

Research Article

Bin Zuo[#], Zhanchao Wang[#], Hua Lu^{*}, Hao Shen

Discovery of lipid metabolism-related diagnostic biomarkers and construction of diagnostic model in steroid-induced osteonecrosis of femoral head

<https://doi.org/10.1515/med-2025-1145>

received July 14, 2024; accepted January 7, 2025

Abstract: This study aimed to identify candidate diagnostic biomarkers for steroid-induced osteonecrosis of the femoral head (SONFH). Two datasets were downloaded from the Gene Expression Omnibus for analyses of differentially expressed genes (DEGs) and lipid scores in SONFH and control groups and a weighted gene co-expression network analysis. Overlap between genes in the disease-related module, DEGs, and lipid metabolism-related genes was evaluated to obtain lipid metabolism-related DEGs. A diagnostic model was constructed and evaluated using a nomogram analysis and calibration curves. Correlations between immune cells and lipid metabolism-related DEGs in the model were evaluated. In total, 1,724 DEGs were screened between the SONFH and control groups. Seven SONFH-related modules and 18 lipid metabolism-related DEGs were identified. A total of six optimized genes, *CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2*, were obtained. A model based on these genes had good diagnostic performance for SONFH in GSE123568 and GSE74089. The six genes were significantly related to two immune cell subsets, myeloid dendritic cell activated and neutrophil. These six genes may be novel biomarkers for SONFH and the combination is a potential diagnostic signature.

Keywords: steroid-induced osteonecrosis of femoral head, lipid metabolism, diagnostic model, immune cells

[#] Co-first authors.

*** Corresponding author: Hua Lu**, Department of Orthopedics, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai, 200092, China, e-mail: luhua@xinhuaamed.com.cn, tel: +86-021-25078999

Bin Zuo, Hao Shen: Department of Orthopedics, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai, 200092, China

Zhanchao Wang: Department of Orthopedics, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Chongming Branch, Shanghai, 202150, China

1 Introduction

Nontraumatic osteonecrosis of the femoral head (NONFH) is a common debilitating orthopedic disease with high disability rate [1]. A large-scale epidemiological study conducted in China estimated that the cumulative number of patients with NONFH has reached 8.12 million [2]. The pain and restricted hip joint mobility resulting from femoral head necrosis can significantly impair the quality of life of affected individuals [3]. Steroid hormones are the primary cause of NONFH [4]. Hormones can lead to an increase in adipocytes in the femoral head and accumulation in the medullary cavity, which increases the pressure in the medullary cavity and compresses the surrounding blood vessels, eventually leading to steroid-induced ONFH (SONFH) [5]. SONFH is a debilitating orthopedic condition characterized by the progressive deterioration of the hip joint in individuals aged 20–50 years [6,7]. As the disease advances, the range of joint motion decreases gradually. When necrotic lesions exceed 30% in volume, the collapse rate of the femoral head can reach 80%, severely impacting quality of life [8]. However, SONFH is challenging to diagnose in its early stages and difficult to reverse in its advanced stages. Therefore, early diagnosis utilizing effective biomarkers and interventions are of great clinical significance.

Abnormal lipid metabolism is a primary pathological process in osteonecrosis [9]. Hormone drugs increase blood lipids, and the abnormal distribution of lipids can cause fat emboli in microvessels, leading to ischemic necrosis of the femoral head. In cases of SONFH, lipid accumulation within osteocytes causes these cells to enlarge, compressing the nucleus and compromising the integrity of the cell membrane, ultimately resulting in cell death and necrosis [10]. Ren et al. confirmed that glucocorticoid-induced ONFH is closely associated with lipid metabolism disorders induced by high-dose steroid administration *in vivo* [11,12]. Jiang et al. [13] found that the occurrence of SONFH could be inhibited by reducing the content of blood lipids, which

could promote the growth of osteoblasts and inhibit adipocyte gene expression. Nozaki *et al.* [14] found that pravastatin can decrease the content of adipocytes within the femoral head of rats by regulating lipid metabolism, which in turn inhibits the occurrence of ONFH. Yeh *et al.* [15] confirmed that the expression of genes related to bone formation can be inhibited by hormones and increase the expression of lipid-related genes. However, the particular molecular mechanism by which lipid metabolism contributes to SONFH needs to be investigated.

In this study, the relationship between SONFH and lipid metabolism-related genes were studied. Genes that are significantly related to lipid metabolism and differentially expressed in SONFH were screened, and a diagnostic model based on these genes was constructed. The results provide a basis for diagnosis as well as potential therapeutic targets for SONFH.

2 Methods

2.1 Dataset and samples

GSE123568 and GSE74089 were downloaded on November 6, 2023 from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database [16] (<http://www.ncbi.nlm.nih.gov/geo/>). GSE123568 [17], used as a training dataset, contained 40 human blood serum samples from 30 patients with SONFH and 10 normal controls. The detection platform for GSE123568 was GPL15207 ([PrimeView] Affymetrix Human Gene Expression Array). GSE74089, used as the validation dataset, contained eight human synovial tissue samples from four patients with NFH and four normal controls (CTRL). The platform for GSE74089 was GPL13497 (Agilent-026652 Whole Human Genome Microarray 4x44K v2).

2.2 Screening of significantly differentially expressed genes (DEGs)

The limma package (version 3.34.7) (<https://bioconductor.org/packages/release/bioc/html/limma.html>) of R3.6.1 language [18] was used to screen DEGs between the control and SONFH groups in the GSE123568 dataset. False discovery rate (FDR) < 0.05 and $|\log_2\text{fold change (FC)}| > 0.5$ were selected as thresholds. Then, based on the DEGs, the pheatmap package (version 1.0.8) [19] (<https://cran.r-project.org/package=pheatmap>) in R3.6.1 was used for a bi-directional

hierarchical clustering analysis of the expression values [20,21]. The results were visualized using a heatmap.

2.3 Assessment of lipid scores

Relevant genes involved in lipid metabolism pathways were downloaded from the Molecular Signatures Database (MSigDB) module of the Gene Set Enrichment Analysis (GSEA) database (<http://software.broadinstitute.org/gsea/downloads.jsp>) [22]. The lipid scores for samples in the SONFH and CTRL groups were evaluated using the GSVA package (version 1.36.3) in R (<http://www.bioconductor.org/packages/release/bioc/html/GSVA.html>) [23]. Subsequently, the distributions of lipid scores in the SONFH and CTRL groups were compared using the Kruskal–Wallis test. Furthermore, the ability of the lipid score to identify SONFH was assessed through the receiver operating characteristic (ROC) curve method using the pROC package (version 1.12.1) (<https://cran.r-project.org/web/packages/pROC/index.html>) in R 3.6.1 [24].

2.4 Identification of disease- and lipid metabolism-related DEGs

A weighed gene co-expression network analysis (WGCNA) is a bioinformatics approach for constructing co-expression networks to identify modules associated with diseases and screening important pathogenic mechanisms or potential therapeutic targets [25]. Modules related to SONFH were screened using the WGCNA package (version 1.61) in R3.6.1 [26] (<https://cran.r-project.org/web/packages/WGCNA/index.html>) based on DEGs. The thresholds for module partitions were as follows: module set containing at least 100 genes and $\text{cutheight} = 0.995$. The correlations between each module and the disease status and lipid score were calculated, and the modules that were significantly correlated with these parameters were retained.

Then, lipid metabolism-related genes obtained from the GSEA database, genes in the modules that were correlated with the disease status and lipid score, and DEGs identified through screening were compared, and the overlapping genes were selected as lipid metabolism-related DEGs. A functional enrichment analysis of lipid metabolism-related DEGs was conducted using the Gene Ontology (GO)-Biology Process (BP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases based on DAVID version 6.8 [27,28] (<https://DAVID.ncifcrf.gov/>). A *p*-value of less than 0.05 was selected as the threshold for significant enrichment.

2.5 Construction of a diagnostic model based on lipid metabolism-related DEGs

Based on the expression levels of lipid metabolism-related DEGs, univariate logistic regression analyses were performed using the *rms* package version 6.3-0 [29] (<https://cran.r-project.org/web/packages/rms/index.html>) in R3.6.1, retaining only those genes with a *p*-value less than 0.05. The lipid metabolism-related DEGs identified through univariate logistic regression analyses were further refined using the LASSO algorithm in the *lars* package (version 1.2) [30] (<https://cran.r-project.org/web/packages/lars/index.html>) in the R 3.6.1 software to select the optimal lipid metabolism-related DEG combination. Subsequently, to assess the diagnostic potential of lipid metabolism-related DEGs in SONFH, three machine learning algorithms – support vector machine (SVM), decision tree, and random forest – were applied. The SVM model was constructed using the *e1071* package in R, the decision tree model was developed using the “*rpart*” package, and the random forest model was created using the “*randomForest*” package. The performance of each model was then evaluated using the ROC curve method in the R 3.6.1 *pROC* package (version 1.12.1), separately assessing the efficacy of model using both the GSE123568 training dataset and GSE74089 validation dataset. Finally, a nomogram model and calibration curves were constructed using the *rms* package (version 5.1-2) (<https://cran.r-project.org/web/packages/rms/index.html>) [31] in R3.6.1.

2.6 Immune correlation analysis

The samples in the GSE123568 dataset were used to evaluate the proportions of immune cell types via CIBERSORT (<https://CIBERSORT.stanford.edu/index.php>) [32]. Differences in immune cell frequencies between SONFH and CTRL were evaluated using the Kruskal–Wallis test in R3.6.1. Finally, correlations between the distribution of immune cells and lipid metabolism-related DEGs in the model were evaluated.

2.7 Cell culture and treatment

The MC3T3-E1 cell line was obtained from the Institute of Cell Bank for Biological Sciences (Shanghai, China). These cells were maintained in Dulbecco's Modified Eagle

Medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum, and incubated at 37°C in a 5% CO₂ atmosphere. To simulate steroid-induced femoral head necrosis at the cellular level, MC3T3-E1 cells were exposed to various concentrations of dexamethasone (DXM) (1, 10, 100, 200, and 300 μM) for 24 h. MC3T3-E1 cells not treated with DXM served as the control group.

2.8 Cell viability

The Cell Counting Kit-8 (CCK-8; Dojindo, Tokyo, Japan) was used to evaluate the viability of MC3T3-E1 cells, both untreated and treated with DXM. Cells were seeded at a density of 5,000 cells per well in 96-well plates and incubated at 37°C with 5% CO₂ for 72 h. Afterward, 10 μL of CCK-8 solution was added to each well and incubated for 2 h. Cell viability was measured at 450 nm using a microplate reader (BioTek, Winooski, VT, USA).

2.9 Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was isolated from the cell lines using RNAiso Plus (Takara, Beijing, China). The RNA was then reverse transcribed into complementary DNA using PrimeScript RT Master Mix (Takara). qRT-PCR was conducted using Power SYBR Green PCR Master Mix (Thermo, Waltham, MA, USA). The thermal cycling protocol included an initial denaturation at 95°C for 2 min, followed by 40 cycles of 95°C for 15 s and 60°C for 60 s. A dissociation step was performed at 95°C for 15 s, 60°C for 60 s, and 95°C for 15 s. Relative gene expression was determined using the $2^{-\Delta\Delta Ct}$ method, with *GAPDH* as the reference gene. Primer sequences are listed in Table 1.

2.10 Statistical analysis

Statistical analyses were performed using GraphPad Prism 9.0 (GraphPad Software), and the results are expressed as the mean ± standard deviation. Differences between two groups were evaluated using Student's *t*-tests, while comparisons among multiple groups were conducted using one-way ANOVA. A *p*-value of less than 0.05 was considered significant.

Table 1: Primers sequences

Name of primer	Sequences (5'-3')
CDK8-F	ACCTGTTGAATACGAGGGCT
CDK8-R	TGCCGACATAGAGATCCAGT
CHPT1-F	CACCGAAGAGGCACCATACTG
CHPT1-R	CCCTAAAGGGGAACAAGATTG
DPEP3-F	GATGCGGAGTTTCCCACTCG
DPEP3-R	GGCACACATGCGGTGAATG
ORMDL3-F	AACACGCGGGTGATGAACAG
ORMDL3-R	AGGGACACTCACAACGGGA
PIP5K1B-F	CTGGGAATAGGATACACAGTGGG
PIP5K1B-R	GCTGGGTAGGAACACATTTTC
PPT2-F	ACAGTGCTCGATCTCTTCGAT
PPT2-R	CAGCCTCTCGAACCCCTTG
OCN-F	CACTCCTCGCCCTATTGGC
OCN-R	CCCTCCTGCTTGGACAAAG
RUNX2-F	TGGTTACTGTCATGGCGGTA
RUNX2-R	TCTCAGATCGTTGAACCTTGCTA
COL1-F	GAGGGCCAAGACGAAGACATC
COL1-R	CAGATCACGTCATCGACAAC
GAPDH-F	TGACAACTTTGGTATCGTGAAGG
GAPDH-R	AGGCAGGGATGATGTTCTGGAGAG

3 Results

3.1 DEGs between SONFH and CTRL groups

A total of 1,724 DEGs (948 downregulated and 776 upregulated) were screened between the SONFH and CTRL

groups. A volcano diagram showing the gene expression differences between the two groups is shown in Figure 1a. A clustering heatmap based on the expression levels of DEGs indicated that genes with similar expression patterns were clustered together and could obviously distinguish SONFH and CTRL samples (Figure 1b).

3.2 Lipid scores

From the MSigDB of GSEA, 738 lipid metabolic pathway-related genes were obtained. The lipid score in the SONFH group was significantly lower than that in the CTRL group ($p = 0.0025$, Figure 2a). We examined the diagnostic value of the lipid score in SONFH. The results suggested that the lipid score effectively discriminates between the SONFH and CTRL groups with an AUC of 0.960, sensitivity of 100%, and specificity of 86.7% (Figure 2b).

3.3 Disease- and lipid metabolism-related DEGs

A soft threshold power of 9 was chosen to create networks with a scale-free topology (R^2 reached 0.9; mean connectivity was equal to 1) (Figure 3a). In the dendrogram, the genes were clustered into 11 modules (Figure 3b). Correlations between these modules and the disease status as well as lipid scores are depicted in Figure 3c. Based on the

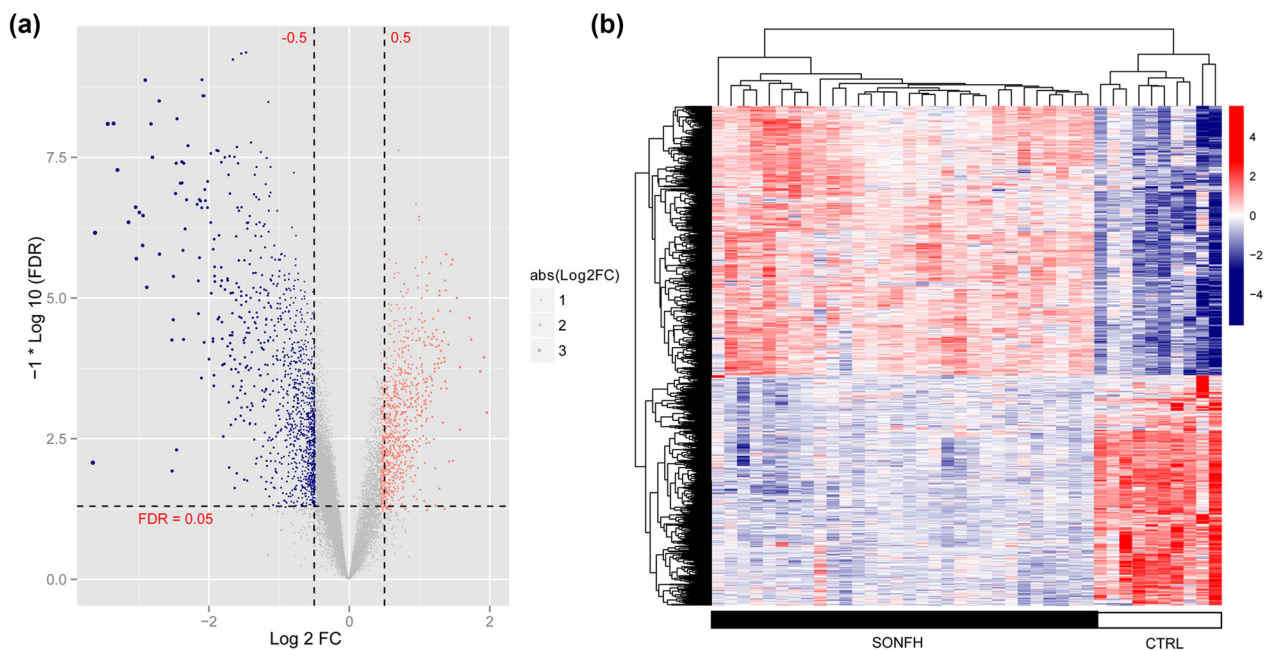


Figure 1: Volcano plot (a) and heat map (b) of DEGs between SONFH and CTRL groups. The gradient color change from blue to red represents expression changes, from downregulation to upregulation. FC: fold change. FDR: false discovery rate.

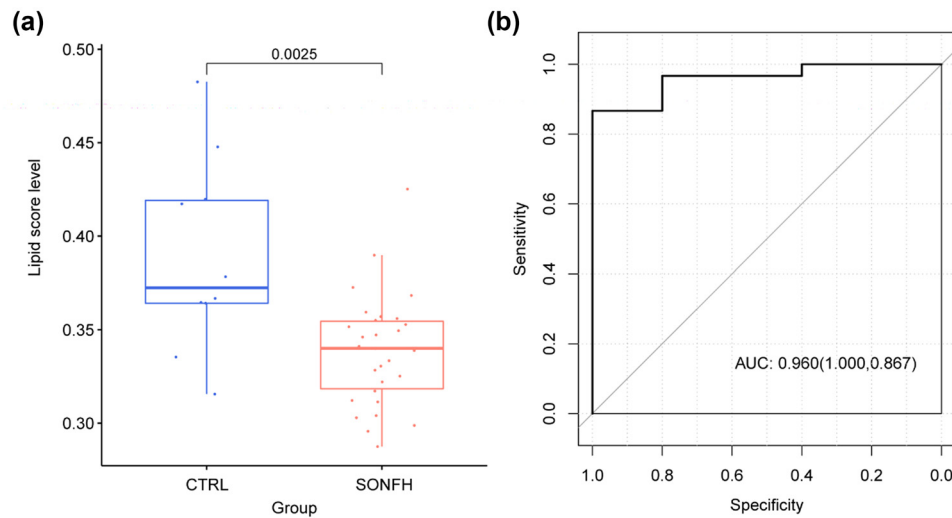


Figure 2: (a) Distribution of lipid scores in SONFH and CTRL groups. (b) ROC curve for the performance of the lipid score in identifying the sample type.

correlation coefficients for relationships between each gene module and phenotypes (SONFH, CTRL, or lipid score), the yellow module exhibited the strongest correlations ($\text{cor} = -0.80$, $p = 0.00$, Figure 3c).

Among the 11 modules, the blue, brown, green, purple, red, turquoise, and yellow modules were significantly associated with both the disease state and lipid score ($p > 0.3$), encompassing a total of 1,929 genes. A Venn diagram analysis revealed 18 overlapping lipid metabolism-related DEGs among the lipid metabolism-related genes obtained from the GSEA database, genes within the modules correlated with the disease status and lipid score, and DEGs identified through screening. These 18 lipid metabolism-related DEGs were as follows: *ACSM3*, *SLC27A3*, *CDK8*, *PTGS2*, *THEM5*, *DPEP1*, *PCTP*, *PIP5K1B*, *CHPT1*, *ORMDL3*, *DPEP3*, *TMEM86B*, *PPT2*, *PTGR1*, *PLA2G4F*, *DPEP2*, *PISD*, and *GPAT3*. GO and KEGG functional analyses of the 18 lipid metabolism-related DEGs revealed 13 GO-BP terms, including lipid metabolic process, leukotriene metabolic process, response to estradiol, aging, and prostaglandin biosynthetic process, and 8 KEGG signaling pathways, including metabolic pathways, glycerophospholipid metabolism, ether lipid metabolism, VEGF signaling pathway, and arachidonic acid metabolism, were identified (Figure 4a and b).

3.4 Diagnostic model based on lipid metabolism-related DEGs

Univariate logistic regression analyses of the 18 lipid metabolism-related DEGs revealed significant associations with

the disease state ($p < 0.05$) (Figure 5a). A LASSO analysis was performed to further screen these genes, resulting in the identification of six optimized lipid metabolism-related DEGs: *CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2* (Figure 5b).

We further visualized the expression levels of these six optimized lipid metabolism-related DEGs in the GSE123568 and GSE74089 datasets. Compared with levels in the CTRL group, the expression levels of *CDK8*, *CHPT1*, *ORMDL3*, *PIP5K1B*, and *PPT2* were significantly lower in the SONFH group, while the expression level of *DPEP3* was significantly higher ($p < 0.05$, Figure 6a and d). The diagnostic value of the optimized lipid metabolism-related DEGs in SONFH was examined utilizing SVM, decision tree, and random forest models. The SVM model achieved AUC values of 0.957 for the GSE123568 dataset and 0.912 for the GSE74089 dataset (Figure 6b and e). The random forest model yielded AUC values of 0.947 for the GSE123568 dataset and 0.873 for the GSE74089 dataset (Figure A1). Similarly, the decision tree model recorded AUC values of 0.943 for the GSE123568 dataset and 0.863 for the GSE74089 dataset (Figure A2). These results indicate that all three models, developed based on the expression levels of six optimized lipid metabolism-related DEGs, demonstrate significant discriminatory ability between the SONFH and CTRL groups. Furthermore, the nomogram model and calibration curves for six genes across two databases are illustrated in Figure 6c and f. The concordance index (C-index) of the nomogram reached 0.9967 in GSE123568 and 0.9669 in GSE74089, indicating a high capacity for distinguishing between the SONFH and CTRL groups.

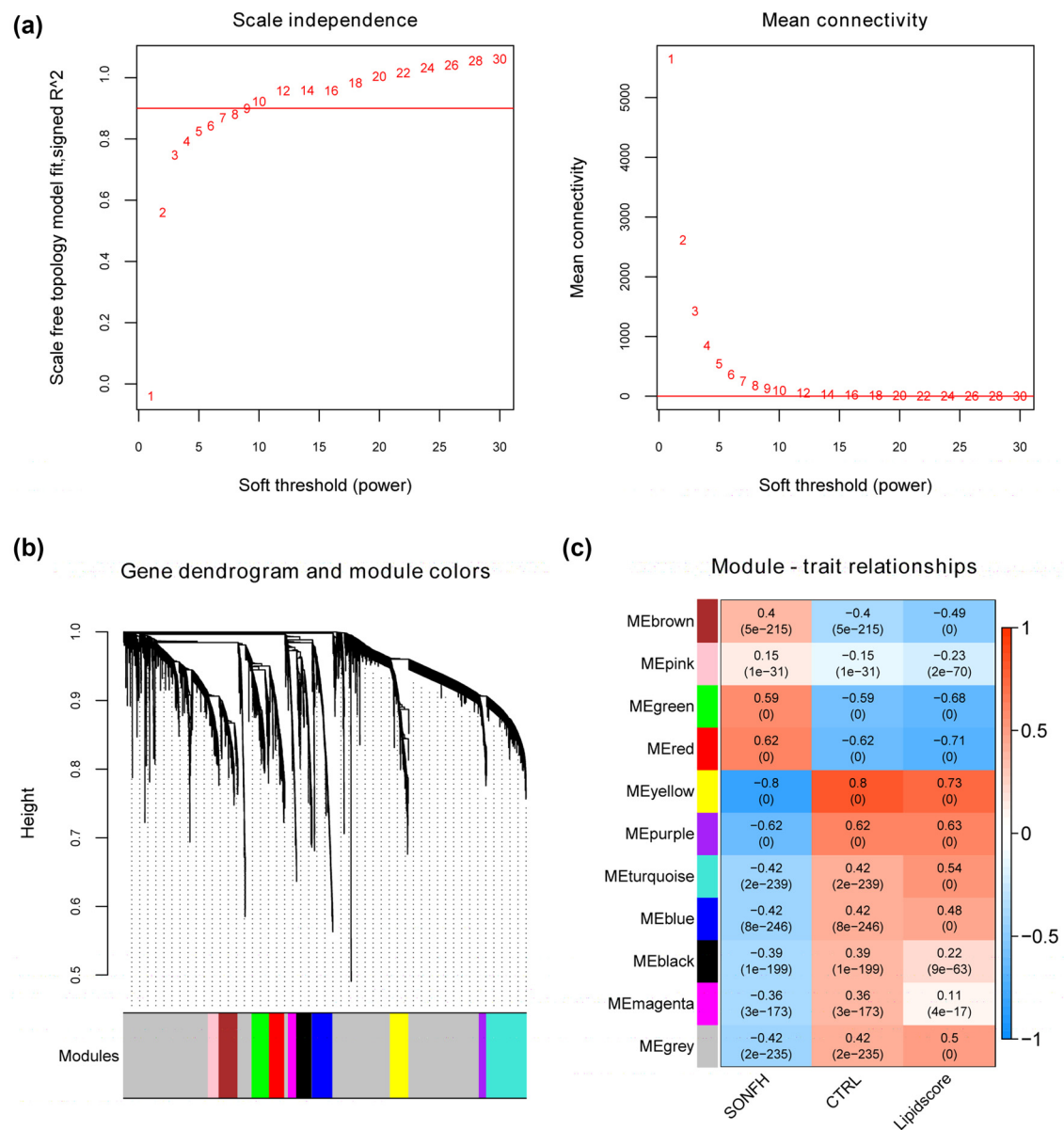


Figure 3: Weighted gene co-expression network analysis. (a) Selection of the soft threshold; (b) co-expression network of DEGs from GSE123568; (c) module-trait relationships: upper numbers for each module indicate the correlation coefficient, and numbers in brackets indicate the p -value.

3.5 Validation of six optimized lipid metabolism-related DEGs

In vitro experiments were conducted to assess the expression of the six optimized DEGs related to lipid metabolism. The MC3T3-E1 cell line was treated with DEX to simulate SONFH. To determine the optimal DEX concentration for induction, cell viability was evaluated using CCK-8 assays after treatment with various DEX concentrations. After 24 h of treatment with 300 μ M DEX, cell viability decreased by over 50% (Figure 7a). Consequently, 200 μ M DEX was

selected for further experiments. Following DEX treatment, mRNA expression analyses revealed that the osteogenesis-related genes *COL1*, *OCN*, and *RUNX2* were downregulated, indicating impaired osteogenic differentiation ($p < 0.05$, Figure 7b). Subsequently, qRT-PCR was used to analyze the mRNA expression levels of the six optimized lipid metabolism-related DEGs in both control and DEX-treated groups. The DEX group exhibited significantly lower levels of *CDK8*, *CHPT1*, *ORMDL3*, *PIP5K1B*, and *PPT2* than those in the control group, while *DPEP3* expression was significantly increased ($p < 0.05$, Figure 7c).

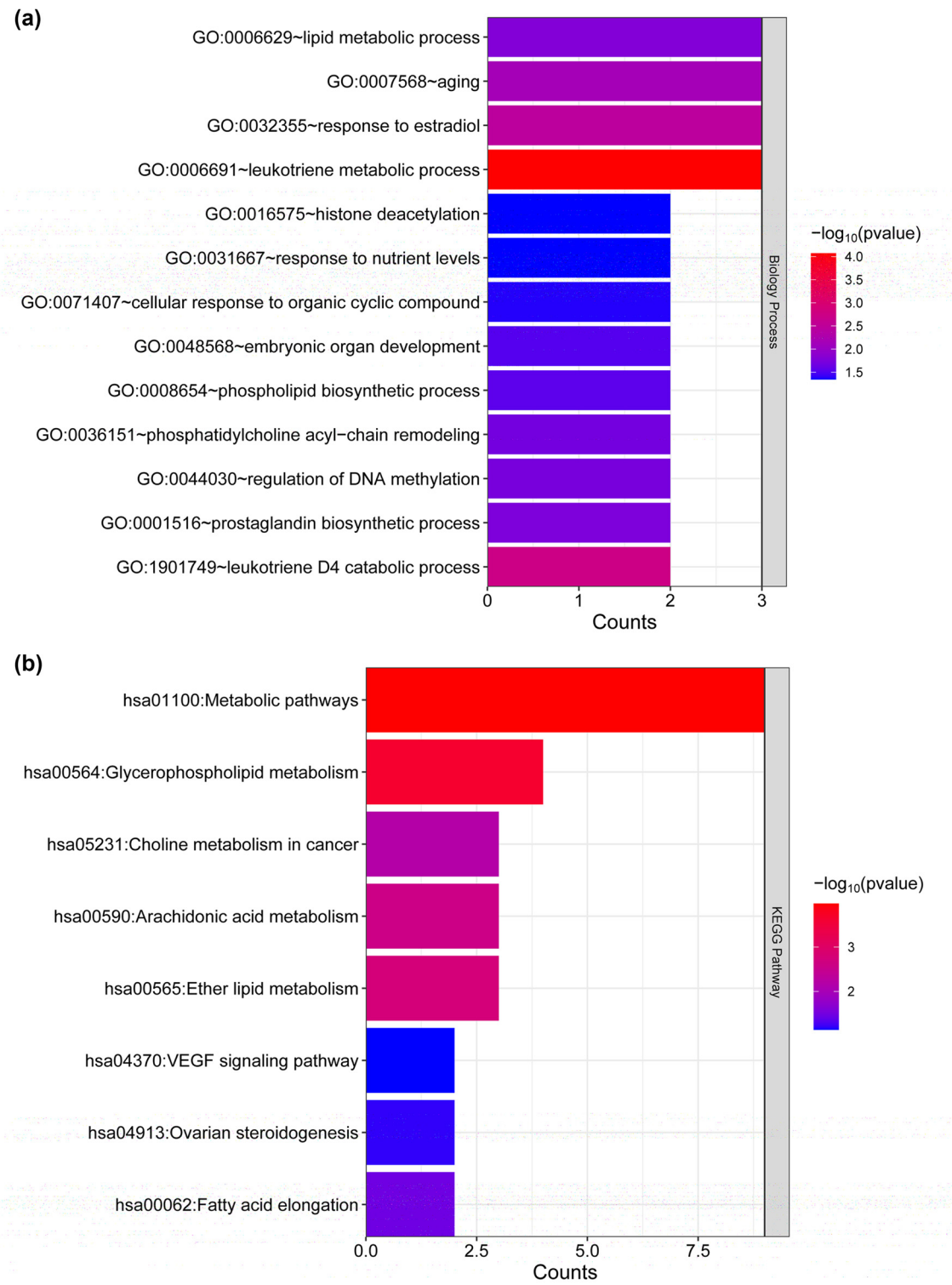


Figure 4: Functional pathway enrichment of lipid metabolism-related DEGs. (a) GO analysis of hub genes and (b) KEGG pathway analysis of hub genes. The horizontal axis represents the number of genes associated with the term, the vertical axis represents the term, and the color represents significance.

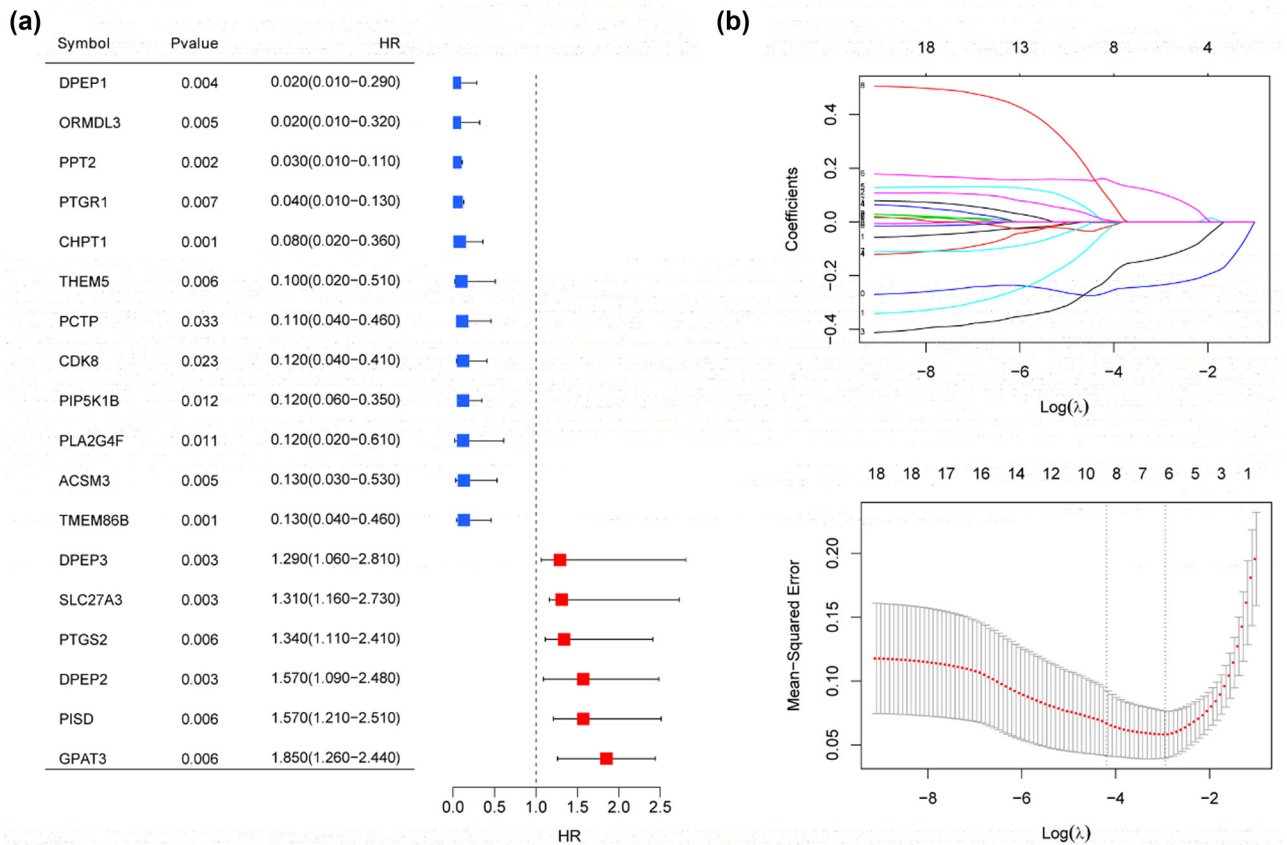


Figure 5: Univariate logistic regression analyses of lipid metabolism-related DEGs. (a) Screening of optimized genes associated with prognosis in GSE123568 by LASSO regression (b).

3.6 Relationships between candidate genes and differential immune cells

Finally, the immune cell distributions in the GSE123568 samples were evaluated using CIBERSORT, as shown in Figure 8a. As determined using the Kruskal–Wallis test, two immune cell types, myeloid dendritic cell activated and neutrophil, differed significantly between the two groups (Figure 8b); the frequency of myeloid dendritic cell activated was significantly higher in CTRL than in SONFH ($p < 0.05$), while that of neutrophil showed the opposite pattern ($p < 0.001$). Immune cell frequencies were positively correlated with the expression levels of *CDK8*, *CHPT1*, *ORMDL3*, *PIP5K1B*, and *PPT2* and *DPEP3* expression was negatively correlated with myeloid dendritic cell activated (Figure 9a). The directions of the correlations between genes and neutrophil and myeloid dendritic cell activated frequencies were opposite (Figure 9b). Additionally, the six genes were significantly correlated with the lipid score (Figure 9c).

4 Discussion

Owing to the high disability rate and difficulty in osteonecrotic tissue repair, SONFH has attracted increasing attention [33]. There is recent evidence for the role of abnormal lipid metabolism in SONFH [5,34]. In this study, we performed a multi-step integrative bioinformatics analysis of microarray data to identify lipid metabolism-related DEGs as diagnostic biomarkers of SONFH. Using gene expression profiles in the GSE123568 dataset, a total of 1,724 DEGs (948 downregulated and 776 upregulated) were identified. Additionally, 738 lipid metabolic pathway-related genes were obtained from the MSigDB of GSEA. A module analysis and optimized gene identification revealed six lipid metabolism-related DEGs. The SVM model was used to predict the potential application of biomarkers in SONFH diagnosis, with high AUC values (i.e., 0.957 for the GSE123568 dataset and 0.912 for the GSE74089 dataset), suggesting that this risk classifier has good discriminatory ability between SONFH and CTRL groups. Our results demonstrated that

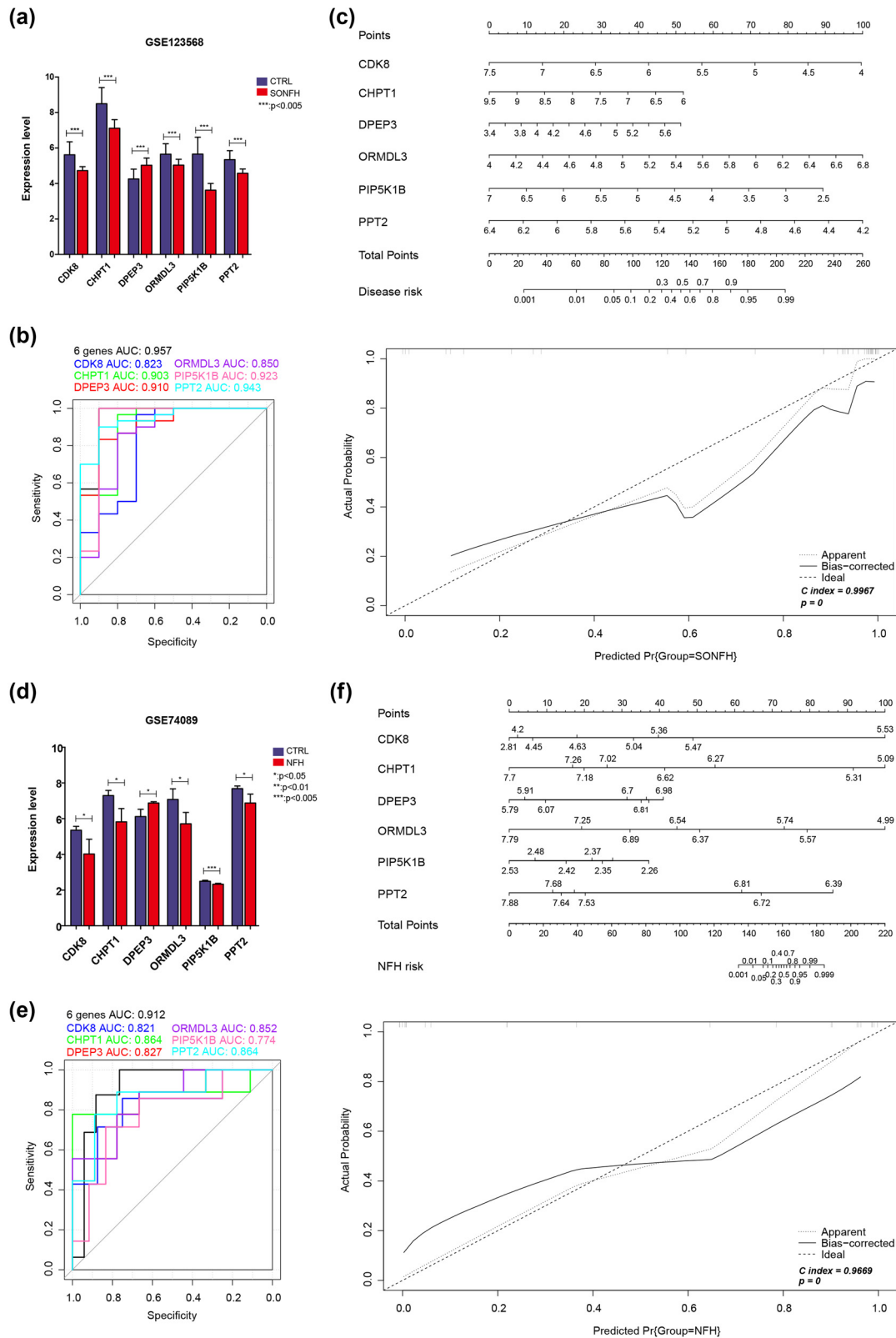


Figure 6: In GSE123568, the expression of six genes in the two groups (a), ROC curves and area under the curve for the candidate genes (b), and nomogram and calibration of the combination of six optimized genes (*CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2*) (c). In GSE74089, the expression of six genes in the two groups (d), ROC curves and area under the curve for the candidate genes (e), and nomogram and calibration of combination of six optimized genes (*CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2*) (f).

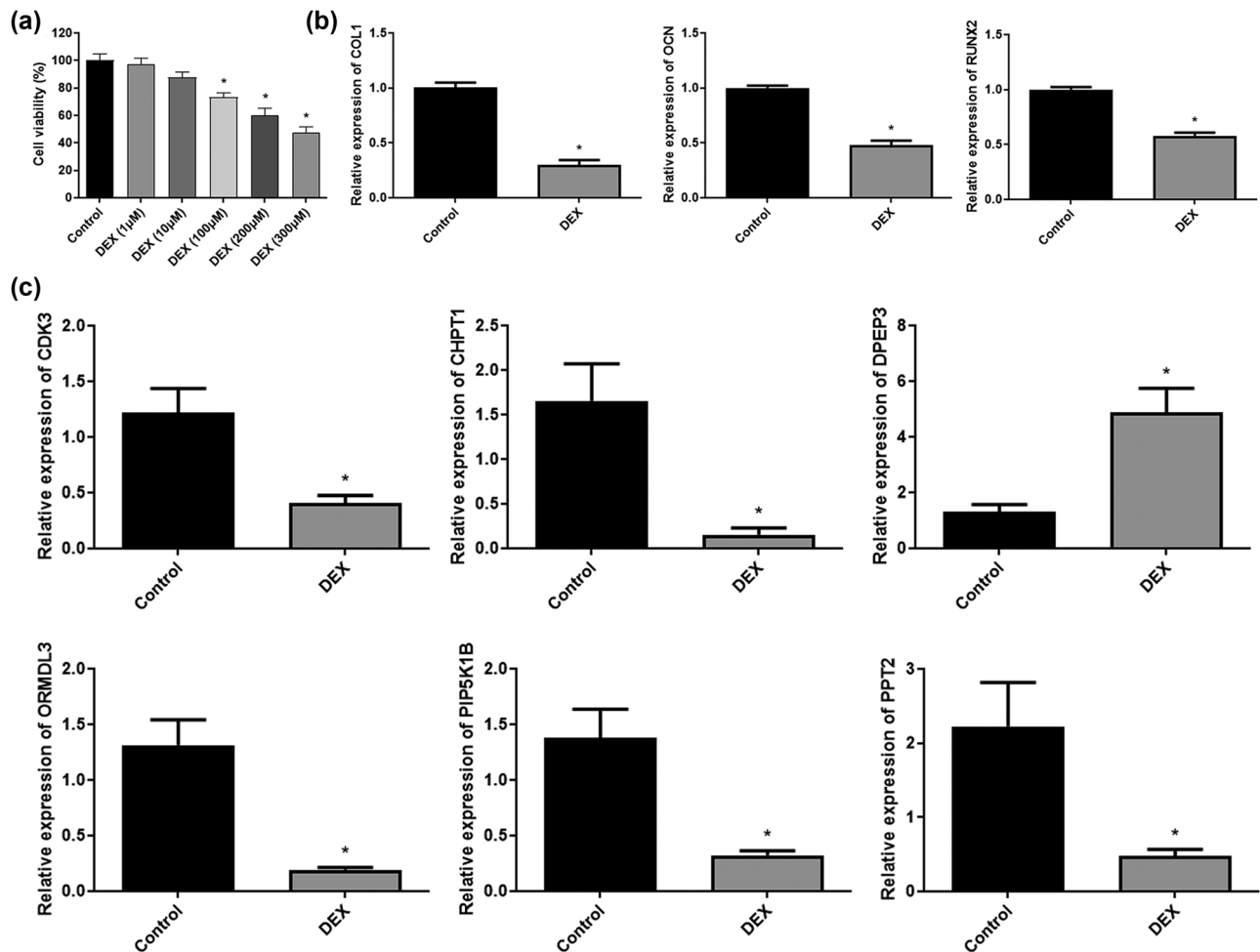


Figure 7: Validation of the six optimized genes. (a) CCK-8 assay was used to detect the viability of MC3T3-E1 cells and DEX-treated MC3T3-E1 cells. $*p < 0.05$, compared with the control. (b) qRT-PCR was used to detect the expression levels of *COL1*, *OCN*, and *RUNX2*. $*p < 0.05$, compared with the control. (c) qRT-PCR was used to detect the expression levels of *CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2*. $*p < 0.05$, compared with the control.

these six optimized lipid metabolism-related DEGs can be used as potential biomarkers for the diagnosis of SONFH.

A total of 18 overlapping lipid metabolism-related, SONFH-related genes were identified. We identified six optimized lipid metabolism-related DEGs among the overlapping DEGs, namely *CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2*. Previous studies have indicated that immune cell infiltration and inflammatory pathway activation significantly impact the occurrence and development of femoral head necrosis [15,34,35]. The correlations between immune cells and SONFH have been investigated extensively [35,36]. A previous study has indicated that circRNA-CDK8 is involved in osteogenic differentiation via influencing autophagy and apoptosis in a hypoxic microenvironment [37]. CDK8 regulates *de novo* lipogenesis and lipid accumulation [38] and is related to lipid production, accumulation, and metabolism [38–40]. In

addition, CDK8 is closely related to immune cells, including NK cells [41] and activated dendritic cells [42]. Dorigheo et al. revealed that CHPT1 contributes to *de novo* phosphatidylcholine synthesis [43]. It may participate in lipid metabolism disorders by regulating glycerophospholipid metabolism in silicosis [44]. DPEP3, a member of the DPEP family, is associated with immune responses and cholesterol clearance in vascular tissues [45,46]. The substrate of DPEP2 is a pro-inflammatory lipid mediator [47]. Lai et al. have shown that DPEP3 expression is positively related to the levels of IL-1 α , IL-1 β , and IL-17A and may be involved in inflammation [48]. ORMDL3 plays important role in metabolic diseases [49]. Members of the ORMDL family are involved in *de novo* sphingolipid synthesis [50], and ORMDL3 contributes to autoimmune/inflammatory diseases [51]. PIP5Ks and PIP4Ks are two lipid kinases families [52]. PIP5K1B encodes PIP5K β and is associated with CD28

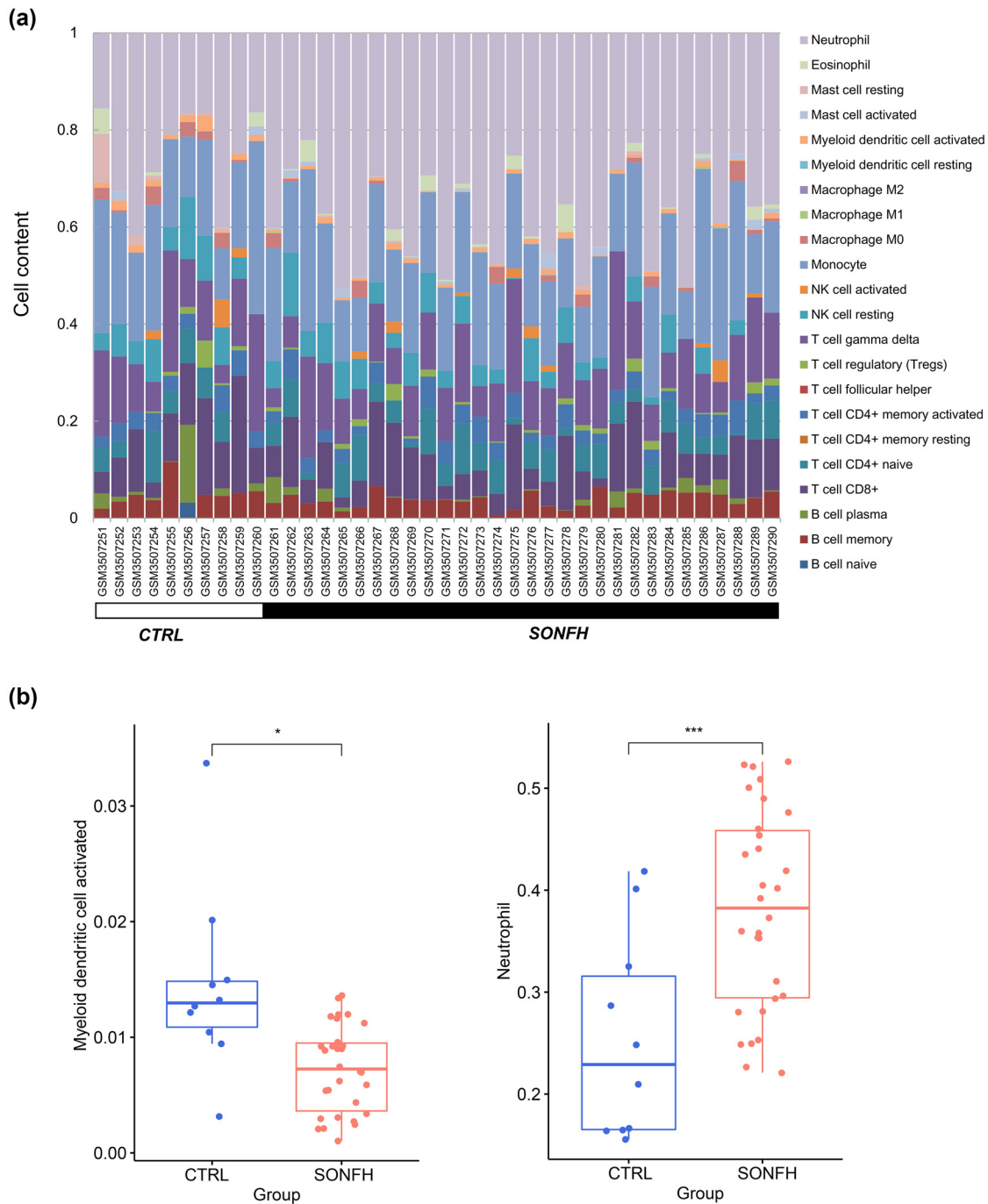


Figure 8: (a) Distribution of immune cells in two samples. (b) Two distinctly distributed immune cells (myeloid dendritic cell activated and neutrophil).

stimulation; it participates in T cell activation and lipid raft accumulation [53]. PPT2, a lipase, may play a special role in lysosomal lipid metabolism [54]. These results indicated that the candidate genes may contribute to the development of SONFH via lipid metabolism and immune responses. In addition, the six-gene combination has relatively high reference value for the diagnosis of SONFH.

The study results should be validated in larger clinical samples using bioinformatics approaches. Furthermore, the clinical diagnostic value of the six-gene model should be evaluated in additional *in vitro* and *in vivo* studies.

Several studies have developed prognostic models for SONFH based on DEGs. For instance, a partial least squares model exhibited high predictive accuracy (81.34–91.20%)

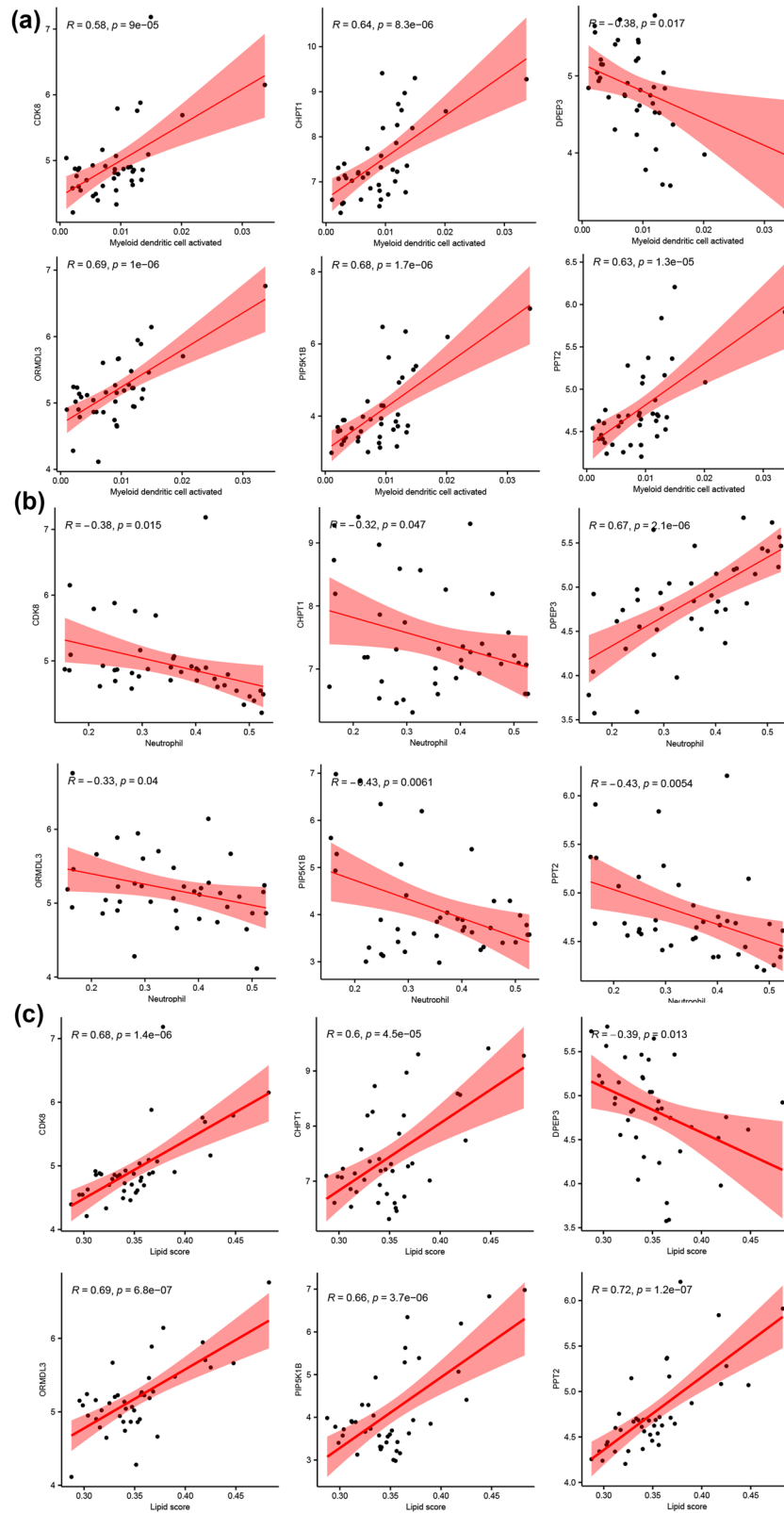


Figure 9: Correlations between expression levels of six genes and myeloid dendritic cell activated (a), neutrophil (b), and lipid score (c).

and AUC (0.804–0.901) when used to distinguish patients with SONFH from controls in a large clinical cohort, suggesting its potential contribution to the diagnosis and prevention of femoral head collapse [55]. Wu et al. conducted an ROC curve analysis based on ten endoplasmic reticulum stress-related genes, finding that all hub genes had AUC values greater than 0.6, confirming the diagnostic potential of these genes for SONFH [56]. Similarly, Liang et al. found 11 hub genes with high diagnostic value in a comparison between SONFH and non-SONFH samples (AUC > 0.9) [57]. In this study, the diagnostic value of six optimized lipid metabolism-related DEGs in SONFH was assessed using the SVM model. An ROC analysis demonstrated good diagnostic performance for SONFH (AUC values for GSE123568 and GSE74089 of 0.957 and 0.912, respectively). These results indicate that the six optimized lipid metabolism-related DEGs possess good diagnostic ability for SONFH.

The study had several limitations. First, the diagnostic efficiency and prognostic value of the critical genes were analyzed and validated solely within the GEO dataset. Second, most of the conclusions were derived from bioinformatics analyses. Therefore, additional experiments are necessary to validate these findings using tumor samples and clinical data. Additionally, *in vivo* and *in vitro* experiments could further elucidate the functions of the critical genes in SONFH.

5 Conclusion

In conclusion, this study showed that *CDK8*, *CHPT1*, *DPEP3*, *ORMDL3*, *PIP5K1B*, and *PPT2* may play important roles in the pathogenesis of SONFH and may be novel biomarkers. A model based on the six genes had relatively high auxiliary diagnostic value for SONFH.

Abbreviations

DEGs	differentially expressed genes
GO	gene ontology
GSEA	Gene Set Enrichment Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
NONFH	nontraumatic osteonecrosis of the femoral head
ROC	receiver operating characteristic
SONFH	steroid-induced osteonecrosis of femoral head
SVM	support vector machine
WGCNA	weighed gene co-expression network analysis

Acknowledgments: None.

Funding information: The authors state no funding involved.

Author contributions: Conception and design of the research: B.Z.; acquisition of data: B.Z.; analysis and interpretation of data: Z.C.W. and H.S.; statistical analysis: H.L.; drafting the manuscript: B.Z. and Z.C.W.; revision of manuscript for important intellectual content: H.L. and H.S. All authors read and approved the final manuscript.

Conflict of interest: The authors state no conflict interest.

Data availability statement: The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

- [1] Agarwala S, Vijayargiya M. Bisphosphonate combination therapy for non-femoral avascular necrosis. *J Orthop Surg Res.* 2019;14(1):112.
- [2] Zhao DW, Yu M, Hu K, Wang W, Yang L, Wang BJ, et al. Prevalence of nontraumatic osteonecrosis of the femoral head and its associated risk factors in the Chinese population: results from a nationally representative survey. *Chin Med J (Engl).* 2015;128(21):2843–50.
- [3] Liu N, Zheng C, Wang Q, Huang Z. Treatment of non-traumatic avascular necrosis of the femoral head (Review). *Exp Ther Med.* 2022;23(5):321.
- [4] Liu F, Wang W, Yang L, Wang B, Zhao D. An epidemiological study of etiology and clinical characteristics in patients with nontraumatic osteonecrosis of the femoral head. *J Res Med Sci.* 2017;22(1):15.
- [5] Lu SB, Peng J, Wang AY, Sun MX, Dong LM. Biomedical engineering in surgical repair of osteonecrosis: the role of imaging technology. In: Qin L, Genant HK, Griffith JF, Leung KS, editors. *Advanced bioimaging technologies in assessment of the quality of bone and scaffold materials.* Berlin, Heidelberg: Springer; 2007.
- [6] Kong L, Zuo R, Wang M, Wang W, Xu J, Chai Y, et al. Silencing microRNA-137-3p, which targets RUNX2 and CXCL12 prevents steroid-induced osteonecrosis of the femoral head by facilitating osteogenesis and angiogenesis. *Int J Biol Sci.* 2020;16(4):655–70.
- [7] Yu H, Liu P, Zuo W, Sun X, Liu H, Lu F, et al. Decreased angiogenic and increased apoptotic activities of bone microvascular endothelial cells in patients with glucocorticoid-induced osteonecrosis of the femoral head. *BMC Musculoskelet Disord.* 2020;21(1):277.
- [8] Chen L, Hong G, Fang B, Zhou G, Han X, Guan T, et al. Predicting the collapse of the femoral head due to osteonecrosis: from basic methods to application prospects. *J Orthop Transl.* 2017;11:62–72.
- [9] Gangji V, De Maertelaer V, Hauzeur J-P. Autologous bone marrow cell implantation in the treatment of non-traumatic osteonecrosis of the femoral head: five year follow-up of a prospective controlled study. *Bone.* 2011;49(5):1005–9.

- [10] Boskey AL, Raggio CL, Bullough PG, Kinnett JG. Changes in the bone tissue lipids in persons with steroid- and alcohol-induced osteonecrosis. *Clin Orthop Relat Res.* 1983;(172):289–95.
- [11] Ren X, Fan W, Shao Z, Chen K, Yu X, Liang Q. A metabolomic study on early detection of steroid-induced avascular necrosis of the femoral head. *Oncotarget.* 2018;9(8):7984–95.
- [12] Ren X, Shao Z, Fan W, Wang Z, Chen K, Yu X. Untargeted metabolomics reveals the effect of lovastatin on steroid-induced necrosis of the femoral head in rabbits. *J Orthop Surg Res.* 2020;15(1):497.
- [13] Jiang Y, Liu D, Kong X, Xu Y, Chen W, Lin N. Huogu I formula prevents steroid-induced osteonecrosis in rats by down-regulating PPAR γ expression and activating Wnt/LRP5/ β -catenin signaling. *J Tradit Chin Med.* 2014;34(3):342–50.
- [14] Nozaki Y, Kumagai K, Miyata N, Niwa M. Pravastatin reduces steroid-induced osteonecrosis of the femoral head in SHRSP rats. *Acta Orthop.* 2012;83(1):87–92.
- [15] Yeh CH, Chang JK, Wang YH, Ho ML, Wang GJ. Ethanol may suppress Wnt/ β -catenin signaling on human bone marrow stroma cells: a preliminary study. *Clin Orthop Relat Res.* 2008;466(5):1047–53.
- [16] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 2012;41(D1):D991–5.
- [17] Jia Y, Zhang Y, Li S, Li R, Li W, Li T, et al. Identification and assessment of novel dynamic biomarkers for monitoring non-traumatic osteonecrosis of the femoral head staging. *Clin Transl Med.* 2023 Jun;13(6):e1295. doi: 10.1002/ctm2.1295.
- [18] Ritchie ME, Belinda P, Di W, Yifang H, Law CW, Wei S, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
- [19] Wang L, Cao C, Ma Q, Zeng Q, Wang H, Cheng Z, et al. RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. *BMC Plant Biol.* 2014;14(1):169.
- [20] Szekely GJ, Rizzo ML. Hierarchical clustering via joint between-within distances: extending Ward's minimum variance method. *J Classif.* 2005;22(2):151–83.
- [21] Press W, Teukolsky S, Vetterling W, Flannery B. Hierarchical clustering by phylogenetic trees. Section 16.4 numerical recipes: The art of scientific computing. 3rd. New York, NY: Cambridge University Press; 2007.
- [22] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci.* 2005;102(43):15545–50.
- [23] Hnzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinf.* 2013;14(1):7.
- [24] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf.* 2011;12(1):77.
- [25] Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell.* 2012;148(6):1293–307.
- [26] Langfelder P, Horvath S, Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 2009;9(559):559.
- [27] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2008;44(1):44–57.
- [28] Wei HD, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37(1):1–13.
- [29] Harrell Jr FE. rms: Regression modeling strategies. R package version. 2016;5(2):1–263.
- [30] Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* 1997;16(4):385–95.
- [31] Wu J, Zhang H, Li L, Hu M, Chen L, Xu B, et al. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: a population-based analysis. *Cancer Commun.* 2020;40(7):301–12.
- [32] Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243–59.
- [33] Wang A, Ren M, Wang J. The pathogenesis of steroid-induced osteonecrosis of the femoral head: a systematic review of the literature. *Gene.* 2018;671:103–9.
- [34] Adapala NS, Yamaguchi R, Phipps M, Aruwajoye O, Kim HKW. Necrotic bone stimulates proinflammatory responses in macrophages through the activation of toll-like receptor 4. *Am J Pathol.* 2016;186(11):2987–99.
- [35] Zhao J, Zhang X, Guan J, Su Y, Jiang J. Identification of key biomarkers in steroid-induced osteonecrosis of the femoral head and their correlation with immune infiltration by bioinformatics analysis. *BMC Musculoskelet Disord.* 2022;23(1):1–12.
- [36] Wang B, Gong S, Shao W, Han L, Li Z, Zhang Z, et al. Comprehensive analysis of pivotal biomarkers, immune cell infiltration and therapeutic drugs for steroid-induced osteonecrosis of the femoral head. *Bioengineered.* 2021;12(1):5971–84.
- [37] Zheng J, Zhu X, He Y, Hou S, Liu T, Zhi K, et al. CircCDK8 regulates osteogenic differentiation and apoptosis of PDLSCs by inducing ER stress/autophagy during hypoxia. *Ann N Y Acad Sci.* 2021;1485(1):56–70.
- [38] Zhao X, Feng D, Wang Q, Abdulla A, Xie XJ, Zhou J, et al. Regulation of lipogenesis by cyclin-dependent kinase 8-mediated control of SREBP-1. *J Clin Invest.* 2012;122(7):2417–27.
- [39] Lindsay AK, Morales DK, Liu Z, Grahl N, Zhang A, Willger SD, et al. Analysis of *Candida albicans* mutants defective in the Cdk8 module of mediator reveal links between metabolism and biofilm formation. *PLOS Genet.* 2014;10(10):e1004567.
- [40] Gu W, Wang C, Li W, Hsu F-N, Tian L, Zhou J, et al. Tumor-suppressive effects of CDK8 in endometrial cancer cells. *Cell Cycle.* 2013;12(6):987–99.
- [41] Sohn GS, Cho S, Kim YM, Cho CH, Kim MR, Lee SR. Current medical treatment of uterine fibroids. *Obstet Gynecol Sci.* 2018;61(2):192–201.
- [42] Johannessen L, Sundberg T, O'Connell D, Kolde R, Berstler J, Billings K, et al. Small-molecule studies identify CDK8 as a regulator of IL-10 in myeloid cells. *Nat Chem Biol.* 2017;13(10):1102–8.
- [43] Dorighello G, McPhee M, Halliday K, Dellaire G, Ridgway ND. Differential contributions of phosphotransferases CEPT1 and CHPT1 to phosphatidylcholine homeostasis and lipid droplet biogenesis. *J Biol Chem.* 2023;299(4):3.
- [44] He H, Qi R, Cui J, Liu M, Guan B, Zhou Y, et al. Lipid characteristics of lung tissue in silicosis rat model were studied based on lipid metabolomics. *Toxicol Lett.* 2024;391:111–9.
- [45] Dabritz J, Weinlage T, Varga G, Wirth T, Walscheid K, Brockhausen A, et al. Reprogramming of monocytes by GM-CSF contributes to regulatory immune functions during intestinal inflammation. *J Immunol.* 2015;194(5):2424.

- [46] Chen HH, Keyhanian K, Zhou X, Vilmundarson RO, Almontashiri NA, Cruz SA, et al. IRF2BP2 reduces macrophage inflammation and susceptibility to atherosclerosis. *Circ Res: A J Am Heart Assoc.* 2015;117(8):671–83.
- [47] Steinke JW, Negri J, Payne SC, Borish L. Biological effects of leukotriene E4 on eosinophils. *Prostaglandins Leukot Essent Fat Acids.* 2014;91(3):105–10.
- [48] Lai W, Wang C, Lai R, Peng X, Luo J. Lycium barbarum polysaccharide modulates gut microbiota to alleviate rheumatoid arthritis in a rat model. *npj Sci Food.* 2022;6(1):34.
- [49] Brown RDR, Spiegel S. ORMDL in metabolic health and disease. *Pharmacol Ther.* 2023;245:108401.
- [50] Curtiss ML, Deshane JS. “Stick a Fork in Me; I’m Done”: epithelial cell expression of ORMDL sphingolipid biosynthesis regulator 3 mediates autophagic cell death. *Am J Respir Cell Mol Biol.* 2022;66(6):593–5.
- [51] Worgall TS. Sphingolipids, ORMDL3 and asthma: what is the evidence? *Curr Opin Clin Nutr Metab Care.* 2017;20(2):99–103.
- [52] Shibasaki Y. Cloning of cDNAs encoding two isoforms of 68-kDa type I phosphatidylinositol 4-phosphate 5-kinase. *J Biol Chem.* 1996;271(39):23611–4.
- [53] Kallikourdis M, Trovato AE, Roselli G, Muscolini M, Porciello N, Tuosto L, et al. Phosphatidylinositol 4-phosphate 5-kinase β controls recruitment of lipid rafts into the immunological synapse. *J Immunol.* 2016;196(4):1955–63.
- [54] Soyombo AA, Hofmann SL. Molecular cloning and expression of palmitoyl-protein thioesterase 2 (PPT2), a homolog of lysosomal palmitoyl-protein thioesterase with a distinct substrate specificity. *J Biol Chem.* 1997;272(43):27456–63.
- [55] Li T, Zhang Y, Wang R, Xue Z, Li S, Cao Y, et al. Discovery and validation an eight-biomarker serum gene signature for the diagnosis of steroid-induced osteonecrosis of the femoral head. *Bone.* 2019;122:199–208.
- [56] Wu T, Shi W, Zhou Y, Guo S, Tian H, Jiang Y, et al. Identification and validation of endoplasmic reticulum stress-related genes in patients with steroid-induced osteonecrosis of the femoral head. *Sci Rep.* 2024;14(1):21634.
- [57] Liang XZ, Liu XC, Li S, Wen MT, Chen YR, Luo D, et al. IRF8 and its related molecules as potential diagnostic biomarkers or therapeutic candidates and immune cell infiltration characteristics in steroid-induced osteonecrosis of the femoral head. *J Orthop Surg Res.* 2023;18(1):27.

Appendix

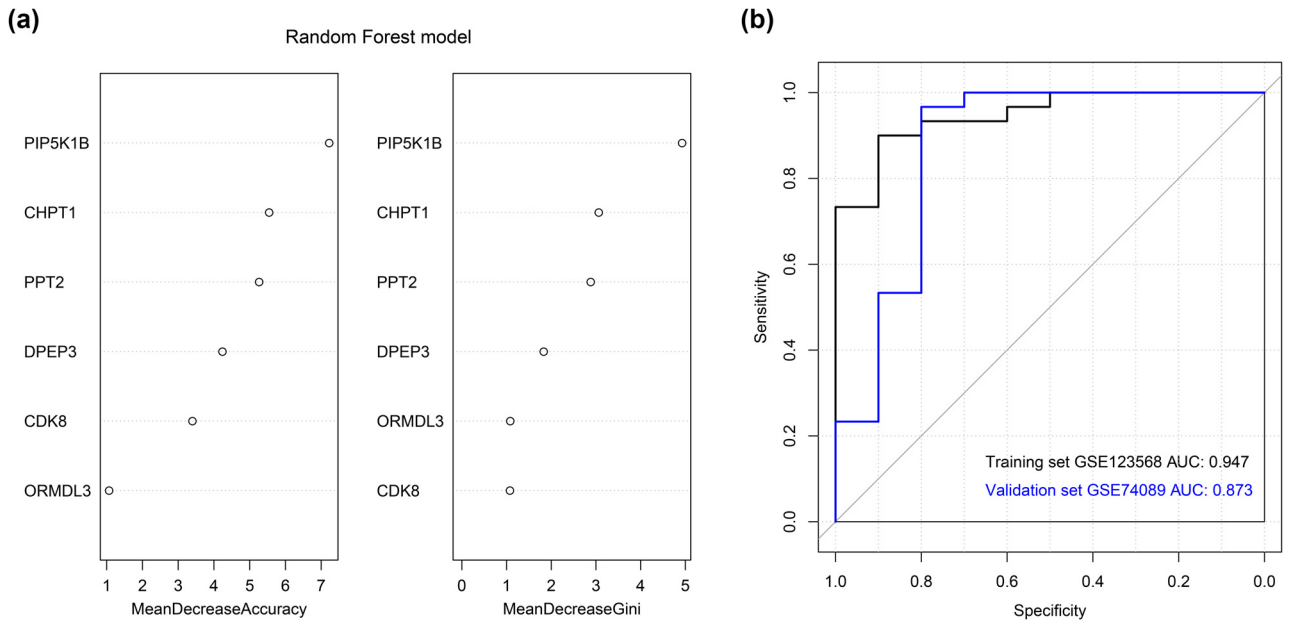


Figure A1: ROC results for the random forest model based on the six optimized genes for discrimination between the SONFH and CTRL groups.

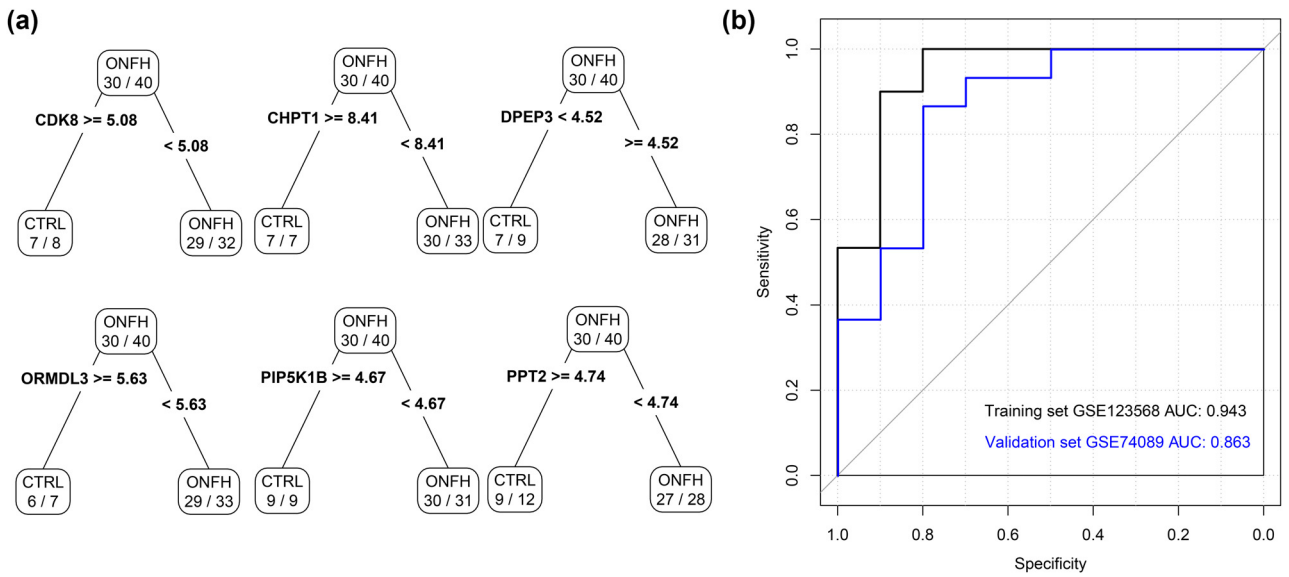


Figure A2: ROC results for the decision tree model based on the six optimized genes for discrimination between the SONFH and CTRL groups.