Journal of Cardiothoracic Surgery

RESEARCH

Open Access

Identification of lipid metabolism-related biomarkers in familial hypercholesterolemia via integrated bioinformatics and machine learning approaches

Linji Long[1], Baosheng Zhu[1,2*] and Tao Lv[1,2*]

Abstract

Background Familial hypercholesterolemia (FH) is a genetic disorder characterized by imbalances in lipid metabolism, markedly increasing cardiovascular risk. Identifying lipid metabolism-related biomarkers is essential for understanding FH pathogenesis and developing therapeutic strategies.

Methods A comprehensive bioinformatics analysis was conducted using 776 lipid metabolism-related genes (LMRGs) from the MSigDB database and the GSE6054 dataset containing 10 FH and 13 healthy samples. Differentially expressed genes (DEGs) intersected with LMRGs to obtain candidate genes, followed by random forest and SVM-RFE machine learning to identify key biomarkers. Functional enrichment, subcellular localization, co-expression, transcription factor (TF) prediction, ceRNA network, and drug screening were subsequently performed.

Results We identified 429 DEGs and 13 candidate genes, refined to six key genes (STAR, GRHL1, BZRAP1, LGMN, PLA2G4D, PLA2G12B). ROC analysis demonstrated that all six key genes exhibited AUC values greater than 0.7, underscoring their diagnostic potential. Functional enrichment further revealed significant associations with the ribosome and spliceosome pathways. Subcellular localization suggested mitochondrial and extracellular functions. Co-expression showed a significant positive correlation between GRHL1 and PLA2G12B. TF prediction revealed 17 TF-gene interactions, while ceRNA analysis outlined regulatory relationships. Drug prediction identified 224 potential therapeutic compounds, with STAR showing the most interactions.

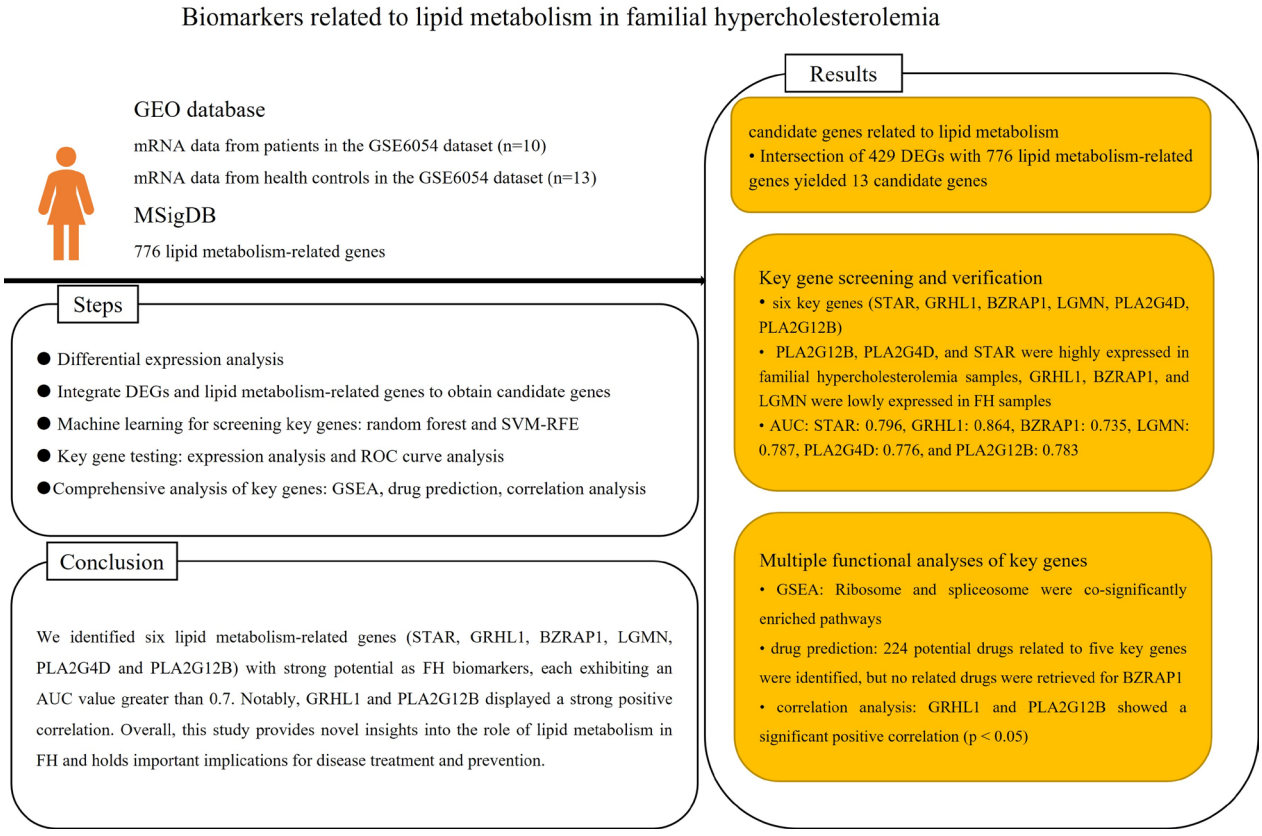
Conclusion This study highlights six lipid metabolism-related biomarkers in FH through integrated bioinformatics and machine learning. These findings provide new insights into FH mechanisms and potential therapeutic targets.

Keywords Familial hypercholesterolemia, Lipid metabolism, Machine learning, Biomarkers

*Correspondence:
Baosheng Zhu
baoshengzhu2024@163.com
Tao Lv
taolv851109@126.com

Full list of author information is available at the end of the article

Graphical abstract



Introduction

Familial hypercholesterolemia (FH) is a common genetic disorder affecting cholesterol metabolism. Unlike many other genetic diseases, FH is relatively prevalent, with an estimated global prevalence of approximately 20 million individuals. However, nearly 90% of affected individuals remain undiagnosed [1, 2]. FH significantly increases the risk of atherosclerotic diseases, often leading to premature mortality [3]. Among these, heterozygous FH (HeFH) is the most common genetic disorder, primarily inherited in an autosomal dominant or co-dominant manner. Its prevalence is approximately 1 in 250 subjects in most countries. Nevertheless, the most common cause of severe hypercholesterolemia is homozygous FH (HoFH). Without treatment, individuals with HoFH typically develop atherosclerotic cardiovascular disease (ASCVD) before the age of 18 and often succumb to coronary heart disease before the age of 30 [4]. The global incidence of HoFH is estimated to be approximately 1 in 300,000 [5, 6], yet the diagnosis rate remains below 5% [7]. FH is characterized by persistently elevated levels of low-density lipoprotein cholesterol (LDL-C) in the blood, leading to its extravascular deposition. This results in tendon and cutaneous xanthomas, xanthomas, corneal arcus, and vascular deposition, which contribute to the progressive development of premature atherosclerosis and coronary heart disease (CHD), ultimately increasing both morbidity and mortality [8].

Lipids play an important role in regulating biological behaviors, such as energy conversion, information recognition, and transmission [9]. As a result, lipid metabolism has garnered significant global attention. Emerging evidence indicates a strong association between lipid metabolism and the progression of chronic age-related diseases [10], as well as cancer proliferation and metastasis [2, 11]. In recent years, lipid metabolism has been increasingly recognized as a potential marker for various malignant tumors [12–14]. Moreover, the pivotal role of lipid metabolism related genes (LMRGs) in tumor occurrence, progression, and treatment has been well proven [15]. Studies have shown that monocytes in FH patients exhibit altered cholesterol metabolism [16]. The differential metabolites between HoFH patients and HeFH or non-FH subjects are mainly enriched in lipid metabolism pathways [17]. Furthermore, evidence suggests that the differential metabolites in HoFH are mainly related to cholesterol homeostasis, oxidative stress, inflammation, and the progression of ASCVD [17]. Therefore, lipid

metabolism may play a key role in the occurrence and progression of FH. However, there remains a significant lack of lipid metabolism-related biomarkers for the effective diagnosis of FH patients.

Biomarkers are important indicators for early disease detection and risk assessment [18]. Identifying key disease-associated biomarkers enables effective screening of high-risk populations and early-stage patients, thereby facilitating early intervention and improving clinical outcomes [19, 20]. Among different categories of biomarkers, lipid metabolism-related genes play important roles in disease development and progression and have been widely applied in diagnosis, prognostic evaluation, and treatment response monitoring [21]. It should be emphasized that the primary objective of this study is to identify biologically meaningful biomarkers and explore underlying mechanisms, rather than to develop or benchmark classification models.

This study integrates bioinformatics analysis and machine learning approaches to identify key genes differentiating FH samples from control samples, based on a dataset comprising six heterozygous FH (HeFH) and four homozygous FH (HoFH) samples from the GSE6077 dataset. Prior cross-condition transcriptomic and systems-level studies have shown that integrating gene expression signatures across biologically distinct yet related conditions can reveal shared pathological mechanisms, and similar end-to-end integrative bioinformatics frameworks have been widely applied and methodologically validated across multiple disease contexts [22, 23]. The expression levels of these key genes were analyzed, and their potential was assessed using Receiver Operating Characteristic (ROC) curve evaluations. Furthermore, the distribution patterns of key gene expression levels were examined. The subcellular localization of hub genes was predicted, and a competing endogenous RNA (ceRNA) network was constructed based on these hub genes. Through this comprehensive analytical framework, our study offers a novel approach to identifying multiple potential biomarkers and therapeutic targets, laying a foundation for future clinical applications and advancing the understanding of FH pathogenesis.

Materials and methods

Accessible datasets

At the outset, 776 lipid metabolism-related genes (LMRGs) were collected in the MSigDB (https://www.broadinstitute.org/msigdb) by screening six human lipid metabolism pathways, including encompass peroxisome proliferator-activated receptor alpha, metabolism of lipids, transcriptional regulation of white adipocyte differentiation, sphingolipid metabolism, glycerophospholipid metabolism and fatty acid metabolism. In addition, the mRNA dataset GSE6054 containing 10 FH patients and 13 healthy samples was obtained from the Gene Expression Omnibus (GEO) database. One of the FH patients consisted of 4 homozygous FH (HoFH) and 6 heterozygous FH (HeFH).

DEGs and candidate genes

At the beginning of the experiment, differentially expressed genes (DEGs) between FH and healthy samples in the GSEE6054 dataset were analyzed with the R package limma (version: 3.58.1) [24], volcano map and heat map were generated with ggplot2 (version: 3.3.6) [25] and pheatmap (version: 1.0.12) [26]. Following this, GO and KEGG enrichment analysis of DEGs was implemented by the R package clusterProfiler (version: 3.18.1) [27]. The biological processes and signaling pathways in which they are involved were visualized according to the R package ggplot2.

The results obtained by taking the intersection of the DEGs obtained from the aforementioned analysis with the 776 LMRGs were defined as candidate genes. Subsequently, a PPI network was constructed for the candidate genes based on the STRING database to view the linkages between the genes. Cytoscape (version: 3.9.1) [28] was used to visualize the results.

Machine learning based screening of key genes

In recent years, machine learning algorithms have been widely used in bioinformatics research and are capable of effectively capturing complex non-linear relationships in data [29]. Therefore, to identify potential key genes, Random Forest (RF) and Support Vector Machine–Recursive Feature Elimination (SVM-RFE) models were constructed based on the candidate genes for feature selection. They were implemented by the R packages randomForest (version: 4.7) [30] and e1071 (version: 1.7–14) [31]. Specifically, RF identified genes through mean decrease Gini, and SVM-RFE screened genes based on the minimum error rate lambda. The results of the two algorithms were intersected to obtain potential key genes. Immediately afterwards, the expression levels of the pre-screened key genes were analyzed between FH samples and healthy samples based on GSE6054. The ggpubr package (version: 0.6.0) (https://CRAN.R-project.org/package=ggpubr) was used for visualization and wilcox for differential expression tests. The significantly expressed genes were identified as key genes. In addition, ROC curve analysis was performed on these key genes based on GSE6054 to verify their effectiveness.

GSEA and subcellular distribution

To investigate the biological pathways in which key genes are involved in disease, Gene Set Enrichment Analysis (GSEA) was performed on these key genes using the clusterProfile package with the aid of the reference gene set

'c2.cp.kegg.v7.4.symbols.gmt'. The screening thresholds of $p < 0.05$ and false discovery rate (FDR) < 0.25 were considered statistically significant [32]. Subsequently, we adopted the Spearman method to analyze the co-expression patterns among these six key genes and generated correlation heat maps with the ggcorrplot package (version 0.1.4.1). Afterwards, based on the GeneCards (https://www.genecards.org/) database [31, 33], we pre dicted the subcellular locations of key genes to facilitate a detailed understanding of the subcellular organelles in which they reside.

Prediction of functionally similar genes and their transcription factors

The GeneMANIA database is a powerful tool for gene network construction and function prediction, enabling multi-dimensional exploration of the associations between genes and their roles in biological processes. We leveraged it to predict genes that are functionally similar to key genes and to generate interaction networks [34]. In addition, transcription factor (TF) prediction analysis provides a novel insight into the study of gene regulatory mechanisms [35]. We predicted transcription factors for the key gene interactions in the NetworkAnalyst database. Meanwhile, the TFs-mRNA relationship network was visualized by Cytoscape.

Competitive endogenous RNA networks and drug prediction

The regulatory network between key gene expression and cellular function was realized through the miRNet database (https://www.mirnet.ca). First, miRNAs predicted to bind and regulate mRNAs were predicted based on database analysis. Next, lncRNAs targeted by these miRNAs were retrieved. Subsequently, the regulatory relationships among mRNAs, miRNAs, and lncRNAs were integrated to construct a ceRNA network [36], which was visualized in Cytoscape. In addition, we retrieved potential therapeutic drugs for these key genes from the Comparative Toxicogenomics Database (CTD), and intuitively demonstrated the drug-gene interaction relationship in the Cytoscape software. Unless otherwise specified, all analyses were conducted using the default parameters of the corresponding R packages in R.

Results

GO and KEGG for 429 DEGs

Under the supervision of the threshold ($|\log2FC| > 1$, $p < 0.05$), 429 differentially expressed genes (DEGs) existed between FH and healthy samples. The heatmap revealed a higher proportion of highly expressed genes in FH samples, while a greater proportion of lowly expressed genes was observed in healthy samples (Fig. 1A). In addition, the volcano plot showed that of these DEGs, 264 were up-regulated DEGs and 165 were down-regulated DEGs (Fig. 1B). Immediately following, at a threshold $p < 0.05$ reconciliation, GO analysis revealed that these DEGs were significantly enriched in 111 terms, and KEGG showed that they were significantly enriched in 18 pathways.

Among the GO terms, 67 were associated with biological processes (BP), 12 with cellular components (CC), and 32 with molecular functions (MF). Bar plots were generated to display the top 10 most significantly enriched terms in each category (Fig. 1C). Of these, pathways such as fatty acid biosynthetic process, lipoprotein lipase activity, 1-acyl-2-lysophosphatidylserine acylhydrolase activity, phosphatidylserine 1-acylhydrolase activity, and phospholipase A1 activity directly influence lipid metabolic balance, and then participate in the pathogenesis of FH. The chord diagram illustrated that 24 DEGs (supplementary material Table S1) were enriched in the top 10 most significant GO terms (Fig. 1D). Additionally, the bubble plot highlighted the top 20 significantly enriched KEGG pathways (Fig. 1E), including cholesterol metabolism, steroid hormone biosynthesis, arachidonic acid metabolism, alpha-linolenic acid metabolism, linoleic acid metabolism, and vascular smooth muscle contraction. These pathways are closely associated with metabolism and inflammation, collectively participating in the pathogenesis of FH. The generated chord diagram showed that 26 DEGs (supplementary material Table S2) were enriched in the top 10 most significant KEGG pathways (Fig. 1F).

13 candidate genes and PPI network

Taking the intersection of 429 DEGs and 776 LMRGs yielded 13 candidate genes (Fig. 2A), including CYP21A2, GRHL1, ACOT12, ELOVL4, PLA2G1B, PLA2G4D, LGMN, CH25H, BZRAP1, HSD17B3, STAR, PLA2G12B, and SPTSSB. Subsequently, a PPI network constructed based on these candidate genes revealed six nodes and five edges (Fig. 2B). Interactions were observed among HSD17B3, STAR, and CYP21A2, as well as between STAR and ACOT12, and between PLA2G12B and PLA2G1B.

Identification of six key genes in machine learning

Machine learning has been recognized as an effective approach for gene selection. Using the SVM-RFE algorithm, it was determined that when the number of feature genes is 13, the 10-fold cross-validation error rate is minimized, yielding the best model performance (Fig. 3A). Furthermore, analysis of the error rate and Gini index distribution in the RF model revealed that as the number of trees increased, the error rate gradually decreased and stabilized (Fig. 3B). This model identified seven genes with a Gini value greater than 0 (Fig. 3C). Afterwards, the

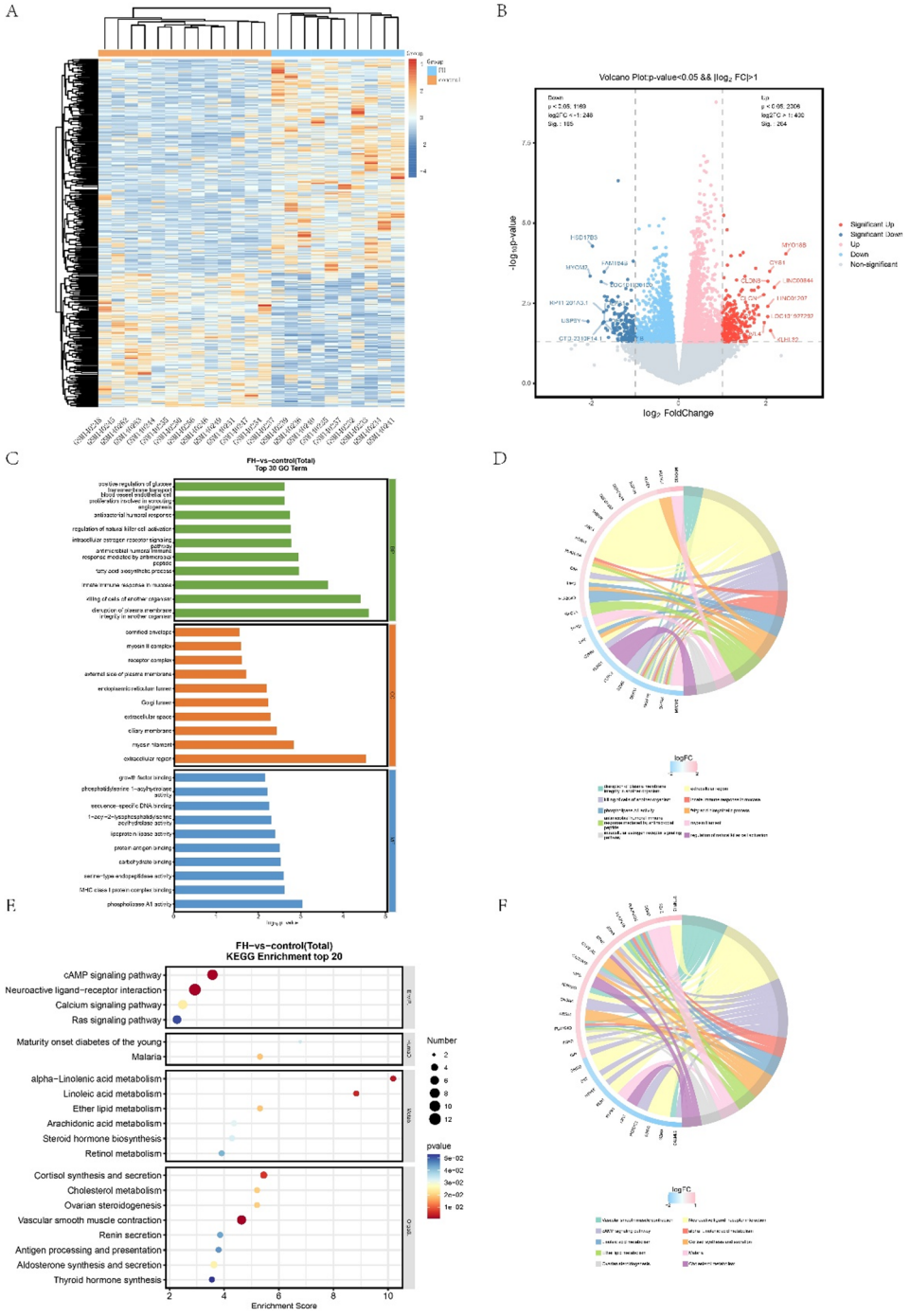


Fig. 1 Expression differences of LMRGs in the FH dataset. (**A**) Heatmap of DEGs. (**B**) Volcano plot of DEGs. (**C**) GO terms significantly enriched in DEGs. (**D**) The distribution of DEGs in the top ten most significant GO term enrichments. (**E**) KEGG pathways significantly enriched in DEGs. (**F**) The distribution of DEGs in the top ten most significant KEGG pathway enrichments

intersection of the results from the two algorithms identified seven key genes (Fig. 3D), including STAR, GRHL1, PLA2G1B, BZRAP1, LGMN, PLA2G4D, and PLA2G12B. However, expression validation of the seven key genes between FH and control samples revealed no significant difference in PLA2G1B (Fig. 3E). Therefore, we identified STAR, GRHL1, BZRAP1, LGMN, PLA2G4D, and PLA2G12B as the key genes. Among them, PLA2G12B, PLA2G4D, and STAR were highly expressed in FH samples, while GRHL1, BZRAP1, and LGMN were lowly expressed in FH samples. The ROC curves of the six key genes confirmed the reliability of the screening results,

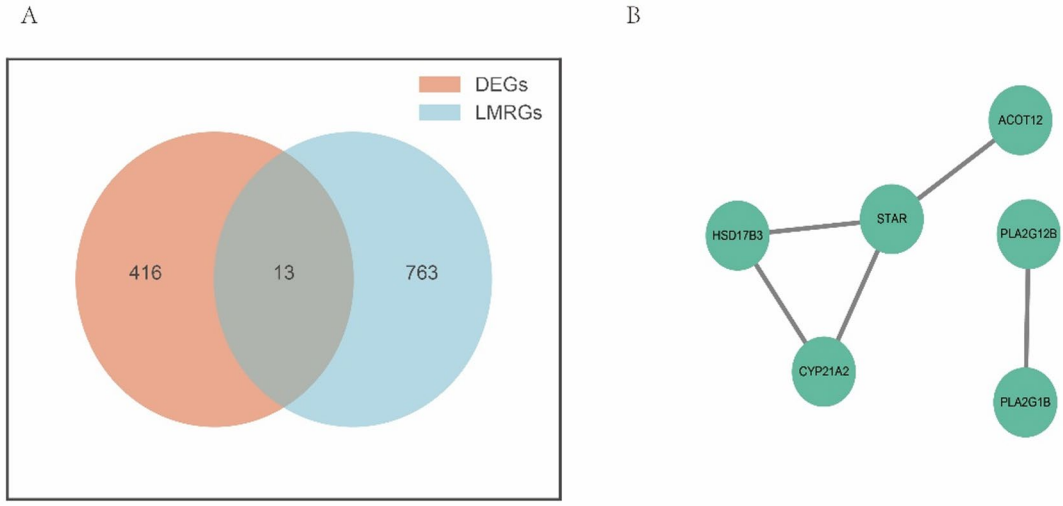


Fig. 2 Candidate genes and their PPI networks. (**A**) Venn diagram of candidate gene screening. (**B**) PPI networks

with AUC values as follows: STAR: 0.796, GRHL1: 0.864, BZRAP1: 0.735, LGMN: 0.787, PLA2G4D: 0.776, and PLA2G12B: 0.783 (Fig. 3F).

TFs associated with key genes and significantly enriched pathways

GSEA analysis revealed the biological pathways associated with the six key genes in FH, highlighting that many pathways were commonly enriched across multiple key genes ($p < 0.05$) (Fig. 4A-F). Strikingly, the RIBOSOME pathway was significantly enriched in BZARP1, PLA2G4D, PLA2G12B, and STAR, suggesting that associated proteins within this pathway may regulate key enzymes involved in cholesterol metabolism. Meanwhile, the SPLICEOSOME pathway showed significant enrichment in GRHL1 and PLA2G4D, with potential abnormalities in this pathway potentially disrupting the proper expression of cholesterol metabolism-related genes. Subsequently, prediction analysis using the NetworkAnalyst database revealed that 15 TFs interacted with the STAR gene and two TFs interacted with GRHL1 (Fig. 4G). Unfortunately, no interacting TFs were found for LGMN, PLA2G4D, PLA2G12B, or BZRAP1 (Table 1). Furthermore, a correlation heatmap of the key genes revealed a significant positive correlation ($p < 0.05$) in expression patterns between GRHL1 and PLA2G12B (Fig. 4H), suggesting that they may jointly contribute to the pathological processes of FH.

Analysis of functionally similar genes and subcellular distribution

Subcellular localization prediction using the GeneCards database identified 11 subcellular sites associated with the six key genes (supplementary material Table S3). Notably, PLA2G12B and LGMN showed high confidence levels in the extracellular, while STAR and BZRAP1 demonstrated high confidence in the mitochondrion (confidence score = 5) (Fig. 5A). These findings highlight their potential pivotal roles in regulating cholesterol metabolism and mitochondrial dysfunction in FH. Furthermore, 20 genes with similar functions to the key genes were found in the GeneMANIA database (Fig. 5B), including GPC5, RGS14, CD99L2, CTBS, SYN1, TM4SF5, MPZL2, AHSG, NRP1, CCHCR1, NFATC2IP, HTR2B, KIF3B, MPRS31, TGM4, ATG2A, SLC22A888, CECR2, HMBS, and VPS45. These findings suggest that these genes may play a synergistic role in disease progression, providing important insights for further investigation into disease-associated networks and potential therapeutic targets.

A ceRNA network based on key gene constructs

To further investigate the regulation of gene expression and cellular functions, the miRNet database predicted 15 upstream miRNAs associated with four key genes. Among these, LGMN and GRHL1 were regulated by the highest number of miRNAs (supplementary material Table S4), including hsa-mir-20a-5p, hsa-mir-20a-3p, and hsa-mir-18a-5p, which were found to co-regulate both genes. This suggests that these miRNAs may be involved in the onset and progression of the disease by modulating

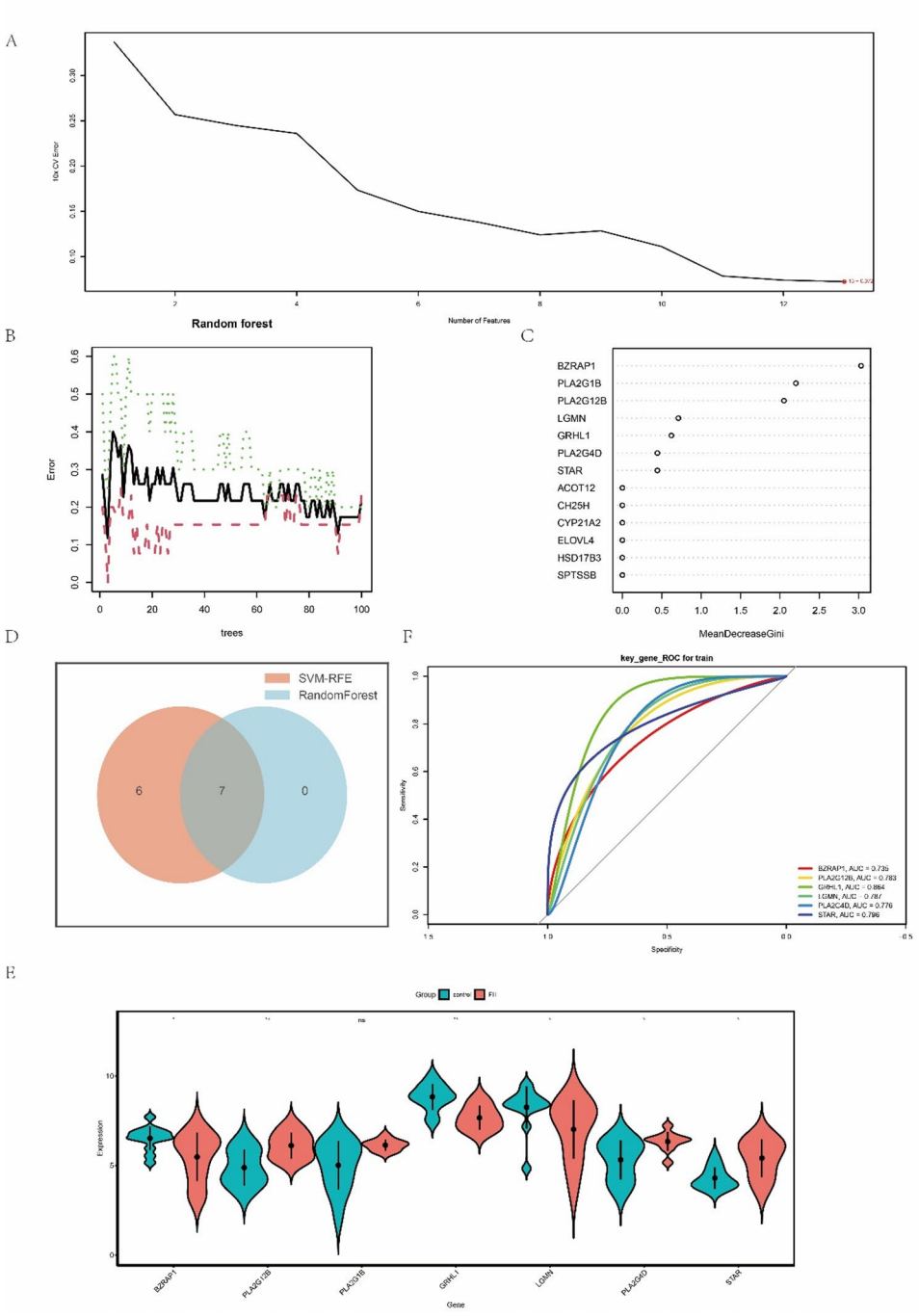


Fig. 3 Screening of key genes and their expression distribution. (**A**) The error rate curve for the 10-fold cross-validation of the SVM-RFE algorithm. (**B**) The error rate curve for the RF algorithm. (**C**) The Gini value generated by the RF algorithm. (**D**) Venn diagram of key gene screening. (**E**) The expression levels of the key genes, *: $p < 0.05$, **: $p < 0.01$. (**F**) ROC curve of key genes

key genes. Unexpectedly, no relevant miRNAs were identified for BZRAP1 and PLA2G12B. Subsequently, based on these 15 miRNAs, 289 targeted lncRNAs were matched in the miRNet database. Meanwhile, the Cytoscape software revealed the abundant and complex regulatory relationships within the mRNA-miRNA-lncRNA network (Fig. 6).

Drug forecasting

Drug prediction is beneficial for identifying potential therapeutic options for FH and providing meaningful references for adjunctive treatment. Using the CTD database, we identified 224 potential drugs associated with five key genes; however, no related drugs were retrieved for BZRAP1. Among the key genes, STAR had the

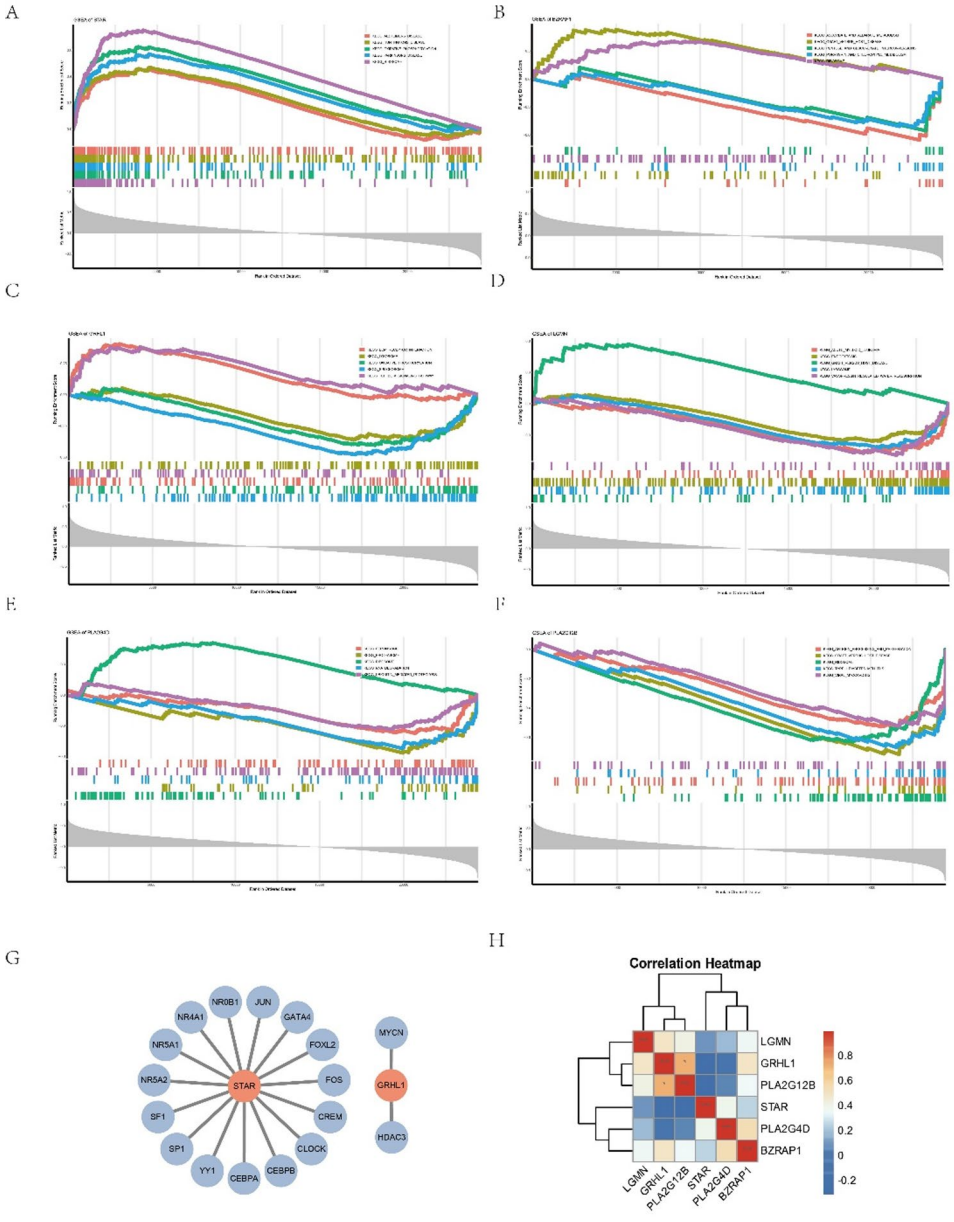


Fig. 4 Key gene analysis. (**A-F**) GSEA enrichment plots. (**G**) The prediction network for TFs of key genes. (**H**) Correlation heatmap of key genes. *: $p < 0.05$, ***: $p < 0.0001$

Table 1 Predicted results of TFs for 6 key genes

Gene	TFs	Gene	TFs
GRHL1	HDAC3	STAR	NR4A1
GRHL1	MYCN	STAR	NR5A1
STAR	CEBPA	STAR	NR5A2
STAR	CEBPB	STAR	SF1
STAR	CLOCK	STAR	SP1
STAR	CREM	STAR	YY1
STAR	FOS	BZRAP1	-
STAR	FOXL2	PLA2G12B	-
STAR	GATA4	PLA2G1B	-
STAR	JUN	LGMN	-
STAR	NR0B1	PLA2G4D	-

highest number of interacting drugs, with 140 identified. This was followed by GRHL1, with 57 drugs, and LGMN, with 56 drugs. In comparison, PLA2G12B and PLA2G4D have 32 and 13 drug interactions, respectively. Additionally, the interaction network between the genes and predicted drugs revealed a significant number of common potential drugs among these key genes, suggesting a degree of functional association between them (Fig. 7).

Discussion

FH is a congenital metabolic disorder primarily caused by mutations in the gene encoding the low-density lipoprotein receptor (LDLR). Serum markers of cholesterol

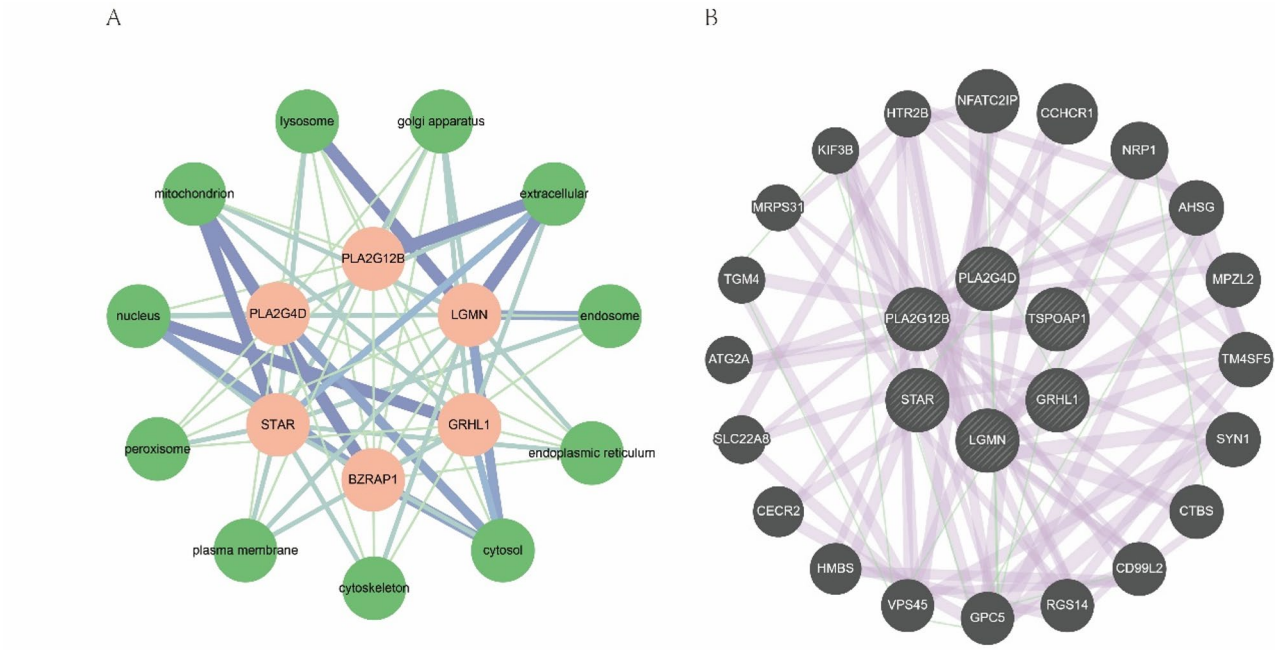


Fig. 5 Integrated Analysis of Subcellular Localization and Key Gene Functions. (**A**) Subcellular localization analysis. (**B**) The prediction of key gene functions

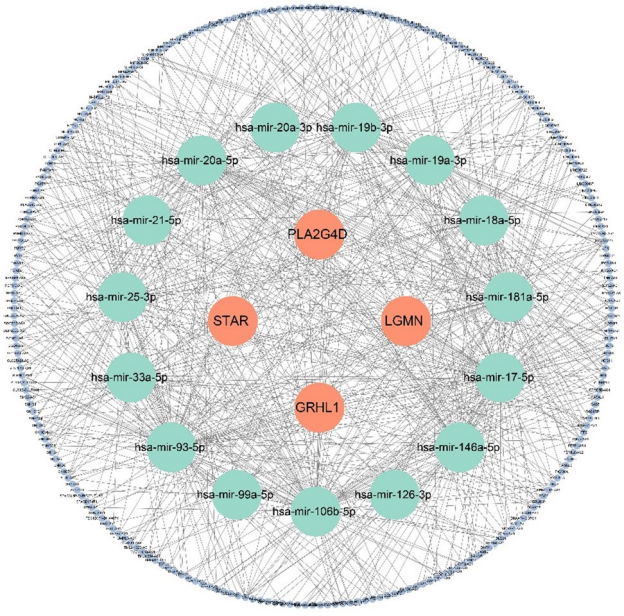


Fig. 6 A ceRNA regulatory network

absorption and synthesis have been associated with cardiovascular risks. However, to date, information on the potential alterations in cholesterol absorption and synthesis in FH patients remains limited [37]. Importantly, numerous studies have demonstrated that multiple metabolite markers can contribute to the development of cardiovascular diseases [38, 39]. Therefore, understanding the correlation between blood lipids in FH patients and lipid metabolism-related biomarkers is of great significance for refining treatment strategies and enhancing preventive measures. For this research, we employed bioinformatics analysis and machine learning techniques to detect biomarkers associated with lipid metabolism in FH. To the best of our knowledge, this is the first experiment to employ transcriptomic analysis for the identification of potential lipid metabolism related biomarkers in FH.

Comparing the FH group with the healthy group, we identified 429 DEGs. GO analysis revealed 111 enriched terms, while KEGG pathway analysis identified 18 significantly enriched pathways. Combined with machine learning, seven key genes (STAR, GRHL1, BZRAP1, LGMN, PLA2G1B, PLA2G4D, and PLA2G12B) were obtained. Expression analysis revealed that PLA2G1B exhibited no significant difference between FH samples and healthy samples. Subsequently, ROC curve analysis demonstrated the potential diagnostic value of the remaining six key genes as biomarkers for FH (STAR: 0.796, GRHL1: 0.864, BZRAP1: 0.735, LGMN: 0.787, PLA2G4D: 0.776, PLA2G12B: 0.783). GSEA revealed that the identified key genes were significantly enriched in multiple pathways, including RIBOSOME and SPLICEOSOME, both of which are closely associated with cholesterol metabolism. Furthermore, 15 TFs interacted with STAR genes in the NetworkAnalyst database, while LGMN and GRHL1 were found to be the most regulated by miRNAs. Notably, we observed a strong positive correlation between GRHL1 and PLA2G12B, suggesting

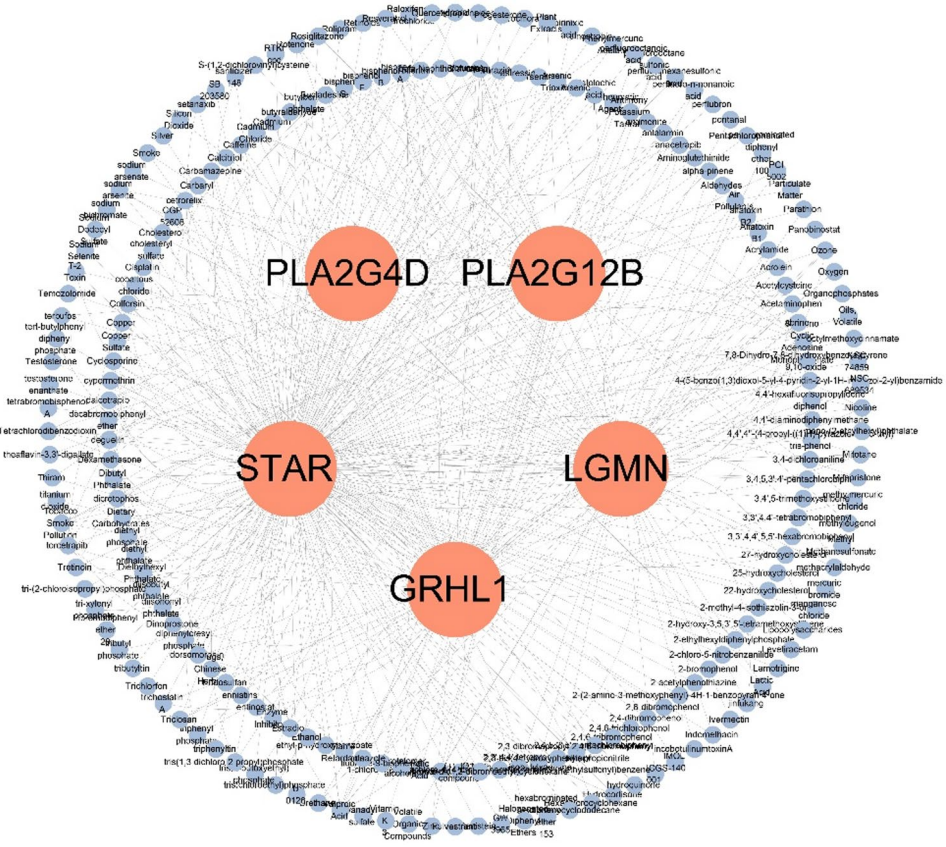


Fig. 7 Drug prediction results

their crucial roles in the initiation and progression of FH. Additionally, data from the GeneCards database indicated that STAR and BZRAP1 exhibit high confidence levels of localization in the mitochondria, while PLA2G12B and LGMN demonstrate high confidence levels of localization in the extracellular.

We performed functional and pathway enrichment analyses of DEGs in FH using two well-established databases KEGG and GO to highlight the functions and pathways associated with FH. Notably, FH-associated biological processes were prominently identified, primarily involving fatty acid biosynthetic process, lipoprotein lipase activity. Previous studies have demonstrated that major cardiovascular events in individuals predominantly share amino acid and lipid metabolism pathways, including fatty acid biosynthesis [40]. Fatty acids serve as a critical energy substrate for tumor cells, with various cancer types preferentially relying on fatty acid metabolism to fuel their uncontrolled proliferative demands [41]. Hypoxia-inducible factor (HIF) serves as a pivotal regulator in fatty acid biosynthesis and has been identified as a central mediator activated under hypoxic conditions, which has been recognized as a major contributing factor to angiogenesis-driven tumor proliferation and metastatic progression [42, 43]. This mechanistic framework suggests that elevated fatty acid levels may contribute to sustaining the hyperproliferative state of FH cells under hypoxic conditions. A comprehensive understanding of fatty acid biosynthesis regulation in FH may, therefore, provide novel therapeutic avenues for managing this refractory cardiovascular disorder. Furthermore, the peripheral lipolytic pathway mediated by lipoprotein lipase (LPL) plays a critical role in systemic lipid metabolism. Angiopoietin-like protein 3 (ANGPTL3), predominantly secreted by the liver, serves as the primary physiological inhibitor of LPL activity [44]. Recent advances have identified ANGPLT3 as a novel therapeutic target for lipid modulation, with mechanistic studies revealing that its N-terminal coiled-coil domain (CCD) mediates potent suppression of LPL activity, thereby regulating triglyceride hydrolysis and subsequent lipoprotein clearance [45]. The same is true for KEGG pathways, with 18 significantly enriched pathways associated with FH, most of which are related to the processes of inflammation, oxidative stress, and atherosclerosis [17]. For example, numerous studies have hightlighted the detrimental role of arachidonic acid and its oxidative metabolic derivatives (oxylipins) in the regulation of atherosclerosis, inflammation, and thrombosis. The elevated metabolites from the arachidonic acid metabolic pathway, along with

extremely high cholesterol levels, may synergistically promote the formation of corneal arcus, xanthomas, and calcified VAS in patients with HOFH [17], thus playing a key role in the emergence and progression of cardiovascular diseases [46, 47]. Cholesterol metabolism constitutes a highly coordinated biological process involving multiple interrelated components. Intestinal epithelial absorption, lipoprotein-mediated transport, peripheral tissue uptake, and hepatic excretion of cholesterol constitute tightly regulated processes that collectively ensure systemic cholesterol homeostasis. Dysregulation of this metabolic network or mutations in pathway-associated regulatory proteins may disrupt cholesterol flux dynamics, potentially contributing to various pathological conditions [48]. Additionally, pathways such as ribosome and spliceosome may represent upstream regulatory disturbances affecting lipid-related gene expression rather than direct lipid metabolic processes.

As essential components of the lipid metabolism machinery, these critical LMRGs have been well-documented for their pivotal roles in tumorigenesis, tumor progression, and cancer treatment. These roles span multiple biological processes, including cell proliferation, apoptosis, cell cycle regulation, metabolism, and signal transduction [15]. Steroidogenic acute regulatory protein (STAR) stands as the first identified member of the START domain family. In recent years, a burgeoning body of research has revealed that, beyond the well-characterized typical steroid-producing tissues like the testis, ovary, and adrenal gland, the STAR protein is also expressed and exerts its functions in a variety of atypical tissues and organs, including the heart, liver, brain, kidney, and adipose tissue [49]. Some investigations have indicated that "Since the function of STAR in heart cells is known to promote cholesterol mobilization between the mitochondrial membranes, a plausible mechanism could be a connection between this activity and the protein's anti-apoptotic effect in heart cells." However, to data, no relevant literature has reported the use of the STAR gene as a biomarker for the diagnosis of FH. Our study revealed that the upregulation of the STAR gene in FH, along with its significant enrichment in the ribosomal pathway, suggests that STAR-associated proteins within this pathway may modulate key enzymes involved in cholesterol metabolism, thereby potentially disrupting the normal expression patterns of cholesterol metabolism-related genes.

Grainyhead like transcription factor 1 (GRHL1) is a 70 kDa protein and belongs to the homologous group of GRH families, which function as transcription factors in warm-blooded animals [50]. The vitamin D receptor (VDR) acts as a transcription factor and affects gene expression by binding to vitamin D response elements (VDREs) in DNA, thereby regulating target genes [51]. These include cell-cycle regulatory factors such as SP1, CTCF, and FOXO4, as well as genes in the cardiovascular system that regulate development, blood pressure, reduce inflammation, and enhance endothelial function. In addition, some studies have shown that GRHL1 may play a potential role in the inheritance effect that favors RAS inhibition [52]. Our investigation identified downregulation of the GRHL1 gene in FH, with GSEA revealing its significant enrichment in the spliceosome pathway. This pathway-specific alteration suggests potential dysregulation in mRNA processing machinery, which may compromise the fidelity of cholesterol metabolism-related gene expression through aberrant splicing mechanisms in the pathological context.

Legumain (LGMN), a lysosomal cysteine protease, has been demonstrated to exert a pivotal role in cardiovascular pathologies, including atherosclerosis and myocardial infarctio [53]. Extensive database searches inquiries have uncovered that LGMN is predominantly expressed in HLA-DRhighCCR2-resident cardiac macrophages in the human body. Notably, among all the genes associated with macrophage phagocytosis, LGMN exhibits the highest level of expression, highlighting its potential significance in macrophage-mediated immune responses and cardiovascular pathology. Accumulating evidence from several studies indicates that the absence or deficiency of LGMN significantly compromises the efferocytosis function of macrophages, leading to an exacerbated loss of cardiomyocytes, a critical factor in the pathophysiology of heart-related conditions. Following a myocardial infarction, the normal expression and proper functioning of LGMN are indispensable for the intricate processes of cardiac repair and remodeling. This study identified significant overexpression of the lysosomal protease LGMN gene in patients with FH. Bioinformatic subcellular localization predictions strongly support its extracellular localization (confidence score > 0.85), suggesting a potential role in extracellular regulatory mechanisms of cholesterol metabolism. The observed spatial distribution pattern, combined with its differential expression profile, indicates that extracellular LGMN may serve as a crucial modulator in FH pathogenesis, potentially through proteolytic processing of key proteins involved in lipid transport or receptor regulation. Consequently, therapeutic strategies aimed at activating LGMN and enhancing efferocytosis hold great promise as effective interventions for facilitating post-myocardial infarction repair [54]. Nevertheless, despite considerable progress in elucidating the role of LGMN, the precise mechanism governing the continuous clearance and degradation of cellular contents by macrophages remains largely elusive, highlighting a critical area for further investigation.

Phospholipase A2 (PLA2) is a family of enzymes that play a crucial role in the synthesis of lipid mediators by

catalyzing the release of fatty acids from cell membranes [55]. Lipid mediators are involved in many physiological processes, such as inflammation, host defense, and barrier function [55]. Analysis of transcriptional changes indicates that PLA2G4D deficiency is associated with the dysregulation of multiple genes involved in proliferation, differentiation, inflammation, and lipid metabolism [56]. The PLA2G4D gene, which is abundantly expressed in the spleen and lymph nodes, has been shown to attenuate Th1 and Th17 immune responses and is associated with β-lactam metabolites and u3 polyunsaturated fatty acids [57]. The clinical significance of Th1/Th2/Th17 cell detection in coronary heart disease (CHD)-related inflammation or stenosis has been well documented [58]. PLA2G4D, identified as a lipid antigen, has been demonstrated to be expressed in psoriatic mast cells. This molecule activates CD1a-restricted T cells, which subsequently secrete substantial quantities of IL-22 and IL-17 A two cytokines recognized as pivotal cytokines driving pathogenesis in inflammatory disorders [59]. It is known that mutations in phospholipase A2 group 12B (PLA2G12B) disrupt lipoprotein homeostasis, yet its mechanistic role in this process remains unclear. Previous studies in mice have demonstrated that disruption of PLA2G12B leads to reduced secretion of hepatic triglycerides and APOB52, as well as decreased high-density cholesterol [60], suggesting that this gene may play a role in lipoprotein secretion. Furthermore, pharmacological inhibition of the secretory pathway using BFA increased the levels of luminal ApoB in PLA2G12B mutants. Additionally, specific research indicates that PLA2G12B mutant mice with compromised lipoprotein uptake due to PCSK9 overexpression still maintain relatively normal plasma triglyceride (TG) levels [61]. Our analysis demonstrated that patients with FH exhibit significantly elevated expression levels of PLA2G4D and PLA2G12B. These phospholipases enzymes may serve as promising therapeutic targets for mitigating the global burden of hyperlipidemia and associated cardiovascular diseases. Importantly, to our knowledge, no previous studies have documented comparable findings, thereby highlighting the novelty and potential clinical relevance of our investigation.

The benzodiazepine binding inhibitor (peripheral) associated protein 1 (BZRAP1, alternatively known as RIMBP1) is an adaptor molecule that is thought to regulate synaptic transmission by linking the vesicle release machinery to voltage-gated Ca2 + channels [62]. Genome-wide analyses have identified three suggestive variants within the BZRAP1 locus on chromosome 17, demonstrating genetic overlap between Alzheimer's disease (AD), C-reactive protein (CRP), and blood lipid profiles through its modulation of cardiovascular-related genetic correlations, which may elevate AD risk [63]. Moreover, BZRAP1 was found to be significantly upregulated in gastrointestinal goblet cell adenocarcinomas [64]. However, in this study, BZRAP1 was found to be significantly downregulated in FH samples and notably enriched in the ribosome pathway. Furthermore, subcellular localization predictions indicate a high confidence level for its mitochondrial localization, suggesting that BZRAP1 may play a pivotal role in regulating cholesterol metabolism and mitochondrial dysfunction. Correlation analysis revealed a significant positive association between BZRAP1 and GRHL1, suggesting that these genes may act in concert to drive the pathological processes underlying FH.

We identified a ceRNA sub-network associated with FH and hightlighted novel regulatory mechanisms. Notably, our investigation revealed that key genes in peripheral blood mononuclear cells (PBMCs) of FH patients were post-transcriptionally regulated by specific miRNAs. Among these, LGMN and GRHL1 emerged as the most prominently regulated targets, with multiple miRNAs including hsa-miR-20a-5p, hsa-miR-20a-3p, and hsa-miR-18a-5p exhibiting regulatory potential. These findings suggest that these specific miRNAs may exert pivotal roles in FH pathogenesis by modulating critical molecular pathways. However, technical limitations of miRNA microarray platforms restrict detection to known miRNAs, potentially overlooking novel regulatory elements. Moreover, the clinical relevance of PBMC miRNA profiles to FH disease progression and activity remains unexplored, highlighting an important gap in current understanding that warrants further investigation. Extensive studies have mechanistically established that statins exert therapeutic effects through competitive inhibition of HMG-CoA reductase, the rate-limiting enzyme in cholesterol biosynthesis. Notably, hsa-miR-20a-5p demonstrates direct binding capacity to three distinct sites within the 3' untranslated region (UTR) of LDL receptor (LDLR) mRNA, with its expression being significantly downregulated following atorvastatin treatment [65]. In addition to its role in lipid metabolism, emerging evidence has implicated hsa-miR-20a-5p in oncogenic processes, particularly through its regulatory effects on cancer cell proliferation and migratory capacity, as substantiated by multiple independent investigations [5, 66]. However, no studies to date have directly linked this miRNA to cholesterol metabolism, and the proposed interaction between hsa-miR-20a-5p and LDLR requires further experimental validation. MiR-20a-3p, another mature microRNA derived from the precursor miR-20a stem-loop structure, was traditionally regarded as non-functional. However, accumulating evidence now indicates that miR-20a-3p indeed exerts functional roles in normal cellular behaviors [67]. Mechanistic studies have experimentally validated the involvement of

hsa-miR-20a-3p in disease-associated cellular hyperproliferation, particularly in models of psoriasis pathogenesis and IL-22-induced keratinocyte proliferation [68]. Of particular interest, the rs17848060 polymorphism, located within the miRNA: mRNA binding interface of hsa-miR-20a-3p and CXCR4, exhibits allele-specific binding characteristics, with the A > T substitution substantially impairing miRNA-3′-UTR interaction efficiency [69]. This co-targeting relationship suggests that members of the miR-17–92 cluster (including miR-20a) may exert coordinated regulatory effects on both genes through synergistic interactions. Nevertheless, functional validation is warranted to substantiate this proposed mechanism. Furthermore, emerging evidence delineates hsa-miR-18a-5p as a potent anti-inflammatory regulator that modulates adaptive immune responses through dual mechanisms: surface receptor-mediated signaling and canonical SMAD/MAPK pathway regulation [70]. This immunomodulatory axis may critically influence both the pathogenesis and clinical severity of FH. Independent investigations have demonstrated that upregulation of that hsa-miR-18a-5p exerts cardioprotective effects by targeting NOTCH2, effectively suppressing high glucose-induced endothelial-mesenchymal transition (EndMT) and subsequent cardiac fibrogenesis through phenotypic reprogramming of stromal cells [71]. Collectively, these findings suggest that dysregulated miRNAs may orchestrate pathogenic convergence through coordinated regulation of shared molecular targets, thereby contributing to both disease initiation and progression in FH.

We integrated bioinformatics and machine learning to identify six significant diagnostic markers for FH and searched for potential therapeutic targets for this disease, providing valuable explorations for the clinical diagnosis of FH. In addition, our study is not without limitations. Firstly, the key genes and pathways we have identified still need to be experimentally validated. Secondly, there is only one dataset, which is insufficient for conducting more comprehensive bioinformatics analyses to further confirm the experimental results. Therefore, this study should be regarded as hypothesis-generating, and the identified candidate genes require further validation in larger independent cohorts and experimental investigations.

Conclusion

In this study, we combined bioinformatics and machine learning approaches to determine seven candidate LMRGs associated with FH. Six of these genes exhibited significantly different expression levels between FH and healthy samples. ROC curve analysis demonstrated their potential as biomarkers for FH, with the following AUC values: STAR, 0.796; GRHL1, 0.864; BZRAP1, 0.735; LGMN, 0.787; PLA2G4D, 0.776; and PLA2G12B, 0.783. Moreover, several pathways-including ribosome and spliceosome signaling-were commonly enriched across multiple genes, both of which are closely linked to lipid metabolism. Notably, GRHL1 and PLA2G12B showed a strong positive correlation, suggesting a potential cooperative role in promoting FH pathogenesis. In addition, we identified 224 potential therapeutic compounds for FH. Collectively, this study provides novel insights into the role of lipid metabolism in FH and holds significant implications for the development of targeted therapies and preventive strategies.

Abbreviations

FH	Familial hypercholesterolemia
HeFH	Heterozygous Familial hypercholesterolemia
HoFH	Homozygous Familial hypercholesterolemia
ASCVD	Atherosclerotic cardiovascular disease
LDL-C	Low-density lipoprotein cholesterol
CHD	Coronary heart disease
GEO	Gene Expression Omnibus
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
GSEA	Gene set enrichment analysis
DEGs	Differentially expressed genes
ROC	Operating characteristic curves
LASSO	Least absolute shrinkage and selection operator
LMRGs	Lipid metabolism-related genes
TF	Transcription factor
ceRNA	Competitive endogenous RNA
RF	Random forest
SVM-RFE	Support vector machine recursive feature elimination
CTD	Comparative Toxicogenomics Database
FDR	False discovery rate
BP	Biological processes
CC	Cellular components
MF	Molecular functions
LDLR	Low-density lipoprotein receptor
EndMT	Endothelial-mesenchymal transition
HIF	Hypoxia-inducible factor
LPL	Lipoprotein lipase
ANGPTL3	Angiopoietin-like protein 3
CCD	Coiled-coil domain
STAR	Steroidogenic acute regulatory protein
GRHL1	Grainyhead like transcription factor 1
VDR	The vitamin D receptor
VDREs	Vitamin D response elements
PBMCs	Peripheral blood mononuclear cells
LGMN	Legumain
PLA2	Phospholipase A2
PLA2G12B	Phospholipase A2 group 12B
AD	Alzheimer's disease
CRP	C-reactive protein
UTR	Untranslated region

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13019-026-04107-8.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Acknowledgements
We thank the mentioned public databases for providing us the data and analytical tools.

Author contributions
LinJi Long: Writing-original draft, Software, Formal analysis, Data curation, Validation.Baosheng Zhu: Writing-review & editing, Methodology, Formal analysis, Supervision.Tao Lv: Writing-original draft, Methodology, Formal analysis, Supervision.

Funding
This work was supported by Yunling Scholar Project of Yunnan Province (YNWR-YLXZ-2019-0005) and Yunnan Provincial "Double First-Class" Discipline Construction Project (SRDP-2023-004).

Data availability
The datasets utilized in this study were all procured from public databases GEO and MSigDB, including GSE6054 and LMRGs. Data will be made available on request.

Declarations

Ethics approval and consent to participate
This study utilized publicly accessible datasets from the GEO database for secondary analysis. As the data were anonymized and complied with the ethical standards of their original sources, no additional institutional ethics approval or individual consent was required. All procedures adhered to the ethical principles outlined in the Declaration of Helsinki for retrospective biomedical research involving pre-existing datasets.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Clinica trial number
Not applicable.

Author details
[1]School of Medicine, Kunming University of Science and Technology, Kunming, China
[2]Department of Medical Genetics, National Health Commission Key Laboratory of Healthy Birth and Birth Defects Prevention and Control in Western China, Key Laboratory of Birth Defects and Genetic Diseases of Yunnan Province, The First People's Hospital of Yunnan Province, Kunming University of Science and Technology Affiliated Hospital, Kunming, China

Received: 5 October 2025 / Accepted: 10 April 2026
Published online: 24 April 2026

References

1. Watts GF, Gidding S, Wierzbicki AS, Toth PP, Alonso R, Brown WV, et al. Integrated guidance on the care of familial hypercholesterolaemia from the International FH Foundation: executive summary. J Atheroscler Thromb. 2014;21(4):368–74.
2. Currie E, Schulze A, Zechner R, Walther TC, Farese RV. Jr. Cellular fatty acid metabolism and cancer. Cell Metabol. 2013;18(2):153–61.
3. Félix-Redondo FJ, Grau M, Fernández-Bergés D. Cholesterol and cardiovascular disease in the elderly. Facts and gaps. Aging disease. 2013;4(3):154–69.
4. Defesche JC, Gidding SS, Harada-Shiba M, Hegele RA, Santos RD, Wierzbicki AS. Familial hypercholesterolaemia. Nat reviews Disease primers. 2017;3:17093.
5. Liao W, He J, Disoma C, Hu Y, Li J, Chen G, et al. Hsa_circ_0107593 Suppresses the Progression of Cervical Cancer via Sponging hsa-miR-20a. Front Oncol. 2020;10:590627.
6. Beheshti SO, Madsen CM, Varbo A, Nordestgaard BG. Worldwide Prevalence of Familial Hypercholesterolemia: Meta-Analyses of 11 Million Subjects. J Am Coll Cardiol. 2020;75(20):2553–66.
7. Tromp TR, Hartgers ML, Hovingh GK, Vallejo-Vaz AJ, Ray KK, Soran H, et al. Worldwide experience of homozygous familial hypercholesterolaemia: retrospective cohort study. Lancet (London England). 2022;399(10326):719–28.
8. Goldstein JL, Brown MS. Familial hypercholesterolemia: pathogenesis of a receptor disease. Johns Hopkins Med J. 1978;143(1):8–16.
9. Xu H, Sun J, Zhou L, Du QC, Zhu HY, Chen Y, et al. Development of a lipid metabolism-related gene model to predict prognosis in patients with pancreatic cancer. World J Clin Cases. 2021;9(35):10884–98.
10. Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. Nature. 2018;561(7721):45–56.
11. Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid metabolic reprogramming in cancer cells. Oncogenesis. 2016;5(1):e189.
12. Guillaumond F, Bidaut G, Ouaissi M, Servais S, Gouirand V, Olivares O, et al. Cholesterol uptake disruption, in association with chemotherapy, is a promising combined metabolic therapy for pancreatic adenocarcinoma. Proc Natl Acad Sci U S A. 2015;112(8):2473–8.
13. Delmas D, Rébé C, Micheau O, Athias A, Gambert P, Grazide S, et al. Redistribution of CD95, DR4 and DR5 in rafts accounts for the synergistic toxicity of resveratrol and death receptor ligands in colon carcinoma cells. Oncogene. 2004;23(55):8979–86.
14. Stratton SA, Barton MC. p53-mediated regulation of hepatic lipid metabolism: forging links between metabolism, atherogenesis, and cancer. J Hepatol. 2012;56(3):518–9.
15. Cheng C, Geng F, Cheng X, Guo D. Lipid metabolism reprogramming and its potential targets in cancer. Cancer Commun. 2018;38(1):27.
16. Mosig S, Rennert K, Büttner P, Krause S, Lütjohann D, Soufi M, et al. Monocytes of patients with familial hypercholesterolemia show alterations in cholesterol metabolism. BMC Med Genomics. 2008;1:60.
17. Du Z, Li F, Jiang L, Li L, Du Y, Yu H, et al. Metabolic systems approaches update molecular insights of clinical phenotypes and cardiovascular risk in patients with homozygous familial hypercholesterolemia. BMC Med. 2023;21(1):275.
18. Nazir A, Nazir A, Afzaal U, Aman S, Sadiq SR, Akah OZ et al. Advancements in Biomarkers for Early Detection and Risk Stratification of Cardiovascular Diseases—A Literature Review. Health Sci Rep. 2025;8(5):e70878.
19. Kohaar I, Hodges NA, Srivastava S. Biomarkers in cancer screening promises and challenges in cancer early detection. Hematol Oncol Clin North Am. 2024;38(4):869–88.
20. Taz TA, Arslanturk S. Unveiling genetic variant-level biomarkers for aggressive prostate cancer. Inf Med Unlocked. 2023;43:101410.
21. Wu JH, Luo JY, Cai HH, Zhu HY, Lei Z, Lu Y, et al. Expression characteristics of lipid metabolism-related genes and correlative immune infiltration landscape in acute myocardial infarction. Sci Rep. 2024;14(1):13.
22. Taz TA, Ahmed K, Paul BK, Kawsar M, Aktar N, Mahmud SMH, et al. Network-based identification genetic effect of SARS-CoV-2 infections to Idiopathic pulmonary fibrosis (IPF) patients. Brief Bioinform. 2021;22(2):1254–66.
23. Taz TA, Ahmed K, Paul BK, Al-Zahrani FA, Mahmud SMH, Moni MA. Identification of biomarkers and pathways for the SARS-CoV-2 infections that make complexities in pulmonary arterial hypertension patients. Brief Bioinform. 2021;22(2):1451–65.
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
25. Gustavsson EK, Zhang D, Reynolds RH, Garcia-Ruiz S, Ryten M. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. Bioinf (Oxford England). 2022;38(15):3844–6.
26. Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics. CPT: pharmacometrics & systems pharmacology. 2013;2(10):e79.
27. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov (Cambridge (Mass)). 2021;2(3):100141.
28. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
29. Alam MNU, Basava K, Chitransh A, Fattah HMA, Garcia-Verdugo HD, Lo SH, et al. Machine learning in biological research: key algorithms, applications, and future directions. BMC Biol. 2025;23(1):20.
30. Hu J, Szymczak S. A review on longitudinal data analysis with random forest. Brief Bioinform. 2023;24(2):bbad002.
31. Yoon S, Kim S. AdaBoost-based multiple SVM-RFE for classification of mammograms in DDSM. BMC Med Inf Decis Mak. 2009;9(Suppl 1):S1.
32. Shi W, Zhang X, Xu C, Pang R, Fan Z, Wan X, et al. Identification of hub genes and pathways associated with oxidative stress of cartilage in

osteonecrosis of femoral head using bioinformatics analysis. Cartilage. 2022;13(1):19476035221074000.
33. Zhang J, Ahn J, Suh Y, Hwang S, Davis ME, Lee K. Identification of CTLA2A, DEFB29, WFDC15B, SERPINA1F and MUP19 as novel tissue-specific secretory factors in mouse. PLoS One. 2015;10(5):e0124962.
34. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol. 2008;9(Suppl 1Suppl 1):S4.
35. Zeng L, Tang Y, Zhang Y, Yue L, Ma G, Ye X, et al. The molecular mechanism underlying dermatomyositis related interstitial lung disease: evidence from bioinformatic analysis and in vivo validation. Front Immunol. 2023;14:1288098.
36. Li T, Huang C, Ma J, Ding R, Zhang Q, Wang W. Identification of inflammation-related genes and exploration of regulatory mechanisms in patients with osteonecrosis of the femoral head. Biomed Res Int. 2022;2022:4501186.
37. Matsuki K, Harada-Shiba M, Hori M, Ogura M, Akiyama Y, Fujii H, et al. Association between familial hypercholesterolemia and serum levels of cholesterol synthesis and absorption markers: the CACHE study FH analysis. J Atheroscler Thromb. 2023;30(9):1152–64.
38. Ussher JR, Elmariah S, Gerszten RE, Dyck JR. The emerging role of metabolomics in the diagnosis and prognosis of cardiovascular disease. J Am Coll Cardiol. 2016;68(25):2850–70.
39. Joshi A, Rienks M, Theofilatos K, Mayr M. Systems biology in cardiovascular disease: a multiomics approach. Nat Rev Cardiol. 2021;18(5):313–30.
40. Lv J, Pan C, Cai Y, Han X, Wang C, Ma J, et al. Plasma metabolomics reveals the shared and distinct metabolic disturbances associated with cardiovascular events in coronary artery disease. Nat Commun. 2024;15(1):5729.
41. Baenke F, Peck B, Miess H, Schulze A. Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. Dis Model Mech. 2013;6(6):1353–63.
42. Khwairakpam AD, Shyamananda MS, Sailo BL, Rathnakaram SR, Padmavathi G, Kotoky J, et al. ATP citrate lyase (ACLY): a promising target for cancer prevention and treatment. Curr Drug Targets. 2015;16(2):156–63.
43. Budhu A, Roessler S, Zhao X, Yu Z, Forgues M, Ji J, et al. Integrated metabolite and gene expression profiles identify lipid biomarkers associated with progression of hepatocellular carcinoma and patient outcomes. Gastroenterology. 2013;144(5):1066–e751.
44. Kersten S. Angiopoietin-like 3 in lipoprotein metabolism. Nat Rev Endocrinol. 2017;13(12):731–9.
45. Ahmad Z, Banerjee P, Hamon S, Chan KC, Bouzelmat A, Sasiela WJ, et al. Inhibition of Angiopoietin-Like Protein 3 with a monoclonal antibody reduces triglycerides in hypertriglyceridemia. Circulation. 2019;140(6):470–86.
46. Badimon L, Vilahur G, Rocca B, Patrono C. The key contribution of platelet and vascular arachidonic acid metabolism to the pathophysiology of atherothrombosis. Cardiovasc Res. 2021;117(9):2001–15.
47. Zordoky BN, El-Kadi AO. Effect of cytochrome P450 polymorphism on arachidonic acid metabolism and their impact on cardiovascular diseases. Pharmacol Ther. 2010;125(3):446–63.
48. Benito-Vicente A, Uribe KB, Jebari S, Galicia-Garcia U, Ostolaza H, Martin C. Familial hypercholesterolemia: the most frequent cholesterol metabolism disorder caused disease. Int J Mol Sci. 2018;19(11):3426.
49. Dinh Cat AN, Friederich-Persson M, White A, Touyz RM. Adipocytes, aldosterone and obesity-related hypertension. J Mol Endocrinol. 2016;57(1):F7–21.
50. Mlacki M, Kikulska A, Krzywinska E, Pawlak M, Wilanowski T. Recent discoveries concerning the involvement of transcription factors from the Grainyhead-like family in cancer. Experimental biology and medicine (Maywood, NJ). 2015;240(11):1396–401.
51. Kechichian E, Ezzedine K. Vitamin D and the skin: an update for dermatologists. Am J Clin Dermatol. 2018;19(2):223–35.
52. Byars SG, Prestes PR, Suphapimol V, Takeuchi F, De Vries N, Maier MC, et al. Four-week inhibition of the renin-angiotensin system in spontaneously hypertensive rats results in persistently lower blood pressure with reduced kidney renin and changes in expression of relevant gene networks. Cardiovasc Res. 2024;120(7):769–81.
53. Clerin V, Shih HH, Deng N, Hebert G, Resmini C, Shields KM, et al. Expression of the cysteine protease legumain in vascular lesions and functional implications in atherogenesis. Atherosclerosis. 2008;201(1):53–66.
54. Jia D, Chen S, Bai P, Luo C, Liu J, Sun A, et al. Cardiac Resident Macrophage-Derived Legumain Improves Cardiac Repair by Promoting Clearance and Degradation of Apoptotic Cardiomyocytes After Myocardial Infarction. Circulation. 2022;145(20):1542–56.
55. Shao S, Chen J, Swindell WR, Tsoi LC, Xing X, Ma F, et al. Phospholipase A2 enzymes represent a shared pathogenic pathway in psoriasis and pityriasis rubra pilaris. JCI Insight. 2021;6(20):e151911.
56. Breithofer J, Bulfon D, Fawzy N, Tischitz M, Zitta C, Hartig L, et al. Phospholipase A2 group IVD mediates the transacylation of glycerophospholipids and acylglycerols. J Lipid Res. 2024;65(12):100685.
57. Miki Y, Yamamoto K, Taketomi Y, Sato H, Shimo K, Kobayashi T, et al. Lymphoid tissue phospholipase A2 group IID resolves contact hypersensitivity by driving antiinflammatory lipid mediators. J Exp Med. 2013;210(6):1217–34.
58. Kidd P. Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease. Altern Med review: J Clin therapeutic. 2003;8(3):223–46.
59. Yamada-Kanazawa S, Mijiddorj MT, Kajihara I, Kanemaru H, Sawamura S, Makino K, et al. Circulating phospholipase A2 group IV D DNA copies are elevated in psoriasis. J Eur Acad Dermatology Venereology: JEADV. 2022;36(4):e302–4.
60. Aljakna A, Choi S, Savage H, Hageman Blair R, Gu T, Svenson KL, et al. Pla2g12b and Hpn are genes identified by mouse ENU mutagenesis that affect HDL cholesterol. PLoS One. 2012;7(8):e43139.
61. Thierer JH, Foresti O, Yadav PK, Wilson MH, Moll TOC, Shen MC, et al. Pla2g12b drives expansion of triglyceride-rich lipoproteins. Nat Commun. 2024;15(1):2095.
62. Wang Y, Sugita S, Sudhof TC. The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins. J Biol Chem. 2000;275(26):20033–44.
63. Tan MS, Yang YX, Xu W, Wang HF, Tan L, Zuo CT, et al. Associations of Alzheimer's disease risk variants with gene expression, amyloidosis, tauopathy, and neurodegeneration. Alzheimers Res Ther. 2021;13(1):15.
64. Lin DL, Wang LL, Zhao P, Ran WW, Wang W, Zhang LX, et al. Gastrointestinal Goblet Cell Adenocarcinomas Harbor Distinctive Clinicopathological, Immune, and Genomic Landscape. Front Oncol. 2021;11:758643.
65. Zambrano T, Hirata RDC, Hirata MH, Cerda Á, Salazar LA. Altered microRNome profiling in statin-induced HepG2 cells: a pilot study identifying potential new biomarkers involved in lipid-lowering treatment. Cardiovasc Drugs Ther. 2015;29(6):509–18.
66. Guo L, Zhu Y, Li L, Zhou S, Yin G, Yu G, et al. Breast cancer cell-derived exosomal miR-20a-5p promotes the proliferation and differentiation of osteoclasts by targeting SRCIN1. Cancer Med. 2019;8(12):5687–701.
67. Li R, Qiao M, Zhao X, Yan J, Wang X, Sun Q. MiR-20a-3p regulates TGF-β1/Survivin pathway to affect keratinocytes proliferation and apoptosis by targeting SFMBT1 in vitro. Cell Signal. 2018;49:95–104.
68. Pink RC, Samuel P, Massa D, Caley DP, Brooks SA, Carter DR. The passenger strand, miR-21-3p, plays a role in mediating cisplatin resistance in ovarian cancer cells. Gynecol Oncol. 2015;137(1):143–51.
69. Wang Y, Li Y, Jia D, Zheng J, Wang G. Correlation between single nucleotide polymorphisms in CXCR4 microRNA binding site and the susceptibility to knee osteoarthritis in Han Chinese population. J Clin Lab Anal. 2021;35(2):e23600.
70. Sivashanmugam PKM, Kandasamy K, Subbiah M, Ravikumar R. Repurposing of histone deacetylase inhibitors: A promising strategy to combat pulmonary fibrosis promoted by TGF-β signalling in COVID-19 survivors. Life Sci. 2021;266:118883.
71. Geng H, Guan J. Mir-18a-5p inhibits endothelial-mesenchymal transition and cardiac fibrosis through the Notch2 pathway. Biochem Biophys Res Commun. 2017;491(2):329–36.

Publisher's note