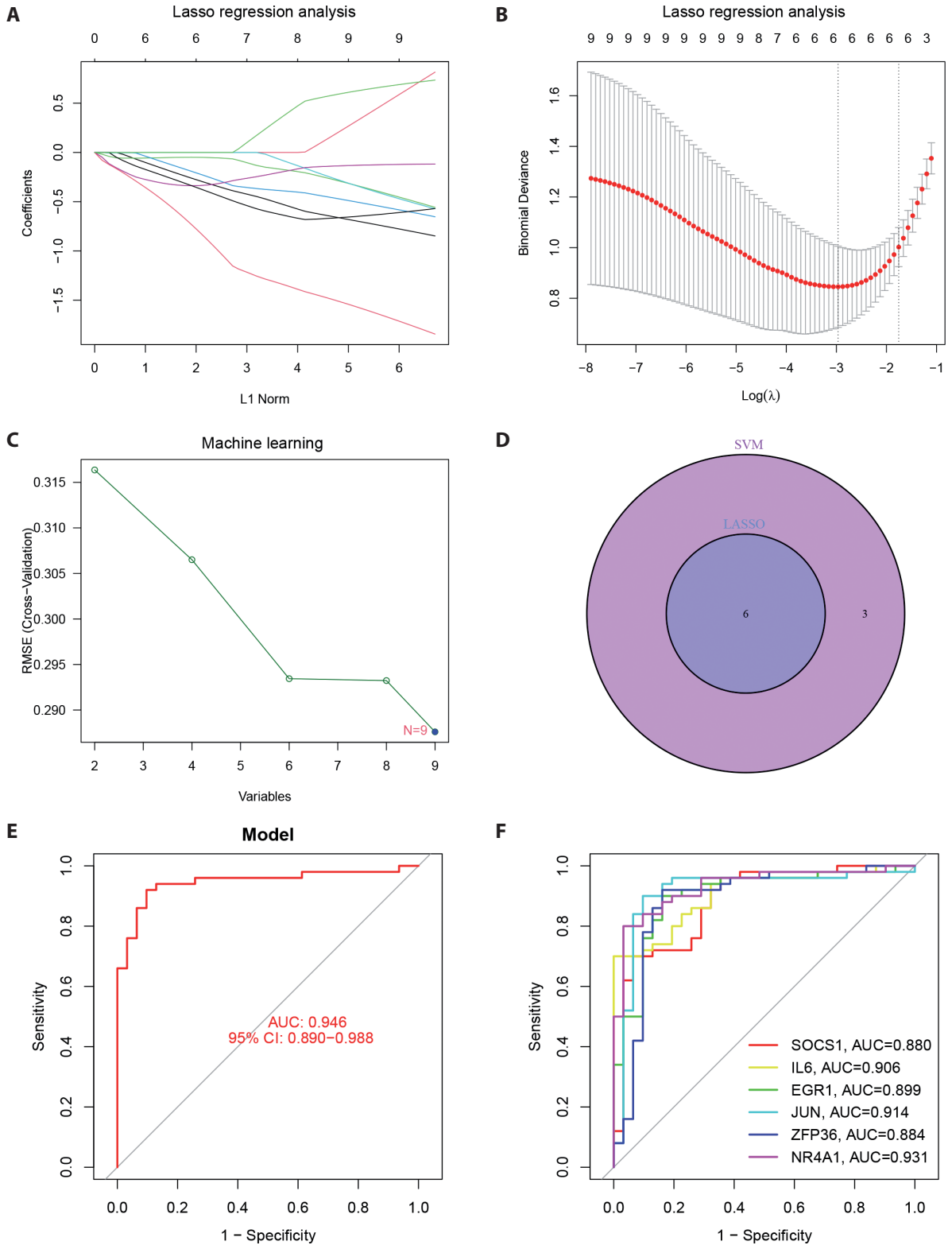

Figure 3. A. LASSO regression of the nine Ferroptosis-related DEGs. **B.** Partial likelihood deviance for the LASSO coefficient profiles. Six genes were selected at the value (lambda.min). **C.** SVM-RFE algorithm filtered 9 DE-FRGS and finally screened all nine genes as the best diagnostic genes. **D.** NAFLD marker genes were obtained from the LASSO model and the SVM-RFE model. **E.** Determination of logistic regression model for AUC of disease samples. **F.** ROC curves of 6 marker genes. (See online version for color figure.)

ing confirmed the establishment of NAFLD in the mouse models (Fig. 5A,B). Subsequently, we validated the marker genes in the NAFLD mouse model using qRT-PCR. The results indicated differential expression of the four marker genes in the livers of HFD-fed mice (Fig. 5C–F). Specifically, HFD-fed mice exhibited significantly lower hepatic expression levels of IL-6, EGR1, SOCS1, and NR4A1 compared to CON mice ($p < 0.05$).

Discussion

Hepatocyte death constitutes a pivotal mechanism in various liver injuries, encompassing multiple cell death pathways, including autophagy, pyroptosis, and programmed necrosis, alongside apoptosis and necrosis (Qian et al. 2021; Knorr et al. 2022). The liver is particularly susceptible to oxidative damage, with many liver disorders marked by an overabundance of iron accumulation (Wu et al. 2021). NAFLD is closely related to the mechanism of ferroptosis lipid peroxidation, and ferroptosis suppression has been shown to substantially alleviate NAFLD (Tsurusaki et al. 2019). Therefore, ferroptosis could be an essential target for treating and preventing NAFLD. In this study, we used bioinformatics approaches to analyze the genetic differences in ferroptosis between NAFLD and normal samples. We screened diagnostic markers, explored the molecular pathogenesis of ferroptosis in NAFLD, and validated these genes in an established NAFLD mouse model. Our findings offer valuable clinical insights for preventing and treating NAFLD.

In our study, we assessed six DEGs implicated in ferroptosis in NAFLD: EGR1, IL-6, JUN, SOCS1, ZFP36, and PTGS2. The AUC values of the ROC curves for these genes exceeded 0.8, indicating their reliable accuracy and specificity in distinguishing NAFLD cases from healthy samples. Due to the challenge of obtaining human liver samples, we induced NAFLD in C57BL/6 mice by feeding them a HFD to further validate the expression of these genes. This model closely mimics the development of human NAFLD and is commonly used in metabolic studies (Eng and Estall 2021). Our findings demonstrate a notable decrease in the expression levels of EGR1, IL-6, NR4A1, and SOCS1 in NAFLD mice compared to the control group, validating the forecasts generated by our analysis.

EGR1 is a transcription factor primarily involved in tissue damage, immune response, and fibrosis (Ma et al. 2023). By targeting and suppressing the expression of miR-15a-5p, the knockdown of EGR1 can increase the protein expression of GPX4 (Fan et al. 2021), which is the primary endogenous inhibitor of ferroptosis (Bersuker et al. 2019). This implies that low levels of EGR1 can contribute to the disease by preventing ferroptosis. Our findings indicate a downregulation of EGR1 expression in NAFLD mice, implying a potential organismal defense mechanism. It is conceivable that an EGR1/GPX4 axis, by modulating ferroptosis, contributes to the regulation of NAFLD development. SOCS1, a member of the cytokine signaling repressor protein family, functions as a negative feedback regulator, inhibiting cytokine signaling pathways within cells (Dai et al. 2023). It specifically inhibits the JAK/STAT signaling pathway by targeting unphosphorylated JAK and blocking JAK phosphorylation through the kinase inhibitory region of SOCS1 (Liau et al. 2018). Furthermore, GSVA analysis of SOCS1 confirmed the upregulation of the JAK/STAT signaling pathway in the low-expression group. SOCS1 plays a critical role in regulating ferroptosis in cancer cells, thereby indirectly enhancing cellular susceptibility to lipid oxidation and ferroptosis (Yan et al. 2023). Remarkably, our model demonstrated downregulation of SOCS1 expression, suggesting its role in inhibiting ferroptosis and its protective function in NAFLD. Additionally, in mice subjected to a high-fat diet, the knockdown of NR4A1 exacerbates insulin resistance and liver steatosis, whereas NR4A1 overexpression mitigates hepatic triglyceride accumulation (Sun et al. 2021). NR4A1 regulates tumor ferroptosis (Ye et al. 2021) and is consequently involved in both ferroptosis and NAFLD. The GSVA analysis of the ferroptosis-related gene NR4R1 showed upregulation of the "JAK-STAT signaling pathway" and "NOD-like signaling pathway" in the low-expression group. This observation suggests that NR4A1 plays a role in NAFLD development by suppressing these pathways. Additionally, IL-6 induces chondrocyte ferroptosis by triggering cellular oxidative stress and perturbing iron homeostasis (Bin et al. 2021). This is because IL-6 has isozyme-specific effects on GPX expression and can reduce its transcript concentration (Bin et al. 2021). Increased IL-6 expression in the presence of hepatic steatosis is associated with the severity of NAFLD as well as insulin resistance in the liver and adipose tissue (Cobbina and Akhlaghi 2017). Nevertheless, our analysis uncovered a reduction in IL-6 expression in NAFLD. This observation indicates the complexity of IL-6 involvement in NAFLD, with the decrease in IL-6 levels potentially attributed to excessive loss induced by ferroptosis in the liver.

The liver harbors numerous innate and adaptive immune cells pivotal for preserving immune homeostasis (Kubes and Jenne 2018), and immune dysregulation is implicated in NAFLD development. Consequently, we investigated immune infiltration in the livers of both NAFLD patients and healthy individuals employing the CIBERSORT algorithm. Our results revealed significantly elevated levels of M1 and M2 macrophages, as well as resting dendritic cells and mast cells, in the livers of NAFLD patients compared to those of healthy controls. Notably, M2 macrophages and γδ T cells predominated in liver tissue, underscoring their importance in the immunological milieu of NAFLD.

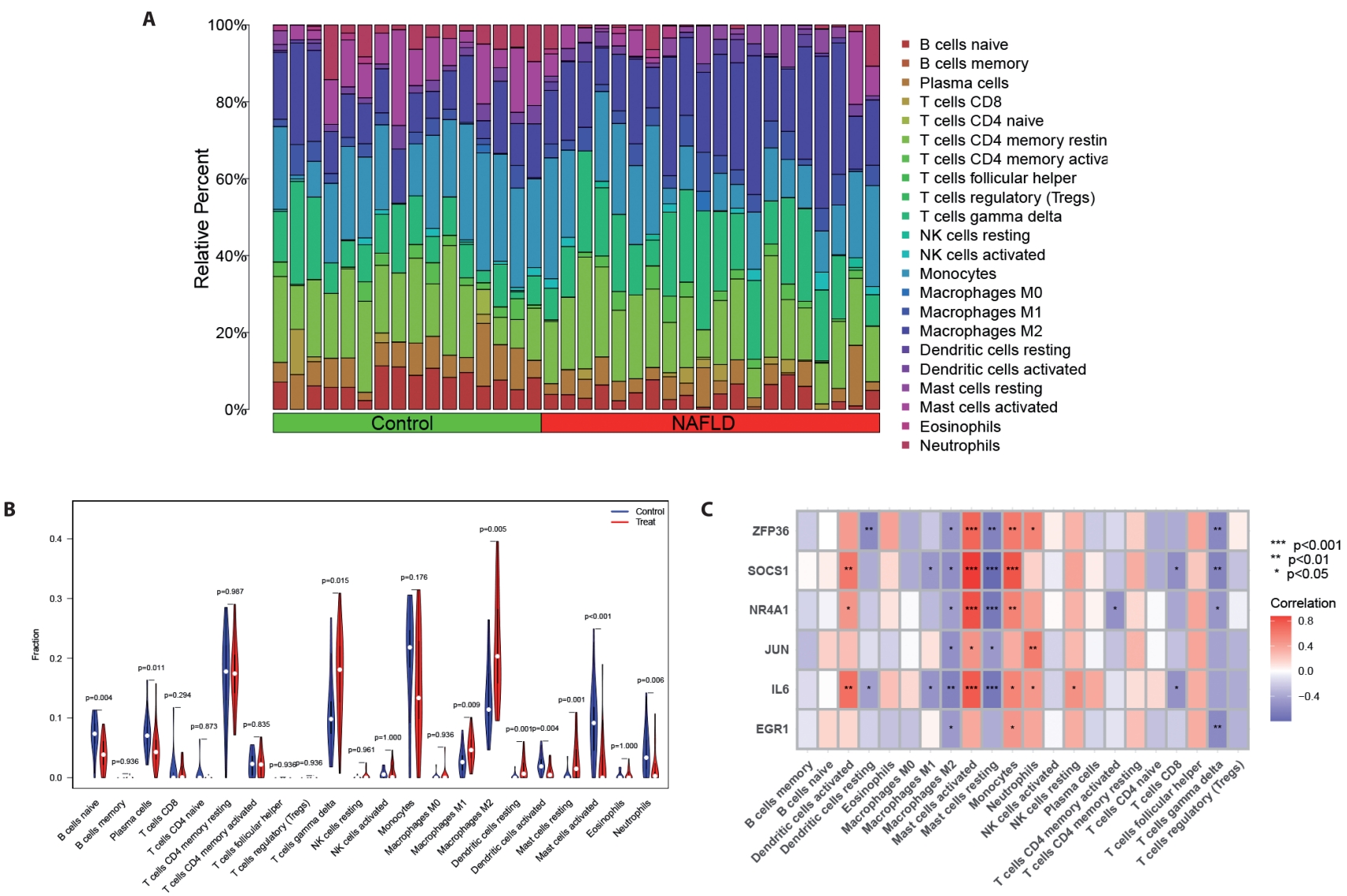

Figure 4. Immune infiltration in NAFLD. **A.** Percentage of 22 immune cell species in NAFLD and normal samples. **B.** CIBERSORT method was used to investigate the variations in the immune microenvironment between normal samples and NAFLD patients. **C.** Pearson correlation analysis demonstrates the association between immune cells and marker genes. Red indicates a positive correlation, blue is a negative correlation, and deeper colors have a more significant correlation. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (See online version for color figure.)

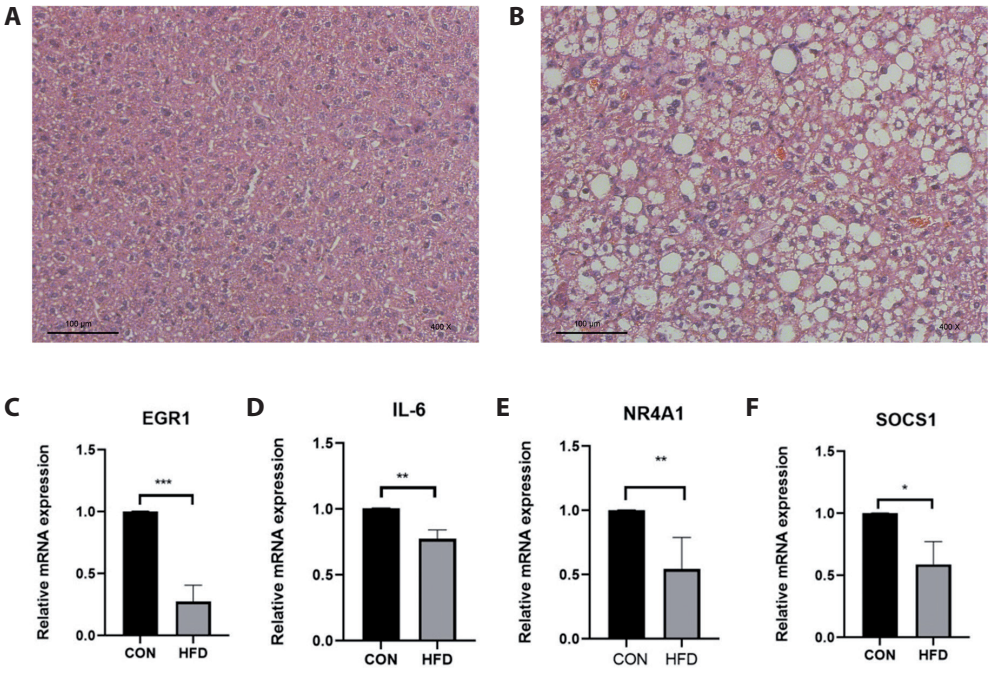

Figure 5. Marker genes expression in high-fat diet (HFD)-fed mice. **A,B.** Hematoxylin and eosin (H&E) staining of liver slices (magnification ×400, scale bar = 100 μm). EGR1 **(C)**, IL-6 **(D),** NR4A1 **(E)**, and SOCS1 **(F)** mRNA levels were considerably lower in the HFD group than in the CON group. Values are shown as the mean ± SD. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Simple steatosis triggers macrophage infiltration in the liver, with the macrophage count escalating alongside the severity of steatosis and liver inflammation (Barreby et al. 2022). Macrophage activation and polarization profoundly influence NAFLD progression (Kazankov et al. 2019). During the reparative phase of NASH, hepatic macrophages polarize towards an inflammatory and pro-fibrotic M2 phenotype, thereby fostering fibrosis (Kazankov et al. 2019). This observation suggests that while dysfunctional immune cells may exacerbate liver damage, an early inflammatory response could facilitate the recovery and regeneration of liver tissue (Ahmed et al. 2021). Nevertheless, our study revealed a notable abundance of hepatic M1 macrophages in NAFLD patients compared to normal individuals. This phenomenon may stem from the plasticity and adaptability of macrophages, allowing them to partially reverse their phenotype (Sica et al. 2014). γδ T cells represent a distinct subset of T lymphocytes crucial for host defense, immune surveillance, and immune system homeostasis. IL-17A, a pivotal pro-inflammatory cytokine in the liver, is predominantly secreted by γδ T cells (Mills 2023). The escalating presence of γδ T cells in the liver during NAFLD progression and exacerbation of steatohepatitis modulates CD4 T cells and augments IL-17 expression (Xi et al. 2019). This finding aligns with our study results, which unveiled a negative correlation between the expression levels of NAFLD ferroptosis-related genes and γδ T cells. KEGG enrichment analysis revealed the predominant enrichment of these genes in the IL-17 signaling pathway, which facilitates NAFLD progression by recruiting neutrophils and stimulating ROS production. Knockdown of the pivotal gene IL-17RA within this pathway impedes NAFLD progression (Harley et al. 2014). Consequently, γδ cells, responsible for IL-17A production, emerge as the principal regulatory cells in NAFLD progression, suggesting that diminishing the hepatic enrichment of γδ cells could offer an effective strategy to halt NAFLD advancement.

The expression levels of various ferroptosis-related genes exhibited consistent correlations with distinct immune cell populations. In our investigation, IL-6, SOCS1, and NR4A1 displayed a robust positive correlation with activated mast cells but a pronounced negative correlation with resting mast cells, while EGR1 showed a significant negative association with γδ T cells. Additionally, the expression of these genes was downregulated in NAFLD patients. These findings suggest that the altered immune microenvironment in NAFLD patients may be associated with the expression of these genes. We hypothesize that the regulation of immune cell aggregation by the ferroptosis-related genes EGR1, IL-6, SOCS1, and NR4A1 occurs through immunological or inflammatory pathways, thereby influencing the progression of NAFLD. Nevertheless, these conclusions rely primarily on the analysis of existing data, and additional clinical trials or animal studies are warranted to corroborate and substantiate our findings.

Our analysis of drugs targeting the marker and ceRNA networks revealed six out of the 20 drugs as IL-6 antagonists: siltuximab, clazakizumab, sirukumab, elsilimomab, pf-04236921, and olokizumab. Although these drugs have been utilized in conditions like rheumatoid arthritis, Crohn's

disease, and myeloma, their efficacy in NAFLD treatment remains unreported. Additionally, genipin, which targets SREBP-1c and reduces hyperlipidemia and hepatic lipid accumulation in mice, exhibits beneficial effects, such as enhancing pancreatic β-cell function in energy metabolism for the treatment of metabolic disorders (Wang et al. 2022). Non-coding RNAs, such as miR-34a-3p, miR-23a-3p, miR-200b, and miR-200c, play essential roles in lipid metabolism in NAFLD (Li et al. 2021; Xu et al. 2021; Goncalves et al. 2023). Nevertheless, the effectiveness of these targeted drugs and non-coding RNAs in treating NAFLD remains uncertain, necessitating additional research to elucidate their exact mechanisms. Consequently, these gene-targeting drugs and non-coding RNAs selected for study merit further exploration as prospective NAFLD treatments. However, this study has several limitations. Firstly, the reliance on mouse models instead of human samples may hinder the validation of disparities in NAFLD marker gene expression. Secondly, the sample size is relatively small; underscoring the need for larger sample sizes in future animal studies and clinical cohort investigations.

Conclusion

In conclusion, this study employed a comprehensive bioinformatics approach to investigate ferroptosis-related genes and their implications in NAFLD. The expression levels of EGR1, IL-6, SOCS1, and NR4A1 were found to be significantly diminished in an established NAFLD model. This observation implies their involvement in ferroptosis in NAFLD and potentially in modulating the immunological milieu within the livers of NAFLD patients. While further experiments are necessary to validate the molecular interactions between these ferroptosis-related genes and NAFLD, our findings are crucial for advancing our understanding of NAFLD pathophysiology and treatment. Consequently, we will continue to explore these genes in future research endeavors to enhance our comprehension of NAFLD management.

Conflict of interest. The authors report no conflicts of interest in this work.

Author contributions. XXY, KY designed study, XXY, KY and ZHF developed methodology, KY, ZHG and ZCD analyzed data, XXY, TTH and CXL wrote the manuscript and ZHF and CL revised the full text.

References

Ahmed O, Robinson MW, O'Farrelly C (2021): Inflammatory processes in the liver: divergent roles in homeostasis and pathology. Cell. Mol. Immunol. **18,** 1375-1386 https://doi.org/10.1038/s41423-021-00639-2

Arendt BM, Comelli EM, Ma DW, Lou W, Teterina A, Kim T, Fung SK, Wong DKH, McGilwray 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

Barreby E, Chen P, Aouadi M (2022): Macrophage functional diversity in NAFLD - more than inflammation. Nat. Rev. Endocrinol. **18,** 461-472 https://doi.org/10.1038/s41574-022-00675-6

Bersuker K, Hendricks JM, Li Z, Magtanong L, Ford B, Tang PH, Roerts MA, Tong B, Maimone TJ, Zoncu R (2019): The CoQ oxidoreductase FSP1 acts parallel to GPX4 to inhibit ferroptosis. Nature **575,** 688-692 https://doi.org/10.1038/s41586-019-1705-2

Bin S, Xin L, Lin Z, Jinhua Z, Rui G, Xiang Z (2021): Targeting miR-10a-5p/IL-6R axis for reducing IL-6-induced cartilage cell ferroptosis. Exp. Mol. Pathol. **118,** 104570 https://doi.org/10.1016/j.yexmp.2020.104570

Chai JL, Lu BW, Du HT, Wen MT, Liang XZ, Wang P (2023): Pyroptosis -related potential diagnostic biomarkers in steroid-induced osteonecrosis of the femoral head. BMC Musculoskelet. Disord. **24,** 609 https://doi.org/10.1186/s12891-023-06729-8

Chen Y, Wang X (2020): miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. **48,** D127-131 https://doi.org/10.1093/nar/gkz757

Chen Z, Tian R, She Z, Cai J, Li H (2020): Role of oxidative stress in the pathogenesis of nonalcoholic fatty liver disease. Free Radic. Biol. Med. **152,** 116-141 https://doi.org/10.1016/j.freeradbiomed.2020.02.025

Cobbina E, Akhlaghi F (2017): Non-alcoholic fatty liver disease (NAFLD) - pathogenesis classification and effect on drug metabolizing enzymes and transporters. Drug Metab. Rev. **49,** 197-211 https://doi.org/10.1080/03602532.2017.1293683

Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I (2016): TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. **44,** e71 https://doi.org/10.1093/nar/gkv1507

Dai L, Han Y, Yang Z, Zeng Y, Liang W, Shi Z, Tao Y, Liang X, Liu W, Zhou S (2023): Identification and validation of SOCS1/2/3/4 as potential prognostic biomarkers and correlate with immune infiltration in glioblastoma. J. Cell Mol. Med. **27,** 2194-2214 https://doi.org/10.1111/jcmm.17807

Doncheva NT, Morris JH, Gorodkin J, Jensen LJ (2019): Cytoscape stringApp: Network analysis and visualization of proteomics data. J. Proteome. Res. **18,** 623-632 https://doi.org/10.1021/acs.jproteome.8b00702

Eng JM, Estall JL (2021): Diet-induced models of non-alcoholic fatty liver disease: food for thought on sugar fat and cholesterol. Cells **10,** 1805 https://doi.org/10.3390/cells10071805

Estes C, Anstee QM, Arias-Loste MT, Bantel H, Bellentani S, Caballeria J, Colombo M, Craxi A, Crespo J, Day CP (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

Fan K, Huang W, Qi H, Song C, He C, Liu Y, Zhang Q, Wang L, Sun H (2021): The Egr-1/miR-15a-5p/GPX4 axis regulates ferroptosis in acute myocardial infarction. Eur. J. Pharmacol. **909,** 174403 https://doi.org/10.1016/j.ejphar.2021.174403

Frades I, Andreasson E, Mato JM, Alexandersson E, Matthiesen R, Martínez-Chantar ML (2015): Integrative genomic signatures of hepatocellular carcinoma derived from nonalcoholic fatty liver disease. PLoS One **10,** e0124544 https://doi.org/10.1371/journal.pone.0124544

Furió-Tarí P, Tarazona S, Gabaldón T, Enright AJ, Conesa A (2016): spongeScan: A web for detecting microRNA binding elements in lncRNA sequences. Nucleic Acids Res. **44,** 176-180 https://doi.org/10.1093/nar/gkw443

Gao G, Xie Z, Li E.W, Yuan Y, Fu Y, Wang P, Qiao Y, Xu J, Holscher C (2021): Dehydroabietic acid improves nonalcoholic fatty liver disease through activating the Keap1/Nrf2-ARE signaling pathway to reduce ferroptosis. J. Nat. Med. **75,** 540-552 https://doi.org/10.1007/s11418-021-01491-4

Goncalves BS, Meadows A, Pereira DG, Puri R, Pillai SS (2023): Insight into the inter-organ crosstalk and prognostic role of liver-derived microRNAs in metabolic disease progression. Biomedicines **11,** 1597 https://doi.org/10.3390/biomedicines11061597

Hänzelmann 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

Harley IT, Stankiewicz TE, Giles DA, Softic S, Flick LM, Cappelletti M, Sheridan R, Xanthakos SA, Steinbrecher KA, Balfour Sartor R (2014): IL-17 signaling accelerates the progression of non-alcoholic fatty liver disease in mice. Hepatology **5,** 1830-1839 https://doi.org/10.1002/hep.26746

He L, Wang J, Tao B, Zhu R, Li C, Ning B (2023): Identification of ferroptosis-related genes in the progress of NASH. Front. Endocrinol. **14,** 1184280 https://doi.org/10.3389/fendo.2023.1184280

Hu K (2020): Become competent within one day in generating boxplots and violin plots for a novice without prior r experience. Methods Protoc. **3,** 64 https://doi.org/10.3390/mps3040064

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004): Human microRNA targets. PLoS Biol. **2,** e363 https://doi.org/10.1371/journal.pbio.0020363

Kazankov K, Jørgensen SMD, Thomsen KL, Møller HJ, Vilstrup H, George J, Schuppan D, Gronbek H. (2019): The role of macrophages in nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Nat. Rev. Gastroenterol. Hepatol. **16,** 145-159 https://doi.org/10.1038/s41575-018-0082-x

Knorr J, Wree A, Feldstein AE (2022): Pyroptosis in steatohepatitis and liver diseases. J. Mol. Biol. **434,** 167271 https://doi.org/10.1016/j.jmb.2021.167271

Kubes P, Jenne C (2018): Immune responses in the liver. Annu. Rev. Immunol. **36,** 247-277 https://doi.org/10.1146/annurev-immunol-051116-052415

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

Li J, Cao F, Yin HL, Huang ZJ, Lin ZT, Mao N, Sun B, Wang G (2020): Ferroptosis: past present and future. Cell Death Dis. **11,** 88 https://doi.org/10.1038/s41419-020-2298-2

Li L, Zhang, Ren H, Huang X, Shen T, Tang W, Dou L, Li J (2021): miR-23a/b-3p promotes hepatic lipid accumulation by regulating Srebp-1c and Fas. J. Mol. Endocrinol. **68,** 35-49 https://doi.org/10.1530/JME-20-0324

Li X, Wang TX, Huang X, Li Y, Sun T, Zang S, Guan KL, Xiong Y, Liu J, Yuan HX (2020): Targeting ferroptosis alleviates methionine-choline deficient (MCD)-diet induced NASH by suppressing liver lipotoxicity. Liver Int. **40,** 1378-1394 https://doi.org/10.1111/liv.14428

Liau NPD, Laktyushin A, Lucet IS, Murphy JM, Yao S, Whitlock E, Callaghan K, Nicola NA, Kershaw NJ, Babon J (2018): The molecular basis of JAK/STAT inhibition by SOCS1. Nat. Commun. **9,** 1558 https://doi.org/10.1038/s41467-018-04013-1

Liu C, Fang Z, Yang K, Ji Y, Yu X, Guo Z, Dong Z, Zhu T, Liu C (2024): Identification and validation of cuproptosis-related molecular clusters in non-alcoholic fatty liver disease. J. Cell. Mol. Med. **28,** e18091 https://doi.org/10.1111/jcmm.18091

Ma ZG, Yuan YP, Fan D, Zhang, Teng T, Song P, Kong CY, Hu C, Wei WY, Tang QZ (2023): IRX2 regulates angiotensin II-induced cardiac fibrosis by transcriptionally activating EGR1 in male mice. Nat. Commun. **14,** 4967 https://doi.org/10.1038/s41467-023-40639-6

Manne V, Handa P, Kowdley KV (2018): Pathophysiology of non-alcoholic fatty liver disease/nonalcoholic steatohepatitis. Clin. Liver Dis. **22,** 23-37 https://doi.org/10.1016/j.cld.2017.08.007

McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, Bartel DP (2019): The biochemical basis of microRNA targeting efficacy. Science **366,** eaav1741 https://doi.org/10.1126/science.aav1741

Mills KHG (2023): IL-17 and IL-17-producing cells in protection versus pathology. Nat. Rev. Immunol. **23,** 38-54 https://doi.org/10.1038/s41577-022-00746-9

Motamedi F, Pérez-Sánchez H, Mehridehnavi A, Fassihi A, Ghasemi F (2022): Accelerating big data analysis through LASSO-random forest algorithm in QSAR studies. Bioinformatics **38,** 469-475 https://doi.org/10.1093/bioinformatics/btab659

Qian H, Chao X, Williams J, Fulte S, Li T, Yang L, Ding WX (2021): Autophagy in liver diseases: A review. Mol. Aspects Med. **82,** 100973 https://doi.org/10.1016/j.mam.2021.100973

Qin X, Yi S, Rong J, Lu H, Ji B, Zhang W, Ding R, Wu L, Chen Z (2023): Identification of anoikis-related genes classification patterns and immune infiltration characterization in ischemic stroke based on machine learning. Front. Aging Neurosci. **15,** 1142163 https://doi.org/10.3389/fnagi.2023.1142163

Sica A, Invernizzi P, Mantovani A (2014): Macrophage plasticity and polarization in liver homeostasis and pathology. Hepatology **59,** 2034-2042 https://doi.org/10.1002/hep.26754

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005): Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA **102,** 15545-15550 https://doi.org/10.1073/pnas.0506580102

Sun B, Zhang R, Liang Z, Fan A, Kang D (2021): Hyperoside attenuates non-alcoholic fatty liver disease through targeting Nr4A1 in macrophages. Int. Immunopharmacol. **94,** 107438 https://doi.org/10.1016/j.intimp.2021.107438

Targher G, Tilg H, Byrne CD (2021): Non-alcoholic fatty liver disease: a multisystem disease requiring a multidisciplinary and holistic approach. Lancet Gastroenterol. Hepatol. **6,** 578-588 https://doi.org/10.1016/S2468-1253(21)00020-0

Tsurusaki S, Tsuchiya Y, Koumura T, Nakasone M, Sakamoto T, Matsuoka M, Imai H, Kok CYY, Okochi H, Nakano H, et al. (2019): Hepatic ferroptosis plays an important role as the trigger for initiating inflammation in nonalcoholic steatohepatitis. Cell Death Dis. **10,** 449 https://doi.org/10.1038/s41419-019-1678-y

Uddin S, Khan A, Hossain ME, Moni MA (2019): Comparing different supervised machine learning algorithms for disease prediction. BMC Med. Inform. Decis. Mak. **19,** 281 https://doi.org/10.1186/s12911-019-1004-8

Van Herck MA, Weyler J, Kwanten WJ, Dirinck EL, De Winter BY, Francque SM, Vonghia L (2019): The differential roles of T cells in non-alcoholic fatty liver disease and obesity. Front. Immunol. **10,** 82 https://doi.org/10.3389/fimmu.2019.00082

Violi F, Cangemi R (2010): Pioglitazone vitamin E or placebo for nonalcoholic steatohepatitis. N. Engl. J. Med. **363,** 1185-1186 https://doi.org/10.1056/NEJMc1006581

Wagner AH, Coffman AC, Ainscough BJ, Spies NC, Skidmore ZL, Campbell KM, Krysiak K, Pan D, McMichael JF, et al. (2016): DGIdb 2.0: mining clinically relevant drug-gene interactions. Nucleic Acids Res. **44,** D1036-1044 https://doi.org/10.1093/nar/gkv1165

Wang L, Chen G, Wu S, Xu Y, Guo C, Wang M, Liang T, Guo Z, Di HJ, Hu Z (2022): Genipin improves lipid metabolism and sperm parametersin obese mice via regulation of miR-132 expression. Acta Biochim. Biophys. Sin. (Shanghai) **54,** 1278-1288 https://doi.org/10.3724/abbs.2022120

Wu J, Wang Y, Jiang R, Xue R, Yin X, Wu M, Meng Q (2021): Ferroptosis in liver disease: new insights into disease mechanisms. Cell Death Discov. **7,** 276 https://doi.org/10.1038/s41420-021-00660-4

Xi C, Jia Z, Xiaoli W, Na Z, He W, Hao J (2019): New aspect of liver IL-17(+)γδ T cells. Mol. Immunol. **107,** 41-43 https://doi.org/10.1016/j.molimm.2018.12.030

Xu Y, Zhu Y, Hu S, Pan X, Bawa FC, Wang HH, Wang DQH, Yin L, Zhang Y (2021): Hepatocyte miR-34a is a key regulator in the development and progression of non-alcoholic fatty liver disease. Mol. Metab. **51,** 101244 https://doi.org/10.1016/j.molmet.2021.101244

Yan P, Cheng M, Wang L, Zhao W (2023): A ferroptosis-related gene in Helicobacter pylori infection SOCS1 serves as a potential prognostic biomarker and corresponds with tumor immune infiltration in stomach adenocarcinoma: In silico approach. Int. Immunopharmacol. **119,** 110263 https://doi.org/10.1016/j.intimp.2023.110263

Ye Z, Zhuo Q, Hu Q, Xu X, Mengqi L, Zhang Z, Xu W, Liu W, Fan G, Qin Y, et al. (2021): FBW7-NRA41-SCD1 axis synchronously regulates apoptosis and ferroptosis in pancreatic cancer cells. Redox Biol. **38,** 101807 https://doi.org/10.1016/j.redox.2020.101807

Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, George J, Bugianesi E (2018): Global burden of NAFLD and NASH: trends predictions risk factors and prevention. Nat. Rev. Gastroenterol. Hepatol. **15,** 11-20 https://doi.org/10.1038/nrgastro.2017.109

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

Yu X, Guo Z, Fang Z, Yang K, Liu C, Dong Z, Liu C (2023): Identification and validation of disulfidptosis-associated molecular clusters in non-alcoholic fatty liver disease. Front. Genet. **14,** 1251999 https://doi.org/10.3389/fgene.2023.1251999

Zhang H, Zhang E, Hu H (2021): Role of ferroptosis in non-alcoholic fatty liver disease and its implications for therapeutic strategies. Biomedicines **9,** 1660 https://doi.org/10.3390/biomedicines9111660

Zhao S, Zhang L, Ji W, Shi Y, Lai G, Chi H, Huang W, Cheng C (2022): Machine learning-based characterization of cuprotosis-related biomarkers and immune infiltration in Parkinson's disease. Front. Genet. **13,** 1010361 https://doi.org/10.3389/fgene.2022.1010361

Zhao S, Chi H, Yang Q, Chen S, Wu C, Lai G, Xu K, Su K, Luo H, Peng G, et al. (2023): Identification and validation of neurotrophic factor-related gene signatures in glioblastoma and Parkinson's disease. Front. Immunol. **14,** 1090040 https://doi.org/10.3389/fimmu.2023.1090040

Zhou N, Bao J (2020): FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. Database **2020,** baaa021 https://doi.org/10.1093/database/baaa021

Received: December 7, 2023
Final version accepted: March 29, 2024

doi: 10.4149/gpb_2024017

Supplementary Material

Bioinformatics analysis of potential ferroptosis and non-alcoholic fatty liver disease biomarkers

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

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

Supplementary Figures

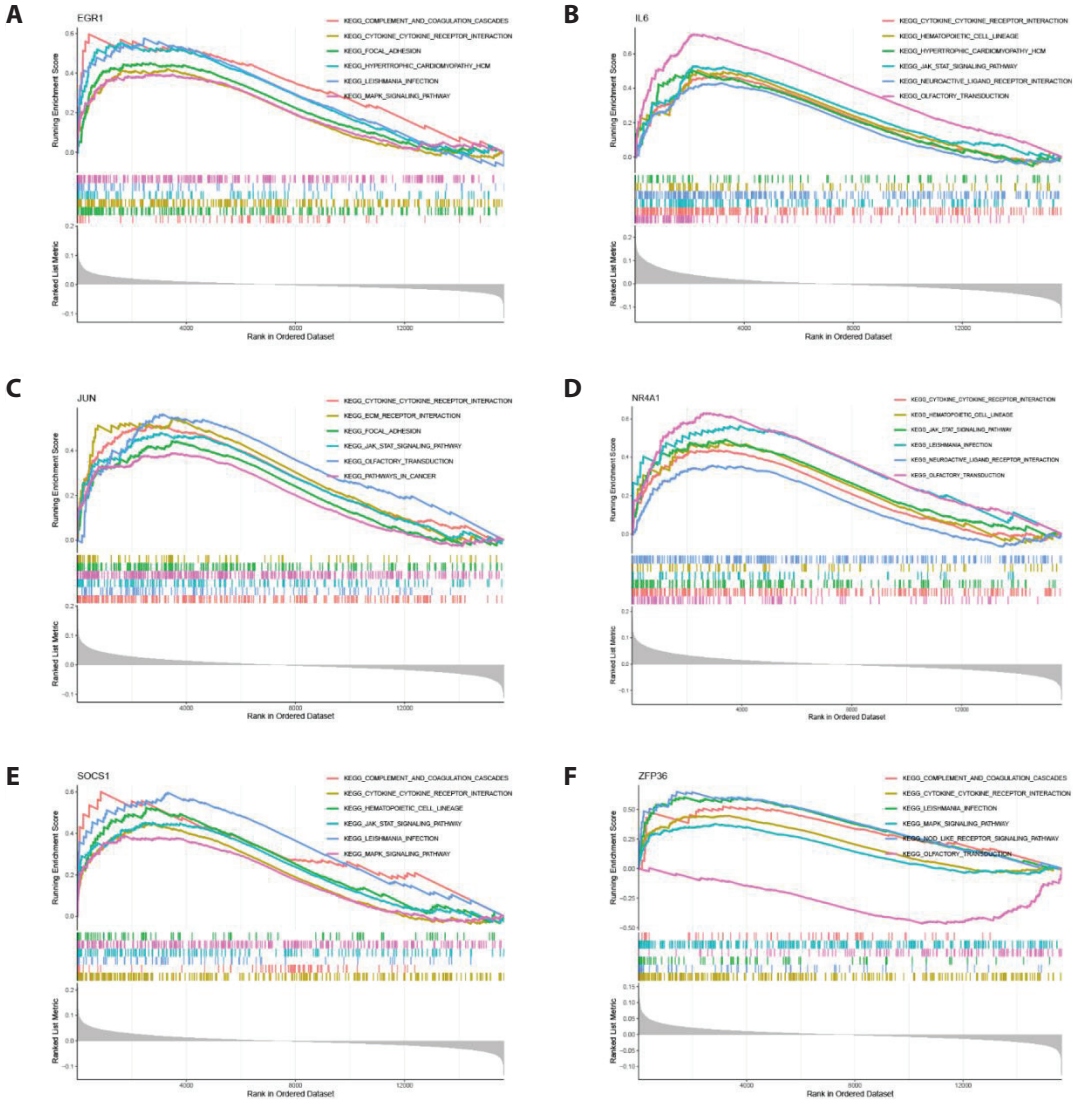

Figure S1. Single-gene GSEA-KEGG pathway analysis of EGR1 **(A),** IL-6 **(B),** JUN **(C),** NR4A1 **(D),** SOCS1 **(E)** and ZFP36 **(F).**

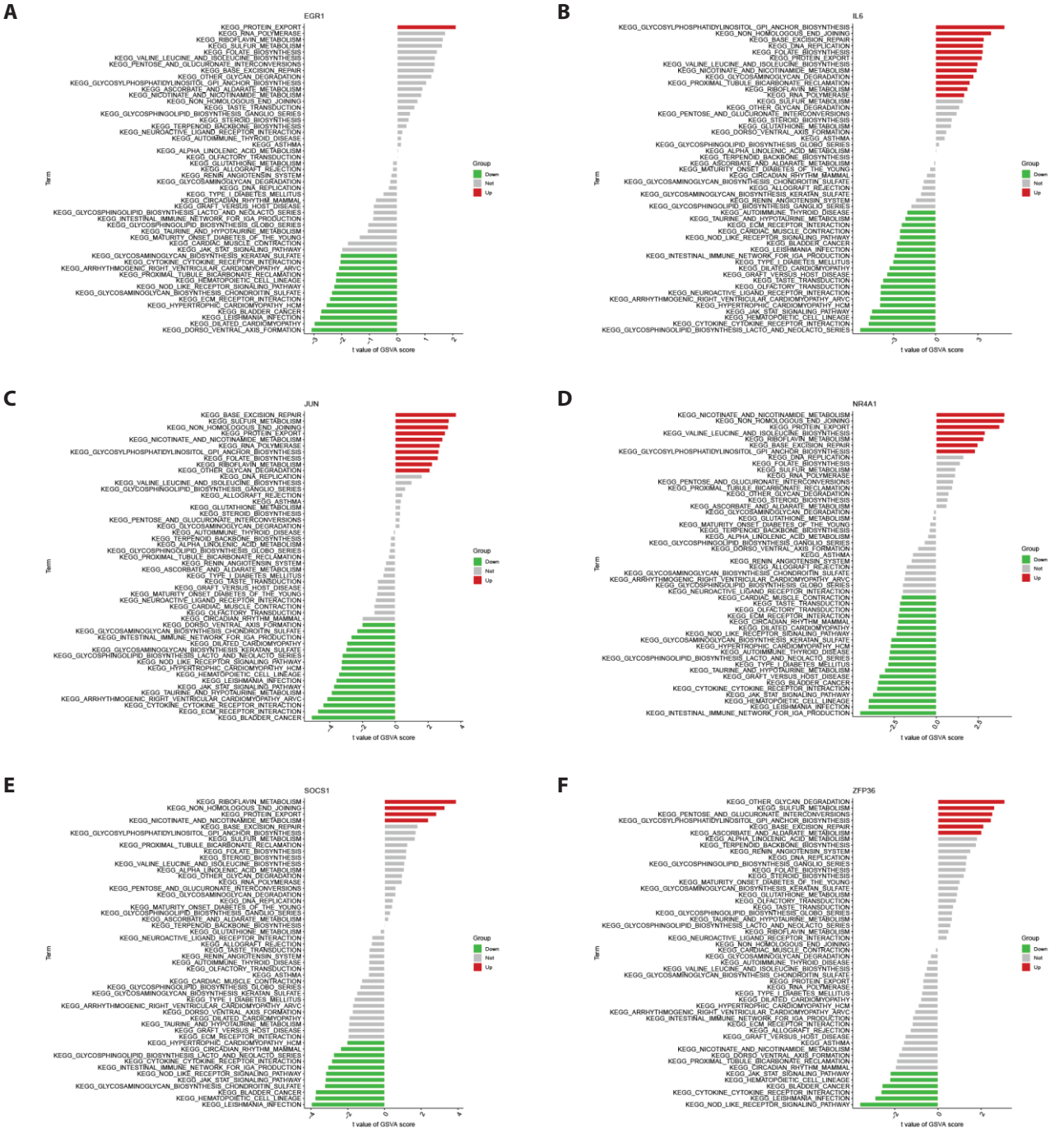

Figure S2. Differential activation pathways were found between groups with high and low expression levels for each marker gene. GSVA in EGR1 **(A)**, IL-6 **(B)**, JUN **(C)**, NR4A1 **(D)**, SOSC1 **(E)**, and ZFP36 **(F)**.

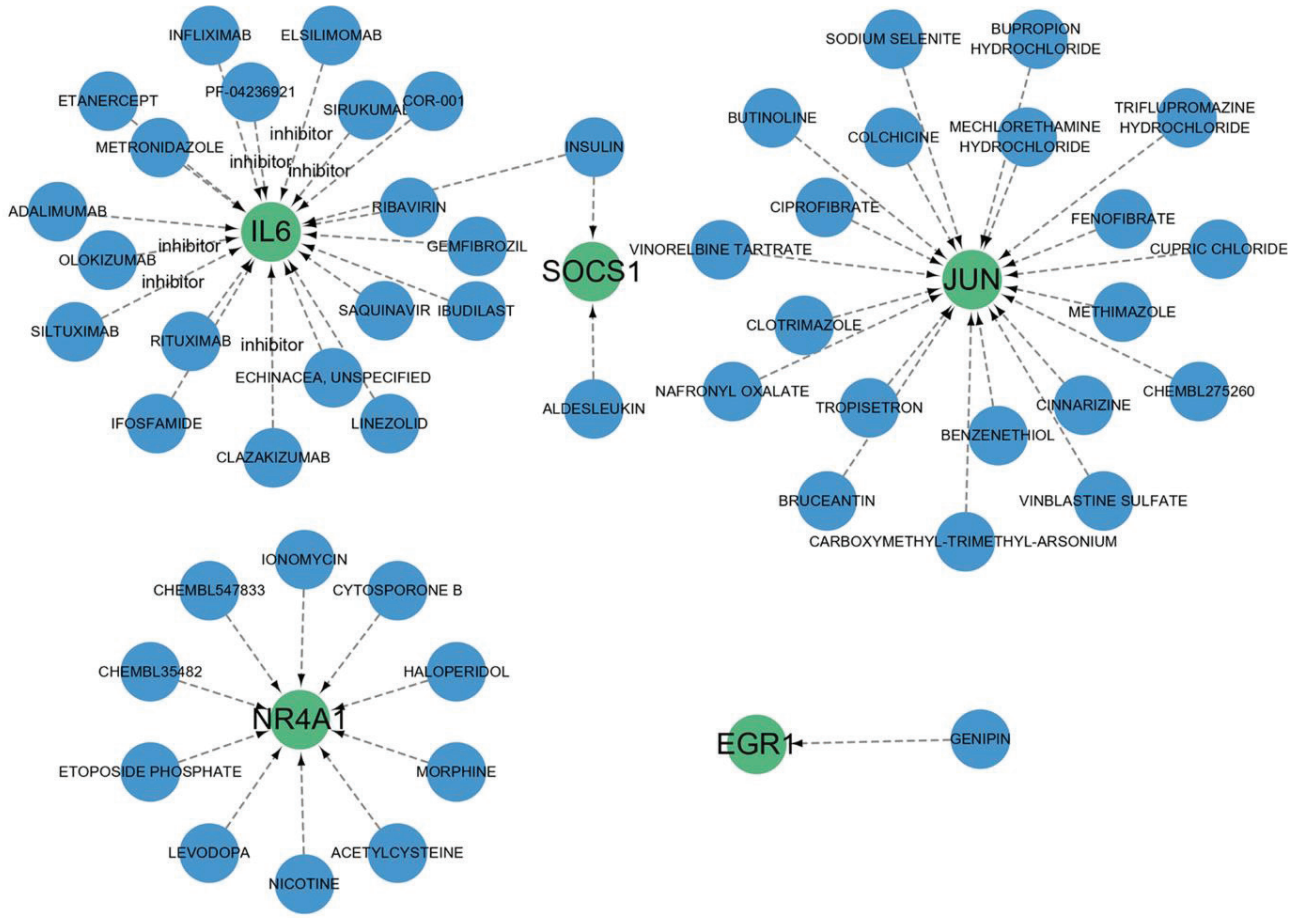

Figure S3. Marker gene-drug prediction. These medicines can target marker genes *via* DGIdb.

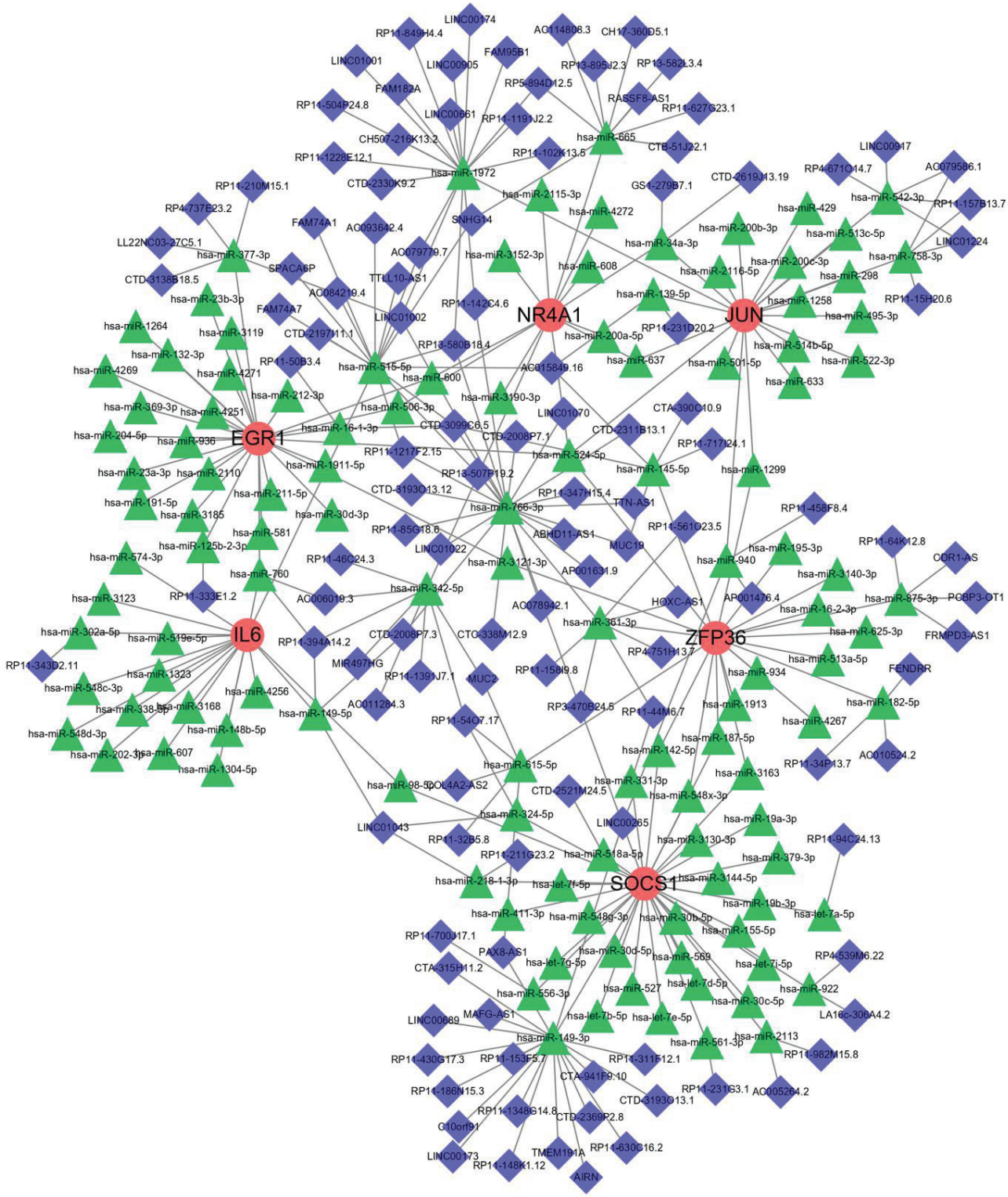

Figure S4. A ceRNA network based on marker genes. The network has 290 edges and 248 nodes (6 marker genes, 119 miRNAs, and 123 lncRNAs).

Supplementary Tables

Table S1. Ferroptosis-related genes from the FerrDb database

RPL8	ACO1	CHAC1	ELOVL5
IREB2	IREB2	MAPK14	FADS1
ATP5MC3	SLC38A1	LINC00472	ALOX12
CS	GLS2	NOX4	FBXW7
EMC2	G6PDX	GOT1	PTEN
ACSF2	ULK1	BECN1	NR1D1
NOX1	ATG3	PRKAA2	NR1D2
CYBB	ATG4D	PRKAA1	TBK1
NOX3	ATG5	ELAVL1	IL6
NOX4	BECN1	BAP1	USP7
NOX5	MAP1LC3A	TP53	miR-182-5p
DUOX1	GABARAPL2	ABCC1	miR-378a-3p
DUOX2	GABARAPL1	ACSL4	CTSB
G6PD	ATG16L1	MIR6852	ACSL4
PGD	WIPI1	ACVR1B	ATF4
VDAC2	WIPI2	TGFBR1	BECN1
PIK3CA	SNX4	BAP1	AQP3
FLT3	ATG13	EPAS1	AQP5
SCP2	ULK2	HILPDA	AQP8
TP53	NCOA4	HIF1A	LINC00618
ACSL4	ACSL4	ALOX12	IREB2
LPCAT3	TP53	ACSL4	MT1DP
NRAS	SAT1	HMOX1	ACSL4
KRAS	ALOX15	IFNG	PEX10
HRAS	ACSL4	ANO6	KEAP1
TF	LPCAT3	LPIN1	AGPAT3
TFRC	ALOX15	HMGB1	PEX12
TFR2	ACSL4	TNFAIP3	CHP1
SLC38A1	KEAP1	TLR4	GPAT4
SLC1A5	EGFR	NOX4	BRPF1
GLS2	NOX4	ATF3	OSBPL9
GOT1	MAPK3	ATM	INTS2
CARS1	MAPK1	YY1AP1	MMD
TP53	BID	EGLN2	CYP4F8
ALOX5	ACSL4	MIOX	MLLT1
KEAP1	ZEB1	TAFAZZIN	TTPA
HMOX1	KEAP1	MTDH	GRIA3
TP53	DPP4	IDH1	EPT1
TP53	ALOX15	SIRT1	POM121L12
GLS2	ALOX12	TAFAZZIN	LIG3
ATG5	CDKN2A	BECN1	AEBP2
ATG7	PEBP1	FBXW7	AGPS
NCOA4	SOCS1	PANX1	CDCA3
TF	CDO1	DNAJB6	PEX2
ALOX5	MYB	BACH1	
ALOX12	HMOX1		LPCAT3
ALOX12B	MAPK8	LONP1	PEX6
ALOX15		CD82	TIMM9
ALOX15B	MAPK9	IL1B	DCAF7
ALOXE3	MAPK1	CTSB	LCE2C
PHKG2	MAPK3	POR	FAR1
TFRC	SLC1A5	CYB5R1	PHF21A

(continued)

Table S1. (continued)

SMAD7	SLC11A2	GSTZ1	TF
LYRM1	IFNA1	TP53	TFRC
AMN	IFNA2	ACO1	FTH1
PEX3	IFNA4	GJA1	GPX4
MTCH1	IFNA5	IREB2	HSPB1
ZEB1	IFNA6	SLC7A11	NFE2L2
SIRT1	IFNA7	PGRMC1	GPX4
ACADSB	IFNA8	CIRBP	FTH1
PVT1	IFNA10	FAR1	SLC7A11
hsa_circ_0008367	IFNA13	circPSEN1	GPX4
SLC39A14	IFNA14	USP11	AKR1C1
NCOA4	IFNA16	STING1	AKR1C2
MAP3K11	IFNA17	YAP1	AKR1C3
GSK3B	IFNA21	HMOX1	GPX4
MAPK8	SMG9	MIR135B	RB1
BRD7	NR1D1	TRIM26	HSPB1
TP53	ACSL4	YAP1	HSF1
SLC25A28	PPARG	NDRG1	SLC7A11
ACSL4	TLR4	MIR302A	GPX4
MFN2	IL6	ASMTL-AS1	GCLC
ACSL4	MIR335	ZFAS1	SLC7A11
SLC11A2	ATF3	FADS2	NFE2L2
ZFAS1	HMOX1	PIEZO1	SQSTM1
SLC38A1	HMGB1	LIFR	NQO1
TSC1	EPAS1	PTPN6	HMOX1
PEBP1	SNX5	MIR15A	FTH1
TGFB1	PAQR3	EGR1	
SNCA	MICU1	ADAM23	SLC3A2
SIRT3	NOX4	ARHGEF26-AS1	MT1G
PRKAA2	TOR2A	ACSL4	NFE2L2
TFRC	MIR375		SLC40A1
CGAS	MAP3K14	lncRNA	SLC7A11
STING1	SIRT3	AABR07017145.1	GPX4
HDDC3	CircKDM4C	TIMP1	SLC7A11
MIR761	MIR324	MIR15A	CISD1
MDM2	QSOX1	KDM6B	SLC7A11
MDM4		NCOA4	FANCD2
ALOX15	KLF2	GSK3B	GPX4
POR	MIR5096	IFNG	NFE2L2
MIR214	TFRC	METTL14	FTMT
DLD	HOTAIR	CHAC1	HSPA5
LONP1	H19	MIB1	ATF4
	FOXO4	KDM5C	SLC7A11
BACH1	ELAVL1	ACSL4	GPX4
DNAJB6	YTHDC2	MEG3	GPX4
WWTR1	DDR2	CCDC6	HMOX1
SIRT1	SLC39A7	ATF3	ATF4
ATM	TRIM46	IREB2	NFE2L2
PRKCA	ACSL1	CFL1	TP53
LGMN	KDM5A	ALOXE3	SLC7A11
ACSL4	TRIM21	MIR539	HELLS
TP53	HMOX1	KMT2D	SCD
IFNG	DPEP1	PTGS2	FADS2
SMPD1	CYGB	CHAC1	SRC
MYCN	IDO1	SLC40A1	STAT3

(continued)

Table S1. (continued)

NFE2L2	CHMP6	BRD4	HSF1
PML	AKR1C1	BRDT	PROM2
MTOR	AKR1C2	SCD	PLA2G6
NFS1	AKR1C3	SLC7A11	HIF1A
TP63	CBS	DECR1	NEAT1
SLC7A11	NFE2L2	NFE2L2	RRM2
TP53	CAV1	GPX4	SLC7A11
CDKN1A	GCH1	SLC7A11	FTMT
MIR137	SIRT3	NFE2L2	
SLC40A1	DAZAP1	GLRX5	PARP1
GPX4	PIR	GPX4	PARP2
GPX4	GCLC	NCOA3	PARP3
ENPP2	FTL		PARP4
VDAC2	HCAR1	GPX4	PARP6
FH	SLC16A1	MTOR	PARP8
CISD2	RRM2	PANX2	PARP9
SLC40A1	SCD	RHEBP1	PARP10
MIR9-1		TFAP2A	PARP11
MIR9-2	PIK3CA	CP	PARP12
MIR9-3	RPTOR	SLC7A11	PARP14
CBS	SREBF1	ARF6	PARP15
NFE2L2	SREBF2	GDF15	PARP16
	FZD7	ABHD12	PDSS2
SQSTM1	NFE2L2	PPP1R13L	TXN
GPX4	NFE2L2	TFAM	SENP1
ISCU	P4HB	KDM3B	PLA2G6
FTH1	NT5DC2	RNF113A	OIP5-AS1
ACSL3	BCAT2	PARK7	MIR190A
OTUB1	HSF1	AHCY	FGF21
CD44	PLA2G6	FXN	CREB1
LINC00336	MIR424	circ-TTBK2	CREB3
STAT3	PARK7	MIR522	CREB5
BRD4	FXN	IDH2	FTMT
PRDX6	SUV39H1	PPARA	GOT1
MIR17	ATF2	NOS2	TFRC
SCD	CDKN1A	SIAH2	GPX4
SESN2	FTH1	RELA	MIR130B
NF2	NFE2L2	PRKAA2	BEX1
ARNTL	STAT3	VDR	ASAH2
HIF1A	ACOT1	NEDD4	SCD
JUN	NFE2L2	FXN	FABP4
CA9	ALDH3A2	AIFM2	AKT1S1
HSPA5	NFE2L2	PRDX1	MLST8
TMBIM4	STK11	AR	MTOR
HSPA5	FNDC5	CBS	RPTOR
PLIN2	CircIL4R	NFE2L2	CDH1
MIR212	CDH1	CHMP5	SIRT1
Fer1HCH	NFE2L2	CHMP6	TYRO3
AIFM2	MIR214	HMOX1	SIRT6
AIFM2	NEDD4L	ZFP36	TMSB4X
LAMP2	SQSTM1	LAMP2	TMSB4Y
ZFP36	TF	MTF1	KIF20A
GPX4	FTMT	COPZ1	ECH1
PROM2	BRD2	NUPR1	circRHOT1
CHMP5	BRD3	USP35	ETV4

(continued)

Table S1. (continued)

MEG8	MIR27A	DHODH	FURIN
VCP	MIR670	SLC7A11	circRHBG
circ_0007142	MEF2C	MIR545	GALNT14
ENPP2	NF2	OTUB1	KLHDC3
	CDH1	PDK4	LINC01833
RBMS1	HSPB1	CircPVT1	circGFRA1
KDM4A	EZH2	MIR9-3HG	MAPKAP1
CBS	PEDS1	ADIPOQ	MLST8
MGST1	SMPD1	circDTL	MTOR
circKIF4A	ADAMTS13	GPX4	PRR5
miR-7-5p	CDC25A	mmu_circRNA_0000309	RICTOR
PRDX6	G6PD		GSTM1
circ_0067934	SRSF9	IL6	TERT
MPC1	CAV1	PTPN18	circ0097009
CHMP1A	CircFNDC3B	FTH1	TMEM161B-DT
CAMKK2	PPARD	FTH1	circEPSTI1
SOX2	CISD2	FTL	MIR18A
SRSF9	ENO3	LCN2	RARRES2
PROK2	SESN2	ABCC5	USP11
MIR4443	LCN2	CISD3	
SIRT2	MARCHF5	MS4A15	
circRNA1615	TRIB2	LCN2	

Table S2. DEGs involved in NAFLD samples

id	logFC	AveExpr	t	P.Value	adj.P.Val	B
FOSB	-3.71737	10.12579	-12.7345	3.38E-21	5.29E-17	37.51716
MYC	-2.17667	11.23325	-11.2139	2.77E-18	2.16E-14	31.05087
JUNB	-2.19622	10.84673	-10.9619	8.58E-18	4.48E-14	29.95585
FOS	-2.43184	11.91821	-10.5913	4.59E-17	1.20E-13	28.33465
WNT5A	1.260966	9.792385	10.53082	6.03E-17	1.35E-13	28.06887
RAB26	1.086493	11.51695	10.47823	7.66E-17	1.42E-13	27.83765
THBS1	-1.61357	11.2616	-10.4649	8.14E-17	1.42E-13	27.77889
FAM107A	-1.63472	10.55895	-10.4146	1.02E-16	1.60E-13	27.55749
SOCS2	-1.61042	11.3142	-10.373	1.24E-16	1.76E-13	27.37426
PPP1R15A	-1.39261	10.68695	-10.1109	4.08E-16	5.32E-13	26.21729
TMEM169	1.156489	9.314684	9.958645	8.19E-16	9.86E-13	25.54279
APOLD1	-2.08094	9.67447	-9.88041	1.17E-15	1.31E-12	25.1958
NAT8B	1.509599	9.901963	9.79192	1.76E-15	1.83E-12	24.80293
CYR61	-1.70267	11.61684	-9.64914	3.38E-15	3.21E-12	24.16824
HBEGF	-1.37925	9.215596	-9.64099	3.51E-15	3.21E-12	24.132
KLF6	-1.19481	11.19339	-9.62986	3.69E-15	3.21E-12	24.08246
GADD45G	-2.03225	10.76135	-9.5499	5.33E-15	4.39E-12	23.72659
P4HA1	-1.18962	11.01287	-9.48327	7.24E-15	5.66E-12	23.42981
GADD45B	-1.08372	12.4328	-9.40936	1.02E-14	7.42E-12	23.10046
FILIP1L	-1.21127	9.979781	-9.40381	1.04E-14	7.42E-12	23.07574
RNF43	1.30893	9.97378	9.31652	1.56E-14	1.06E-11	22.68652
EPHA2	-1.39975	10.01729	-9.30523	1.64E-14	1.07E-11	22.63618
NR4A2	-1.44078	8.908895	-9.23526	2.26E-14	1.31E-11	22.324
ADAMTS1	-1.86824	10.68568	-9.2125	2.51E-14	1.40E-11	22.22246
SIK1	-1.09196	9.015107	-9.2056	2.59E-14	1.40E-11	22.19167

(continued)

Table S2. (continued)

SOCS3	-1.35832	8.991903	-8.85323	1.31E-13	5.86E-11	20.61829
PPRC1	-1.10482	11.25239	-8.8508	1.33E-13	5.86E-11	20.60741
FOSL2	-1.47649	10.23076	-8.84484	1.36E-13	5.86E-11	20.58082
MAP3K8	-1.14189	10.70128	-8.84153	1.39E-13	5.86E-11	20.56603
CYP7A1	2.299719	10.53905	8.789267	1.76E-13	7.26E-11	20.33259
PIM1	-1.52117	11.59116	-8.68052	2.91E-13	1.08E-10	19.84692
GPRC5A	-1.31432	8.760031	-8.66853	3.07E-13	1.12E-10	19.79337
PHLDA1	-1.44048	10.71129	-8.60014	4.21E-13	1.45E-10	19.48801
ARL14	-1.59334	8.587361	-8.54446	5.43E-13	1.77E-10	19.23944
GINS2	1.256891	9.710579	8.508509	6.41E-13	2.01E-10	19.07896
PNRC1	-1.53511	10.62952	-8.44794	8.47E-13	2.60E-10	18.80868
NR4A1	-1.30722	8.684865	-8.43716	8.90E-13	2.68E-10	18.76059
IL6	-2.35007	9.069118	-8.34785	1.34E-12	3.82E-10	18.36226
C17orf96	-1.14053	9.971487	-8.31168	1.58E-12	4.35E-10	18.20102
FMO1	1.426674	9.13203	8.300543	1.67E-12	4.37E-10	18.15136
JUN	-1.11399	12.39631	-8.30013	1.67E-12	4.37E-10	18.14952
RGS1	-1.08806	9.930366	-8.29957	1.67E-12	4.37E-10	18.14701
FOSL1	-1.79873	8.95223	-8.29397	1.72E-12	4.41E-10	18.12205
PTGS2	-1.4921	9.75136	-8.28569	1.79E-12	4.45E-10	18.08517
RRS1	-1.2678	10.85954	-8.23254	2.28E-12	5.48E-10	17.84835
IGFBP2	-1.40787	11.57056	-8.22588	2.35E-12	5.49E-10	17.81866
IL1RL1	-1.58356	8.842817	-8.19846	2.66E-12	6.04E-10	17.69654
PEG10	1.42138	9.446904	8.112393	3.95E-12	8.25E-10	17.31345
CRISPLD2	-1.17879	10.97731	-8.10746	4.04E-12	8.33E-10	17.29151
KRT222	1.407107	9.127814	7.998881	6.65E-12	1.28E-09	16.80875
FRAT1	1.042989	9.833934	7.947413	8.42E-12	1.60E-09	16.58015
SLITRK3	-2.04001	10.23091	-7.89825	1.05E-11	1.92E-09	16.36196
SOCS1	-1.47636	9.36709	-7.85896	1.26E-11	2.24E-09	16.18768
ZFP36	-1.25894	10.35613	-7.85572	1.28E-11	2.25E-09	16.17332
BCL6	-1.05441	11.10534	-7.69318	2.69E-11	4.05E-09	15.45358
MOGAT2	1.069046	9.292723	7.693023	2.69E-11	4.05E-09	15.45289
EGR1	-1.41884	12.62904	-7.66067	3.12E-11	4.60E-09	15.30991
SLC25A34	1.35093	10.20203	7.650587	3.26E-11	4.73E-09	15.26533
AGPAT9	-1.0472	11.07115	-7.58813	4.34E-11	6.17E-09	14.98959
NAGS	1.233252	10.20879	7.519561	5.92E-11	8.06E-09	14.68724
CCL2	-1.40749	11.91881	-7.49932	6.49E-11	8.57E-09	14.5981
IL4R	-1.01569	10.87308	-7.45902	7.79E-11	9.92E-09	14.42069
SLC2A3	-1.14262	10.88932	-7.43001	8.89E-11	1.10E-08	14.29311
CLCF1	-1.00011	8.986573	-7.29514	1.64E-10	1.84E-08	13.70112
FOXC1	-1.18151	8.876078	-7.24508	2.05E-10	2.18E-08	13.48192
THBD	-1.37212	10.04276	-7.20084	2.51E-10	2.57E-08	13.28849
DBP	1.370326	9.166775	7.198497	2.53E-10	2.57E-08	13.27823
KCNK1	-1.18268	10.33526	-7.1898	2.63E-10	2.64E-08	13.24023
RGS2	-1.18032	11.66866	-7.13521	3.37E-10	3.23E-08	13.00193
EMP1	-1.43843	10.61761	-7.12459	3.53E-10	3.32E-08	12.95565
MAFF	-1.02445	8.541991	-7.12064	3.60E-10	3.35E-08	12.93841
PIM3	-1.02626	10.12722	-7.11842	3.63E-10	3.36E-08	12.92877
ANKS4B	1.051835	9.999815	7.110194	3.77E-10	3.46E-08	12.89289
IGFBP1	-1.45743	10.76558	-7.10627	3.83E-10	3.49E-08	12.87581
PADI4	-1.07857	8.70138	-7.05877	4.75E-10	4.15E-08	12.66895

(continued)

Table S2. (continued)

PLAUR	-1.02246	9.487603	-7.0461	5.02E-10	4.37E-08	12.61387
RTP3	1.38638	12.07905	6.908341	9.31E-10	7.51E-08	12.01615
C14orf80	1.201282	10.14715	6.870586	1.10E-09	8.54E-08	11.85285
LOC73010	1.046043	9.31037	6.845053	1.24E-09	9.47E-08	11.74254
C2CD4B	-1.04511	9.113979	-6.84048	1.26E-09	9.62E-08	11.72281
IER3	-1.41272	10.62984	-6.78533	1.61E-09	1.19E-07	11.48493
C10orf10	-1.25472	11.44701	-6.78355	1.62E-09	1.19E-07	11.47727
KIAA0040	-1.33136	9.646372	-6.77792	1.67E-09	1.21E-07	11.45301
KLF4	-1.08931	9.428132	-6.7501	1.88E-09	1.33E-07	11.33329
TGFB3	-1.35435	10.61449	-6.72184	2.14E-09	1.47E-07	11.21177
CEBPA	1.271243	12.30586	6.700904	2.34E-09	1.59E-07	11.12184
RALGDS	-1.12556	10.20209	-6.65117	2.92E-09	1.91E-07	10.90853
TNFRSF12	-1.51369	10.02705	-6.61356	3.45E-09	2.18E-07	10.74753
PTX3	-1.40955	8.258027	-6.60887	3.52E-09	2.20E-07	10.72748
CRYAA	1.307361	10.26304	6.515115	5.33E-09	3.15E-07	10.32739
MYBPH	-1.17212	8.533135	-6.51265	5.39E-09	3.17E-07	10.3169
IL1B	-1.32246	10.34258	-6.50533	5.56E-09	3.26E-07	10.28574
PRSS3	1.038414	8.660752	6.502674	5.63E-09	3.27E-07	10.27443
CISH	-1.18097	11.00277	-6.37136	1.00E-08	5.35E-07	9.717377
S100P	-1.38336	10.31997	-6.34566	1.12E-08	5.86E-07	9.608814
C5AR1	-1.09339	10.12948	-6.31798	1.26E-08	6.46E-07	9.491984
LIF	-1.06317	8.813163	-6.27784	1.51E-08	7.48E-07	9.322925
MMP19	-1.0004	8.940009	-6.27506	1.52E-08	7.55E-07	9.311236
ABCC6P1	1.0071	9.770297	6.269816	1.56E-08	7.65E-07	9.289181
OSMR	-1.02132	10.60639	-6.25713	1.65E-08	7.96E-07	9.23584
FPR1	-1.10705	9.920089	-6.24222	1.76E-08	8.39E-07	9.173215
ACTG2	-1.31272	9.66197	-6.23236	1.84E-08	8.70E-07	9.131838
LGALS4	1.226828	10.3704	6.231008	1.85E-08	8.73E-07	9.126148
SPSB1	-1.36716	10.63129	-6.1996	2.12E-08	9.85E-07	8.994479
IER5L	-1.02771	9.871274	-6.15456	2.57E-08	1.15E-06	8.806042
GFPT2	-1.13491	9.130288	-6.13704	2.77E-08	1.24E-06	8.732857
LOC15476	-1.15654	10.14418	-6.06822	3.73E-08	1.57E-06	8.446205
RASD1	-1.53849	9.857731	-6.00462	4.91E-08	1.97E-06	8.182321
SERPINE1	-1.20041	10.92924	-5.9923	5.17E-08	2.04E-06	8.131337
SLC7A1	-1.18456	9.018894	-5.98292	5.38E-08	2.12E-06	8.09255
FAM124B	1.033475	9.558248	5.940316	6.46E-08	2.42E-06	7.916632
PROK2	-1.34674	9.214529	-5.9137	7.24E-08	2.64E-06	7.80699
RND1	-1.3873	10.33941	-5.85269	9.39E-08	3.28E-06	7.556381
NFE2	-1.00478	9.905769	-5.79417	1.20E-07	4.01E-06	7.316963
FAM169B	-1.03682	8.894854	-5.51489	3.89E-07	1.06E-05	6.188566
C2orf82	1.082272	10.0155	5.351491	7.63E-07	1.84E-05	5.54003
CDH15	1.181789	9.058184	5.27981	1.02E-06	2.39E-05	5.258431
RGS16	-1.01054	9.320187	-5.12185	1.94E-06	4.06E-05	4.644495
CNN1	-1.00385	8.998817	-5.08551	2.25E-06	4.56E-05	4.504585
AVPR1A	-1.13181	10.87159	-4.98583	3.34E-06	6.23E-05	4.123495
S100A12	-1.2928	10.23392	-4.65708	1.21E-05	0.000174	2.895781
AKR1B10	1.412058	9.01886	4.266948	5.24E-05	0.00056	1.502928

Table S3. GSEA enrichment analysis results of marker genes

ID	NES	pvalue
KEGG_FOCAL_ADHESION	1.942953	2.92E-07
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	1.868545	7.75E-07
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	2.219586	1.17E-06
KEGG_HYPERTROPHIC_CARDIOMYOPATHY_HCM	2.140582	1.23E-06
KEGG_LEISHMANIA_INFECTION	2.140898	8.91E-06
KEGG_MAPK_SIGNALING_PATHWAY	1.751111	1.09E-05
KEGG_DILATED_CARDIOMYOPATHY	2.009846	1.21E-05
KEGG_PATHWAYS_IN_CANCER	1.626635	6.65E-05
KEGG_CARDIAC_MUSCLE_CONTRACTION	1.961385	7.27E-05
KEGG_ECM_RECEPTOR_INTERACTION	1.894468	0.00017
KEGG_JAK_STAT_SIGNALING_PATHWAY	1.732594	0.000179
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY	1.938346	0.000222
KEGG_HEMATOPOIETIC_CELL_LINEAGE	1.819244	0.000486
KEGG_RIBOSOME	-1.79725	0.000644
KEGG_TGF_BETA_SIGNALING_PATHWAY	1.795346	0.000726
KEGG_EPITHELIAL_CELL_SIGNALING_IN_HELICOBACTER_PYLORI_INFECTION	1.826502	0.001695
KEGG_AXON_GUIDANCE	1.641564	0.002772
KEGG_P53_SIGNALING_PATHWAY	1.695529	0.003021
KEGG_BLADDER_CANCER	1.768668	0.003591
KEGG_ARACHIDONIC_ACID_METABOLISM	1.708824	0.003728
KEGG_ARRHYTHMOGENIC_RIGHT_VENTRICULAR_CARDIOMYOPATHY_ARVC	1.629059	0.004725
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	1.468294	0.005398
KEGG_PROXIMAL_TUBULE_BICARBONATE_RECLAMATION	1.784299	0.006858
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	1.609481	0.006989
KEGG_VIRAL_MYOCARDITIS	1.630695	0.007132
KEGG_COLORECTAL_CANCER	1.648278	0.009028
KEGG_RENAL_CELL_CARCINOMA	1.66676	0.010064
KEGG_WNT_SIGNALING_PATHWAY	1.443352	0.0115
KEGG_CHEMOKINE_SIGNALING_PATHWAY	1.444012	0.012821
KEGG_CELL_CYCLE	1.496342	0.013071
KEGG_PEROXISOME	-1.39908	0.015793
KEGG_TIGHT_JUNCTION	1.490507	0.01762
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS	1.450568	0.019727
KEGG_ARGININE_AND_PROLINE_METABOLISM	1.621661	0.020159
KEGG_PROTEIN_EXPORT	-1.55876	0.023471
KEGG_ALDOSTERONE_REGULATED_SODIUM_REABSORPTION	1.503404	0.026848
KEGG_PRION_DISEASES	1.599667	0.026876
KEGG_RNA_POLYMERASE	-1.46094	0.028834
KEGG_TAURINE_AND_HYPOTAURINE_METABOLISM	1.631103	0.029518
KEGG_SMALL_CELL_LUNG_CANCER	1.456026	0.029649
KEGG_NOTCH_SIGNALING_PATHWAY	1.513394	0.033097
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY	1.387971	0.03418
KEGG_TASTE_TRANSDUCTION	-1.46974	0.036857
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION	1.431274	0.037142
KEGG_PROSTATE_CANCER	1.354674	0.040185
KEGG_ACUTE_MYELOID_LEUKEMIA	1.466222	0.045967
KEGG_PATHOGENIC_ESCHERICHIA_COLI_INFECTION	1.408733	0.048739

Table S4. Immune infiltration analysis results of the study

id	B cells naive	B cells memory	Plasma cells
GSM1539891	0.070791582	0	0.051561206
GSM1539893	0	0	0.090541988
GSM2385759	0.061058317	0	0.0632446
GSM2385761	0.057256702	0	0.075938632
GSM2385763	0.057399112	0	0.076427674
GSM2385764	0.023116413	0	0.021375858
GSM2385766	0.113277229	0	0.060310119
GSM2385767	0.110335179	0	0.064767336
GSM2385768	0.088829568	0	0.083640338
GSM2385769	0.10717875	0	0.083021368
GSM2385770	0.083192459	0	0.057812708
GSM2385772	0.096032067	0	0.03906588
GSM2385775	0.060304762	0	0.163461848
GSM2385776	0.076451298	0	0.092065997
GSM2385779	0.051186756	0	0.1078879
GSM2385781	0.082360992	0	0.045425706
GSM1539877	0.038930058	0	0.02820127
GSM1539878	0.038307103	0	0.064388861
GSM1539879	0.028983741	0	0.0494745
GSM1539880	0.0635875	0	0.072987865
GSM1539881	0.022571379	0	0.050479731
GSM1539882	0.043306238	0	0.03645072
GSM1539883	0.077302866	0	0.03182737
GSM2385726	0.025527627	0	0.059429691
GSM2385728	0.036644156	0	0.031735407
GSM2385731	0.005660617	0	0.102617481
GSM2385735	0.040293639	0	0.033973871
GSM2385737	0.0662848	0	0.062651318
GSM2385738	0	0.006683332	0.023488476
GSM2385739	0.064973561	0	0.05271381
GSM2385741	0.089826078	0	0.005517916
GSM2385744	0.060010725	0	0.065046364
GSM2385747	0	0	0
GSM2385750	0.020096543	0	0.034267883
GSM2385773	0.009201353	0	0.157561792
GSM2385777	0.049451183	0	0.022539481

Table S4. (continued)

id	cells CD4 memory restin	T cells CD4 memory activated	T cells follicular helper
GSM1539891	0.223420593	0.037784177	0
GSM1539893	0.113729762	0.003560686	0
GSM2385759	0.201147491	0	0
GSM2385761	0.168393568	0	0
GSM2385763	0.231656045	0.004578058	0
GSM2385764	0.236727245	0.050528201	0
GSM2385766	0.17608986	0.031329294	0
GSM2385767	0.17954754	0	0
GSM2385768	0.221035903	0.03557498	0
GSM2385769	0.136268974	0.055188537	0
GSM2385770	0.28512492	0.02663323	0
GSM2385772	0.18743392	0.019592963	0
GSM2385775	0	0.017348366	0
GSM2385776	0.071070388	0.027141243	0
GSM2385779	0.074543822	0.054295019	0
GSM2385781	0.13539238	0.008549487	0
GSM1539877	0.161406017	0.004272449	0
GSM1539878	0.188200594	0	0
GSM1539879	0.290295202	0.012876254	0
GSM1539880	0.233826942	0.028734099	0
GSM1539881	0.185018855	0.048601436	0
GSM1539882	0.217931666	0	0
GSM1539883	0.203780814	0.060533029	0
GSM2385726	0.130584306	0.068205712	0
GSM2385728	0.183621617	0.038919789	0
GSM2385731	0.071518144	0	0
GSM2385735	0.166474149	0.024645877	0
GSM2385737	0.210423014	0.024252927	0
GSM2385738	0.076847986	0.023219426	0
GSM2385739	0.263905403	0.030306477	0
GSM2385741	0.155763433	0.040575919	0
GSM2385744	0.135662994	0.01721717	0
GSM2385747	0.105727169	0.002751352	0
GSM2385750	0.174598106	0.006230625	0
GSM2385773	0.174160996	0.02100275	0
GSM2385777	0.144123242	0	0.000946596

Table S4. (continued)

id	T cells gamma delta	NK cells resting	NK cells activated
GSM1539891	0.131530027	0	0.005373215
GSM1539893	0.267805764	0	0.006147203
GSM2385759	0.214459393	0	0
GSM2385761	0.079505362	0	0
GSM2385763	0.068730239	0.002288686	0
GSM2385764	0.096429996	0	0.0178989
GSM2385766	0.101281579	0	0.011921649
GSM2385767	0.179042793	0	0.002208011
GSM2385768	0.126246448	0	0
GSM2385769	0.068234014	0	0.020724033
GSM2385770	0.099895721	0	0
GSM2385772	0.08700979	0.006133335	0.005737064
GSM2385775	0.031611076	0	0
GSM2385776	0.110576801	0	0.008180495
GSM2385779	0.017951715	0.003348638	0.008271023
GSM2385781	0.075673371	0	0.021312387
GSM1539877	0.082242674	0	0.024698527
GSM1539878	0.131705212	0	0.023850051
GSM1539879	0.263312462	0	0.000101988
GSM1539880	0.177424446	0.016906401	0
GSM1539881	0.200756482	0	0
GSM1539882	0.131305627	0	0
GSM1539883	0.066859921	0.015033257	0
GSM2385726	0.218473745	0	0.021814082
GSM2385728	0.24056683	0	0
GSM2385731	0.308980321	0	0
GSM2385735	0.207406705	0	0.001380741
GSM2385737	0.14639098	0	0.013263313
GSM2385738	0.205233816	0	0.029160277
GSM2385739	0.112187016	0	0
GSM2385741	0.224928617	0	0
GSM2385744	0.241311413	0	0.001356968
GSM2385747	0.184394742	0	0.046232399
GSM2385750	0.164724511	0	0.002624469
GSM2385773	0.007289046	0.025227388	0
GSM2385777	0.081497611	0	0.020937138

Table S4. (continued)

id	Macrophages M1	Macrophages M2	Dendritic cells resting
GSM1539891	0.018852097	0.174326391	0.004982996
GSM1539893	0.079659562	0.264314242	0.008305843
GSM2385759	0.051547152	0.236814885	0.01871418
GSM2385761	0.025388179	0.109314221	0
GSM2385763	0.02331105	0.113274747	0
GSM2385764	0.034739059	0.104507499	0
GSM2385766	0.030741842	0.114966315	0
GSM2385767	0	0.1408852	0
GSM2385768	0.030446746	0.101816492	0
GSM2385769	0.04422553	0.069809603	0
GSM2385770	0.027007721	0.099697076	0
GSM2385772	0.0038878	0.174089625	0
GSM2385775	0.015668015	0.046489915	0
GSM2385776	0.002347153	0.187849306	0
GSM2385779	0.058744121	0.108972194	0
GSM2385781	0.003188739	0.139680513	0
GSM1539877	0.036023303	0.139178154	0.022392675
GSM1539878	0.025658299	0.204048072	0.00239425
GSM1539879	0.061556571	0.17018565	0.048732191
GSM1539880	0.018151587	0.095311588	0.00426687
GSM1539881	0.032856267	0.147341285	0
GSM1539882	0.056658466	0.220517101	0.005018648
GSM1539883	0.046426591	0.105331384	0
GSM2385726	0.03260889	0.308742026	0
GSM2385728	0.079758046	0.202433435	0.013851194
GSM2385731	0.101109016	0.208170923	0.022184415
GSM2385735	0.046139427	0.263465525	0.039064221
GSM2385737	0.03792031	0.278459344	0.060570037
GSM2385738	0.045715521	0.360961131	0.031008417
GSM2385739	0.070145254	0.165854717	0.002292974
GSM2385741	0.073256817	0.161960713	0.005245199
GSM2385744	0.015419416	0.292709401	0.031543171
GSM2385747	0.058606803	0.395652503	0.007912597
GSM2385750	0.08000862	0.341829661	0.013373487
GSM2385773	0.006180817	0.137660521	0
GSM2385777	0.053079968	0.170007449	0

Table S4. (continued)

id	Mast cells resting	Mast cells activated	Eosinophils
GSM1539891	0.035792063	0	0.014996543
GSM1539893	0	0.038546927	0
GSM2385759	0	0.02432442	0.010220834
GSM2385761	0	0.11660769	0
GSM2385763	0	0.122231385	0.017442175
GSM2385764	0	0.089045696	0.017571196
GSM2385766	0	0.08006466	0
GSM2385767	0	0.249120716	0
GSM2385768	0	0.094308786	0
GSM2385769	0	0.111722663	0
GSM2385770	0	0.070280244	0
GSM2385772	0.013820858	0.029651092	0
GSM2385775	0	0.155793108	0
GSM2385776	0	0.048302299	0
GSM2385779	0	0.16729165	0.002313027
GSM2385781	0	0.114460524	0
GSM1539877	0	0.057882559	0.012810558
GSM1539878	0	0.061173086	0
GSM1539879	0.038030954	0	0
GSM1539880	0.008137408	0.008427534	0
GSM1539881	0	0.032248662	0.015832289
GSM1539882	0	0.068166577	0.013347225
GSM1539883	0	0.014337509	0.020800322
GSM2385726	0.046585865	0	0.03184925
GSM2385728	0.01557953	0	0
GSM2385731	0.098819582	0	0
GSM2385735	0.037179561	0	0
GSM2385737	0.014300865	0	0
GSM2385738	0.048924798	0	0
GSM2385739	0.075076198	0	0
GSM2385741	0.109415568	0	0
GSM2385744	0.009387474	0.001017477	0
GSM2385747	0.062728521	0	0.001114667
GSM2385750	0.028472719	0	0
GSM2385773	0	0.189693742	0
GSM2385777	0	0.077188862	0

Table S4. (continued)

id	T cells CD4 naïve	Macrophages M0	T cells CD8
GSM1539891	0	0	0
GSM1539893	0	0	0.117525508
GSM2385759	0	0	0.012318602
GSM2385761	0	0	0
GSM2385763	0	0	0
GSM2385764	0	0	0
GSM2385766	0	0	0.025266467
GSM2385767	0	0	0.000438604
GSM2385768	0	0	0
GSM2385769	0	0	0
GSM2385770	0	0	0
GSM2385772	0	0	0
GSM2385775	0.064695369	0.021833691	0.023710265
GSM2385776	0	0	0
GSM2385779	0	0	0
GSM2385781	0	0	0
GSM1539877	0	0	0
GSM1539878	0	0	0.001155125
GSM1539879	0	0	0.027294636
GSM1539880	0	0	0
GSM1539881	0	0	0
GSM1539882	0	0	0
GSM1539883	0	0	0
GSM2385726	0	0	0.010635507
GSM2385728	0	0	0.040260132
GSM2385731	0.00550047	0.051311474	0.021921473
GSM2385735	0	0	0.042503902
GSM2385737	0	0	0
GSM2385738	0	0	0
GSM2385739	0	0	0.018173212
GSM2385741	0	0	0.034525796
GSM2385744	0	0	0.003068266
GSM2385747	0	0	0.014303565
GSM2385750	0	0	0
GSM2385773	0	0	0
GSM2385777	0	0	0

Table S4. (continued)

id	Monocytes	T cells regulatory	Dendritic cells activated
GSM1539891	0.215279279	0	0.015309831
GSM1539893	0.009862515	0	0
GSM2385759	0.093285632	0	0.00947854
GSM2385761	0.207111124	0	0.018439626
GSM2385763	0.242848411	0	0.018554839
GSM2385764	0.210343136	0	0.014826106
GSM2385766	0.221084673	0	0.012326927
GSM2385767	0	0	0.061407445
GSM2385768	0.135388285	0	0.019400879
GSM2385769	0.242623848	0	0.029029269
GSM2385770	0.201303107	0	0.013867214
GSM2385772	0.301517549	0	0.020561394
GSM2385775	0.305850965	0	0.043744368
GSM2385776	0.278171388	0	0.034863427
GSM2385779	0.258403031	0	0.029191527
GSM2385781	0.231162635	0	0.047652831
GSM1539877	0.314722832	0	0.014824125
GSM1539878	0.226941897	0	0.032177449
GSM1539879	0	0	0.002229866
GSM1539880	0.233602053	0	0.037656058
GSM1539881	0.236331543	0	0.0216182
GSM1539882	0.205378731	0	0.001919001
GSM1539883	0.28275574	0	0.011042235
GSM2385726	0.040694183	0	0.002846863
GSM2385728	0.11330435	0	0
GSM2385731	0	0	0
GSM2385735	0.097472383	0	0
GSM2385737	0.062020313	0	0.013954984
GSM2385738	0.148756821	0	0
GSM2385739	0.137974659	0	0.00639672
GSM2385741	0.098983944	0	0
GSM2385744	0.111661136	0	0.014588025
GSM2385747	0.107214271	0.003624444	0
GSM2385750	0.12898634	0	0
GSM2385773	0.224508281	0	0.030631287
GSM2385777	0.262670252	0	0.010491942

Table S4. (continued)

id	Neutrophils
GSM1539891	0
GSM1539893	0
GSM2385759	0.003385956
GSM2385761	0.142044897
GSM2385763	0.021257581
GSM2385764	0.082890695
GSM2385766	0.021339386
GSM2385767	0.012247177
GSM2385768	0.063311575
GSM2385769	0.031973412
GSM2385770	0.035185601
GSM2385772	0.015466664
GSM2385775	0.049488251
GSM2385776	0.062980206
GSM2385779	0.057599575
GSM2385781	0.095140437
GSM1539877	0.062414797
GSM1539878	0
GSM1539879	0.006925985
GSM1539880	0.000979649
GSM1539881	0.00634387
GSM1539882	0
GSM1539883	0.063968964
GSM2385726	0.002002254
GSM2385728	0.003325514
GSM2385731	0.002206085
GSM2385735	0
GSM2385737	0.009507794
GSM2385738	0
GSM2385739	0
GSM2385741	0
GSM2385744	0
GSM2385747	0.009736968
GSM2385750	0.004787035
GSM2385773	0.016882027
GSM2385777	0.107066275

Table S5. Sequences of PCR primers used in the study

Gene	Species	Forward Primer	Reverse Primer
EGR1	mouse	TGACCAATCCTCCGACCTCT	AGATGGGACTGCTGTCGTTG
IL6	mouse	GCAGAAAAAGGTGGGTGTGTC	GGAAGTGGCATTGCATCCCT
NR4A1	mouse	TCCCCGAGCCAGACTTATGA	GCATGGAATAGCTCTCCCCC
SOCS1	mouse	CAACGGAACTGCTTCTTCGC	AGCTCGAAAAGGCAGTCGAA
ZFP36	mouse	CCGATCCTGATGACTACGCC	ATTGAAGATGGGGAGACGCC
JUN	mouse	TGGGCACATCACCACTACAC	TCTGGCTATGCAGTTCAGCC

Table S6. The relationship between marker genes and marker gene-targeted drugs

search_term	match_term	drug	interaction_types
SOCS1	SOCS1	INSULIN	
SOCS1	SOCS1	ALDESLEUKIN	
IL6	IL6	SILTUXIMAB	inhibitor
IL6	IL6	CLAZAKIZUMAB	inhibitor
IL6	IL6	SIRUKUMAB	inhibitor
IL6	IL6	ELSILIMOMAB	inhibitor
IL6	IL6	PF-04236921	inhibitor
IL6	IL6	OLOKIZUMAB	inhibitor
IL6	IL6	RITUXIMAB	
IL6	IL6	SAQUINAVIR	
IL6	IL6	IFOSFAMIDE	
IL6	IL6	LINEZOLID	
IL6	IL6	GEMFIBROZIL	
IL6	IL6	INFLIXIMAB	
IL6	IL6	METRONIDAZOLE	
IL6	IL6	ETANERCEPT	
IL6	IL6	ECHINACEA,UNSPECIFIED	
IL6	IL6	ADALIMUMAB	
IL6	IL6	INSULIN	
IL6	IL6	RIBAVIRIN	
IL6	IL6	IBUDILAST	
IL6	IL6	COR-001	
IL6	IL6	CISPLATIN	
IL6	IL6	NELFINAVIR	
IL6	IL6	FENTANYL	
IL6	IL6	LEVOFLOXACIN	
IL6	IL6	FENOFIBRATE	
EGR1	EGR1	GENIPIN	
JUN	JUN	BUPROPIONHYDROCHLORIDE	
JUN	JUN	BUTINOLINE	
JUN	JUN	CLOTRIMAZOLE	
JUN	JUN	BENZENETHIOL	
JUN	JUN	CARBOXYMETHYL-TRIMETHYL-ARSONIUM	
JUN	JUN	CHEMBL275260	
JUN	JUN	FENOFIBRATE	
JUN	JUN	MECHLORETHAMINEHYDROCHLORIDE	
JUN	JUN	SODIUMSELENITE	
JUN	JUN	VINORELBINETARTRATE	
JUN	JUN	TROPISETRON	
JUN	JUN	CINNARIZINE	
JUN	JUN	VINBLASTINESULFATE	
JUN	JUN	CUPRICCHLORIDE	
JUN	JUN	TRIFLUPROMAZINEHYDROCHLORIDE	
JUN	JUN	COLCHICINE	
JUN	JUN	CIPROFIBRATE	
JUN	JUN	NAFRONYLOXALATE	
JUN	JUN	BRUCEANTIN	
JUN	JUN	METHIMAZOLE	

(continued)

Table S6. (continued)

search_term	match_term	drug	interaction_types
JUN	JUN	SERGEOLIDE	
JUN	JUN	DIPHENHYDRAMINEHYDROCHLORIDE	
JUN	JUN	IRISOLIDONE	
JUN	JUN	ROTENONE	
JUN	JUN	CRIDANIMOD	
JUN	JUN	BENZO[B]FLUORANTHENE	
JUN	JUN	2-MERCAPTOPYRIMIDINE	
JUN	JUN	SANGIVAMYCIN	
JUN	JUN	ANTHRACENE-9-CARBOXYLICACID	
JUN	JUN	AMINEPTINE	
JUN	JUN	AZELASTINEHYDROCHLORIDE	
JUN	JUN	PATULIN	
JUN	JUN	CLOFIBRATE	
JUN	JUN	QUINAPRILHYDROCHLORIDE	
JUN	JUN	RETINYLRETINOATE	
JUN	JUN	LIPOICACID,ALPHA	
JUN	JUN	ISOLIQUIRITIGENIN	
JUN	JUN	ATOMOXETINEHYDROCHLORIDE	
JUN	JUN	NEOCHAMAEJASMINA	
JUN	JUN	GEMFIBROZIL	
JUN	JUN	SERTRALINE	
JUN	JUN	(-)-CAMPHOR	
JUN	JUN	CHEMBL477052	
JUN	JUN	HOLACANTHONE	
NR4A1	NR4A1	ACETYLCYSTEINE	
NR4A1	NR4A1	HALOPERIDOL	
NR4A1	NR4A1	IONOMYCIN	
NR4A1	NR4A1	CHEMBL547833	
NR4A1	NR4A1	CHEMBL35482	
NR4A1	NR4A1	ETOPOSIDEPHOSPHATE	
NR4A1	NR4A1	LEVODOPA	
NR4A1	NR4A1	NICOTINE	
NR4A1	NR4A1	MORPHINE	
NR4A1	NR4A1	CYTOSPORONEB	

Table S7. Detailed information on the ceRNA network

Node1	Node2	Interaction
IL6	hsa-miR-519e-5p	mRNA
IL6	hsa-miR-149-5p	mRNA
EGR1	hsa-miR-760	mRNA
NR4A1	hsa-miR-3152-3p	mRNA
SOCS1	hsa-miR-3163	mRNA
EGR1	hsa-miR-3119	mRNA
NR4A1	hsa-miR-342-5p	mRNA
ZFP36	hsa-miR-187-5p	mRNA
SOCS1	hsa-miR-411-3p	mRNA
EGR1	hsa-miR-3185	mRNA
SOCS1	hsa-let-7i-5p	mRNA
NR4A1	hsa-miR-637	mRNA
JUN	hsa-miR-2116-5p	mRNA
IL6	hsa-miR-1304-5p	mRNA
SOCS1	hsa-miR-548x-3p	mRNA
EGR1	hsa-miR-132-3p	mRNA
EGR1	hsa-miR-4269	mRNA
IL6	hsa-miR-548d-3p	mRNA
ZFP36	hsa-miR-934	mRNA
JUN	hsa-miR-34a-3p	mRNA
ZFP36	hsa-miR-142-5p	mRNA
ZFP36	hsa-miR-1299	mRNA
JUN	hsa-miR-940	mRNA
NR4A1	hsa-miR-4272	mRNA
EGR1	hsa-miR-936	mRNA
JUN	hsa-miR-513c-5p	mRNA
EGR1	hsa-miR-23a-3p	mRNA
SOCS1	hsa-let-7d-5p	mRNA
EGR1	hsa-miR-1264	mRNA
NR4A1	hsa-miR-2115-3p	mRNA
NR4A1	hsa-miR-3190-3p	mRNA
ZFP36	hsa-miR-361-3p	mRNA
EGR1	hsa-miR-191-5p	mRNA
SOCS1	hsa-miR-149-3p	mRNA
ZFP36	hsa-miR-1913	mRNA
SOCS1	hsa-let-7e-5p	mRNA
IL6	hsa-miR-4256	mRNA
SOCS1	hsa-miR-2113	mRNA
EGR1	hsa-miR-4251	mRNA
EGR1	hsa-miR-1911-5p	mRNA
IL6	hsa-miR-338-5p	mRNA
JUN	hsa-miR-542-3p	mRNA
IL6	hsa-miR-302a-5p	mRNA
SOCS1	hsa-miR-30c-5p	mRNA
ZFP36	hsa-miR-3163	mRNA
JUN	hsa-miR-1258	mRNA
SOCS1	hsa-miR-187-5p	mRNA
SOCS1	hsa-miR-3130-3p	mRNA
ZFP36	hsa-miR-548x-3p	mRNA
SOCS1	hsa-miR-556-3p	mRNA
JUN	hsa-miR-514b-5p	mRNA
JUN	hsa-miR-633	mRNA
IL6	hsa-miR-3123	mRNA
JUN	hsa-miR-522-3p	mRNA
SOCS1	hsa-miR-142-5p	mRNA
JUN	hsa-miR-524-5p	mRNA
IL6	hsa-miR-98-5p	mRNA
SOCS1	hsa-miR-518a-5p	mRNA
ZFP36	hsa-miR-3121-3p	mRNA
SOCS1	hsa-miR-922	mRNA
NR4A1	hsa-miR-600	mRNA
SOCS1	hsa-let-7g-5p	mRNA
SOCS1	hsa-miR-19b-3p	mRNA
SOCS1	hsa-miR-324-5p	mRNA
IL6	hsa-miR-574-3p	mRNA
NR4A1	hsa-miR-608	mRNA
SOCS1	hsa-miR-527	mRNA
SOCS1	hsa-miR-19a-3p	mRNA
IL6	hsa-miR-202-3p	mRNA
IL6	hsa-miR-3168	mRNA
EGR1	hsa-miR-23b-3p	mRNA
SOCS1	hsa-miR-3144-5p	mRNA
EGR1	hsa-miR-204-5p	mRNA
IL6	hsa-miR-760	mRNA
SOCS1	hsa-miR-766-3p	mRNA
NR4A1	hsa-miR-506-3p	mRNA
EGR1	hsa-miR-211-5p	mRNA
ZFP36	hsa-miR-16-2-3p	mRNA
JUN	hsa-miR-200c-3p	mRNA
JUN	hsa-miR-200b-3p	mRNA
ZFP36	hsa-miR-524-5p	mRNA
JUN	hsa-miR-501-5p	mRNA
JUN	hsa-miR-637	mRNA
EGR1	hsa-miR-4271	mRNA
IL6	hsa-miR-548c-3p	mRNA
EGR1	hsa-miR-524-5p	mRNA
IL6	hsa-miR-515-5p	mRNA
SOCS1	hsa-miR-155-5p	mRNA
EGR1	hsa-miR-377-3p	mRNA
NR4A1	hsa-miR-524-5p	mRNA
SOCS1	hsa-miR-30d-5p	mRNA
SOCS1	hsa-let-7b-5p	mRNA
JUN	hsa-miR-1299	mRNA
SOCS1	hsa-miR-98-5p	mRNA
EGR1	hsa-miR-369-3p	mRNA
EGR1	hsa-miR-212-3p	mRNA

(continued)

Table S7. (continued)

Node1	Node2	Interaction
ZFP36	hsa-miR-615-5p	mRNA
SOCS1	hsa-miR-218-1-3p	mRNA
ZFP36	hsa-miR-513a-5p	mRNA
ZFP36	hsa-miR-195-3p	mRNA
SOCS1	hsa-miR-30b-5p	mRNA
JUN	hsa-miR-1972	mRNA
SOCS1	hsa-miR-548g-3p	mRNA
ZFP36	hsa-miR-182-5p	mRNA
ZFP36	hsa-miR-3140-3p	mRNA
EGR1	hsa-miR-506-3p	mRNA
NR4A1	hsa-miR-665	mRNA
NR4A1	hsa-miR-200a-5p	mRNA
NR4A1	hsa-miR-34a-3p	mRNA
JUN	hsa-miR-495-3p	mRNA
SOCS1	hsa-miR-331-3p	mRNA
EGR1	hsa-miR-2110	mRNA
ZFP36	hsa-miR-145-5p	mRNA
JUN	hsa-miR-298	mRNA
IL6	hsa-miR-607	mRNA
IL6	hsa-miR-148b-5p	mRNA
SOCS1	hsa-miR-561-3p	mRNA
JUN	hsa-miR-429	mRNA
ZFP36	hsa-miR-4267	mRNA
EGR1	hsa-miR-600	mRNA
ZFP36	hsa-miR-875-3p	mRNA
JUN	hsa-miR-758-3p	mRNA
EGR1	hsa-miR-125b-2-3p	mRNA
IL6	hsa-miR-1323	mRNA
ZFP36	hsa-miR-625-3p	mRNA
SOCS1	hsa-let-7f-5p	mRNA
EGR1	hsa-miR-581	mRNA
SOCS1	hsa-miR-379-3p	mRNA
SOCS1	hsa-miR-569	mRNA
EGR1	hsa-miR-16-1-3p	mRNA
JUN	hsa-miR-139-5p	mRNA
EGR1	hsa-miR-30d-3p	mRNA
SOCS1	hsa-let-7a-5p	mRNA
EGR1	hsa-miR-3121-3p	mRNA
MUC2	hsa-miR-615-5p	lncRNA
CDR1-AS	hsa-miR-875-3p	lncRNA
RP5-894D12.5	hsa-miR-665	lncRNA
TTLL10-AS1	hsa-miR-515-5p	lncRNA
PAX8-AS1	hsa-miR-615-5p	lncRNA
C10orf91	hsa-miR-149-3p	lncRNA
MUC19	hsa-miR-145-5p	lncRNA
AC079779.7	hsa-miR-1972	lncRNA
LINC01043	hsa-miR-149-5p	lncRNA
AC079779.7	hsa-miR-515-5p	lncRNA
RP11-102K13.5	hsa-miR-1972	lncRNA
LL22NC03-27C5.1	hsa-miR-377-3p	lncRNA
RP11-627G23.1	hsa-miR-665	lncRNA
RP5-894D12.5	hsa-miR-1972	lncRNA
RP13-580B18.4	hsa-miR-1972	lncRNA
LINC01070	hsa-miR-766-3p	lncRNA
RP3-470B24.5	hsa-miR-361-3p	lncRNA
RP11-333E1.2	hsa-miR-125b-2-3p	lncRNA
RP4-737E23.2	hsa-miR-377-3p	lncRNA
RP11-982M15.8	hsa-miR-2113	lncRNA
FAM182A	hsa-miR-1972	lncRNA
RP13-507P19.2	hsa-miR-766-3p	lncRNA
RP13-507P19.2	hsa-miR-515-5p	lncRNA
LINC01002	hsa-miR-1972	lncRNA
RP11-1228E12.1	hsa-miR-1972	lncRNA
MUC19	hsa-miR-766-3p	lncRNA
CTD-3138B18.5	hsa-miR-377-3p	lncRNA
LINC01001	hsa-miR-1972	lncRNA
RP13-507P19.2	hsa-miR-361-3p	lncRNA
CTA-941F9.10	hsa-miR-149-3p	lncRNA
RP11-394A14.2	hsa-miR-149-5p	lncRNA
CTD-2008P7.3	hsa-miR-766-3p	lncRNA
LINC01043	hsa-miR-324-5p	lncRNA
LINC00174	hsa-miR-1972	lncRNA
RP11-210M15.1	hsa-miR-377-3p	lncRNA
RP13-580B18.4	hsa-miR-766-3p	lncRNA
AIRN	hsa-miR-149-3p	lncRNA
AC078942.1	hsa-miR-766-3p	lncRNA
RP11-46C24.3	hsa-miR-342-5p	lncRNA
PCBP3-OT1	hsa-miR-875-3p	lncRNA
CTB-51J22.1	hsa-miR-665	lncRNA
RP11-94C24.13	hsa-let-7a-5p	lncRNA
LINC01002	hsa-miR-377-3p	lncRNA
RP13-580B18.4	hsa-miR-515-5p	lncRNA
RP4-671O14.7	hsa-miR-542-3p	lncRNA
AC079586.1	hsa-miR-542-3p	lncRNA
RP11-231G3.1	hsa-miR-561-3p	lncRNA
RP11-54O7.17	hsa-miR-324-5p	lncRNA
LINC00265	hsa-miR-149-3p	lncRNA
RP11-311F12.1	hsa-miR-149-3p	lncRNA
COL4A2-AS2	hsa-miR-615-5p	lncRNA
RP11-1217F2.15	hsa-miR-766-3p	lncRNA
CTD-3193O13.12	hsa-miR-766-3p	lncRNA
RP11-157B13.7	hsa-miR-542-3p	lncRNA
LINC00689	hsa-miR-149-3p	lncRNA
AC005264.2	hsa-miR-2113	lncRNA
CTD-3099C6.5	hsa-miR-145-5p	lncRNA
CTD-3099C6.5	hsa-miR-16-1-3p	lncRNA

(continued)

Table S7. (continued)

Node1	Node2	Interaction
AC011284.3	hsa-miR-342-5p	lncRNA
SPACA6P	hsa-miR-515-5p	lncRNA
AC093642.4	hsa-miR-515-5p	lncRNA
LINC01043	hsa-miR-218-1-3p	lncRNA
MUC2	hsa-miR-342-5p	lncRNA
LINC00917	hsa-miR-542-3p	lncRNA
RP11-85G18.6	hsa-miR-766-3p	lncRNA
CH507-216K13.2	hsa-miR-1972	lncRNA
RP11-394A14.2	hsa-miR-342-5p	lncRNA
LINC01022	hsa-miR-766-3p	lncRNA
FAM74A1	hsa-miR-515-5p	lncRNA
CTD-2330K9.2	hsa-miR-1972	lncRNA
LINC00905	hsa-miR-1972	lncRNA
RP11-64K12.8	hsa-miR-875-3p	lncRNA
RP11-343D2.11	hsa-miR-302a-5p	lncRNA
FRMPD3-AS1	hsa-miR-875-3p	lncRNA
RP11-142C4.6	hsa-miR-766-3p	lncRNA
LINC01002	hsa-miR-766-3p	lncRNA
CTA-315H11.2	hsa-miR-149-3p	lncRNA
RP11-153F5.7	hsa-miR-149-3p	lncRNA
LINC00173	hsa-miR-149-3p	lncRNA
AC079586.1	hsa-miR-758-3p	lncRNA
RP13-895J2.3	hsa-miR-665	lncRNA
RP11-1391J7.1	hsa-miR-342-5p	lncRNA
AC015849.16	hsa-miR-139-5p	lncRNA
RP11-54O7.17	hsa-miR-342-5p	lncRNA
TMEM191A	hsa-miR-149-3p	lncRNA
RP11-32B5.8	hsa-miR-615-5p	lncRNA
RP11-50B3.4	hsa-miR-16-1-3p	lncRNA
RP11-186N15.3	hsa-miR-149-3p	lncRNA
RP13-582L3.4	hsa-miR-665	lncRNA
RP11-142C4.6	hsa-miR-1972	lncRNA
LINC01002	hsa-miR-515-5p	lncRNA
RP11-34P13.7	hsa-miR-182-5p	lncRNA
LINC01224	hsa-miR-542-3p	lncRNA
CTD-2311B13.1	hsa-miR-766-3p	lncRNA
CTA-390C10.9	hsa-miR-145-5p	lncRNA
RP11-44M6.7	hsa-miR-331-3p	lncRNA
RP11-504P24.8	hsa-miR-1972	lncRNA
CTC-338M12.9	hsa-miR-766-3p	lncRNA
RP11-849H4.4	hsa-miR-1972	lncRNA
AC006019.3	hsa-miR-760	lncRNA
RP11-211G23.2	hsa-miR-218-1-3p	lncRNA
RP11-231D20.2	hsa-miR-139-5p	lncRNA
FENDRR	hsa-miR-182-5p	lncRNA
SNHG14	hsa-miR-665	lncRNA
AP001631.9	hsa-miR-766-3p	lncRNA
RP11-54O7.17	hsa-miR-615-5p	lncRNA
CTD-3193O13.1	hsa-miR-149-3p	lncRNA
HOXC-AS1	hsa-miR-361-3p	lncRNA
LINC01224	hsa-miR-758-3p	lncRNA
CH17-360D5.1	hsa-miR-665	lncRNA
AC010524.2	hsa-miR-182-5p	lncRNA
TTN-AS1	hsa-miR-766-3p	lncRNA
CTD-2619J13.19	hsa-miR-34a-3p	lncRNA
RP11-394A14.2	hsa-miR-760	lncRNA
GS1-279B7.1	hsa-miR-34a-3p	lncRNA
CTD-2197I11.1	hsa-miR-515-5p	lncRNA
MIR497HG	hsa-miR-342-5p	lncRNA
RP11-1217F2.15	hsa-miR-515-5p	lncRNA
MAFG-AS1	hsa-miR-149-3p	lncRNA
RP11-717I24.1	hsa-miR-145-5p	lncRNA
RP11-158I9.8	hsa-miR-361-3p	lncRNA
CTD-2008P7.3	hsa-miR-149-5p	lncRNA
AC114808.3	hsa-miR-665	lncRNA
RP11-1191J2.2	hsa-miR-1972	lncRNA
AC015849.16	hsa-miR-145-5p	lncRNA
CTD-3099C6.5	hsa-miR-766-3p	lncRNA
CTD-3099C6.5	hsa-miR-515-5p	lncRNA
RP11-148K1.12	hsa-miR-149-3p	lncRNA
AC006019.3	hsa-miR-342-5p	lncRNA
SNHG14	hsa-miR-515-5p	lncRNA
RP11-430G17.3	hsa-miR-149-3p	lncRNA
RP11-1348G14.8	hsa-miR-149-3p	lncRNA
LINC00661	hsa-miR-1972	lncRNA
AC084219.4	hsa-miR-515-5p	lncRNA
AC015849.16	hsa-miR-515-5p	lncRNA
RP11-44M6.7	hsa-miR-361-3p	lncRNA
RP11-561O23.5	hsa-miR-361-3p	lncRNA
PAX8-AS1	hsa-miR-149-3p	lncRNA
CTD-2008P7.1	hsa-miR-766-3p	lncRNA
RP4-751H13.7	hsa-miR-361-3p	lncRNA
FAM74A7	hsa-miR-515-5p	lncRNA
CTD-2521M24.5	hsa-miR-518a-5p	lncRNA
RASSF8-AS1	hsa-miR-665	lncRNA
RP11-630C16.2	hsa-miR-149-3p	lncRNA
RP11-157B13.7	hsa-miR-758-3p	lncRNA
ABHD11-AS1	hsa-miR-766-3p	lncRNA
RP11-700J17.1	hsa-miR-556-3p	lncRNA
RP11-15H20.6	hsa-miR-758-3p	lncRNA
FAM95B1	hsa-miR-1972	lncRNA
CTD-2369P2.8	hsa-miR-149-3p	lncRNA
LA16c-306A4.2	hsa-miR-922	lncRNA
RP11-458F8.4	hsa-miR-940	lncRNA
RP11-347H15.4	hsa-miR-766-3p	lncRNA
RP4-539M6.22	hsa-miR-922	lncRNA
AP001476.4	hsa-miR-940	lncRNA
LINC00265	hsa-miR-940	lncRNA