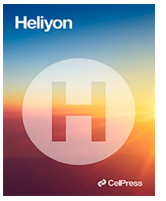
Heliyon 11 (2025) e41705

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Identification of tacrolimus-related genes in familial combined hyperlipidemia and development of a diagnostic model using bioinformatics analysis

Yuan Xu [a], Hongfei He [b], Haiyang Li [b,*]

[a] Department of Organ Transplantation, Affiliated Hospital of Guizhou Medical University, China
[b] Guizhou Organ Transplant Centre, China

ARTICLE INFO

Keywords:
Familial combined hyperlipidemia
Hypertriglyceridemia
Lipid metabolism
Immune infiltration
Tacrolimus
Co-expressed genes

ABSTRACT

Background: Clinical observations have revealed that patients undergoing organ transplantation administered tacrolimus often experience abnormal lipid metabolism with serious consequences. Thus, the intricate interplay between tacrolimus and lipid metabolism must be addressed to develop targeted therapeutic interventions. Our ongoing research aims to develop precision medicine approaches that not only alleviate the immediate repercussions for organ transplant patients but also enhance their long-term outcomes. To this end, we investigated the potential genes associated with tacrolimus metabolism in familial combined hyperlipidemia (FCHL) to identify relevant biomarkers of FCHL, develop predictive diagnostic models for hyperlipidemia, and reveal potential therapeutic targets for FCHL.

Methods: Dataset GSE1010 containing information on patients diagnosed with FCHL was obtained from the Gene Expression Omnibus (GEO), and an ensemble of tacrolimus-related genes (TRGs) was retrieved from the GeneCards, STITCH, and Molecular Signatures Database databases. A thorough weighted gene co-expression network analysis was conducted, including a differential expression analysis of the GSE1010 and TRG datasets, to identify intricate patterns of gene co-expression and provide insights on the underlying molecular dynamics within the datasets. Key genes were screened, diagnostic models were constructed, and all genes associated with logFC values were assessed using gene set variation and enrichment analyses. Upregulated genes were identified by a positive logFC (>0) and $P < 0.05$, while downregulated genes were characterized by a negative logFC (<0) and $P < 0.05$. These criteria facilitated a more nuanced categorization of gene expression changes within the analyzed datasets. Given tacrolimus's immunosuppressive impact, the gene expression matrix data obtained from dataset GSE1010 was submitted to CIBERSORT to assess immune cell infiltration outcomes. Finally, we examined the regulatory network of screened key genes that interact with RNA-binding proteins, potential drugs, small-molecule compounds, and transcription factors.

Results: We screened 14 statistically significant key genes, built a reliable risk model, and grouped the dataset into categories at high and low risk for hyperlipidemia development. FCHL was linked to memory B and immature B immune cells. The gene set variation analysis revealed two pathways associated with cholesterol homeostasis and the complement system that were closely associated with the potential functions of FCHL and tacrolimus-related differentially expressed genes.

* Corresponding author.
E-mail address: xuyuan@gmc.edu.cn (H. Li).

https://doi.org/10.1016/j.heliyon.2025.e41705
Received 27 April 2024; Received in revised form 3 October 2024; Accepted 3 January 2025
Available online 19 January 2025

Conclusions: Our research offers a better understanding of FCHL and the TRGs involved in lipid metabolism. Additionally, it provides research directions for identifying potential targets for clinical therapies.

1. Background

Hyperlipidemia commonly develops subsequent to renal transplantation and is associated with a deterioration in the patient's prognosis [1]. Hypertriglyceridemia (HTG) is a common cause of atherosclerosis and cardiovascular disease [2,3]. Familial combined hyperlipidemia (FCHL) stands out as the predominant form of genetic dyslipidemia (11.3 %) in individuals below the age of 60 diagnosed with coronary heart disease, and it presents a prevalence of 1–2% within the general population [4]. However, the pathogenesis of abnormal lipid metabolism remains uncertain and may be connected to diet, obesity, and genetic factors [5–9].Dyslipidemia has emerged as a significant contributor to atherosclerotic cardiovascular disease, which leads to critical life-threatening conditions [10]. Reports have suggested that FCHL is brought on by changes in the lipoprotein lipase (LPL), ApoC2, or ApoA5 genes, all of which are involved in triglyceride (TG) metabolism [11] and usually characterized by multiple mutations [12].Triglycerides are a type of glyceride produced by the esterification of three hydroxyl groups in glycerol with three fatty acid molecules. A recent Mendelian randomized study found that genetic variants of LPL, which is responsible for TG hydrolysis, and the low-density lipoprotein (LDL) receptor, which is involved in LDL metabolism, are closely associated with lowering the risk of heart attacks [13]. The principal advantage of current lipid-lowering medications, as exemplified by statins, lies in their ability to decrease LDL-C levels. However, adverse consequences, including liver function abnormalities, and rhabdomyolysis, are often a concern in clinical practice. To address these challenges, lipid-lowering drugs that act on alternative targets are constantly being introduced. For example, the US FDA approved the following two noteworthy agents for the treatment of homozygous familial hypercholesterolemia as early as 2012 and 2013: lomitapide, an inhibitor of microsomal TG transfer protein, and mipomersen, an inhibitor of ApoB100 synthesis [14,15]. Moreover, angiopoietin-like protein 3 plays a pivotal role in very-low-density lipoprotein (VLDL) metabolism, assuming a regulatory function by impeding the activity of LPL. VLDL is composed of intermediate metabolites of celiac granule remnants, bile acids, fatty acids, sugars, proteins and apolipoproteins synthesized in the liver. Its density is 0.96–1.063 mg/dl and contains 60 % triacylglycerols. Phase 2 and 3 clinical trials are currently ongoing for evinacumab, a monoclonal antibody designed for humans that specifically targets angiopoietin-like protein 3 [16]. Apolipoprotein C3 (ApoC3), a key player in the regulation of chylomicron and VLDL metabolism, exerts its influence by suppressing LPL and hepatic lipase activities, thus underscoring its critical role in these intricate metabolic pathways. Volanesorsen, a 2nd-generation antisense oligonucleotide of ApoC3, induced a significant 77 % decrease in TG levels in a phase 3 clinical investigation; however, it was not approved by the US FDA because of the risk of severe thrombocytopenia [17]. These studies underscore the importance of research and development to identify effective targets for regulating lipid metabolism and lowering TG levels. In this study, we concentrated on the metabolism of triglycerides, specifically the dysregulation of lipid metabolism following the administration of tacrolimus post-transplantation, leading to hypertriglyceridemia.

Tacrolimus (TAC) is a macrolide antibiotic isolated from *Streptomyces* spp. That is widely used as a potent immunosuppressant in solid organ transplant recipients. TAC inhibits the release of interleukin-2 (IL-2), which comprehensively suppresses T lymphocytes. Moreover, TAC has a biphasic distribution after absorption in the human body, with most of the TAC (98.8 %) in plasma bound to plasma proteins, namely, serum albumin and α-1-acid glycoprotein, and TAC in erythrocytes bound to TAC-binding proteins (FKBP) (mainly FKBP-12) [18,19]. In addition, TAC metabolism is dependent on CYP3A4 and CYP3A5 within the hepatic and intestinal walls. A comprehensive CYP3A4/5 whole-gene analysis was previously conducted on a cohort of 2595 kidney transplant recipients with diverse ancestral backgrounds, including European, African, Native American, and Asian backgrounds, and it revealed an association between the CYP3A5*3 variant (rs776746) and elevated doses of standardized tacrolimus (TAC) troughs across all ethnic groups. Notably, this association was manifested by varying allele frequencies and effect sizes [20]. Furthermore, certain genes implicated in carbohydrate metabolism, such as the LEPR rs1137101 G allele and CDKAL1 rs10946398 cellular components (CC) genotype, have been connected to a higher likelihood of developing diabetes mellitus in patients administered TAC treatment during renal transplantation [21,22]. A study investigating the disruption in lipid metabolism resulting from the synergy of TAC and hormones identified CircFOXN2 as a promising intervention. This circRNA is a regulator that alleviates the adverse effects of disrupted lipid metabolism by modulating the polypyrimidine tract binding protein 1 (PTBP1) and fatty acid synthase (FASN) axis [23]. Kidney transplant recipients require long-term immunosuppressive medications and present a high prevalence of HTG [24], which is experienced by 27%–71 % of organ transplant patients [25]. Moreover, cardiovascular disease is the primary cause of mortality among heart and kidney transplant recipients [26] and second most prevalent cause of death among liver transplant recipients [27]. Beyond immediate issues related to the transplant, cardiovascular disease persists as a significant contributor to long-term mortality among transplant patients. With the advent of solid organ transplantation, new-onset metabolic syndrome after organ transplantation has been increasingly observed as a common complication that is significantly associated with a higher prevalence of heart-related conditions [28].An increasing number of transplant physicians have noted a correlation between post-transplant hyperlipidemia and cardiovascular disease, suggesting a potential link to tacrolimus and its molecular partner, FK506-binding protein 51 (FKBP51) [29–35].Although a multitude of risk factors increase the possibility of developing cardiovascular disease following transplantation, dyslipidemia remains one of the most common and modifiable risks [36]. In our earlier single-center study of kidney transplant recipients, we observed a consistent correlation between TGs and abnormally high TAC drug concentrations. Additionally, a significantly higher incidence of postoperative HTG in renal transplant recipients with blood type O than in those with other blood types was observed [37]. Researchers have

highlighted a correlation between genetic polymorphisms and an elevated susceptibility to developing metabolic syndrome in recipients of solid organ transplants [38] and have suggested that the prevalence of HTG among kidney transplant recipients might be functionally regulated by genes or related to immunosuppressive drug administration or immune cell function; thus, they made an effort to investigate the connection among the three in terms of gene pooling.

We used GSE1010 data in the manuscript from a study of familial mixed hyperlipidemia (FCHL). Metabolic disorders caused by FCHL and TAC (tacrolimus) involve similar changes in lipid metabolic pathways, with hypertriglyceridemia (HTG) being a typical manifestation of both.The administration of TAC may elicit metabolic derangements akin to FCHL. Therefore, elucidating the gene regulatory mechanisms in FCHL is crucial for comprehending the lipid metabolic alterations induced by TAC. Our examination of the GSE1010 data facilitated the identification of key genes implicated in FCHL that could be subject to altered regulation during prolonged TAC therapy, potentially contributing to lipid metabolism disorders.Although the GSE1010 dataset is centered on familial combined hyperlipidemia (FCHL), uncovering the gene expression patterns and regulatory mechanisms within lipid metabolism offers a foundation for extrapolating insights to tacrolimus (TAC)-related metabolic research. Leveraging the broad utility of bioinformatics, we can harness FCHL data to uncover the mechanisms behind TAC-induced lipid metabolism disorders, potentially pinpointing novel targets for intervention. Consequently, we anticipate that this study will yield fresh perspectives and therapeutic strategies aimed at the clinical management and prevention of lipid metabolic disturbances in organ transplant patients treated with TACs.

In this study, TAC-related genes (TRGs) retrieved from expression profile data of FCHL patients from different databases, including GEO, GeneCards [39], STITCH, and online Molecular Signatures Database (MSigDB) [40], were analyzed to further investigate the pathogenesis of TAC action, immune cell function, and hyperlipidemia development. The results presented here are important for preserving human health as well as preventing and managing cardiovascular disease.

2. Materials and methods

2.1. Data download and pre-processing

Our initial exploration involved querying the GEO database [41] based on the R package GEOquery for efficient retrieval [42]. The expression profiling dataset GSE1010 of human patients with FCHL was downloaded.

The GSE1010 dataset was sampled from lymphoblastoid cells or cell lines from patients with FCHL and healthy donors, and it contained a total of 24 data samples. A set of 12 data samples was obtained from patients diagnosed with FCHL, which constituted the FCHL subgroup, and another set of 12 data samples was collected from healthy donors, which constituted the control subgroup. The entire sample was included in the analysis.

The GSE1010 dataset utilized the GPL96 [HG-U133A] Affymetrix Human Genome U133A Array as its data platform. Annotation of dataset probe names was performed using the corresponding chip GPL platform file. See Table 1 for specific dataset information.

The GeneCards [39] database (https://www.genecards.org/) was searched using the phrase "Tacrolimus," and only "Protein Coding" results were retained. In total, 1635 TRGs were identified.

The public Comparative Toxicogenomics Database (CTD) [43] (http://ctdbase.org/) was searched using the term "Tacrolimus," and 475 TRGs were identified.The Comparative Toxicogenomics Database (CTD) facilitates the screening of potential drugs or small molecule compounds that interact with key genes. It evaluates mRNA-drug interactions by referencing the associated literature count. Furthermore, We employs Cytoscape software to visualize these interactions within a network framework.

The STITCH [44] online website (http://stitch.embl.de/) was searched using the term "Tacrolimus" and the default parameters, and 10 TRGs were identified.

The ENCORI database, known as starBase 3.0, is engineered to predict interactions between microRNAs (miRNAs), RNA-binding proteins (RBPs), and key genes. It enables the screening of interacting pairs involving miRNA-mRNA and mRNA-RBP. Additionally,we utilizes Cytoscape software to visualize the interaction networks, providing a comprehensive view of the molecular relationships [45].

CHIPBase 3.0 utilizes ChIP-seq data to identify binding sites and elucidate the transcriptional regulatory relationships between transcription factors and their target genes. We also employs Cytoscape software to visualize the interaction networks, offering a graphical representation of these molecular associations [46].Related gene sets in the MSigDB [40] were looked up utilizing the tacrolimus alias "FK506," and the BIOCARTA_CALCINEURIN_PATHWAY gene set containing 18 TRGs was identified.

Finally, published literature on the PubMed website [47] was searched, and 58 TRGs were identified. After combining and de-emphasizing all the TRGs, we obtained 1967 TRGs, as shown in Table 2.

Table 1
Familial combined hyperlipemia data set information list.

	GSE1010
Platform	GPL96
Species	*Homo sapiens*
Tissue	lymphoblastic cells or cell lines
Samples in FCHL group	12
Samples in Control group	12
Reference	–

FCHL, familial combined hyperlipemia; GEO, Gene Expression Omnibus.

2.2. Analysis of differential expression

To reveal the functional modalities of genes and associated biological traits and pathways in FCHL, differentially expressed genes (DEGs) were identified using limma software [48]. This analysis was conducted across diverse subgroups (FCHL/control) within the GSE1010 dataset, employing stringent criteria of $|\log FC| > 0$ and $P < 0.05$. The genes meeting these criteria were selectively designated as DEGs for subsequent in-depth analyses. Notably, genes with logFC >0 and $P < 0.05$ were classified as upregulated, indicating heightened regulatory levels, while those with logFC <0 and $P < 0.05$ were classified as downregulated. This foundational step laid the groundwork for a more nuanced exploration of the intricate molecular landscape associated with FCHL.

Tacrolimus-induced differential gene expression was determined by identifying genes meeting the criteria of $|\log FC| > 0$ and $P < 0.05$ in both the GSE1010 and TRG datasets. Identifying overlaps among these genes provided a robust foundation for subsequent analyses, and such connections were visually depicted through Venn diagrams. The differential analysis outcomes were comprehensively presented using volcano charts via the R tool ggplot2, and gene expression patterns were displayed through heat maps generated with the R package pheatmap. This methodological approach enhances the clarity and depth of the insights gained into the genomic landscape influenced by tacrolimus across datasets.

2.3. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis

To gain comprehensive functional enrichment insights, a GO analysis was performed [49], which included biological processes (BP), molecular functions (MF), and CC categories. Simultaneously, a KEGG [50] analysis of genome-related information, pathways, diseases, and drugs was performed. Using the R package clusterProfiler [51], we conducted GO and KEGG annotation analyses of tacrolimus-related differentially expressed genes (TRDEGs). Rigorous criteria of P.adj <0.05 and FDR (q.value) < 0.05 were applied to meticulously screen entries for statistical significance. This strategic approach minimized redundancy and provided a more comprehensive insights on the functional landscape of TRGs.

2.4. Weighted gene co-expression network analysis (WGCNA)

WGCNA [52] uses the expression correlation coefficients normalized for each gene to evaluate the co-expression relationships between genes. Significant links between genes and co-expression are defined as modules. We used the WGCNA package in R [53] for the analysis, with an RsquaredCut parameter of 0.90, a minimum threshold of 30 genes for module inclusion, and a module merge shear height of 0.2. The co-expression modules associated with genes of different subgroups (FCHL/control) from the GSE1010 dataset samples were thus identified. The intersection between TRDEGs obtained from the differential analysis and module genes whose correlations met the requirements ($|\text{cor}| > 0.3$, $P < 0.05$) was determined to obtain the intersection genes (model genes).

2.5. Screening of key genes

To screen key genes and construct diagnostic models, we first performed a one-way logistic (logistic) regression on TRDEGs, using $P < 0.05$ as the screening criterion, and included the screened TRDEGs in the subsequent random forest (RF) analysis.

RF [54] is an algorithm that integrates multiple decision trees via integration learning. It is a bagging integration method, which also includes self-sampling integration and bootstrap aggregation. We used the randomForest package(https://posit.cloud/) [55] to construct a model with TRDEG expression obtained from one-factor logistic regression screening of the expression matrix of the dataset GSE1010, with the parameter set.seed (234) and ntree = 500. MeanDecreaseGini is the average decrease in the Gini coefficient, which indicates the impurity of the node. As the Gini coefficient increases, the purity decreases and impurity increases. Therefore, MeanDecreaseGini represents the average decrease in impurities of the variable separating nodes for all trees, with higher MeanDecreaseGini values indicating variables with greater importance to the grouping. A tradeoff of the number of variables was then conducted by performing five 10-fold cross-validations in conjunction with the cross-validation curves. We used the training set to perform the cross-validation and only retained a limited number of variables in comparison to relatively small errors, and the important variables were selected for subsequent analyses based on MeanDecreaseGini.

Then, we applied the R package glmnet [56] for a Least Absolute Shrinkage and Selection Operator (LASSO) regression, with the parameter set.seed(500) guided by the outcomes of the RF screening [57]. The regression analysis was performed with a run period of 200 to avoid overfitting. LASSO regression operates on the principles of linear regression but introduces a penalty term (lambda × absolute value of slope) to curb overfitting, which simultaneously improves the model's generalization capabilities. The findings from the LASSO regression analysis were effectively visualized through diagnostic model plots and variable trajectory plots. TRDEGs included in the final LASSO regression model were the key genes in our subsequent analyses.

2.6. Key genes for constructing diagnostic logistic regression models

Logistic models are frequently employed to examine how independent and dependent variables are related when the dependent variable is a dichotomous variable. We included all key genes, constructed a multifactorial logistic model to obtain the coefficients of each key gene, and then multiplied by the matching phrase to determine each sample's risk score. Next, using the median value of the RiskScore (high, low), we carefully divided the illness groups within the dataset into distinct high- and low-risk categories. The intricate RiskScore, a pivotal metric, was then calculated utilizing a sophisticated formula designed to encapsulate the multifaceted

nature of the dataset:

$$\text{Risk}Score = \sum_i Coefficient\ (gene_i) * mRNA\ Expression\ (gene_i)$$

A nomogram [58] is a graphical representation in a planar rectangular coordinate system, consisting of non-intersecting line segments, and it can reveal the intricate relationships among various independent variables. To illustrate the interconnections among pivotal genes incorporated in the multifactor logistic model, a nomogram was constructed using the results derived from the multifactorial logistic model. This visualization was accomplished through the implementation of the R package rms.

Next, using the outcomes of the multifactor logistic regression analysis, we conducted a calibration analysis and used a calibration curve as a visual aid to gauge the model's precision and discriminative ability.

A straightforward technique for assessing biological markers, diagnostic tests, and clinical prediction models is the decision curve analysis (DCA). DCA [59] plots were constructed based on key genes in dataset GSE1010 using the R package ggDCA.

Finally, a receiver operating characteristic (ROC) curve for the logistic risk score (RiskScore) in the GSE1010 dataset was generated using the R package pROC. Additionally, the area under the ROC curve (AUC) was computed to evaluate the diagnostic efficacy of the logistic risk score (RiskScore) in predicting the onset of FCHL. The AUC, which typically ranges from 0.5 to 1, serves as a metric for diagnostic accuracy, with a higher AUC value indicating a more favorable diagnostic effect. AUC values close to 1 indicate extremely high accuracy, values between 0.7 and 0.9 indicate moderate accuracy, and values from 0.5 to 0.7 indicate low accuracy.

2.7. Differential expression validation and functional similarity analysis of key genes

Differences in the expression patterns of key genes between different subgroups (FCHL/control) within the FCHL dataset GSE1010 were then examined based on the Mann–Whitney *U* test (Wilcoxon rank sum test). The findings were then compared using the R package ggplot2.

The pROC package of R [60] has emerged as an indispensable coordinate graphical analysis tool that can identify the optimal threshold for a given model and eliminate second-best models. Thus, it is vital for selecting optimal models.

Subsequently, the Spearman algorithm was applied to identify connections between key genes and their expression in the GSE1010 dataset. Using the R package pheatmap, a correlation heatmap was created to display the correlation analysis findings. Absolute values of correlation coefficients below 0.3 were considered weak or uncorrelated; values between 0.3 and 0.5 were considered weakly correlated; values between 0.5 and 0.8 were considered moderately correlated; and those above 0.8 were considered highly correlated.

Semantic comparisons of GO annotations offer a quantitative method of calculating the similarity of genes and genomes, and they are now a crucial part of many bioinformatic research techniques. We used the GOSemSim package [61] to determine the GO semantic similarity among important genes, further calculate the geometric mean of key genes at the BP, CC, and MF levels to obtain the final scores, and determine the mean value of the similarity scores between each key gene and others, arranging them in a descending order. The ggplot tool was used to illustrate the functional similarity analysis findings and highlight the key genes.

Finally, the R package RCircos [62] was utilized for of chromosomal localization mapping to obtain the locations of key genes on the chromosome.

2.8. Gene set enrichment analysis (GSEA)

In this study, we initially assessed genes based on their logFC values and then performed an enrichment analysis using the clusterProfiler package to identify statistically significant differences within a predetermined set of genes across two distinct biological states. GSEA [63] is a computational approach widely acknowledged for its efficacy in identifying shifts in the activity levels of pathways and biological processes within expansive expression datasets. The following GSEA parameters were used: a seed value of 2022, 5000 permutations for calculations, and a minimum of 10 genes and a maximum of 500 genes in each gene set. The gene set utilized in this analysis, denoted as "c2.all.v2022.1.Hs.symbols.gmt [Curated/Pathway] (6449)," was sourced from the MSigDB [40]. Significant enrichment was defined by the dual thresholds of P.adj <0.05 and FDR value (q.value) < 0.05. This stringent screening process ensured robust identification of gene sets exhibiting substantial enrichment in the analyzed biological contexts.

2.9. Gene set variation analysis (GSVA)

GSVA [64] was applied to the gene set "c2.cp.all.v2022.1.Hs.symbols.gmt" from the MSigDB database to identify subtle variations in gene set activity and reveal the intricate connections encoded within the gene sets curated within the database. All genes in the different subgroups of the FCHL dataset were analyzed, and functional enrichment differences in the genes among the different subgroups were calculated. $P < 0.05$ was applied as the criterion for enrichment screening.

2.10. Immune infiltration analysis

CIBERSORT [65] is an algorithm for immune infiltration analysis. It was used to analyze gene expression matrix data from the GSE1010 dataset. This procedural step is akin to providing a key to unlock biological insights, and it provided a comprehensive evaluation of immune cell composition and abundance within the intricate genomic landscape. Subsequently, we integrated this dataset with the LM22 feature gene matrix to enhance the precision and comprehensiveness of immune cell composition assessment in

the context of the provided gene expression data. Data points were selectively retained, and the outcomes of this filtration process were illustrated. The percentage of immune cell infiltration percentages was visualized through a stacked histogram, showcasing variations between the subgroups categorized. The Wilcoxon rank sum test was used to identify differences in immune cell infiltration abundance associated with varying logistic regression risk scores among the distinct subgroups within the GSE1010 dataset, and the results were visually depicted using boxplot plots. Significant differences in high and low logistic regression risk scores among the subgroups and key genes were determined, and the R program ggplot2 was used to create correlation point plots to illustrate these differences, with correlations calculated by Pearson's correlation analysis.

For the quantification of relative immune cell infiltration abundance, we employed the single-sample gene set enrichment analysis (ssGSEA) algorithm [66]. Utilizing the R package GSVA, our ssGSEA algorithm calculated enrichment scores to effectively represent the extent of infiltration for various immune cell types. Boxplots were used to visually represent the variations in infiltration abundance of the 28 immune cell types between the subgroups characterized by high and low logistic regression risk scores. Next, we computed the connection between immune cells with significant differences in the high and low subgroups of logistic regression risk score and key genes by combining the gene expression matrix of dataset GSE1010, and the outcomes were demonstrated via correlation dot plotting with the R package ggplot2.

2.11. Regulatory network analysis of key genes

Version 3.0 of the starBase database, the ENCORI database [67], provides a number of visual interfaces for miRNA target exploration. We searched miRNA–mRNA interaction pairings using pancancerNum >10 as a screening threshold and predicted miRNAs interacting with key genes using the ENCORI database. At the same time, we seamlessly integrated the visual tool Cytoscape, which

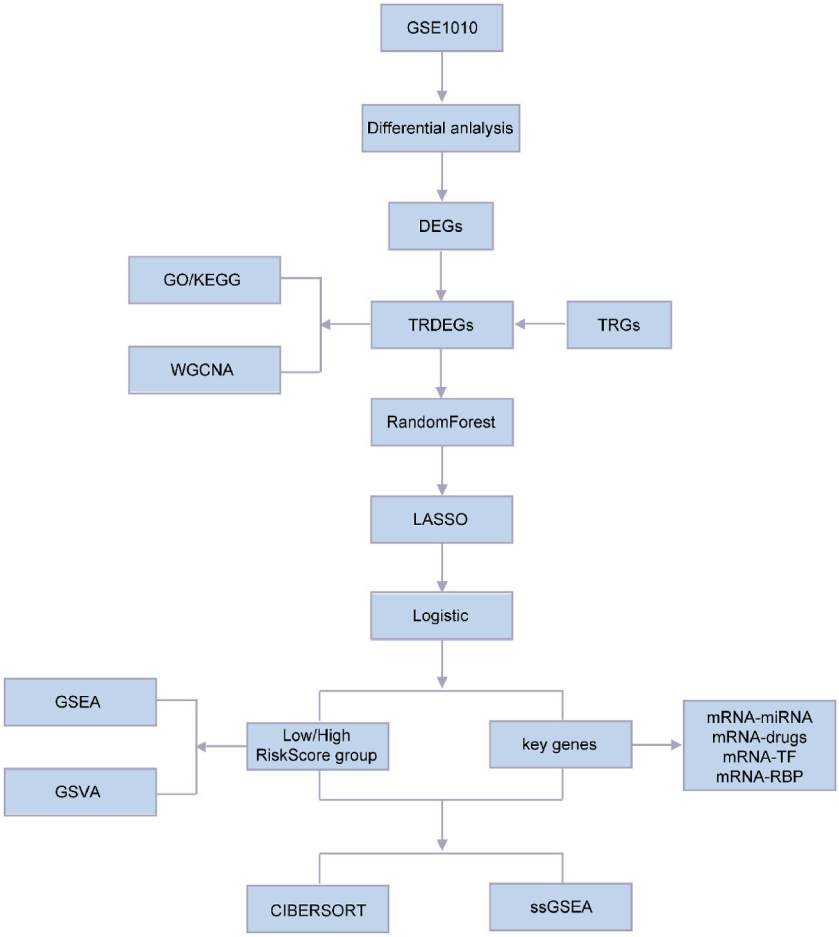


Fig. 1. Technology roadmap.
FCHL, familial combined hyperlipidemia; DEGs, differentially expressed genes; GSEA, gene set enrichment analysis; GSVA, gene set variation analysis; TRGs, tacrolimus-related genes; TRDEGs, tacrolimus-related differentially expressed genes; LASSO, Least Absolute Shrinkage and Selection Operator; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; ssGSEA, single-sample gene set enrichment analysis; TF, transcription factor; RBP, RNA-binding protein.

provides a digital canvas for molecular artistry, to artistically render and illuminate the intricate web of interactions within the miRNA–mRNA network.

Additionally, mRNA–RNA-binding protein (RBP) interaction dynamics were screened using the stringent criterion clusterNum >6 to identify the most relevant interactions. Leveraging the expansive knowledge within the ENCORI database, we explored the intricate network between RBPs and pivotal genes to provide a better understanding of the dynamic interplay between mRNA and RBPs. Cytoscape was then used to visualize the molecular relationships among these complex data.

The publicly accessible CTD was then used to predict potential pharmaceutical interventions or the connections between small-molecule substances and pivotal genes [43], and potential pharmacological targets of the key genes under scrutiny were assessed. The mRNA–drug interaction pairings were screened based on the criterion "Reference Count" > 2. Subsequently, Cytoscape software was used to illuminate these potential molecular connections. Through this visually immersive approach, we aspired to transcend the conventional boundaries of data representation.

The comprehensive CHIPBase database (version 3.0) was then assessed to reveal the transcriptional regulators that bind to pivotal genes [68] and identify associated transcription factors (TFs). mRNA–TF interaction pairings were filtered by aggregating both upstream and downstream samples. The stringent inclusion criterion required that cumulative samples in both directions exceed a threshold of 6.

2.12. Statistical analyses

Data processing and analytical methods were performed based on the robust capabilities of R software (Version 4.2.2). Continuous variables were expressed by mean values and standard deviations. The Wilcoxon rank sum test was used to analyzed differences between two groups, while the Kruskal–Wallis test was used for comparisons involving three or more groups. In addition, Spearman's correlation analysis was used to identify connections between distinct molecular entities. To maintain a comprehensive statistical framework, all P-values were derived from two-sided analyses. A significance threshold of $P < 0.05$ was applied across all analyses (see Fig. 2).

3. Results

Fig. 1.

3.1. Data pre-processing

The datasets in Fig. 2A and B were compared before and after normalization using distribution box-and-line plots. The distribution boxplots showed that the distribution of expression across the samples in the dataset GSE1010 was rather consistent following the process of standardization.

3.2. Differential expression analysis of TRGs

A differential expression analysis was performed using the limma program to identify DEGs within the diverse subgroups of the GSE1010 dataset (FCHL/control) and reveal differential molecular dynamics between the FCHL and control groups. Dataset GSE1010

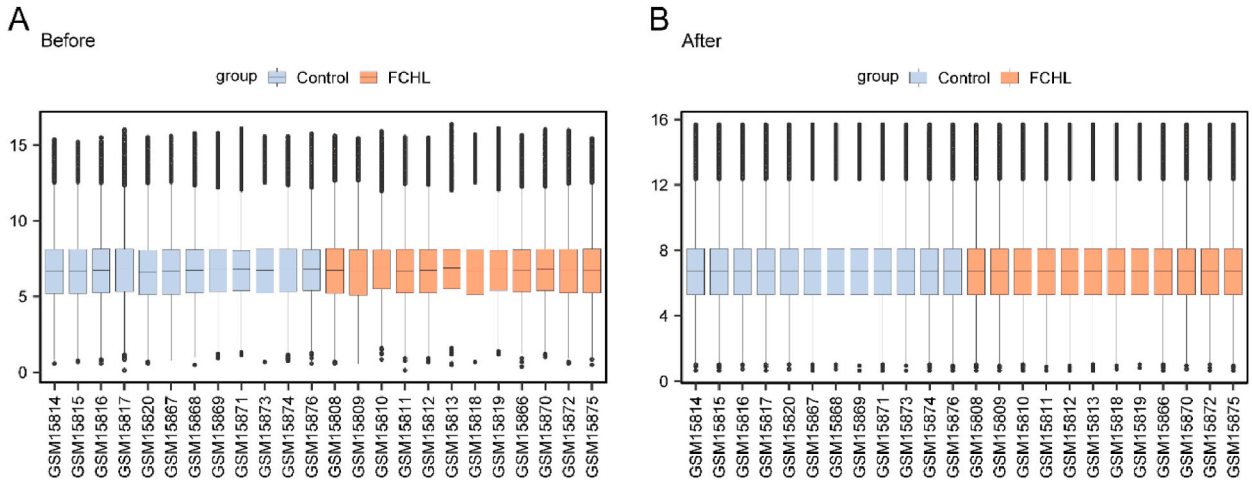


Fig. 2. Data pre-processing.
A-B. Box line plots of the GSE1010 distribution of the dataset before standardization (A) and after standardization (B). FCHL, familial combined hyperlipidemia.

had a total of 12,548 genes, of which 451 DEGs met the threshold of $|logFC| > 0$ and $P < 0.05$. Adhering to this criterion, a total of 248 genes exhibited heightened expression in the FCHL group, indicating low expression in the control group with a positive logFC, characterizing them as upregulated genes. Additionally, 203 genes were identified as upregulated in the FCHL group, signifying elevated expression relative to that in the control group, as evidenced by the negative logFC.

We employed the intersection set of DEGs in dataset GSE1010 to generate TRDEGs that satisfied the thresholds of $|logFC| > 0$ and $P < 0.05$ and TRGs to acquire 43 TRDEGs (Table S1). The findings are shown as a Venn diagram (Fig. 3A). Additionally, to display the

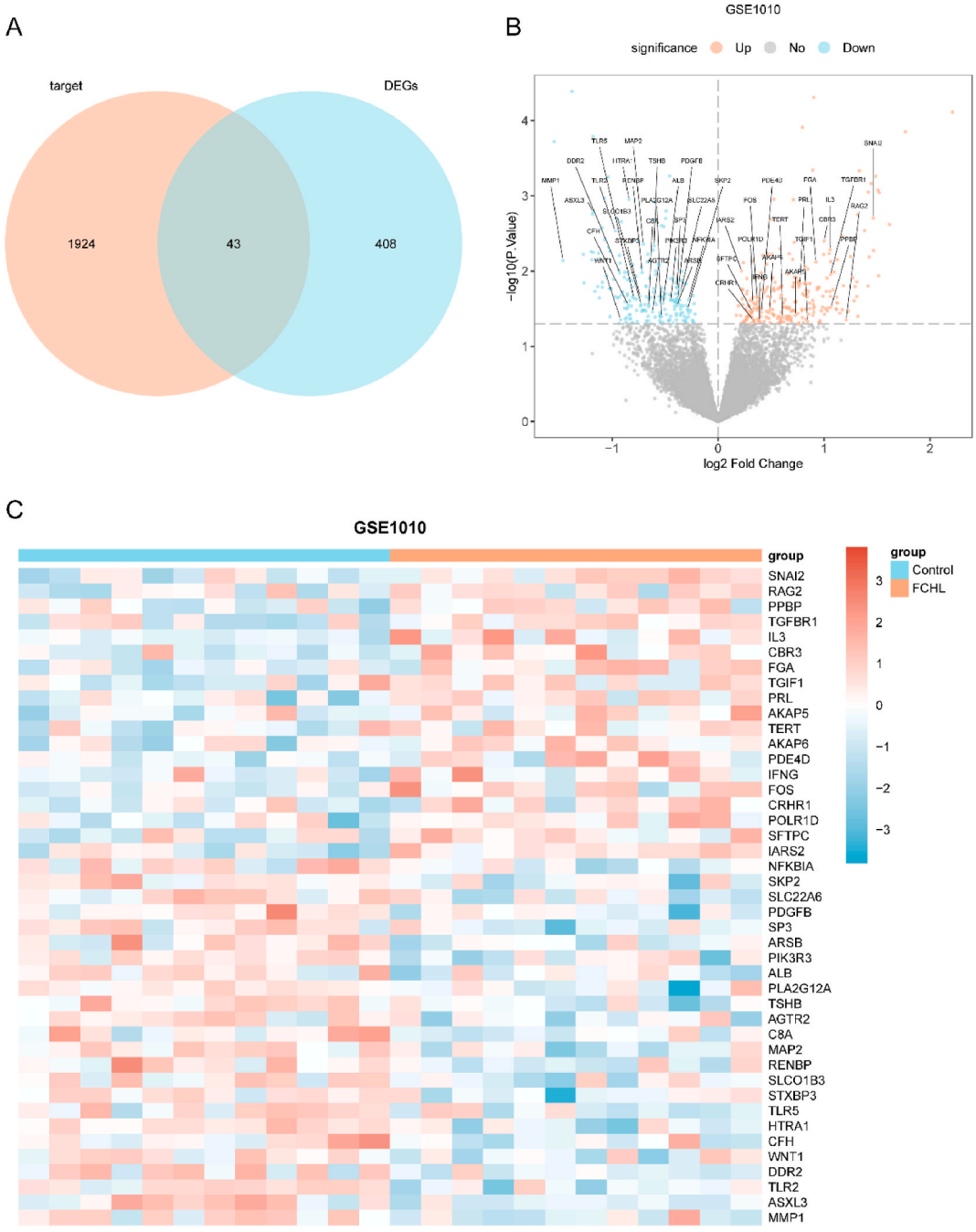


Fig. 3. Differential expression analysis of tacrolimus-related genes.
A. Venn plots of DEGs and TRGs in dataset GSE1010. b. Volcano plots of DEGs analyzed between different subgroups of dataset GSE1010 (FCHL/control), with TRDEGs labeled in the plots. c. Expression heatmap of TRDEGs between different subgroups of dataset GSE1010 (FCHL/control) expression heatmap of TRDEGs. FCHL, familial combined hyperlipidemia; DEGs, differentially expressed genes; TRGs, tacrolimus-related genes; TRDEGs, tacrolimus-related differentially expressed genes.

findings of the difference analysis of the dataset GSE1010, a volcano plot was created, with the 43 TRDEGs appropriately labeled in the plot (Fig. 3B).

Leveraging insights from the Venn diagrams, the expression dynamics of the 43 TRDEGs within the diverse subgroups of GSE1010 (FCHL/control) were assessed. The outcomes of this analysis are portrayed through a heatmap generated using the R package pheatmap (Fig. 3C).

3.3. GO and KEGG enrichment analysis

The relationships between the BP, CC, and MF categories and biological pathways (pathways) of the 43 TRDEGs and FCHL were further explored using GO and KEGG enrichment analyses, as shown in Tables 2 and 3. According to the findings, the 43 TRDEGs were primarily enriched in BPs, including ameboidal-type cell migration, response to peptides, extrinsic apoptotic signaling pathway, epithelial cell migration, epithelial migration; in CCs, such as platelet alpha granule, endoplasmic reticulum lumen, secretory granule lumen, cytoplasmic vesicle lumen, and membrane raft; and in MFs, such as receptor ligand activity, signaling receptor activator activity, cytokine receptor binding, adenylate cyclase binding, and sodium-independent organic anion transmembrane transporter activity. Pathways enriched in DEGs included the cAMP signaling pathway, PD-L1 expression and PD-1 checkpoint pathways in cancer, Toll-like receptor signaling pathway, coronavirus disease (COVID-19), Chagas disease, and other biological pathways. The results of the KEGG and GO enrichment analyses are shown as bubble plots (Fig. 4A). Intricate network diagrams of the BP, CC, and MF categories and biological pathways were generated based on the comprehensive GO and KEGG enrichment analyses (Fig. 4B–E). The size of each node within the diagram serves as a visual indicator, with larger nodes representing a higher count of molecules within the respective entry. Interconnecting lines delineate the relationships between molecules, thereby annotating the associated entries within the dynamic landscape of the network.

3.4. WGCNA

To examine the variations in gene expression across the GSE1010 dataset's various subgroups (FCHL/control), we performed a WGCNA on all genes within these subgroups to look for modules that co-express. We initially selected the genes with a median absolute deviation (MAD) of the top 5000 among all genes as the total input genes for the analysis. Then, we clustered the samples of the GSE1010 dataset using a clustering tree and set the screening criterion as 0.9 for determining the optimal power threshold (Fig. 5A). Fig. 5A shows that the optimal power threshold for our WGCNA analysis was 3. The genes in the GSE1010 dataset (top 5000 genes)

Table 2
GO enrichment Analysis results of TRDEGs.

ONTOLOGY	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0001667	ameboidal-type cell migration	7/43	480/18800	9.84044e−05	0.022214784	0.015045509
BP	GO:1901652	response to peptide	7/43	491/18800	0.000113303	0.022736196	0.015398648
BP	GO:0097191	extrinsic apoptotic signaling pathway	6/43	221/18800	1.04529e−05	0.004758238	0.003222633
BP	GO:0010631	epithelial cell migration	6/43	358/18800	0.00015377	0.023145224	0.015675672
BP	GO:0090132	epithelium migration	6/43	361/18800	0.000160903	0.023145224	0.015675672
CC	GO:0031091	platelet alpha granule	5/43	91/19594	1.61923e−06	0.000215358	0.000165332
CC	GO:0005788	endoplasmic reticulum lumen	5/43	311/19594	0.000572762	0.010616701	0.008150534
CC	GO:0034774	secretory granule lumen	5/43	322/19594	0.000670223	0.010616701	0.008150534
CC	GO:0060205	cytoplasmic vesicle lumen	5/43	325/19594	0.000698841	0.010616701	0.008150534
CC	GO:0045121	membrane raft	5/43	326/19594	0.000708581	0.010616701	0.008150534
MF	GO:0048018	receptor ligand activity	7/43	489/18410	0.000125734	0.009702606	0.006648261
MF	GO:0030546	signaling receptor activator activity	7/43	496/18410	0.000137301	0.009702606	0.006648261
MF	GO:0005126	cytokine receptor binding	6/43	272/18410	3.79227e−05	0.008039606	0.005508767
MF	GO:0008179	adenylate cyclase binding	2/43	12/18410	0.00034652	0.018365566	0.012584152
MF	GO:0015347	sodium-independent organic anion transmembrane transporter activity	2/43	16/18410	0.000626309	0.026555505	0.018195927

GO: Gene Ontology; BP: biological process; CC: cellular component; MF: molecular function. TRDEGs, Tacrolimus related differentially expressed genes.

Table 3
KEGG enrichment Analysis results of TRDEGs.

ONTOLOGY	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
KEGG	has05171	Coronavirus disease – COVID-19	7/33	232/8164	3.107 e−05	0.0019056	0.00127549
KEGG	has05142	Chagas disease	6/33	102/8164	2.7644 e−06	0.00050865	0.00034046
KEGG	has04024	cAMP signaling pathway	6/33	221/8164	0.00022072	0.00559456	0.00374464
KEGG	has05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	5/33	89/8164	2.5652 e−05	0.0019056	0.00127549
KEGG	has04620	Toll-like receptor signaling pathway	5/33	104/8164	5.4439 e−05	0.00250421	0.00167616

KEGG, Kyoto Encyclopedia of Genes and Genomes. TRDEGs, Tacrolimus related differentially expressed genes.

were clustered in the following 14 modules: MEturquoise, MEbrown, MEblack, MEred, MEtan, MEblue, MEmagenta, MEsalmon, MEyellow, MEgreenyellow, MEpink, MEgreen, MEpurple, and MEgrey (Fig. 5B). We then set the module merging shear height to 0.2; that is, modules with shear heights below 0.2 were merged by shear (Fig. 5B). The top 5000 genes in the GSE1010 dataset were grouped, and connections between the genes and associated modules were examined again (Fig. 5C). Specific correlation values between the 14 modules in GSE1010 were then determined (Fig. 5D).

We selected the 13 modules (excluding the MEgrey module) with statistically significant variations ($P < 0.05$, absolute correlation value ≥ 0.3) between different subgroups (HF/control) in the GSE1010 dataset. Genes from module MEmagenta ($|COR| = 0.45$, $P = 0.03$) were selected. Next, we took the intersection of the 43 TRDEGs in the GSE1010 dataset with the module genes included in the MEmagenta module (Fig. 5E) and plotted a Venn diagram to obtain three TRDEGs: TGFBR1, AKAP5, and RAG2.

3.5. Screening of key genes

To further identify key genes, we initially screened the 43 TRDEGs for significance ($P < 0.05$) using one-way logistic regression, resulting in 24 TRDEGs (Table S2). We then analyzed the expression of these 24 TRDEGs in the disease-control (FCHL/control) subgroup of GSE1010 using the RF algorithm. A total of 234 seeds and 500 decision trees were assigned, and then the decision-tree error curve was plotted (Fig. 6A). The results showed that the error reached a minimum and leveled off when approximately 300 decision trees were present. Subsequently, we plotted the MeanDecreaseGini scatter plot (Fig. 6B) for the 24 TRDEGs identified via gene screening. Next, we traded off the number of genes by performing five 10-fold cross-validations and plotting the cross-validation error plot (Fig. 6C). The graph shows that the model error was reduced when there were 24 genes. This finding was then combined with MeanDecreaseGini to select specific genes for subsequent analyses. The results showed (Fig. 6B–C) that the 24 TRDEGs with a significant impact on the diagnosis of FCHL were as follows (in descending order of importance): HTRA1, TLR2, NFKBIA, CBR3, ASXL3, SP3, SNAI2, IARS2, SFTPC, MAP2, RENBP, MMP1, arsb, fos, tlr5, prl, ddr2, tgfbr1, tshb, stxbp3, slc22a6, alb, fga, and polr1d.

After that, we created a LASSO risk model using the 24 TRDEGs screened by the RF algorithm using a LASSO regression analysis, and the results were visualized by plotting a LASSO regression model graph (Fig. 6D) and LASSO variable trajectory graph (Fig. 6E). The LASSO risk model included 14 TRDEGs: HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3, ASXL3, and IARS2. These genes were considered key genes for our subsequent study, and the key genes were plotted on a forest map (Fig. 6F).

3.6. Diagnostic logistic regression models for key gene set construction

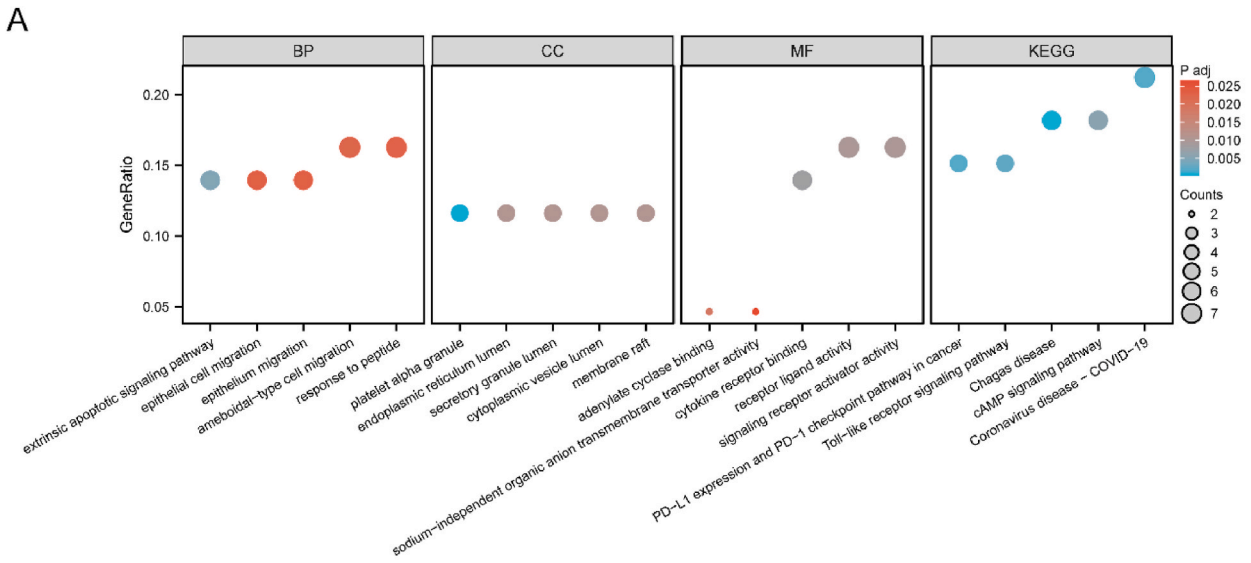
To obtain the FCHL diagnostic model, we constructed a diagnostic multifactorial logistic model based on the 14 key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3, ASXL3, and IARS2) and obtained each key gene's coefficient. Then, based on the RiskSore formula, the expression and coefficients of the 14 key genes were substituted in the dataset GSE1010 to obtain the RiskSore for each sample. The FCHL group was then sorted based on the median RiskSore values, which indicate high and low risk ratings. The formula for calculating the RiskSore is as follows:

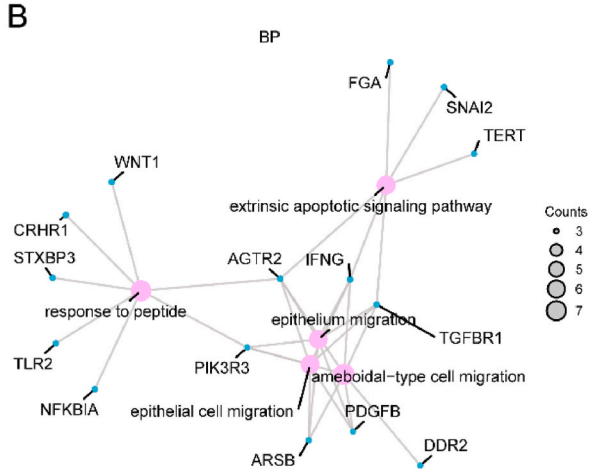
$$\begin{aligned} \text{Risk}Score = &-1.844{*}HTRA1 - 10.38{*}NFKBIA - 1.273{*}MMP1 - 3.508{*}DDR2 + 0.44{*}CBR3 + 3.182{*}PRL + 0.682{*}FGA \\ &+ 3.541{*}SFTPC - 5.216{*}ARSB + 3.412{*}FOS + 1.277{*}SNAI2 - 4.386{*}SP3 - 1.987{*}ASXL3 - 6.228{*}IARS2 \end{aligned}$$

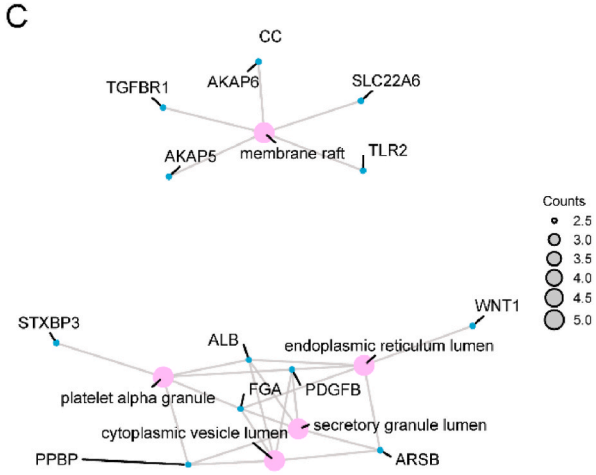
We then plotted the nomograph (Fig. 7A) to show the linkage between the 14 key genes. The graph revealed that the expression of NFKBIA contributes the most to the multifactorial logistic model.

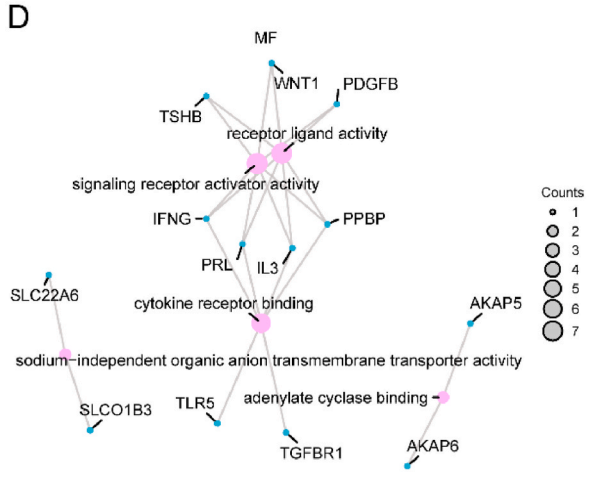
To evaluate the multifactor logistic model's accuracy and discriminative power, a calibration analysis was performed to plot a calibration curve graph. The model's predictive accuracy was then evaluated by analyzing the fit between the actual and predicted probabilities for each of the various scenarios plotted in the graph (Fig. 7B). The calibration curve plot of the multifactor logistic model demonstrated that the dashed calibration line and the diagonal line of the ideal model coincided, indicating that the model had high accuracy and discriminatory power.

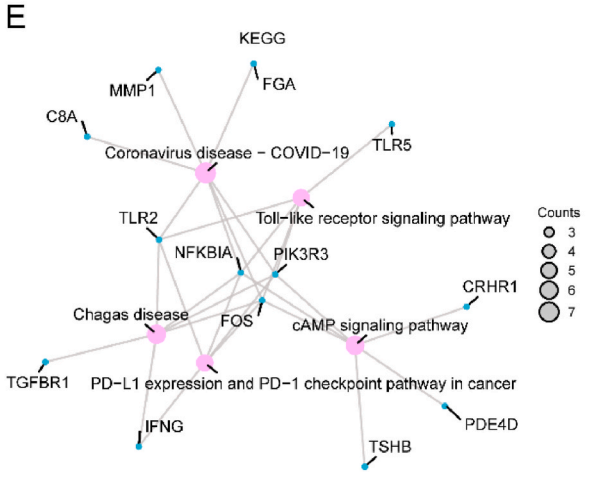
We then performed a DCA to evaluate the multifactorial logistic model for FCHL diagnosis based on the GSE1010 (Fig. 7C). The findings demonstrated that within a particular range, the model's line was continuously higher than the All and None lines and that its











(caption on next page)

Fig. 4. GO and KEGG enrichment analysis.
The GO and KEGG enrichment analysis results for TRDEGs are visualized using a bubble plot. B-E. GO and KEGG enrichment analysis results for TRDEGs illustrated through network diagrams of the BP (B), CC (C), and MF categories (D) and KEGG pathways (E). In the bubble plot (A), the horizontal axis represents GO terms while the vertical coordinates indicate the enrichment factor GeneRatio values for the different GO terms. In the network diagrams, specific genes are denoted by blue dots and pathways are represented by pink dots (B, C, D, E). The selection criteria for GO-enriched entries were P.adj <0.05 and FDR <0.05. GO, Gene Ontology; BP, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes; TRDEGs, tacrolimus-related differentially expressed genes.

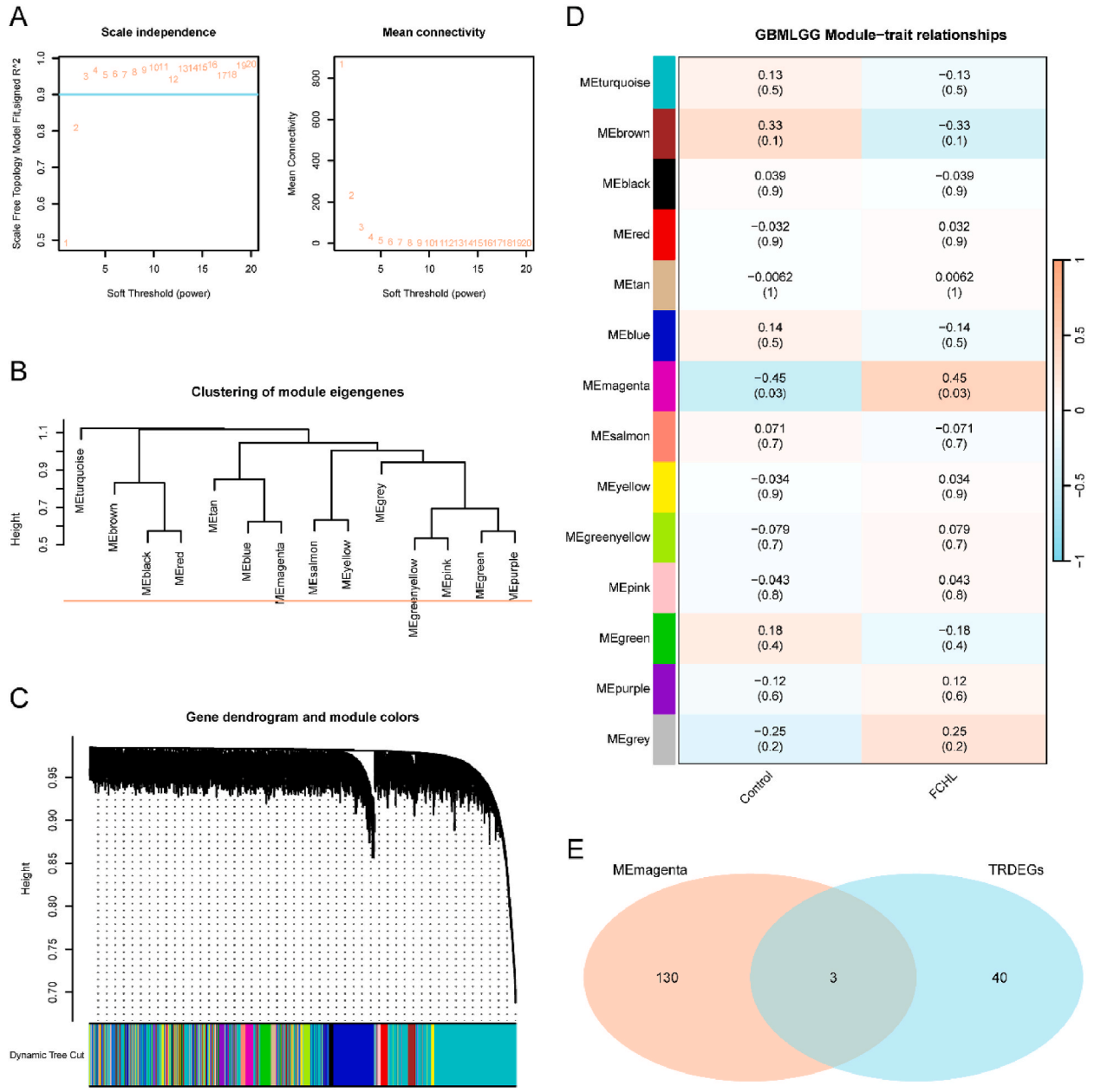


Fig. 5. WGCNA
A. Sample module screening threshold scale-free network display in the GSE1010 dataset. B. Gene(s) module aggregation result display in the GSE1010 dataset. C. Gene(s) and module correspondence in the GSE1010 dataset. D. Correlation analysis of genes in the clustering module with different subgroups in the GSE1010 dataset. E. Venn diagram of TRDEGs with genes in the MEmagenta module. WGCNA, Weighted gene co-expression network analysis; TRDEGs, tacrolimus-related differentially expressed genes.

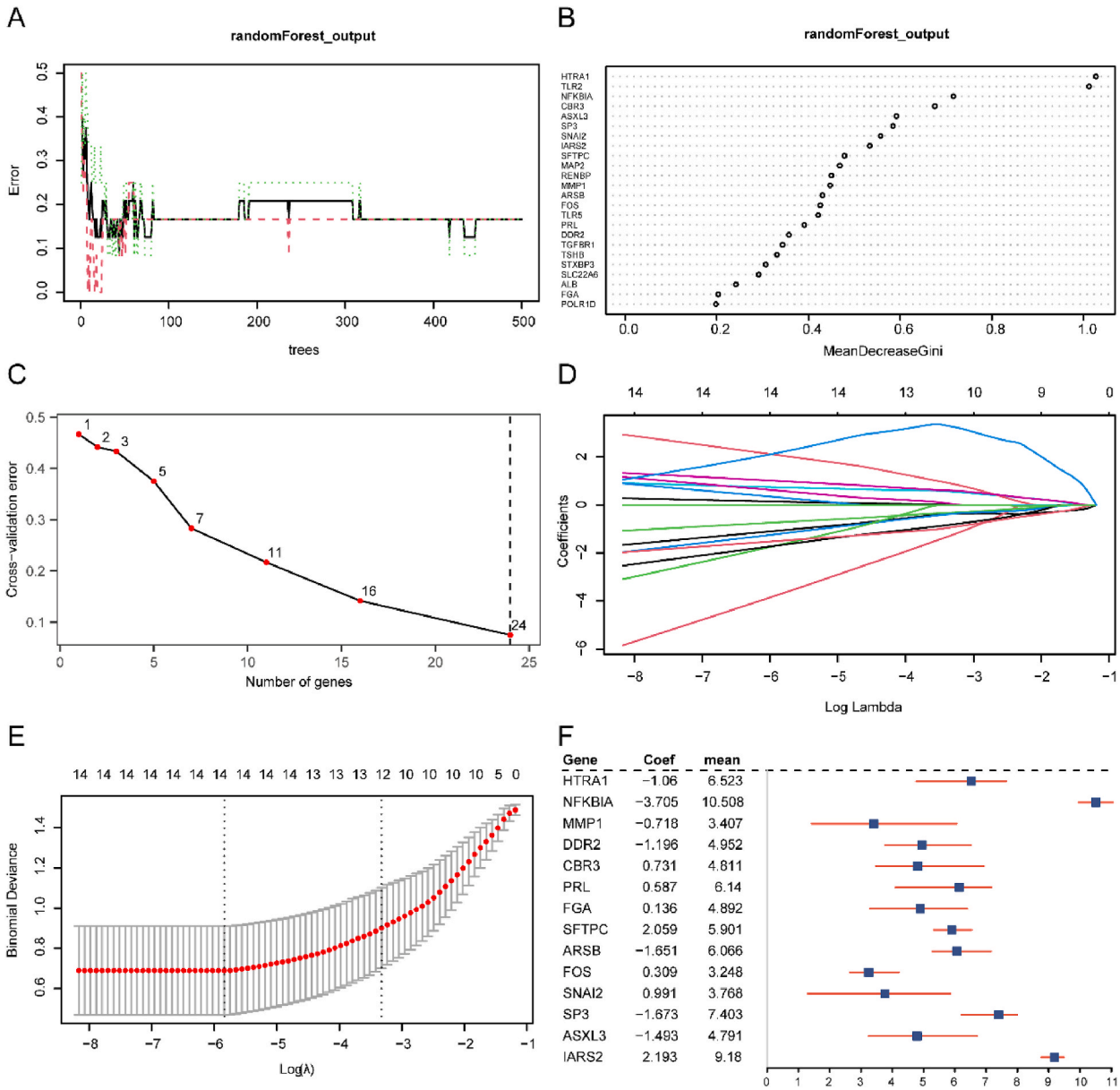


Fig. 6. Screening of key genes.
A. RF algorithm model training error plot. B. MeanDecreaseGini scatterplot of TRDEGs (in descending order of MeanDecreaseGini). C. Cross-validation error curve plot. D. Diagnostic model plot for the LASSO regression model. E. Variable trajectory plot. F. Forest plots of key genes in the LASSO regression model. Forest plots of key genes. TRDEGs, tacrolimus-related differentially expressed genes; LASSO, Least Absolute Shrinkage and Selection Operator.

net gain was larger, indicating that the model was effective for diagnosis.

Finally, the diagnostic multifactorial logistic model was validated for the diagnostic utility of the diagnostic multifactorial logistic model for FCHL by plotting ROC curves for dataset GSE1010 using the R package pROC. The pROC package in R is a vital tool for coordinate graph analysis, designed to identify the optimal threshold for a given model and to discard suboptimal ones. It mitigates sample bias in sequencing data by ensuring accurate model selection and rigorously evaluating performance through receiver operating characteristic (ROC) curve analysis.The ROC values showed that the multifactorial logistic model presented a high degree of accuracy for the diagnosis of FCHL (Fig. 7D, AUC = 1).

3.7. Differential expression validation and functional similarity analysis of key genes

To verify the differences in the 14 key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3,

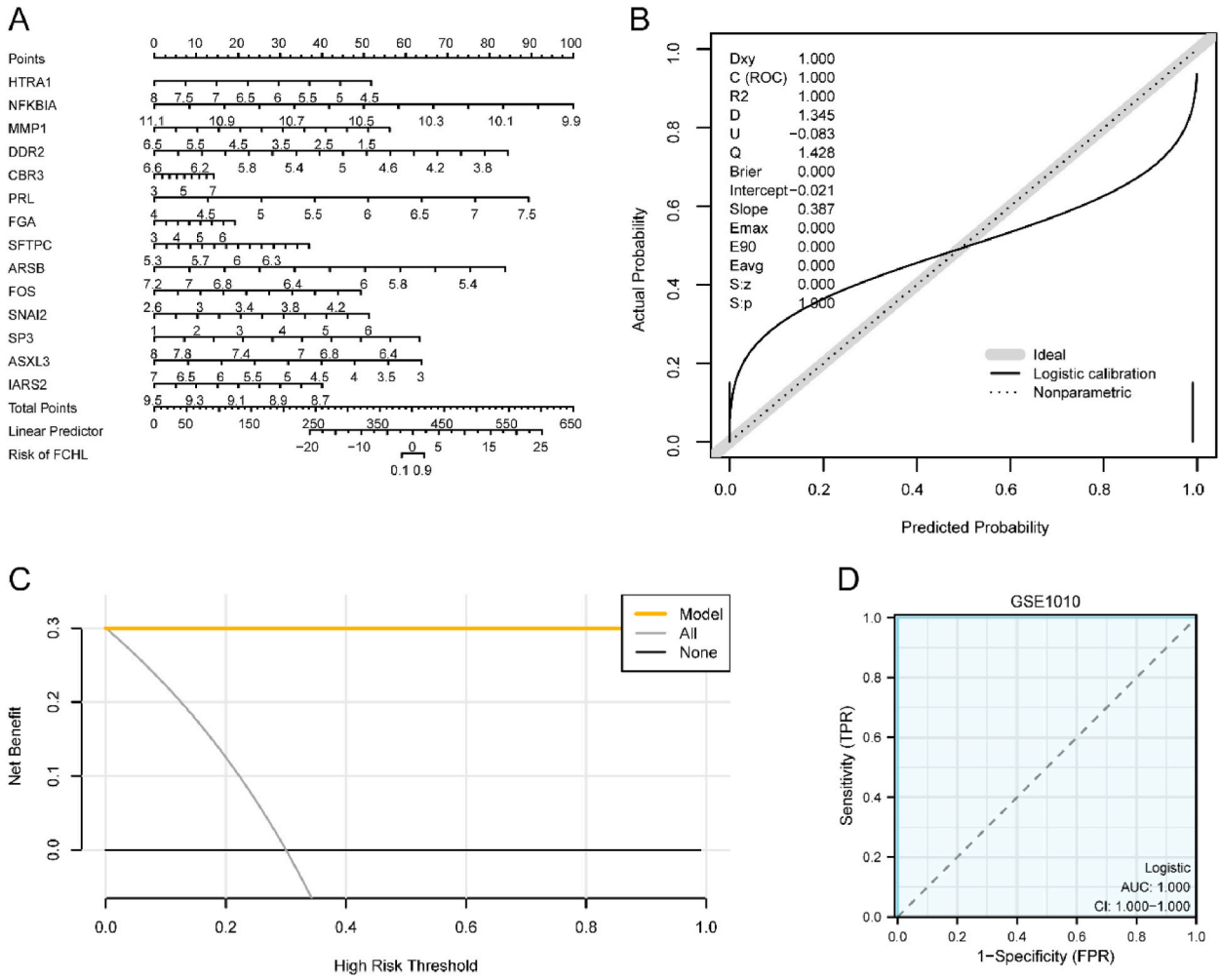


Fig. 7. Diagnostic logistic regression model for key gene construction.
A. Nomogram of the key genes in the diagnostic multifactor logistic model based on dataset GSE1010. B. Calibration curve plot of the key genes in the diagnostic multifactor logistic model based on dataset GSE1010. C. DCA plot of the key genes in the diagnostic multifactor logistic model based on dataset GSE1010. D. Diagnostic ROC curve of the risk score of the diagnostic multifactor logistic model based on dataset GSE1010. DCA plot of key genes in the model. The vertical coordinate of the DCA plot is the net return, and the horizontal coordinate is the probability threshold, or threshold probability. DCA: decision curve analysis. ROC, receiver operating characteristic curve; AUC, area under the curve. FCHL, familial combined hyperlipidemia.

ASXL3, and IARS2) within the different subgroups of dataset GSE1010 (FCHL/control), their expression was analyzed using the test of Wilcoxon rank sum. Variations in expression are presented as group violin plots (Fig. 8A), in which all 14 key genes except SFTPC showed significant differences ($P < 0.05$).

We then correlated the expression of the 14 key genes in the dataset GSE1010 and plotted a correlation heatmap (Fig. 8B), in which the strongest positive correlation was found for SNAI2 and PRL ($r = 0.59$, $P < 0.05$) and the strongest negative correlation was found for IARS2 and ARSB ($r = -0.70$, $P < 0.05$).

The diagnostic performance of the 14 key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3, ASXL3, and IARS2) in dataset GSE1010 for FCHL was determined. We generated ROC curves for each of the 14 key genes in the dataset GSE1010 disease-control subgroup (FCHL/control) and presented the results in Fig. 8C–I. The ROC plots reveal that in dataset GSE1010, the expression of HTRA1 (AUC = 0.875, Fig. 8C), NFKBIA (AUC = 0.771, Fig. 8C), MMP1 (AUC = 0.792, Fig. 8D), DDR2 (AUC = 0.743, Fig. 8D), CBR3 (AUC = 0.847, Fig. 8E), PRL (AUC = 0.778, Fig. 8E), FGA (AUC = 0.778, Fig. 8F), SFTPC (AUC = 0.729, Fig. 8F), ARSB (AUC = 0.764, Fig. 8G), FOS (AUC = 0.750, Fig. 8G), SNAI2 (AUC = 0.854, Fig. 8H), SP3 (AUC = 0.764, Fig. 8H), ASXL3 (AUC = 0.792, Fig. 8I), and IARS2 (AUC = 0.826, Fig. 8I) presented a degree of accuracy in the diagnostic results for FCHL.

To study the functional similarity relationship among the 14 key genes, we calculated the GO semantic similarity of key genes using the GOSemSim package and further calculated the geometric mean of key genes at the BP, CC, and MF levels to obtain the final score. The similarity scores between each key gene and other key genes were averaged and sorted in descending order. The data from the

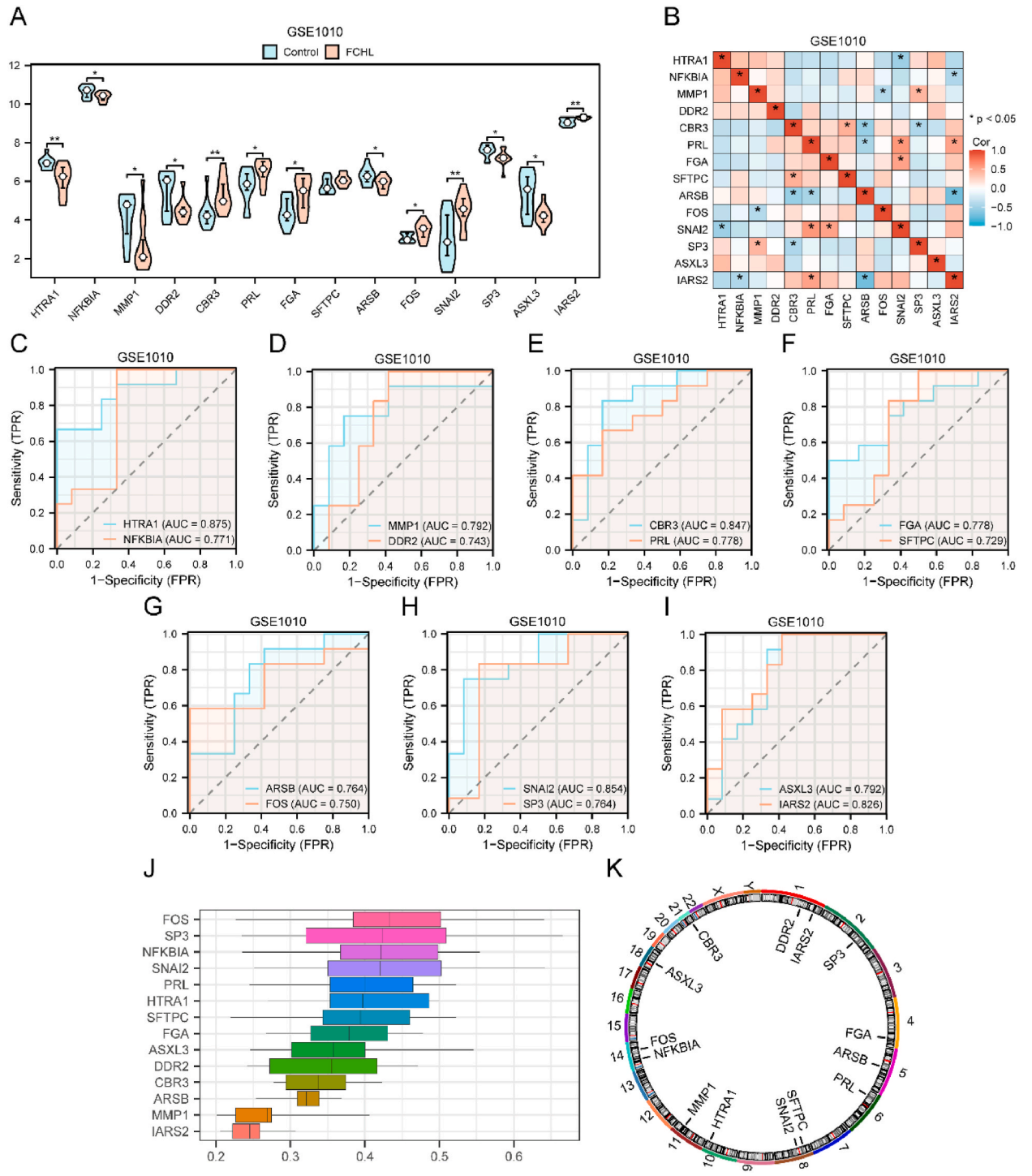


Fig. 8. Differential expression validation and functional similarity analysis of key genes.
A. Subgroup comparison plot of expression differences between different subgroups (FCHL/control) of key genes in dataset GSE1010.B. Heatmap of correlation between key genes in dataset GSE1010.C-I. key genes: HTRA1 (C), NFKBIA (C), MMP1 (D), DDR2 (D). CBR3 (E), PRL (E), FGA (F), SFTPC (F), ARSB (G), FOS (G), SNAI2 (H), SP3 (H), ASXL3 (I), IARS2 (I). ROC curves among different subgroups (FCHL/control) of dataset GSE1010. J. Functionality of key gene similarity plots. K. Chromosomal localization plots of key genes. *P < 0.05 represents statistically significant; **P < 0.01 represents highly statistically significant; ***P < 0.001 represents extremely highly statistically significant. The higher the diagnostic impact, the closer the AUC is to 1. Between 0.5 and 0.7, the AUC has poor accuracy; between 0.7 and 0.9, it has moderate accuracy; and at 0.9 or more, it has great accuracy. The strength of the correlation is as follows: $|r| > 0.95$: significant correlation; $|r| \geq 0.8$: high correlation; $0.5 \leq |r| < 0.8$: moderate correlation; $0.3 \leq |r| < 0.5$: low correlation; $|r| < 0.3$: weak correlation.

functional similarity analysis were plotted in a box-and-line diagram based on the ggplot package (Fig. 8J). The results showed that FOS and other key genes had the greatest functional similarity.

Finally, chromosomal localization maps (Fig. 8K) of the 14 key genes were drawn to show their locations on the chromosomes.

3.8. GSEA based on high and low logistic risk score groupings

We analyzed all 12,584 genes between the high- and low-risk-score subgroups (high/low) using GSEA. The associations between expression and BPs, CCs, and MFs were analyzed, and P.adj <0.05 and FDR value (q.value) < 0.05 were employed as significant enrichment screening criteria. The significantly enriched pathways screened by GSE1010 were visualized using a mountain range diagram (Fig. 9A). The outcomes (Fig. 9B–E, Table 4) showed significant enrichment of genes among the different subgroups (high/low) in the GSE1010 dataset in the pathways BLANCO_MELO_BRONCHIAL_EPITHELIAL_CELLS_INFLUENZA_A_DEL_NS1_INFECTION_DN (Fig. 9B), CROONQUIST_IL6_ DEPRIVATION_DN (Fig. 9C), SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP (Fig. 9D), and MANALO_HYPOXIA_DN (Fig. 9E).

3.9. GSVA enrichment analyses based on high- and low-logistic risk score groupings

GSVA was applied to all dataset genes to determine the h.all.v7.4.symbols.gmt gene set variations between the high and low-risk score subgroups in the FCHL group of GSE1010 (Table 5). The differential expression among four significantly enriched pathways (P < 0.05) was analyzed, and the results were as visually represented through a heatmap (Fig. 10A) and a subgroup comparative plot (Fig. 10B). The subgroup comparison plot showed that the following two pathways remained significant (P < 0.05) in the high/low-risk-score subgroup: HALLMARK_CHOLESTEROL_HOMEOSTASIS and HALLMARK_COMPLEMENT.

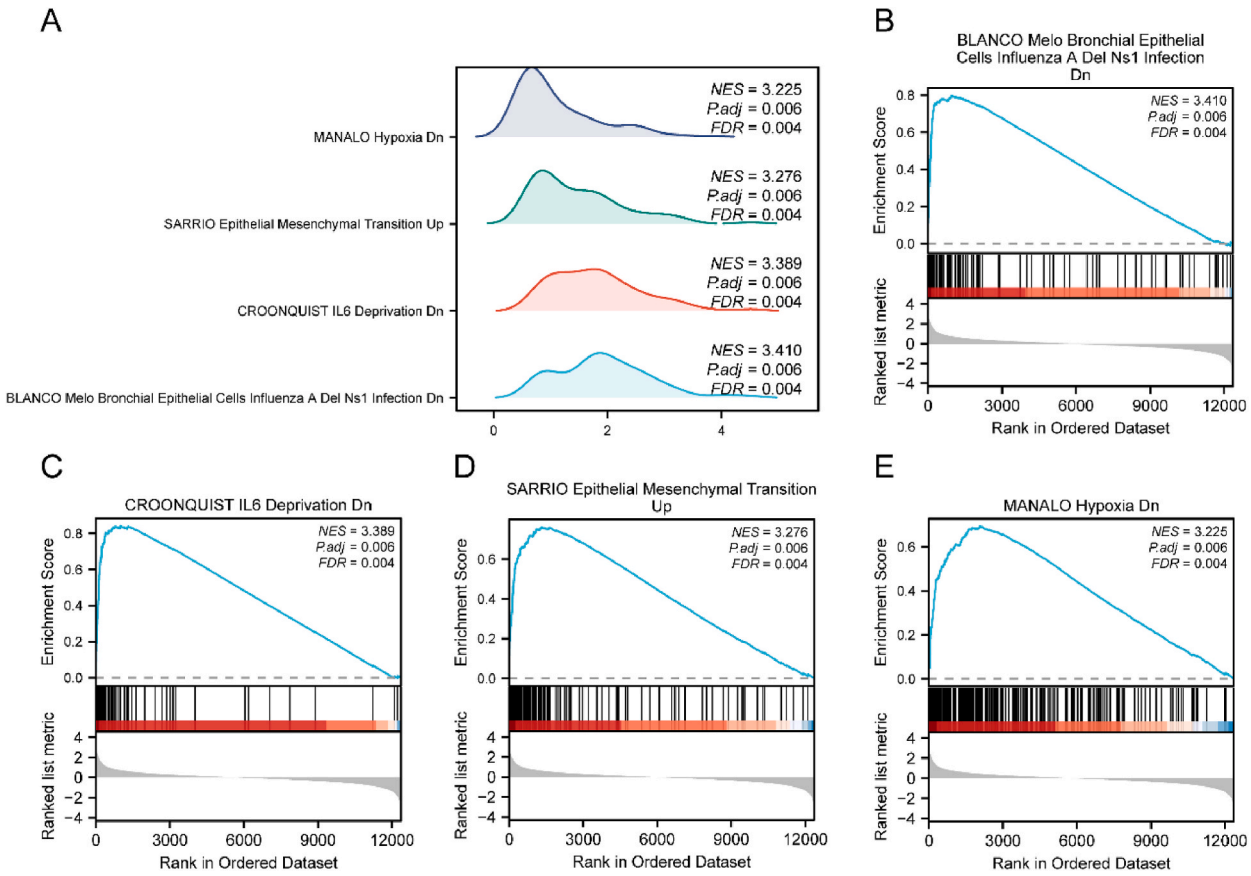


Fig. 9. Enrichment analysis of the GSEA results based on high and low logistic risk score groupings.
A. Mountain range diagrams of the GSEA enrichment analysis results for the different subgroups (high/low) in dataset GSE1010 for the primary biological characteristics. B-E. Genes in dataset GSE1010 were markedly enhanced in the pathways REACTOME_AUF1_HNRNP_D0_BINDS_AND_DESTABILIZES_MRNA (B), WP_OXIDATIVE_PHOSPHORYLATION (C), KEGG_OXIDATIVE_PHOSPHORYLATION (D), REACTOME_INFLUENZA_INFECTION (E). GSEA, gene set enrichment analysis; FCHL, familial combined hyperlipidemia.

Table 4
GSEA enrichment analysis results of FCHL dataset GSE1010.

Description	setSize	enrichmentScore	NES	pvalue	p.adjust	qvalue
BLANCO_MELO_BRONCHIAL_EPITHELIAL_CELLS_INFLUENZA_A_DEL_NS1_INFECTION_DN	137	0.79839326	3.40969979	0.00061312	0.00568501	0.00400179
CROONQUIST_IL6_DEPRIVATION_DN	95	0.8421434	3.38860511	0.00056786	0.00568501	0.00400179
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP	144	0.76153056	3.27631397	0.00062112	0.00568501	0.00400179
MANALO_HYPOXIA_DN	281	0.6921115	3.22505087	0.0007764	0.00600669	0.00422823
TANG_SENESCENCE_TP53_TARGETS_DN	55	0.83914039	3.05770588	0.00052715	0.00568501	0.00400179
WU_APOPTOSIS_BY_CDKN1A_VIA_TP53	50	0.81963946	2.92921109	0.00051733	0.00568501	0.00400179
VILLANUEVA_LIVER_CANCER_KRT19_UP	138	0.67992286	2.90453388	0.00061652	0.00568501	0.00400179
LI_WILMS_TUMOR_VS_FETAL_KIDNEY_1_DN	158	0.66451717	2.87728239	0.00064226	0.00568501	0.00400179
BUFFA_HYPOXIA_METAGENE	46	0.80808409	2.84156298	0.00050607	0.00568501	0.00400179
WIELAND_UP_BY_HBV_INFECTION	97	0.68242569	2.7595773	0.00056593	0.00568501	0.00400179
NAKAMURA_CANCER_MICROENVIRONMENT_DN	39	0.78685688	2.65845552	0.00050582	0.00568501	0.00400179
WILCOX_RESPONSE_TO_PROGESTERONE_UP	145	0.5968719	2.57275641	0.00061767	0.00568501	0.00400179
SCIAN_CELL_CYCLE_TARGETS_OF_TP53_AND_TP73_DN	22	0.85453002	2.51473272	0.00048216	0.00568501	0.00400179
LI_WILMS_TUMOR_ANAPLASTIC_UP	17	0.88866097	2.45801057	0.00047371	0.00568501	0.00400179
REACTOME_NUCLEAR_EVENTS_MEDIATED_BY_NFE2L2	77	0.61830466	2.39669696	0.00054705	0.00568501	0.00400179
WP_METABOLIC_REPROGRAMMING_IN_COLON_CANCER	41	0.69568686	2.3737988	0.00051256	0.00568501	0.00400179
REACTOME_METABOLISM_OF_POLYAMINES	55	0.65066448	2.37092699	0.00052715	0.00568501	0.00400179
REACTOME_AUF1_HNRNP_D0_BINDS_AND_DESTABILIZES_MRNA	50	0.66120627	2.36300571	0.00051733	0.00568501	0.00400179
REACTOME_NEGATIVE_REGULATION_OF_NOTCH4_SIGNALING	53	0.65207952	2.3537702	0.00052798	0.00568501	0.00400179
ALTEMEIER_RESPONSE_TO_LPS_WITH_MECHANICAL_VENTILATION	105	0.55726592	2.27677516	0.00057737	0.00568501	0.00400179

GSEA, Gene Set Enrichment Analysis; FCHL, familial combined hyperlipemia.

Table 5
GSVA enrichment analysis results of FCHL dataset GSE1010.

id	logFC	AveExpr	t	P.Value	adj.P.Val	B
HALLMARK_CHOLESTEROL_HOMEOSTASIS	0.348116	0.00429	2.654589	0.008194	0.409681	2.91391
HALLMARK_FATTY_ACID_METABOLISM	0.285759	0.000221	2.188363	0.029105	0.470829	3.60394
HALLMARK_COMPLEMENT	0.274637	0.004913	2.092879	0.036863	0.470829	3.72887
HALLMARK_NOTCH_SIGNALING	0.27297	0.042363	2.03664	0.042212	0.470829	3.79985

GSVA, Gene Set Variation Analysis; FCHL, familial combined hyperlipemia.

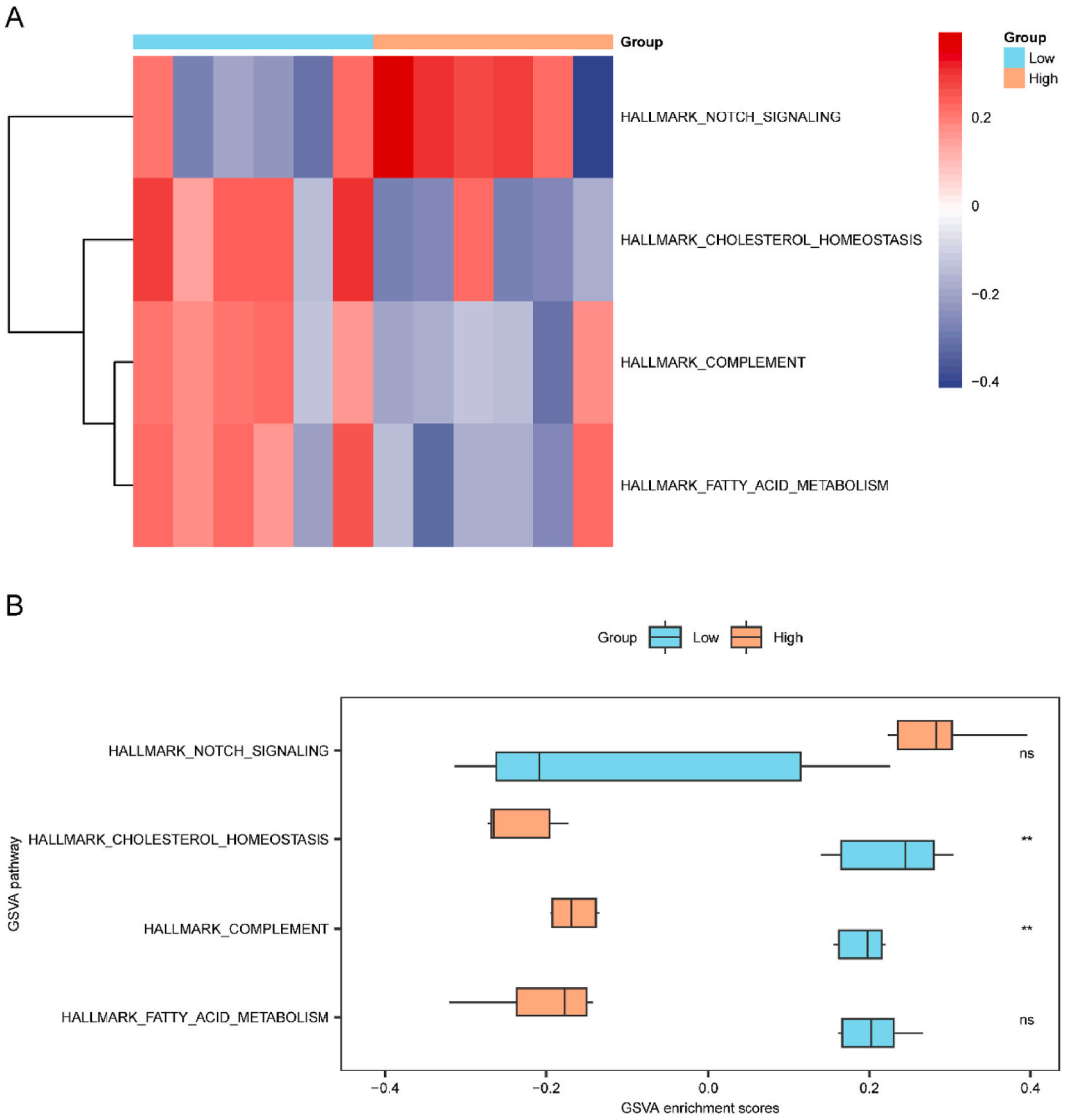


Fig. 10. GSVA-based enrichment analysis of the high and low logistic risk score groupings.
A-B. Complex numerical heatmap (A) and group comparison box line plot (B) of the GSVA results in the high/low-risk-score subgroups of dataset GSE1010. GSVA, gene set variation analysis. ns, not statistically significant ($P \geq 0.05$); *$P < 0.05$, statistically significant; **$P < 0.01$, highly statistically significant; ***$P < 0.001$, extremely highly statistically significant.

3.10. Analysis of CIBERSORT immune infiltration based on high and low logistic risk score groupings

The CIBERSORT algorithm was applied to the samples of dataset GSE1010 to determine the infiltration quantity of the 22 immune cells between the high and low subgroups within the FCHL group. In addition, a stacked bar chart (boxplot) was used to illustrate the infiltration fractional occupancy of the immune cells (Fig. 11A). The results revealed the presence of 17 immune cells (memory B cells,

naïve B cells, dendritic cells activated, eosinophils, macrophages M0, macrophages M1, macrophages M2, mast cells activated, monocytes, neutrophils, NK cells activated, plasma cells, T cells CD4 memory activated, T cells CD8, T cells follicular helper, T cells gamma delta, and T cell regulatory [Tregs]), in the samples of the GSE1010 dataset with non-zero infiltration abundance. We also analyzed the variations in the percentage of infiltration abundance of the 17 immune cells in the samples between the high and low subgroups of the FCHL group from dataset GSE1010 and plotted the subgroups in relation to the results (Fig. 11B).

The Pearson technique was utilized to determine connections between the infiltration abundance of memory B immune cells and expression of 14 key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3, ASXL3, and IARS2). A correlation heatmap was generated (Fig. 11C–D), which revealed that in the low-risk group samples, there were no significant

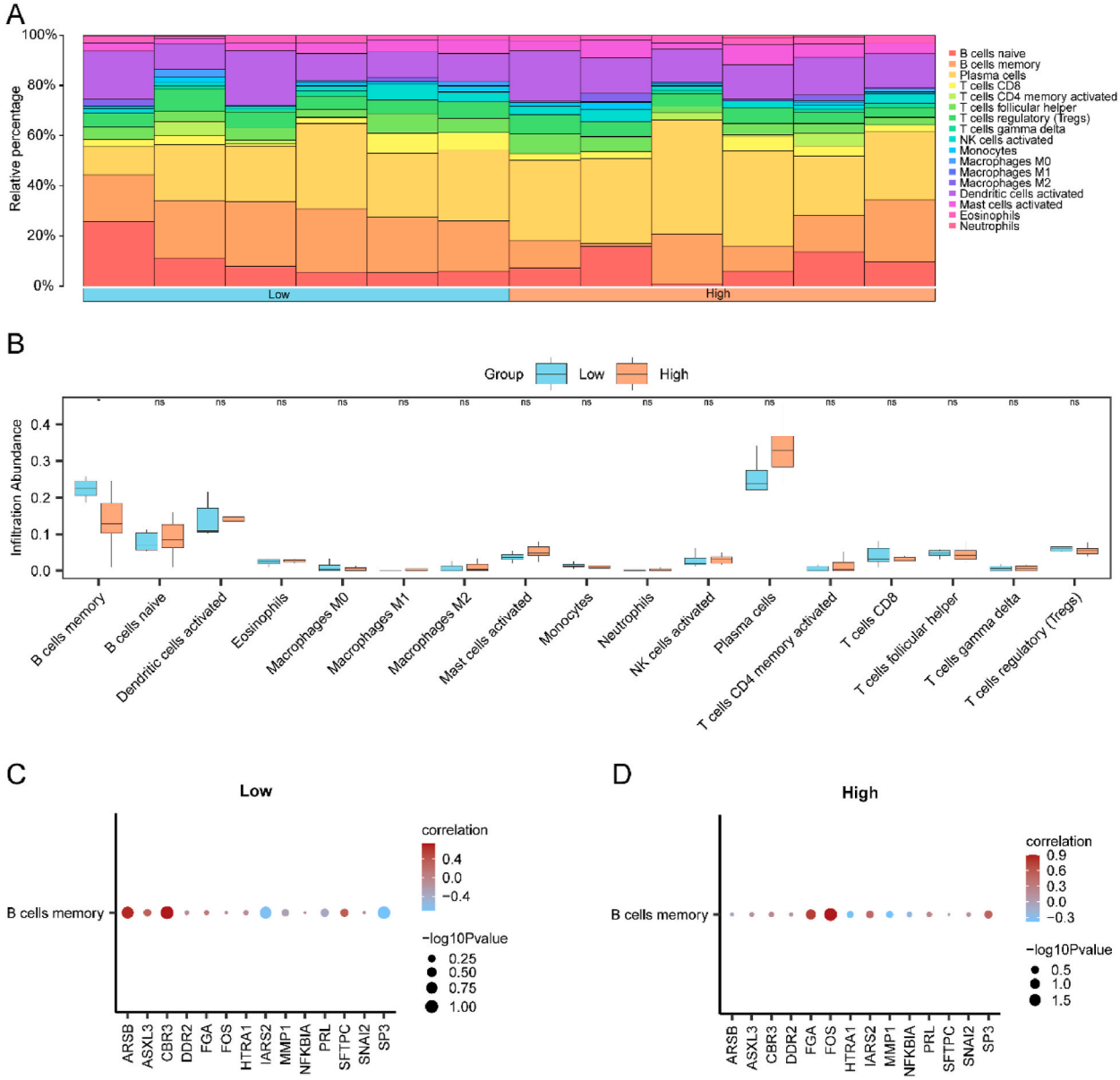


Fig. 11. Analysis of CIBERSORT immune infiltration based on high and low logistic risk score groupings.
A. Results of the CIBERSORT immune infiltration study shown as a stacked histogram between the different subgroups of the GSE1010 dataset. B. Subgroup comparison plot of immune cells between different subgroups of the GSE1010 dataset. C. Correlation heatmap between immune cells with significant differences in the subgroup comparison plot (B) and key genes in the low-risk subgroup of the GSE1010 FCHL samples. D. Correlation heatmap between immune cells with significant differences and key genes in the subgroup comparison plot (B) in the FCHL high-risk subgroup grouping samples of the GSE1010 dataset. ns represents not statistically significant ($P \geq 0.05$); symbols * $P < 0.05$, statistically significant; symbols **$P < 0.01$, highly statistically significant; and symbols ***$P < 0.001$, extremely highly statistically significant. FCHL, familial combined hyperlipidemia.

correlations between memory B immune cells and key genes. However, a noteworthy positive correlation was identified in the high-risk group, specifically with the key gene FOS ($r = 0.90$, $P < 0.05$).

3.11. Analysis of ssGSEA immune infiltration based on high- and low logistic risk score groupings

Utilizing the ssGSEA algorithm, we estimated the infiltration abundance of 28 immune cells in both the high- and low-risk groups among the FCHL patient samples. The Mann–Whitney *U* test was employed to identify infiltration differences, and the findings were visually presented through a group comparison plot (Fig. 12A). A statistically significant difference ($P < 0.05$) was observed in the infiltration abundance of immature B cells within the high- and low-risk groups of dataset GSE1010.

Then, we utilized the "pearson" algorithm to separately calculate the correlation between the infiltration abundance of immature B cells and the expression levels of the 14 key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3, ASXL3, IARS2) in the FCHL low and high-risk group (low/high) samples within the GSE1010 dataset. The correlation results are presented in a heatmap (Fig. 12B–C). In the FCHL low-risk group (low) samples of the GSE1010 dataset, a significant positive correlation was observed between immature B cells and the key gene FOS ($r = 0.94$, $p < 0.05$). Conversely, in the high-risk group (high) samples, a significant positive correlation was observed between immature B cells and the key gene SP3 ($r = 0.90$, $p < 0.05$).

3.12. Regulatory network analysis of key genes

To investigate the interactions between the 14 key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS,

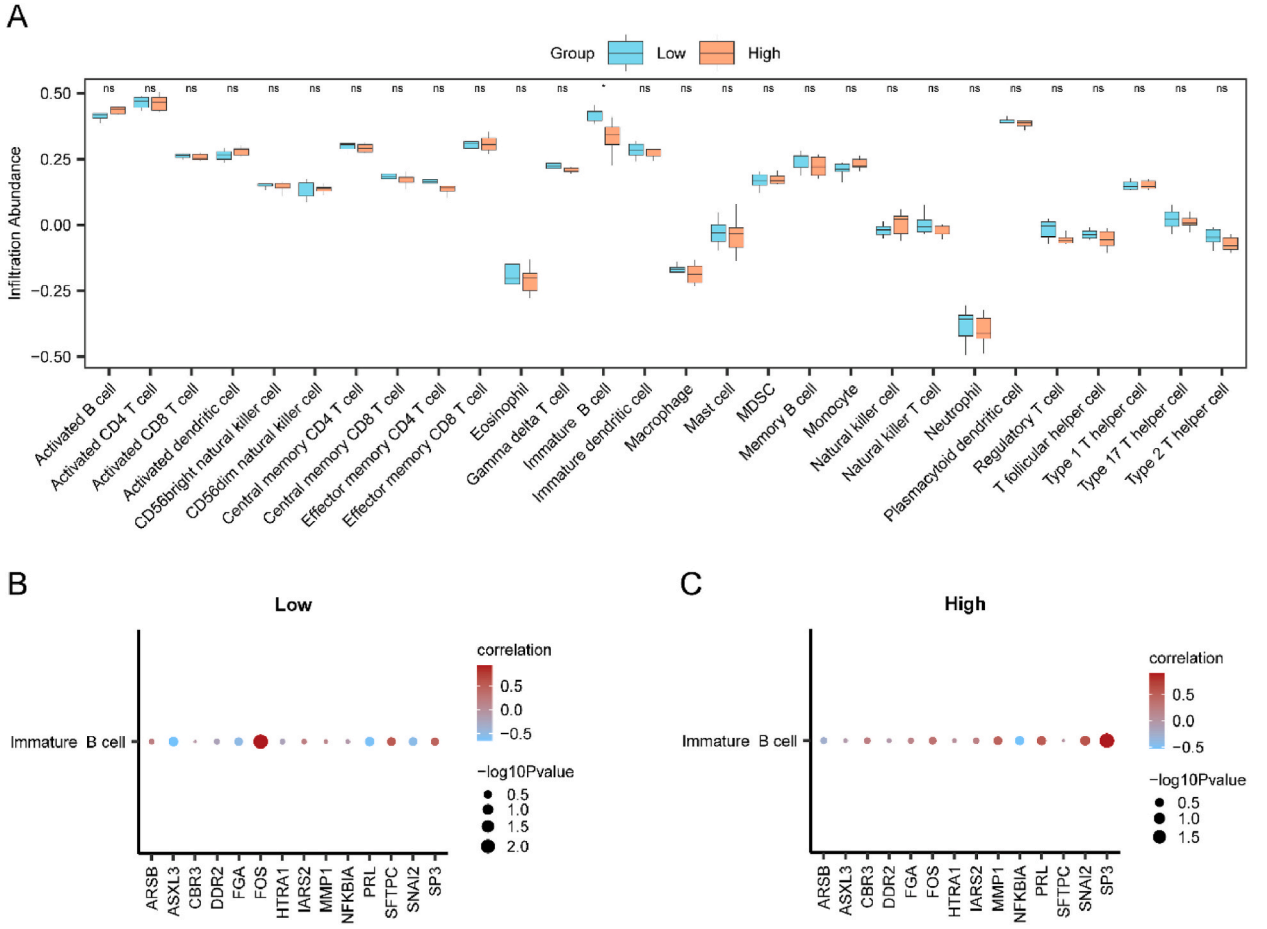


Fig. 12. Analysis of ssGSEA immune infiltration based on high and low logistic risk score groupings.
A. Subgroup comparison plot presentation of the ssGSEA immune infiltration analysis findings relating to the high- and low-risk subgroups of dataset GSE1010. B Heatmap of the correlation between immune cells with significant differences and key genes in the subgroup comparison plot in the FCHL low-risk subgroup samples of dataset GSE1010. C. Heatmap of the correlation between immune cells with significant differences in group comparison plots and key genes in the FCHL high-risk subgroup samples of the GSE1010 dataset. ns indicates not statistically significant ($P \geq 0.05$); *$P < 0.05$, statistically significant; **$P < 0.01$, highly statistically significant; and ***$P < 0.001$, extremely highly statistically significant. FCHL, familial combined hyperlipidemia.

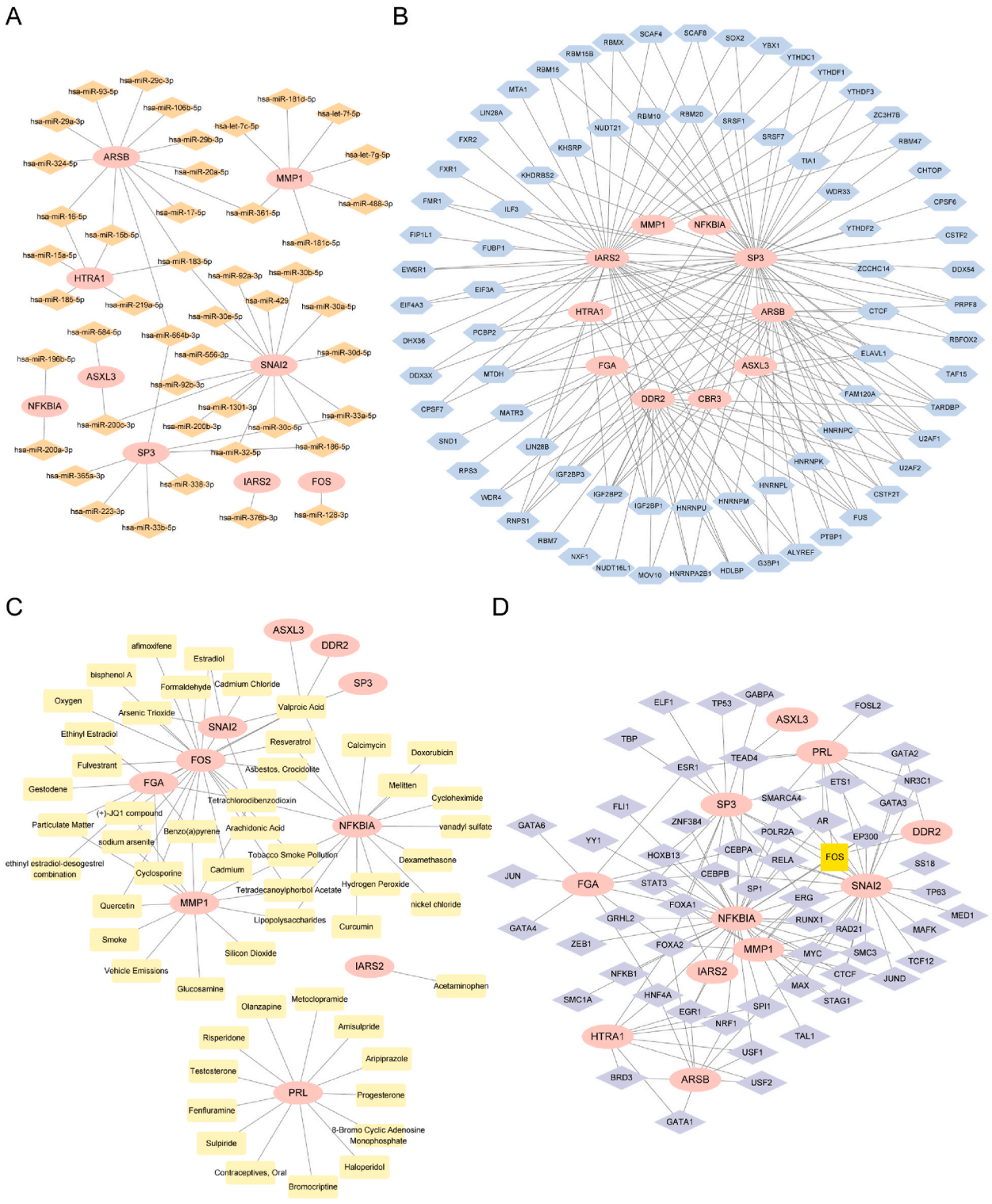


Fig. 13. Regulatory network analysis of key genes.
A. mRNA–miRNA interaction network diagram for key genes: pink oval blocks are mRNAs and orange diamond blocks are miRNAs. B. mRNA–RBP interaction network diagram for key genes: pink oval blocks are mRNAs, and blue polygonal blocks are RBPs. C. mRNA–drug interaction network diagram for key genes: pink oval blocks are mRNAs, and yellow squares are drugs. D. mRNA–RBP interaction network diagram of key genes: pink oval blocks are mRNAs and purple rhombic blocks are TFs. TF, transcription factor; RBP, RNA-binding protein.

SNAI2, SP3, ASXL3, and IARS2) and other molecules, we initially predicted the miRNAs interacting with the 14 key genes utilizing the ENCORI database and screened the mRNA–miRNA interaction pairs using pancancerNum >10 as the screening criterion. The mRNA–miRNA interaction network was visualized utilizing Cytoscape software (Fig. 13A). The mRNA–miRNA interaction network consisted of 9 mRNAs (ARSB, ASXL3, FOS, HTRA1, IARS2, MMP1, NFKBIA, SNAI2, and SP3) and 46 miRNA molecules, which comprise 55 pairs of mRNA–miRNA interaction relationships in total. Table 6 displays the particular mRNA–miRNA interaction relationships.

Additionally, using the ENCORI database, we predicted the RBP interactions with the 14 key genes.The mRNA–RBP interaction pairs were screened using clusterNum >6 as the screening criterion, and the mRNA–RBP interaction network (Fig. 13B) was visualized using the Cytoscape software. The mRNA–RBP interaction network included 10 mRNAs (ARSB, ASXL3, CBR3, DDR2, FGA, HTRA1, IARS2, MMP1, NFKBIA, and SP3) and 77 RBPs, with a total of 155 pairs of mRNA–RBP interactions. Table 7 shows the specific

Table 6
mRNA-miRNA interaction network nodes.

mRNA	miRNA
ARSB	hsa-miR-16–5p
ARSB	hsa-miR-17–5p
ARSB	hsa-miR-20a-5p
ARSB	hsa-miR-29a-3p
ARSB	hsa-miR-93–5p
ARSB	hsa-miR-29b-3p
ARSB	hsa-miR-183–5p
ARSB	hsa-miR-15b-5p
ARSB	hsa-miR-106b-5p
ARSB	hsa-miR-29c-3p
ARSB	hsa-miR-361–5p
ARSB	hsa-miR-324–5p
ARSB	hsa-miR-664b-3p
ASXL3	hsa-miR-200c-3p
ASXL3	hsa-miR-584–5p
FOS	hsa-miR-128–3p
HTRA1	hsa-miR-15a-5p
HTRA1	hsa-miR-16–5p
HTRA1	hsa-miR-183–5p
HTRA1	hsa-miR-219a-5p
HTRA1	hsa-miR-15b-5p
HTRA1	hsa-miR-185–5p
IARS2	hsa-miR-376b-3p
MMP1	hsa-let-7c-5p
MMP1	hsa-let-7f-5p
MMP1	hsa-miR-181c-5p
MMP1	hsa-let-7g-5p
MMP1	hsa-miR-361–5p
MMP1	hsa-miR-181d-5p
MMP1	hsa-miR-488–3p
NFKBIA	hsa-miR-200a-3p
NFKBIA	hsa-miR-196b-5p
SNAI2	hsa-miR-30a-5p
SNAI2	hsa-miR-32–5p
SNAI2	hsa-miR-33a-5p
SNAI2	hsa-miR-92a-3p
SNAI2	hsa-miR-30c-5p
SNAI2	hsa-miR-30d-5p
SNAI2	hsa-miR-183–5p
SNAI2	hsa-miR-200b-3p
SNAI2	hsa-miR-30b-5p
SNAI2	hsa-miR-186–5p
SNAI2	hsa-miR-200c-3p
SNAI2	hsa-miR-30e-5p
SNAI2	hsa-miR-429
SNAI2	hsa-miR-92b-3p
SNAI2	hsa-miR-556–3p
SNAI2	hsa-miR-1301–3p
SP3	hsa-miR-33a-5p
SP3	hsa-miR-223–3p
SP3	hsa-miR-186–5p
SP3	hsa-miR-365a-3p
SP3	hsa-miR-338–3p
SP3	hsa-miR-33b-5p
SP3	hsa-miR-664b-3p

mRNA–RBP interactions.

In addition, the screening criterion for mRNA–drug interaction pairs was set at a "Reference Count" greater than 2. Subsequently, the mRNA–drug interaction network was visualized using Cytoscape software, as depicted in Fig. 13C. The mRNAs are shown as pink ellipses, and drugs are shown as yellow squares. The network of mRNA–drug interactions included 10 mRNAs (ASXL3, DDR2, FGA, FOS, IARS2, MMP1, NFKBIA, PRL, SNAI2, and SP3) and 53 drug molecules, with a total of 79 pairs of mRNA–drug interactions (Table 8).

Finally, TFs that bind to the 12 important genes were identified (version 3.0), and the screening criterion was a sample number (upstream and downstream) greater than six. The mRNA–TF interaction pairs were screened, and the mRNA–TF interaction network was visualized with Cytoscape (Fig. 13D). In the network, pink elliptical blocks symbolize mRNAs and purple diamond blocks symbolize TFs. The mRNA–TF interaction network consisted of 12 mRNAs (HTRA1, ARSB, ASXL3 DDR2, FGA, FOS, IARS2, MMP1, NFKBIA, PRL, SNAI2, and SP3) and 57 TF molecules. The specific mRNA–TF interactions are shown in Table 9.

4. Discussion

The pathogeneses of abnormalities in human lipid metabolism have not been clarified, although most abnormalities are believed to be genetically related. A growing number of transplant physicians have found that TAC is connected to post-transplant hyperlipidemia, and studies have revealed the mechanisms by which post-transplant medications and immune status affect lipid levels [43,68–70]. MicroRNAs (miRNAs) and their associated circular RNAs (circRNAs) are noncoding RNA species that have emerged as pivotal regulators of genes implicated in lipid metabolism. Despite their significance, the transcriptional activities and functional mechanisms of these molecules in tacrolimus-associated dyslipidemia are not well understood. Chenzhi Zhang posits that the circFASN/miR-33a regulatory axis plays a crucial role in the dysregulation of lipid homeostasis induced by tacrolimus. MiR-33a may be a risk factor for dyslipidemia associated with tacrolimus use, suggesting its potential as a therapeutic target to ameliorate lipid abnormalities following liver transplantation [71].Xiao Xu et al. have reported that tacrolimus (TAC) downregulates fibroblast growth factor 21 (FGF21), consequently exacerbating lipid accumulation through the disruption of the autophagy-lysosome pathway. Consequently, the administration of recombinant FGF21 protein has the potential to reverse TAC-induced lipid accumulation and hypertriglyceridemia by enhancing autophagy [72].

In this research, the examination of the FCHL gene expression profile dataset GSE1010 and TRDEGs retrieved from the GeneCards, STITCH, and MSigDB online databases indicated that certain genes serve crucial functional functions in the genesis of hyperlipidemia. For the first time, we screened 14 statistically significant key genes (HTRA1, NFKBIA, MMP1, DDR2, CBR3, PRL, FGA, SFTPC, ARSB, FOS, SNAI2, SP3, ASXL3, and IARS2) and used these genes to build a more reliable model of the FCHL risk connected to TAC. In the future, these genes could be used for predicting the risk of developing hyperlipidemia and may also serve as target genes for the treatment of hyperlipidemia. In addition, our findings reveal for the first time that memory and immature B cells in immunosuppressed states may play a role in lipid metabolism.

FOS (proto-oncogene, AP-1 transcription factor subunit) is a gene that codes for proteins linked to diseases such as congenital extensive lipodystrophy and osteoblastoma, and related pathways include the myd88-dependent cascade initiated by endosomes and prolactin signaling. FOS proteins play a pivotal role as regulators influencing cell proliferation, differentiation, and transformation. Additionally, in certain instances, the expression of FOS has been correlated with apoptotic cell death. Our study found that among the 14 key genes, FOS possessed the greatest functional resemblance to other key genes and was positively correlated with immune B-cell memory in the immune infiltration analysis ($r = 0.90$, $P < 0.05$). Memory B cells revealed significant differences between the high- and low-FCHL subgroups. However, the manner in which these two factors influenced each other remains unknown.

NFKBIA, which encodes the protein NFKB inhibitor α, is implicated in various diseases. Conditions associated with NFKBIA include ectodermal dysplasia and immunodeficiency 2. The related pathways include the endosomal-initiated MyD88-dependent cascade reaction and the TNFR1 pathway. The feasibility of NFKBIA as a biomarker was confirmed in an immunotoxicity study in zebrafish [73], in which the expression of NFKBIA had the highest contribution in a multifactorial logistic model (−10.38), which agrees with the findings of Zhou et al. [74]. Thus, NFKBIA could be used as a novel diagnostic biomarker for immune-related diseases and potential therapeutic targets [74].

SP3 is a member of the Sp1-related gene family and functions as a transcription factor that modulates gene expression by binding to GC-rich elements. Acting in a bifunctional manner, SP3 can either stimulate or repress the transcription of numerous genes. Diseases associated with SP3 include tanystomycosis and Coffin-Siris syndrome 1. The associated pathways include SLBP-independent transport of mature mRNA and protein metabolism. A significant positive correlation was found between immature B cells and SP3 in the high-risk group for FCHL in our immune infiltration analysis ($r = 0.90$, $P < 0.05$). For CD4CD25 regulatory T cell development, Foxp3 is required, and SP3 inhibits proximal FOXP3 promoter transactivation [75]. However, the function of SP3 in the differentiation of immature B cells requires further investigation.

Prolactin is an anterior pituitary hormone that regulates growth in a variety of organs, including immune cells, and it is encoded by the PRL gene. PRL's function in immunomodulation has received increased attention in recent years. PRL regulates the proliferation and survival of lymphocytes and myeloid cells and influences T cell pool selection by affecting the thymic microenvironment. Treg activity in autoimmune diseases is disrupted by PRL. By reducing the activation threshold of defective B cells, it also has an impact on B-cell tolerance [76,77]. The mechanism underlying the association between hyperprolactinemia and cardiovascular disease risk is currently controversial, and attention has been focused on the specific consequences of hyperprolactinemia on the regulation of food intake, body weight, glucose insulinemia, and lipid levels. These effects have been suggested to be directly related to PRL and hypogonadotropic hypogonadism, although it has also been suggested that these clinical manifestations may be related to dopamine

Table 7
mRNA-RBP interaction network nodes.

mRNA	RBP
ARSB	CTCF
ARSB	ELAVL1
ARSB	FAM120A
ARSB	HNRNPC
ARSB	HNRNPK
ARSB	HNRNPL
ARSB	HNRNPM
ARSB	HNRNPU
ARSB	IGF2BP1
ARSB	IGF2BP2
ARSB	IGF2BP3
ARSB	LIN28B
ARSB	MATR3
ARSB	MTDH
ARSB	PCBP2
ARSB	PRPF8
ARSB	RBFOX2
ARSB	TAF15
ARSB	TARDBP
ARSB	U2AF1
ARSB	U2AF2
ASXL3	CSTF2T
ASXL3	ELAVL1
ASXL3	FUS
ASXL3	PTBP1
ASXL3	TARDBP
ASXL3	U2AF1
CBR3	IGF2BP2
DDR2	ALYREF
DDR2	CSTF2T
DDR2	CTCF
DDR2	ELAVL1
DDR2	FUS
DDR2	G3BP1
DDR2	HDLBP
DDR2	HNRNPA2B1
DDR2	IGF2BP2
DDR2	IGF2BP3
DDR2	MOV10
DDR2	NUDT16L1
DDR2	NXF1
DDR2	RBM7
DDR2	RNPS1
DDR2	TARDBP
DDR2	U2AF1
DDR2	U2AF2
DDR2	WDR4
FGA	G3BP1
FGA	IGF2BP1
FGA	LIN28B
FGA	MTDH
FGA	RPS3
FGA	SND1
FGA	U2AF2
HTRA1	FUS
HTRA1	HDLBP
HTRA1	IGF2BP2
IARS2	ALYREF
IARS2	CPSF7
IARS2	CSTF2T
IARS2	CTCF
IARS2	DDX3X
IARS2	DHX36
IARS2	EIF4A3
IARS2	ELAVL1
IARS2	EWSR1
IARS2	FIP1L1
IARS2	FMR1

(*continued on next page*)

Table 7 (*continued*)

mRNA	RBP
IARS2	FUS
IARS2	FXR1
IARS2	FXR2
IARS2	G3BP1
IARS2	HDLBP
IARS2	HNRNPA2B1
IARS2	HNRNPC
IARS2	IGF2BP1
IARS2	LIN28A
IARS2	LIN28B
IARS2	MTA1
IARS2	MTDH
IARS2	PRPF8
IARS2	RBM15
IARS2	RBM15B
IARS2	RBMX
IARS2	RNPS1
IARS2	SCAF4
IARS2	SCAF8
IARS2	SOX2
IARS2	TARDBP
IARS2	U2AF1
IARS2	U2AF2
IARS2	YBX1
IARS2	YTHDC1
IARS2	YTHDF1
IARS2	YTHDF3
IARS2	ZC3H7B
MMP1	RBM47
NFKBIA	PRPF8
NFKBIA	RNPS1
SP3	ALYREF
SP3	CHTOP
SP3	CPSF6
SP3	CPSF7
SP3	CSTF2
SP3	CSTF2T
SP3	CTCF
SP3	DDX3X
SP3	DDX54
SP3	EIF3A
SP3	EIF4A3
SP3	ELAVL1
SP3	EWSR1
SP3	FMR1
SP3	FUBP1
SP3	FUS
SP3	G3BP1
SP3	HNRNPA2B1
SP3	HNRNPC
SP3	IGF2BP1
SP3	IGF2BP2
SP3	IGF2BP3
SP3	ILF3
SP3	KHDRBS2
SP3	KHSRP
SP3	MOV10
SP3	MTDH
SP3	NUDT21
SP3	PCBP2
SP3	PRPF8
SP3	PTBP1
SP3	RBFOX2
SP3	RBM10
SP3	RBM15
SP3	RBM15B
SP3	RBM20
SP3	RBM47
SP3	RBMX

(*continued on next page*)

Table 7 (*continued*)

mRNA	RBP
SP3	RNPS1
SP3	SCAF4
SP3	SCAF8
SP3	SOX2
SP3	SRSF1
SP3	SRSF7
SP3	TARDBP
SP3	TIA1
SP3	U2AF1
SP3	U2AF2
SP3	WDR33
SP3	WDR4
SP3	YTHDC1
SP3	YTHDF1
SP3	YTHDF2
SP3	YTHDF3
SP3	ZC3H7B
SP3	ZCCHC14

RBP, RNA binding protein.

agonist therapy [78]. Therefore, the specific role of PRL in the immune system and lipid metabolism should be studied in depth in renal transplant patients taking immunosuppressants such as TAC over long periods.

Recombinant Activation 2 (RAG2) is a gene that codes for proteins linked to Omenn's syndrome and granulomatous combined cellular and humoral immunodeficiency. The pathways involved include cytokine signaling in the immune system and the RAF/MAP kinase cascade. Innate immunodeficiencies, which include more than 480 disorders, are primarily caused by pathogenic variants of the RAG gene involved in immune function [79]. The utilization of CRISPR-Cas9 technology to engineer the RAG2 locus, involving the complete replacement of the coding sequence, presents a potential therapeutic avenue for the in vitro manipulation of hematopoietic stem and progenitor cells derived from the patient. This innovative approach is a viable alternative to allogeneic hematopoietic stem cell transplantation [80]. It has also been observed that immunodeficient PFP/Rag2 mice have higher adipose tissue mass and hepatic lipid accumulation with age compared to wild-type C57BL/6N mice [81]. As an important recombinase-activating gene, RAG1/RAG2 recombination is limited and controlled by complex motifs, various TFs, and cell cycle protein-dependent kinases. However, whether RAG2 is a key gene was not determined in our study. Immature B lymphocytes may play a role in lipid metabolism based on the KEGG FoxO signaling pathway and primary immunodeficiency pathways. However, further validation is required. In forthcoming research, we intend to utilize human hepatocytes and immune cell lines to perform overexpression and knockdown experiments of the RAG2 gene, thereby assessing its impact on lipid metabolism indicators. Additionally, we will develop in vivo mouse models featuring RAG2 gene knockout or overexpression and induce hyperlipidemia through a high-fat diet to investigate RAG2's specific role in this condition. This will be accompanied by histological assessments and molecular biological analyses. Furthermore, we aim to employ bioinformatics tools to construct maps of RAG2-related signaling pathways. We anticipate that these studies will elucidate RAG2's precise functions in lipid metabolism and immune regulation, potentially offering novel targets and therapeutic strategies for hyperlipidemia diagnosis and management.

In addition, to understand the effect of immunosuppressive status on hyperlipidemia, we computed the infiltration abundance of immune cells between the high and low subgroups of FCHL and identified two types of immune cells associated with FCHL: memory B cells and immature B cells. A significant positive correlation was observed between memory B cells and the key gene FOS in the high-risk group ($r = 0.90$, $P < 0.05$), between immature B cells and the key gene FOS in the low-risk group ($r = 0.94$, $P < 0.05$), and between immature B cells and the key gene SP3 in the high-risk group (high) ($r = 0.90$, $P < 0.05$). These findings imply that the functional state of immune cells may be influenced by lipid metabolism and that the correlation between the two may be regulated by some key genes. Research has demonstrated a significant interplay between the immune system and lipid metabolism. Melissa D. Lempicki and colleagues propose that B cell-activating factor (BAFF), a cytokine within the TNF family, plays a pivotal role in governing the homeostasis and peripheral tolerance of B2 cells. Their studies indicate that in white adipose tissue (WAT) of mice with neutralized BAFF, both B cell activation and phagocytosis pathways are compromised [82].T-lymphocyte are crucial in mediating inflammation within adipose tissue (AT). Studies have shown that obesity is correlated with an increased presence of activated $CD20^{+}$ T cells [83]. Additionally, adipose stem cells regulate the infiltration of T cells into adipose tissue under obese conditions, underscoring their pivotal role in modulating adaptive immunity and mitigating adipose inflammation associated with obesity [84].These studies underscore the pivotal role of immune cells in metabolic regulation and the development of hyperlipidemia.

However, the particular functions of these key genes in immune regulation and lipid metabolism require further investigation.

To reveal the potential functional mechanisms of FCHL and TRDEGs, we analyzed two pathways that were statistically significant ($P < 0.05$) in the GSVA: cholesterol homeostasis, and complementation. Finally, we predicted the molecular functional role relationships of the 14 key genes according to mRNA, RBPs, small-molecule compounds/drugs, and TFs using different types of databases.

Cholesterol homeostasis is a crucial parameter that influences lipid metabolism within the human body. Despite its paramount importance, the specific chemical reaction network responsible for regulating cellular cholesterol homeostasis remains elusive [85].

Table 8
mRNA-drugs interaction network nodes.

mRNA	Drug
ASXL3	Valproic Acid
DDR2	Valproic Acid
FGA	Benzo(a)pyrene
FGA	Cyclosporine
FGA	Ethinyl Estradiol
FGA	ethinyl estradiol-desogestrel combination
FGA	Gestodene
FGA	Tetrachlorodibenzodioxin
FGA	Valproic Acid
FOS	afimoxifene
FOS	Arachidonic Acid
FOS	Arsenic Trioxide
FOS	Asbestos, Crocidolite
FOS	Benzo(a)pyrene
FOS	bisphenol A
FOS	Cadmium
FOS	Cadmium Chloride
FOS	Cyclosporine
FOS	Estradiol
FOS	Formaldehyde
FOS	Fulvestrant
FOS	(+)-JQ1 compound
FOS	Oxygen
FOS	Particulate Matter
FOS	Quercetin
FOS	Resveratrol
FOS	sodium arsenite
FOS	Tetrachlorodibenzodioxin
FOS	Tetradecanoylphorbol Acetate
FOS	Tobacco Smoke Pollution
FOS	Valproic Acid
IARS2	Acetaminophen
MMP1	Benzo(a)pyrene
MMP1	Glucosamine
MMP1	(+)-JQ1 compound
MMP1	Lipopolysaccharides
MMP1	Particulate Matter
MMP1	Quercetin
MMP1	Silicon Dioxide
MMP1	Smoke
MMP1	sodium arsenite
MMP1	Tetrachlorodibenzodioxin
MMP1	Tetradecanoylphorbol Acetate
MMP1	Tobacco Smoke Pollution
MMP1	Vehicle Emissions
NFKBIA	Asbestos, Crocidolite
NFKBIA	Calcimycin
NFKBIA	Curcumin
NFKBIA	Cycloheximide
NFKBIA	Dexamethasone
NFKBIA	Doxorubicin
NFKBIA	Hydrogen Peroxide
NFKBIA	Lipopolysaccharides
NFKBIA	Melitten
NFKBIA	nickel chloride
NFKBIA	Resveratrol
NFKBIA	Tetrachlorodibenzodioxin
NFKBIA	Tetradecanoylphorbol Acetate
NFKBIA	Tobacco Smoke Pollution
NFKBIA	Valproic Acid
NFKBIA	vanadyl sulfate
PRL	8-Bromo Cyclic Adenosine Monophosphate
PRL	Amisulpride
PRL	Aripiprazole
PRL	Bromocriptine
PRL	Contraceptives, Oral
PRL	Fenfluramine
PRL	Haloperidol

(*continued on next page*)

Table 8 (*continued*)

mRNA	Drug
PRL	Metoclopramide
PRL	Olanzapine
PRL	Progesterone
PRL	Risperidone
PRL	Sulpiride
PRL	Testosterone
SNAI2	Arsenic Trioxide
SNAI2	Estradiol
SNAI2	Tetrachlorodibenzodioxin
SNAI2	Valproic Acid
SP3	Valproic Acid

The maintenance of cholesterol homeostasis hinges on intricate processes, such as biosynthesis, uptake, efflux, transport, storage, utilization, and/or excretion of cholesterol [86]. The SREBP transcription factor family is implicated in orchestrating the transcription of critical rate-limiting cholesterogenic and lipogenic proteins, thereby governing cholesterol production. Immune cells undergo activation, differentiation, and proliferation in response to danger signals. This intricate process is crucial for eliminating threats and involves metabolic reprogramming of both catabolic and anabolic pathways, thereby generating metabolites pivotal in regulating the immune response [87]. Additionally, transintestinal cholesterol excretion, a non-biliary pathway, facilitates the elimination of excess cholesterol from the body through feces [88]. Targeting this pathway holds promise for stimulating cholesterol elimination and mitigating the risk of cardiovascular diseases [89].

The complement protein C1q regulates changes in the lipid metabolism of cholesterol and TGs: C1q regulates the expression of genes connected to Janus kinase and JAK-STAT signaling, PPAR signaling, and TLR signaling, contributing favorably to the inflammatory response during early atherosclerotic lipoproteins [90]. Cholesterol crystals are identified by the complement system's classical pathways and lectin, which cause inflammatory mediators to be released and C3 and C5 to become activated, which has also been of interest [91]. In patients with the HNF1A variant, the dominant-negative HNF1α mutant promotes hepatic steatosis through regulation of hepatic complement factor D and inflammation [92]. Pathway analyses revealed that the complement and coagulation pathways are the primary biological processes connected to changes in omega-3 polyunsaturated fatty acids [93]. Our study also provides new evidence of the possible role of the complement system in lipid metabolism.

This study has several strengths. First, In our manuscript, we utilized data from the GSE1010 dataset, which was derived from a study on familial combined hyperlipidemia (FCHL). Metabolic disorders triggered by FCHL and tacrolimus (TAC) exhibit analogous disruptions in lipid metabolic pathways, with hypertriglyceridemia (HTG) emerging as a characteristic feature of both conditions. Although the GSE1010 dataset is centered on FCHL, uncovering the gene expression patterns and regulatory mechanisms associated with lipid metabolism in this context can inform and extend our understanding to TAC-related metabolic research. By harnessing the extensive capabilities of bioinformatics analysis, we can elucidate the underlying mechanisms of TAC-induced lipid metabolic disorders using FCHL data as a reference. This approach may facilitate the identification of potential therapeutic targets for intervention. Second, Our research plays a significant role in pinpointing key genes that are implicated in both tacrolimus metabolism and lipid metabolism. This work is instrumental in facilitating the early detection of risk genes associated with hypertriglyceridemia in organ transplant recipients. By leveraging these findings, we aim to rationally tailor lipid-lowering drug therapies for patients identified at risk for hypertriglycerides, thereby enhancing their long-term outcomes and overall quality of life. Finally, tour multi-tiered molecular network analysis offers profound insights into the underlying molecular mechanisms of hypertriglyceridemia. This comprehensive understanding paves the way for the development of innovative lipid-lowering therapeutics and identifies potential targets for gene therapy approaches.

However, this study also has some shortcomings. First, the lack of specificity of the sequencing data due to the analysis of different database sources may have introduced sample bias. To mitigate the issue of sample bias stemming from the lack of specificity in sequencing data, we employed the pROC package to optimize the specificity, thereby minimizing this bias.Second, in vitro and in vivo experiments to validate the potential functional roles of the target genes on HTG are lacking, and further validation of the mechanistic roles of the 14 target genes on lipid metabolism under the influence of TAC must be performed in the future. The absence of experimental validation in transplanted animal models has constrained our capacity to elucidate the precise mechanisms underlying tacrolimus (TAC)-induced lipid metabolism disorders and to ascertain the actual functions of key genes within living systems.

Finally, direct evidence is lacking on the influence of TAC on lipid metabolism in patients with solid organ transplants. The expression of the 14 target genes should be verified using a hyperlipidemic mouse model, and whether the target genes under the immunosuppressed state undergo changes in immune cell function and lipid metabolism through both cholesterol homeostatic and complementary pathways should be determined. Such work will help confirm the role of these genes. In the future, we may be able to provide better interventions for the immune cell infiltration state and lipid metabolism abnormalities in humans.

5. Conclusion

The development of HTG may be functionally regulated by genes or associated with the application of immunosuppressive drugs and immune cell dysfunction. In addition, some genes have been suggested as potential targets for clinical therapies. Still, additional

Table 9
mRNA-TF interaction network nodes.

mRNA	TF
HTRA1	BRD3
HTRA1	EGR1
HTRA1	FOXA2
HTRA1	GATA1
HTRA1	HNF4A
HTRA1	NRF1
HTRA1	SPI1
HTRA1	USF1
HTRA1	USF2
ARSB	BRD3
ARSB	EGR1
ARSB	FOXA2
ARSB	GATA1
ARSB	HNF4A
ARSB	NRF1
ARSB	SPI1
ARSB	USF1
ARSB	USF2
ASXL3	TEAD4
DDR2	FOS
DDR2	GATA2
DDR2	EP300
FGA	CEBPA
FGA	CEBPB
FGA	ESR1
FGA	FOXA1
FGA	FOXA2
FGA	GATA4
FGA	GATA6
FGA	HNF4A
FGA	JUN
FOS	EP300
IARS2	EGR1
IARS2	ERG
IARS2	FOXA1
IARS2	SP1
MMP1	CEBPB
MMP1	CTCF
MMP1	ERG
MMP1	FOS
MMP1	FOXA1
MMP1	FOXA2
MMP1	JUND
MMP1	RAD21
MMP1	SMC3
MMP1	STAG1
MMP1	STAT3
MMP1	TAL1
NFKBIA	AR
NFKBIA	CEBPA
NFKBIA	CEBPB
NFKBIA	CTCF
NFKBIA	EGR1
NFKBIA	EP300
NFKBIA	ERG
NFKBIA	FLI1
NFKBIA	FOS
NFKBIA	FOXA1
NFKBIA	FOXA2
NFKBIA	GRHL2
NFKBIA	HNF4A
NFKBIA	HOXB13
NFKBIA	MAX
NFKBIA	MYC
NFKBIA	NFKB1
NFKBIA	NRF1
NFKBIA	POLR2A
NFKBIA	RAD21

(*continued on next page*)

Table 9 (*continued*)

mRNA	TF
NFKBIA	RELA
NFKBIA	RUNX1
NFKBIA	SMARCA4
NFKBIA	SMC1A
NFKBIA	SMC3
NFKBIA	SPI1
NFKBIA	STAG1
NFKBIA	STAT3
NFKBIA	TEAD4
NFKBIA	USF1
NFKBIA	YY1
NFKBIA	ZEB1
NFKBIA	ZNF384
PRL	TEAD4
PRL	AR
PRL	EP300
PRL	ESR1
PRL	FOS
PRL	FOSL2
PRL	GATA2
PRL	GATA3
PRL	NR3C1
PRL	SMARCA4
SNAI2	EP300
SNAI2	ERG
SNAI2	ETS1
SNAI2	FOS
SNAI2	GATA2
SNAI2	GATA3
SNAI2	JUND
SNAI2	MAFK
SNAI2	MAX
SNAI2	MED1
SNAI2	MYC
SNAI2	NR3C1
SNAI2	POLR2A
SNAI2	RAD21
SNAI2	RELA
SNAI2	RUNX1
SNAI2	SMARCA4
SNAI2	SMC3
SNAI2	SPI1
SNAI2	SS18
SNAI2	STAG1
SNAI2	TCF12
SNAI2	TP63
SNAI2	AR
SNAI2	CEBPA
SNAI2	CEBPB
SNAI2	CTCF
SP3	ELF1
SP3	ERG
SP3	ESR1
SP3	ETS1
SP3	FOS
SP3	FOXA1
SP3	FOXA2
SP3	GABPA
SP3	HOXB13
SP3	POLR2A
SP3	SP1
SP3	STAT3
SP3	CEBPA
SP3	TBP
SP3	CEBPB
SP3	TP53

TF: Transcription factors.

investigations are required to elucidate the potential routes that profoundly influence HTG development. Through our study, we screened 14 key genes with statistically significant FHCL family genes and DRGs; these 14 key genes were employed in the creation of a risk model for the development of hyperlipidemia. We identified two types of immune cells associated with FCHL, namely, memory B cells and immature B cells, and two pathways related to the potential function of FCHL and TRDEGs, namely, cholesterol homeostasis, and complementation. This study is poised to enhance our comprehension of the roles played by familial combined hyperlipidemia (FCHL) and tacrolimus-associated genes in lipid metabolism. It will also offer guidance for future research aimed at elucidating the functions of these pivotal genes. Notably, the predictive model we have developed lays the groundwork for assessing the pivotal roles of key genes in hyperlipidemia, potentially identifying them as promising therapeutic targets for the treatment of hypertriglyceridemia in the future.

CRediT authorship contribution statement

Yuan Xu: Writing – original draft. **Hongfei He:** Methodology. **Haiyang Li:** Writing – review & editing.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The scRNA-seq analysis datasets utilized in this study have been deposited in the Gene Expression Omnibus (GEO) repository under the accession number GSE1010 (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1010). Additional datasets can be obtained from the corresponding authors upon reasonable request.

Funding

Financial support for this study was graciously provided by Guizhou Provincial Basic Research Programme (Natural Science Category) Project (Qiankehe Basic-ZK[2023]376); the Science and Technology Fund Project of Guizhou Provincial Health Commission (gzwkj2023-164); and the Cultivation Fund of the National Natural Science Foundation of Guizhou Medical University (gyfynsfc-2021-36).

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:YUAN XU reports financial support was provided by Affiliated Hospital of Guizhou Medical University. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Abbreviations

AUC	Area under curve
BP	Biological processes
CC	Cellular components
CTD	Comparative Toxicogenomics Database
DCA	Decision Curve Analysis
DEG	Differentially expressed genes
FCHL	Familial combined hyperlipidemia
GO	Gene Ontology
GSEA	Gene set enrichment analysis
GSVA	Gene Set Variation Analysis
HTG	Hypertriglyceridemia
KEGG	Kyoto Encyclopedia of Genes and Genomes
LASSO	Least Absolute Shrinkage and Selection Operator

LDL	Low-density lipoprotein
LPL	Lipoprotein lipase
MF	Molecular functions
MSigDB	Molecular Signatures Database
RBP	RNA-binding proteins
ROC	Receiver operating characteristic
SsGSEA	Single-sample gene set enrichment analysis
TAC	Tacrolimus
TF	Transcription factors
TRDEG	Tacrolimus-related differentially expressed genes
VLDL	Very-low-density lipoprotein
WGCNA	Weighted gene co-expression network analysis

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e41705.

References

[1] N. Ichimaru, S. Takahara, Y. Kokado, J.D. Wang, M. Hatori, H. Kameoka, T. Inoue, A. Okuyama, Changes in lipid metabolism and effect of simvastatin in renal transplant recipients induced by cyclosporine or tacrolimus, Atherosclerosis 158 (2) (2001 Oct) 417–423, https://doi.org/10.1016/s0021-9150(01)00438-5.

[2] C. Stătescu, L. Anghel, L.C. Benchea, B.S. Tudurachi, A. Leonte, A. Zăvoi, et al., A systematic review on the risk modulators of myocardial infarction in the "young"-implications of lipoprotein (a), Int. J. Mol. Sci. 24 (2023) 5927, https://doi.org/10.3390/ijms24065927.

[3] R. Aggarwal, R.W. Yeh, K.E. Joynt Maddox, R.K. Wadhera, Cardiovascular risk factor prevalence, treatment, and control in US adults aged 20 to 44 years, 2009 to March 2020, JAMA 329 (2023) 899–909, https://doi.org/10.1001/jama.2023.2307.

[4] Antonio Gaddi, AFG Cicero, FO Odoo, A Poli A, R Paoletti & On behalf of the Atherosclerosis and Metabolic Diseases Study Group (2007) Practical guidelines for familial combined hyperlipidemia diagnosis: an up-date, Vasc. Health Risk Manag., 3:6, 877-886, doi: 10.2147/vhrm.s12187425.

[5] M. Li, Z. Cui, S. Meng, T. Li, T. Kang, Q. Ye, et al., Associations between dietary glycemic index and glycemic load values and cardiometabolic risk factors in adults: findings from the China health and nutrition survey, Nutrients 13 (2020) 116, https://doi.org/10.3390/nu13010116.

[6] V. Mela, P. Ruiz-Limón, M. Balongo, H. Motahari Rad, A. Subiri-Verdugo, A. Gonzalez-Jimenez, et al., Mitochondrial homeostasis in obesity-related hypertriglyceridemia, J. Clin. Endocrinol. Metab. 107 (2022) 2203–2215, https://doi.org/10.1210/clinem/dgac332.

[7] K. Murawska, M. Krintus, M. Kuligowska-Prusinska, L. Szternel, A. Stefanska, G. Sypniewska, Relationship between serum angiopoietin-like proteins 3 and 8 and atherogenic lipid biomarkers in non-diabetic adults depends on gender and obesity, Nutrients 13 (2021) 4339, https://doi.org/10.3390/nu13124339.

[8] I. Jialal, A.T. Remaley, B. Adams-Huet, The triglyceride-waist circumference index is a valid biomarker of metabolic syndrome in African Americans, Am. J. Med. Sci. 365 (2023) 184–188, https://doi.org/10.1016/j.amjms.2022.11.003.

[9] K. Lee, E.L. Giovannucci, J. Kim, The effect of smoking and sex on the association between long-term alcohol consumption and metabolic syndrome in a middle-aged and older population, J. Epidemiol. 31 (2021) 249–258, https://doi.org/10.2188/jea.JE20190328.

[10] G.A. Roth, G.A. Mensah, C.O. Johnson, G. Addolorato, E. Ammirati, L.M. Baddour, et al., Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study, J. Am. Coll. Cardiol. 76 (2020) 2982–3021, https://doi.org/10.1016/j.jacc.2020.11.010 [published correction appears in J Am Coll Cardiol. 2021 Apr 20;77:1958-9].

[11] C.J. Willer, E.M. Schmidt, S. SenGupta, G.M. Peloso, S. Gustafsson, S. Kanoni, et al., Discovery and refinement of loci associated with lipid levels, Nat. Genet. 45 (2013) 1274–1283, https://doi.org/10.1038/ng.2797.

[12] R.A. Hegele, H.N. Ginsberg, M.J. Chapman, B.G. Nordestgaard, J.A. Kuivenhoven, M. Averna, et al., The polygenic nature of hypertriglyceridaemia: implications for definition, diagnosis, and management, Lancet Diabetes Endocrinol. 2 (2014) 655–666, https://doi.org/10.1016/S2213-8587(13)70191-8.

[13] B.A. Ference, J.J.P. Kastelein, K.K. Ray, H.N. Ginsberg, M.J. Chapman, C.J. Packard, et al., Association of triglyceride-lowering LPL variants and LDL-C-lowering LDLR variants with risk of coronary heart disease, JAMA 321 (2019) 364–373, https://doi.org/10.1001/jama.2018.20045.

[14] M. Cuchel, E.A. Meagher, H. du Toit Theron, D.J. Blom, A.D. Marais, R.A. Hegele, et al., Efficacy and safety of a microsomal triglyceride transfer protein inhibitor in patients with homozygous familial hypercholesterolaemia: a single-arm, open-label, phase 3 study, Lancet 381 (2013) 40–46, https://doi.org/10.1016/S0140-6736(12)61731-0.

[15] F.J. Raal, R.D. Santos, D.J. Blom, A.D. Marais, M.J. Charng, W.C. Cromwell, et al., Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial, Lancet 375 (2010) 998–1006, https://doi.org/10.1016/S0140-6736(10)60284-X.

[16] F.J. Raal, R.S. Rosenson, L.F. Reeskamp, G.K. Hovingh, J.J.P. Kastelein, P. Rubba, et al., Evinacumab for homozygous familial hypercholesterolemia, N. Engl. J. Med. 383 (2020) 711–720, https://doi.org/10.1056/NEJMoa2004215.

[17] J.L. Witztum, D. Gaudet, S.D. Freedman, V.J. Alexander, A. Digenio, K.R. Williams, et al., Volanesorsen and triglyceride levels in familial chylomicronemia syndrome, N. Engl. J. Med. 381 (2019) 531–542, https://doi.org/10.1056/NEJMoa1715944.

[18] B.T. Thommen, J.M. Dziekan, F. Achcar, S. Tjia, A. Passecker, K. Buczak, et al., Genetic validation of PfFKBP35 as an antimalarial drug target, Elife 12 (2023) RP86975, https://doi.org/10.7554/eLife.86975.

[19] D.D. Zhang, X.P. Duan, K. Mutig, F. Rausch, Y. Xiao, J.Y. Zheng, et al., Calcineurin inhibitors stimulate Kir4.1/Kir5.1 of the distal convoluted tubule to increase NaCl cotransporter, JCI Insight 8 (2023) e165987, https://doi.org/10.1172/jci.insight.165987.

[20] M.E. Mohamed, D.P. Schladt, W. Guan, B. Wu, J. van Setten, B.J. Keating, et al., Tacrolimus troughs and genetic determinants of metabolism in kidney transplant recipients: a comparison of four ancestry groups, Am. J. Transplant. 19 (2019) 2795–2804, https://doi.org/10.1111/ajt.15385.

[21] V. Dziedziejko, K. Safranow, M. Kijko-Nowak, D. Malinowski, L. Domanski, A. Pawlik, Leptin receptor gene polymorphisms in kidney transplant patients with post-transplant diabetes mellitus treated with tacrolimus, Int Immunopharmacol 124 (2023) 110989, https://doi.org/10.1016/j.intimp.2023.110989.

[22] V. Dziedziejko, K. Safranow, M. Kijko-Nowak, J. Sieńko, D. Malinowski, K. Szumilas, et al., The association between CDKAL1 gene rs10946398 polymorphism and post-transplant diabetes in kidney allograft recipients treated with tacrolimus, Genes 14 (2023) 1595, https://doi.org/10.3390/genes14081595.

[23] L.J. Chen, Y. Xin, M.X. Yuan, C.Y. Ji, Y.M. Peng, Q. Yin, CircFOXN2 alleviates glucocorticoid- and tacrolimus-induced dyslipidemia by reducing FASN mRNA stability by binding to PTBP1 during liver transplantation, Am J Physiol Cell Physiol 325 (2023) C796–C806, https://doi.org/10.1152/ajpcell.00462.2022.

[24] H.L.M. Steffen, J.L.C. Anderson, M.L. Poot, Y. Lei, M.A. Connelly, S.J.L. Bakker, et al., Proteoglycan binding as proatherogenic function metric of apoB-containing lipoproteins and chronic kidney graft failure, J. Lipid Res. 62 (2021) 100083, https://doi.org/10.1016/j.jlr.2021.100083.

[25] G. Iannuzzo, G. Cuomo, A. Di Lorenzo, M. Tripaldella, V. Mallardo, P. Iaccarino Idelson, et al., Dyslipidemia in transplant patients: which therapy? J. Clin. Med. 11 (2022) 4080, https://doi.org/10.3390/jcm11144080.
[26] A.K. Israni, J.J. Snyder, M.A. Skeans, Y. Peng, J.R. Maclean, E.D. Weinhandl, et al., Predicting coronary heart disease after kidney transplantation: patient Outcomes in Renal Transplantation (PORT) Study, Am. J. Transplant. 10 (2010) 338–353, https://doi.org/10.1111/j.1600-6143.2009.02949.x.
[27] M. Albeldawi, A. Aggarwal, S. Madhwal, J. Cywinski, R. Lopez, B. Eghtesad, et al., Cumulative risk of cardiovascular events after orthotopic liver transplantation, Liver Transpl 18 (2012) 370–375, https://doi.org/10.1002/lt.22468.
[28] W.H. Lim, C. Tan, J. Xiao, D.J.H. Tan, C.H. Ng, J.N. Yong, et al., De novo metabolic syndrome after liver transplantation: a meta-analysis on cumulative incidence, risk factors, and outcomes, Liver Transpl 29 (2023) 413–421, https://doi.org/10.1097/LVT.0000000000000004.
[29] A. Mehta, Managing dyslipidemia in solid organ transplant patients, Indian Heart J. 76 (Suppl 1) (2024 Mar) S93–S95, https://doi.org/10.1016/j.ihj.2024.01.004. Suppl 1.
[30] L.J. Chen, Y. Xin, M.X. Yuan, C.Y. Ji, Y.M. Peng, Q. Yin, CircFOXN2 alleviates glucocorticoid- and tacrolimus-induced dyslipidemia by reducing FASN mRNA stability by binding to PTBP1 during liver transplantation, Am J Physiol Cell Physiol 325 (3) (2023 Sep 1) C796–C806, https://doi.org/10.1152/ajpcell.00462.2022.
[31] K.B. Smedlund, E.R. Sanchez, T.D. Hinds Jr., FKBP51 and the molecular chaperoning of metabolism, Trends Endocrinol Metab 32 (11) (2021 Nov) 862–874, https://doi.org/10.1016/j.tem.2021.08.003.
[32] J. Sułowicz, D. Cieniawski, E. Ignacak, A. Bętkowska-Prokop, M. Kuźniewski, W. Sułowicz, Comparison of kidney transplant function, lipid metabolism disorders, and glucose and hemoglobin concentration in transplant patients treated with proliferation signal inhibitor (everolimus) or calcineurin inhibitor (tacrolimus), Transplant. Proc. 52 (8) (2020 Oct) 2347–2351, https://doi.org/10.1016/j.transproceed.2020.02.122.
[33] C. Ponticelli, L. Arnaboldi, G. Moroni, A. Corsini, Treatment of dyslipidemia in kidney transplantation, Expert Opin Drug Saf 19 (3) (2020 Mar) 257–267, https://doi.org/10.1080/14740338.2020.1732921.
[34] D.R. Migliozzi, N.J. Asal, Clinical controversy in transplantation: tacrolimus versus cyclosporine in statin drug interactions, Ann. Pharmacother. 54 (2) (2020 Feb) 171–177, https://doi.org/10.1177/1060028019871891.
[35] B.A. Warden, P.B. Duell, Management of dyslipidemia in adult solid organ transplant recipients, J Clin Lipidol 13 (2) (2019 Mar-Apr) 231–245, https://doi.org/10.1016/j.jacl.2019.01.011.
[36] B.A. Warden, P.B. Duell, Management of dyslipidemia in adult solid organ transplant recipients, J Clin Lipidol 13 (2019) 231–245, https://doi.org/10.1016/j.jacl.2019.01.011.
[37] Y. Xu, B. Yang, C.X. Chen, K.J. Zhu, Y.L. Niu, H.Y. Li, Risk factors of new-onset hypertriglyceridemia in kidney transplant recipients: a single-centre analysis, Organ Transplant 14 (2023) 691–699, https://doi.org/10.3969/j.issn.1674-7445.2023055.
[38] L. Quteineh, P.Y. Bochud, D. Golshayan, S. Crettol, J.P. Venetz, O. Manuel, et al., CRTC2 polymorphism as a risk factor for the incidence of metabolic syndrome in patients with solid organ transplantation, Pharmacogenomics J. 17 (2017) 69–75, https://doi.org/10.1038/tpj.2015.82.
[39] G. Stelzer, N. Rosen, I. Plaschkes, S. Zimmerman, M. Twik, S. Fishilevich, et al., The GeneCards suite: from gene data mining to disease genome sequence analyses, Curr Protoc Bioinformatics 54 (2016), https://doi.org/10.1002/cpbi.5, 1.30.1-1.30.33.
[40] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J.P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst 1 (2015) 417–425, https://doi.org/10.1016/j.cels.2015.12.004.
[41] T. Barrett, D.B. Troup, S.E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, et al., NCBI GEO: mining tens of millions of expression profiles–database and tools update, Nucleic Acids Res. 35 (2007 Jan) D760–D765, https://doi.org/10.1093/nar/gkl887.
[42] S. Davis, P.S. Meltzer, GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor, Bioinformatics 23 (2007 Jul 15) 1846–1847, https://doi.org/10.1093/bioinformatics/btm254.
[43] A.P. Davis, C.J. Grondin, R.J. Johnson, D. Sciaky, J. Wiegers, T.C. Wiegers, et al., Comparative Toxicogenomics database (CTD): update 2021, Nucleic Acids Res. 49 (2021) D1138–D1143, https://doi.org/10.1093/nar/gkaa891.
[44] D. Szklarczyk, A. Santos, C. von Mering, L.J. Jensen, P. Bork, M. Kuhn, Stitch 5: augmenting protein-chemical interaction networks with tissue and affinity data, Nucleic Acids Res. 44 (2016) D380–D384, https://doi.org/10.1093/nar/gkv1277.
[45] J.H. Li, S. Liu, H. Zhou, L.H. Qu, J.H. Yang, starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data, Nucleic Acids Res. 42 (Database issue) (2014 Jan) D92–D97, https://doi.org/10.1093/nar/gkt1248.
[46] J. Huang, W. Zheng, P. Zhang, Q. Lin, Z. Chen, J. Xuan, C. Liu, D. Wu, Q. Huang, L. Zheng, S. Liu, K. Zhou, L. Qu, B. Li, J. Yang, ChIPBase v3.0: the encyclopedia of transcriptional regulations of non-coding RNAs and protein-coding genes, Nucleic Acids Res. 51 (D1) (2023 Jan 6) D46–D56, https://doi.org/10.1093/nar/gkac1067.
[47] L. Chen, Y. Peng, C. Ji, M. Yuan, Q. Yin, Network pharmacology-based analysis of the role of tacrolimus in liver transplantation, Saudi J. Biol. Sci. 28 (2021) 1569–1575, https://doi.org/10.1016/j.sjbs.2020.12.050.
[48] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47, https://doi.org/10.1093/nar/gkv007.
[49] G. Yu, Gene ontology semantic similarity analysis using GOSemSim, Methods Mol. Biol. 2117 (2020) 207–215, https://doi.org/10.1007/978-1-0716-0301-7_11.
[50] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30, https://doi.org/10.1093/nar/28.1.27.
[51] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (2012) 284–287, https://doi.org/10.1089/omi.2011.0118.
[52] B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis, Stat. Appl. Genet. Mol. Biol. 4 (2005) Article17, https://doi.org/10.2202/1544-6115.1128.
[53] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559, https://doi.org/10.1186/1471-2105-9-559.
[54] H.E. Gruber, G.L. Hoelscher, J.A. Ingram, E.N. Hanley Jr., Genome-wide analysis of pain-, nerve- and neurotrophin -related gene expression in the degenerating human annulus, Mol. Pain 8 (2012) 63.
[55] Y. Liu, H. Zhao, Variable importance-weighted random forests, Quant Biol 5 (2017) 338–351.
[56] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, Clin Epigenetics 11 (2019) 123, https://doi.org/10.1186/s13148-019-0730-1.
[57] W. Cai, M. van der Laan, Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator, Int. J. Biostat. 16 (2020) 20170070, https://doi.org/10.1515/ijb-2017-0070.
[58] J. Wu, H. Zhang, L. Li, M. Hu, L. Chen, B. Xu, et al., A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: a population-based analysis, Cancer Commun. 40 (2020) 301–312, https://doi.org/10.1002/cac2.12067.
[59] B. Van Calster, L. Wynants, J.F.M. Verbeek, J.Y. Verbakel, E. Christodoulou, A.J. Vickers, et al., Reporting and interpreting decision curve analysis: a guide for investigators, Eur. Urol. 74 (2018) 796–804, https://doi.org/10.1016/j.eururo.2018.08.038.
[60] J.N. Mandrekar, Receiver operating characteristic curve in diagnostic test assessment, J. Thorac. Oncol. 5 (2010) 1315–1316, https://doi.org/10.1097/JTO.0b013e3181ec173d.
[61] G. Yu, F. Li, Y. Qin, X. Bo, Y. Wu, S. Wang, GOSemSim: an R package for measuring semantic similarity among GO terms and gene products, Bioinformatics 26 (2010) 976–978, https://doi.org/10.1093/bioinformatics/btq064.
[62] H. Zhang, P. Meltzer, S. Davis, RCircos: an R package for Circos 2D track plots, BMC Bioinf. 14 (2013) 244, https://doi.org/10.1186/1471-2105-14-244.
[63] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.
[64] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7, https://doi.org/10.1186/1471-2105-14-7.

[65] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (2015) 453–457, doi: 0.1038/nmeth.3337.
[66] B. Xiao, L. Liu, A. Li, C. Xiang, P. Wang, H. Li, et al., Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma, Front. Oncol. 10 (2020) 607622, https://doi.org/10.3389/fonc.2020.607622.
[67] J.H. Li, S. Liu, H. Zhou, L.H. Qu, J.H. Yang, starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data, Nucleic Acids Res. 42 (2014) D92–D97, https://doi.org/10.1093/nar/gkt1248.
[68] K.R. Zhou, S. Liu, W.J. Sun, L.L. Zheng, H. Zhou, J.H. Yang, et al., ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data, Nucleic Acids Res. 45 (2017) D43–D50, https://doi.org/10.1093/nar/gkw965.
[69] L.Y. Mak, A.C. Chan, T.C. Wong, W.C. Dai, W.H. She, K.W. Ma, et al., High prevalence of de novo metabolic dysfunction-associated fatty liver disease after liver transplantation and the role of controlled attenuation parameter, BMC Gastroenterol. 23 (2023) 307, https://doi.org/10.1186/s12876-023-02940-y.
[70] Z. Zhang, L. Xu, X. Qiu, X. Yang, Z. Lian, X. Wei, et al., Fibroblast growth factor 21 (FGF21) attenuates tacrolimus-induced hepatic lipid accumulation through transcription factor EB (TFEB)-regulated lipophagy, J. Zhejiang Univ. - Sci. B 24 (2023) 485–495, https://doi.org/10.1631/jzus.B2200562.
[71] C. Zhang, K. Chen, R. Wei, G. Fan, X. Cai, L. Xu, B. Cen, J. Wang, H. Xie, S. Zheng, X. Xu, The circFASN/miR-33a pathway participates in tacrolimus-induced dysregulation of hepatic triglyceride homeostasis, Signal Transduct Target Ther 5 (1) (2020 Mar 27) 23, https://doi.org/10.1038/s41392-020-0105-2.
[72] Z. Zhang, L. Xu, X. Qiu, X. Yang, Z. Lian, X. Wei, D. Lu, X. Xu, Fibroblast growth factor 21 (FGF21) attenuates tacrolimus-induced hepatic lipid accumulation through transcription factor EB (TFEB)-regulated lipophagy, J. Zhejiang Univ. - Sci. B 24 (6) (2023 Jun 15) 485–495, https://doi.org/10.1631/jzus.B2200562.
[73] B. Luckner, F. Essfeld, S.U. Ayobahan, E. Richling, E. Eilebrecht, S. Eilebrecht, Transcriptomic profiling of TLR-7-mediated immune-challenge in zebrafish embryos in the presence and absence of glucocorticoid-induced immunosuppression, Ecotoxicol. Environ. Saf. 266 (2023) 115570, https://doi.org/10.1016/j.ecoenv.2023.115570.
[74] J. Zhou, J. Huang, Z. Li, Q. Song, Z. Yang, L. Wang, et al., Identification of aging-related biomarkers and immune infiltration characteristics in osteoarthritis based on bioinformatics analysis and machine learning, Front. Immunol. 14 (2023) 1168780, https://doi.org/10.3389/fimmu.2023.1168780.
[75] P. Eckerstorfer, M. Novy, S. Burgstaller-Muehlbacher, W. Paster, H.B. Schiller, H. Mayer, et al., Proximal human FOXP3 promoter transactivated by NF-κB and negatively controlled by feedback loop and SP3, Mol. Immunol. 47 (2010) 2094–2102, https://doi.org/10.1016/j.molimm.2010.04.002.
[76] V.V. Borba, G. Zandman-Goddard, Y. Shoenfeld, Prolactin and autoimmunity, Front. Immunol. 9 (2018) 73, https://doi.org/10.3389/fimmu.2018.00073.
[77] M.V. Legorreta-Haquet, P. Santana-Sánchez, L. Chávez-Sánchez, A.K. Chávez-Rueda, The effect of prolactin on immune cell subsets involved in SLE pathogenesis, Front. Immunol. 13 (2022) 1016427, https://doi.org/10.3389/fimmu.2022.1016427.
[78] A. Glezer, M.R. Santana, M.D. Bronstein, J. Donato, R.S. Jallad, The interplay between prolactin and cardiovascular disease, Front. Endocrinol. 13 (2022) 1018090, https://doi.org/10.3389/fendo.2022.1018090.
[79] N. Haque, T. Kawai, B.D. Ratnasinghe, J.B. Wagenknecht, R. Urrutia, L.D. Notarangelo, et al., RAG genomic variation causes autoimmune diseases through specific structure-based mechanisms of enzyme dysregulation, iScience 26 (2023) 108040, https://doi.org/10.1016/j.isci.2023.108040.
[80] D. Allen, O. Knop, B. Itkowitz, N. Kalter, M. Rosenberg, O. Iancu, et al., CRISPR-Cas9 engineering of the RAG2 locus via complete coding sequence replacement for therapeutic applications, Nat. Commun. 14 (2023) 6771, https://doi.org/10.1038/s41467-023-42036-5.
[81] N. Jabbari, M. Nawaz, J. Rezaie, Ionizing radiation increases the activity of exosomal secretory pathway in MCF-7 human breast cancer cells: a possible way to communicate resistance against radiotherapy, Int. J. Mol. Sci. 20 (2019) 3649, https://doi.org/10.3390/ijms20153649.
[82] A.C.O. Pinho, P. Barbosa, A. Lazaro, J.G. Tralhão, M.J. Pereira, A. Paiva, P. Laranjeira, E. Carvalho, Identification and characterization of circulating and adipose tissue infiltrated CD20+T cells from subjects with obesity that undergo bariatric surgery, Immunol. Lett. 269 (2024 Oct) 106911, https://doi.org/10.1016/j.imlet.2024.106911.
[83] M.D. Lempicki, J.A. Gray, G. Abuna, R.M. Murata, S. Divanovic, C.A. McNamara, A.K. Meher, BAFF neutralization impairs the autoantibody-mediated clearance of dead adipocytes and aggravates obesity-induced insulin resistance, Front. Immunol. 15 (2024 Aug 9) 1436900, https://doi.org/10.3389/fimmu.2024.1436900.
[84] A. Balcerczyk, A. Eljaafari, L. Pirola, Adipose stem cells drive T cell infiltration in obesity, Trends Endocrinol Metab (2024 Jun 29), https://doi.org/10.1016/j.tem.2024.06.013. S1043-2760(24)00174-7.
[85] R. Scheepers, R.P. Araujo, Robust homeostasis of cellular cholesterol is a consequence of endogenous antithetic integral control, Front. Cell Dev. Biol. 11 (2023) 1244297, https://doi.org/10.3389/fcell.2023.1244297.
[86] Y. Duan, K. Gong, S. Xu, F. Zhang, X. Meng, J. Han, Regulation of cholesterol homeostasis in health and diseases: from mechanisms to targeted therapeutics, Signal Transduct Target Ther 7 (2022) 265, https://doi.org/10.1038/s41392-022-01125-5.
[87] R. Kumar, N. Chhillar, D.S. Gupta, G. Kaur, S. Singhal, T. Chauhan, Cholesterol homeostasis, mechanisms of molecular pathways, and cardiac health: a current outlook, Curr. Probl. Cardiol. 49 (2023) 102081, https://doi.org/10.1016/j.cpcardiol.2023.102081.
[88] Y. Huang, Y.F. Wang, X.Z. Ruan, C.W. Lau, L. Wang, Y. Huang, The role of KLF2 in regulating hepatic lipogenesis and blood cholesterol homeostasis via the SCAP/SREBP pathway, J. Lipid Res. (2023) 100472, https://doi.org/10.1016/j.jlr.2023.100472.
[89] H. Xu, Y. Xin, J. Wang, Z. Liu, Y. Cao, W. Li, et al., The TICE pathway: mechanisms and potential clinical applications, Curr Atheroscler Rep 25 (2023) 653–662, https://doi.org/10.1007/s11883-023-01147-6.
[90] M.M. Ho, A. Manughian-Peter, W.R. Spivia, A. Taylor, D.A. Fraser, Macrophage molecular signalling and inflammatory responses during ingestion of atherogenic lipoproteins are modulated by complement protein C1q, Atherosclerosis 253 (2016) 38–46, https://doi.org/10.1016/j.atherosclerosis.2016.08.019.
[91] N. Niyonzima, B. Halvorsen, B. Sporsheim, P. Garred, P. Aukrust, T.E. Mollnes, et al., Complement activation by cholesterol crystals triggers a subsequent cytokine response, Mol. Immunol. 84 (2017) 43–50, https://doi.org/10.1016/j.molimm.2016.09.019.
[92] M. Liu, L. Liu, H. Guo, X. Fan, T. Liu, C. Xu, et al., Dominant-negative HNF1α mutant promotes liver steatosis and inflammation by regulating hepatic complement factor D, iScience 26 (2023) 108018, https://doi.org/10.1016/j.isci.2023.108018.
[93] S.R. Susai, C. Healy, D. Mongan, M. Heurich, J.F. Byrne, M. Cannon, et al., Evidence that complement and coagulation proteins are mediating the clinical response to omega-3 fatty acids: a mass spectrometry-based investigation in subjects at clinical high-risk for psychosis, Transl. Psychiatry 12 (2022) 454, https://doi.org/10.1038/s41398-022-02217-0.