



Research Article

Screening of Potential Biomarkers in the Peripheral Serum for Steroid-Induced Osteonecrosis of the Femoral Head Based on WGCNA and Machine Learning Algorithms

Jian Zhang ¹, Chi Huang,¹ Zehan Liu,¹ Shuai Ren,¹ Zilong Shen,¹ Kecheng Han,² Weiguang Xin,³ Guanyi He,¹ and Jianyu Liu ¹

¹Department of Orthopedics, The Second Affiliated Hospital of Harbin Medical University, Harbin, 150001 Heilongjiang, China

²Department of Orthopedics, The Fifth Hospital of Harbin, Harbin, 150040 Heilongjiang, China

³Department of Orthopedics, The Second Hospital of Heilongjiang Province, Harbin, 150000 Heilongjiang, China

Correspondence should be addressed to Jianyu Liu; liujianyu4@163.com

Received 3 December 2021; Accepted 27 December 2021; Published 10 February 2022

Academic Editor: Fu Wang

Copyright © 2022 Jian Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Steroid-induced osteonecrosis of the femoral head (SONFH) has produced a substantial burden of medical and social experience. However, the current diagnosis is still limited. Thus, this study is aimed at identifying potential biomarkers in the peripheral serum of patients with SONFH. **Methods.** The expression profile data of SONFH (number: GSE123568) was acquired from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) in SONFH were identified and used for weighted gene coexpression network analysis (WGCNA). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to investigate the biological functions. The protein-protein interaction (PPI) network and machine learning algorithms were employed to screen for potential biomarkers. Gene set enrichment analysis (GSEA), transcription factor (TF) enrichment analysis, and competing endogenous RNA (ceRNA) network were used to determine the biological functions and regulatory mechanisms of the potential biomarkers. **Results.** A total of 562 DEGs, including 318 upregulated and 244 downregulated genes, were identified between SONFH and control samples, and 94 target genes were screened based on WGCNA. Moreover, biological function analysis suggested that target genes were mainly involved in erythrocyte differentiation, homeostasis and development, and myeloid cell homeostasis and development. Furthermore, GYPA, TMCC2, and BPGM were identified as potential biomarkers in the peripheral serum of patients with SONFH based on machine learning algorithms and showed good diagnostic values. GSEA revealed that GYPA, TMCC2, and BPGM were mainly involved in immune-related biological processes (BPs) and signaling pathways. Finally, we found that GYPA might be regulated by hsa-miR-3137 and that BPGM might be regulated by hsa-miR-340-3p. **Conclusion.** GYPA, TMCC2, and BPGM are potential biomarkers in the peripheral serum of patients with SONFH and might affect SONFH by regulating erythrocytes and immunity.

1. Introduction

Steroid-induced osteonecrosis of the femoral head (SONFH), a chronic and progressive femoral head disease mainly induced by long-term exposure to excessive glucocorticoids, can cause hip joint damage and dysfunction and ultimately affect the quality of life [1]. The worldwide morbidity of SONFH is increasing year by year, and it is estimated that 20,000-30,000 patients are diagnosed with SONFH that does

not depend on transmission each year in United States [2]. In particular, it is expected that the incidence of SONFH will rise in the next few years around the world because of the continuing impacts of the novel coronavirus pandemic since 2019 (COVID-19) [3]. Currently, although the diagnosis of SONFH, especially joint imaging techniques, is well established, patients with an early stage of SONFH are difficult to find because of the lack of effective and specific biomarkers [4]. Moreover, the cost and inconvenience of joint imaging

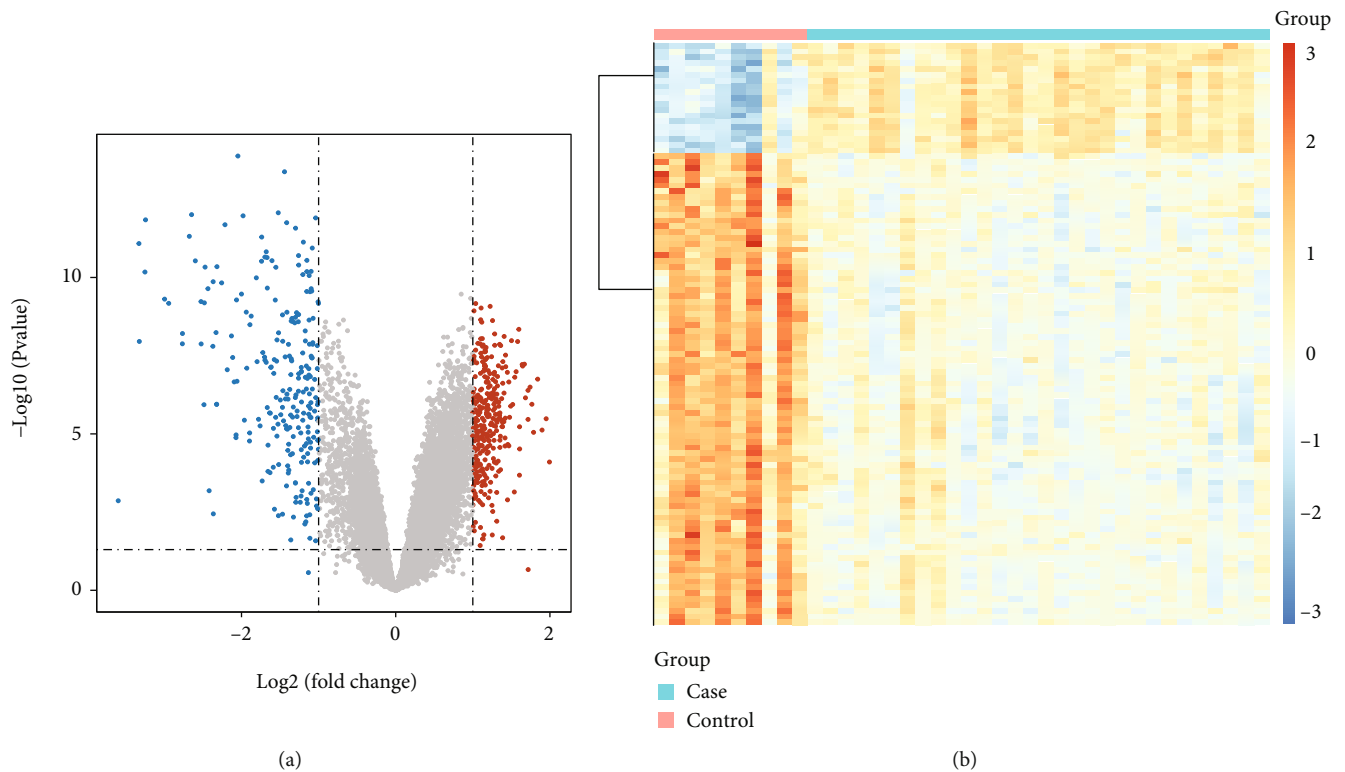


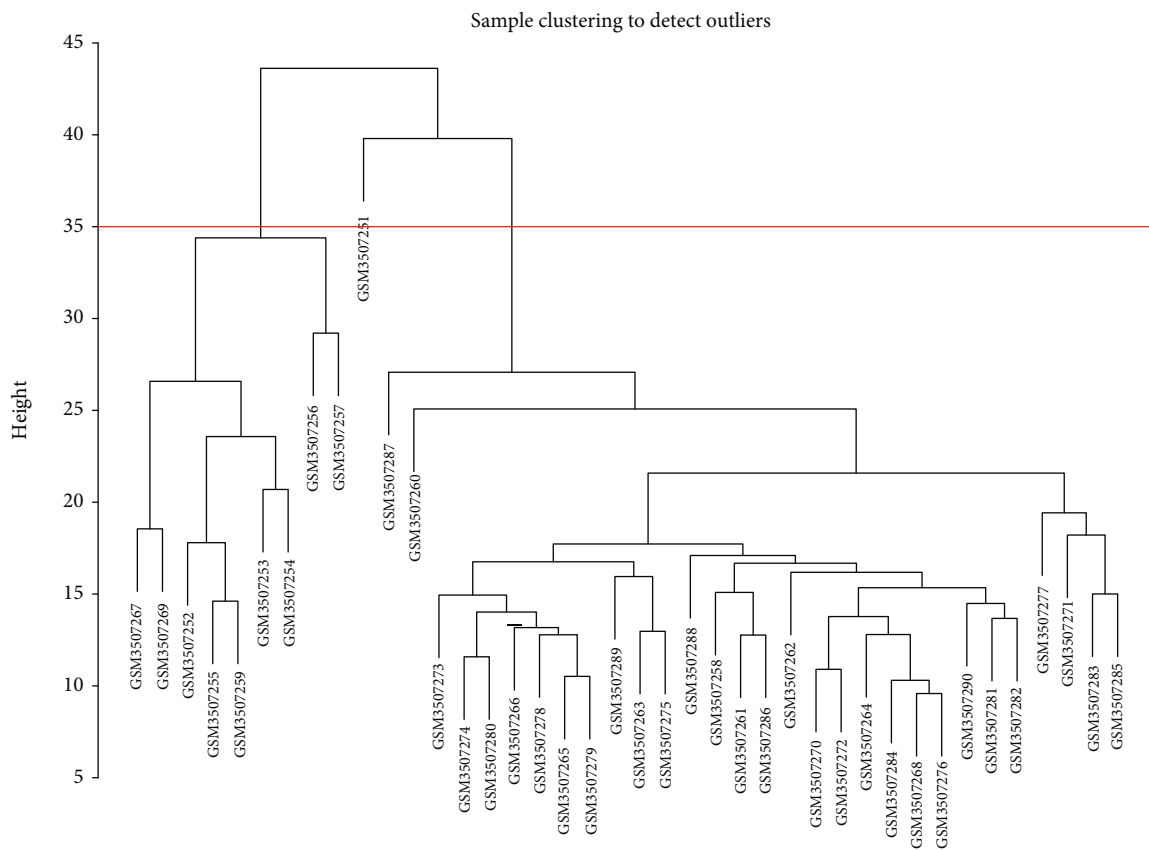
FIGURE 1: DEGs between SONFH patients and control samples. (a) Volcano plot showing the expression levels of DEGs. Red dots indicate upregulated genes in SONFH patients compared with control samples, while blue dots indicate downregulated genes in SONFH patients compared with control samples and gray dots indicate nonsignificantly different genes between SONFH patients and control samples. (b) Heat map showing the expression levels of top100 DEGs. Red indicates high expression, while blue indicates low expression.

techniques severely limit their application, specifically in China [5, 6]. Furthermore, total hip arthroplasty is effective in improving the quality of life, but most patients still endure the mental and financial pressure [7]. Therefore, the development of new biomarkers for SONFH diagnosis and treatment is urgently needed.

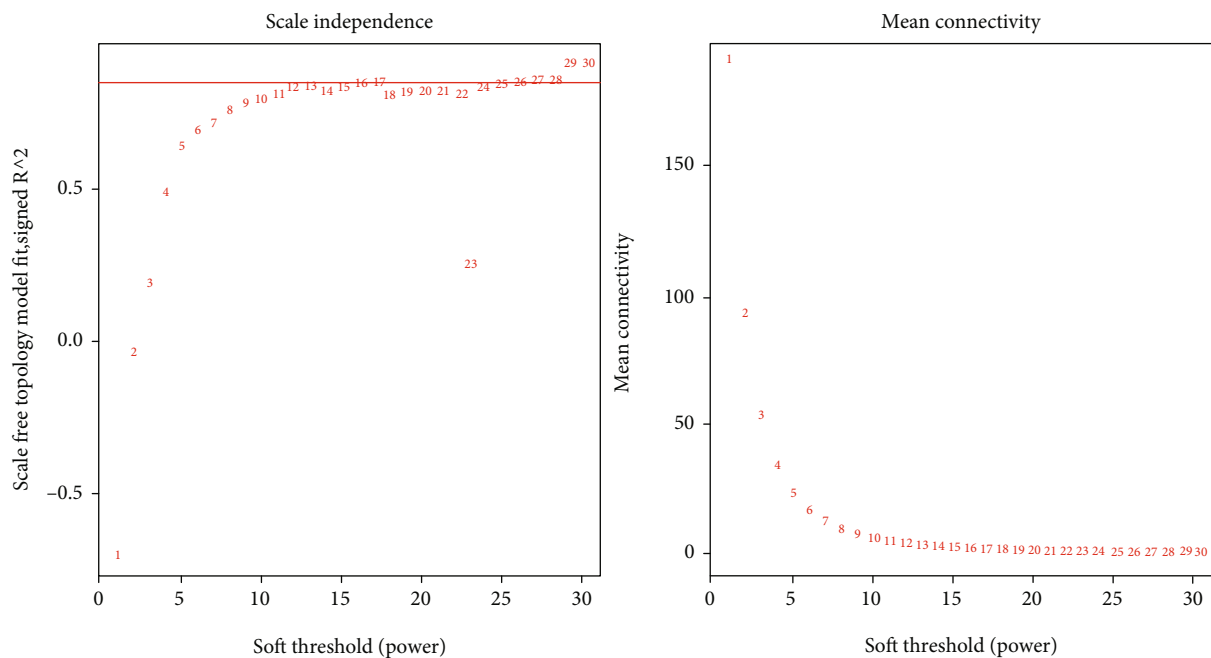
With the ongoing progression in sequencing technologies, bioinformatics analysis is an emerging and promising tool for screening potential biomarkers in a variety of neoplastic and nonneoplastic diseases [8, 9]. Data mining by using bioinformatics analysis based on public databases has facilitated the screening of new biomarkers for nonneoplastic diseases [10, 11]. In particular, combining weighted gene coexpression network analysis (WGCNA) and machine learning algorithms has greatly improved the accurate identification of disease-related biomarkers [12–15]. For example, FADD, CXCL2, and CXCL8 were identified as immune-related biomarkers of rheumatoid arthritis by integrating WGCNA and least absolute shrinkage and selection operator (LASSO) logistic regression and support vector machine recursive feature elimination (SVM-RFE) algorithms [12]. MACROD1 was found to contribute to the early diagnosis of tendinopathy based on the LASSO model, SVM-REF, and Gaussian mixture model (GMMs) algorithms [13]. Moreover, LSP1, GNLY, and MEOX2 are

likely to aid in the diagnosis and treatment of rheumatoid arthritis based on WGCNA machine learning strategies. Furthermore, CDK1, TOP2A, ADRA1A, FANCI, XRCC1, TPX2, CCNB2, CDK4, GLYATL1, and CFHR3 were identified to be core hub genes as potential biomarkers and treatment targets for hepatoblastoma by integrating WGCNA and random forest (RF) algorithm [14]. However, most bioinformatics analyses for SONFH have not been performed by integrating WGCNA and machine learning algorithms.

Therefore, the present study is aimed at identifying potential biomarkers in the peripheral serum for SONFH by integrating WGCNA and machine learning algorithms, including LASSO logistic regression, SVM-RFE, and RF. First, we identified differentially expressed genes (DEGs) between SONFH and control samples in the GSE123568 dataset downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) in NCBI. Moreover, WGCNA was performed based on the expression matrix of DEGs to screen target genes, and machine learning algorithms were used to identify potential biomarkers. Finally, we investigated the possible functions and regulatory mechanisms of the potential biomarkers. This study may provide new biomarkers for the diagnosis and treatment of SONFH and contribute to clarifying the pathogenesis of SONFH.

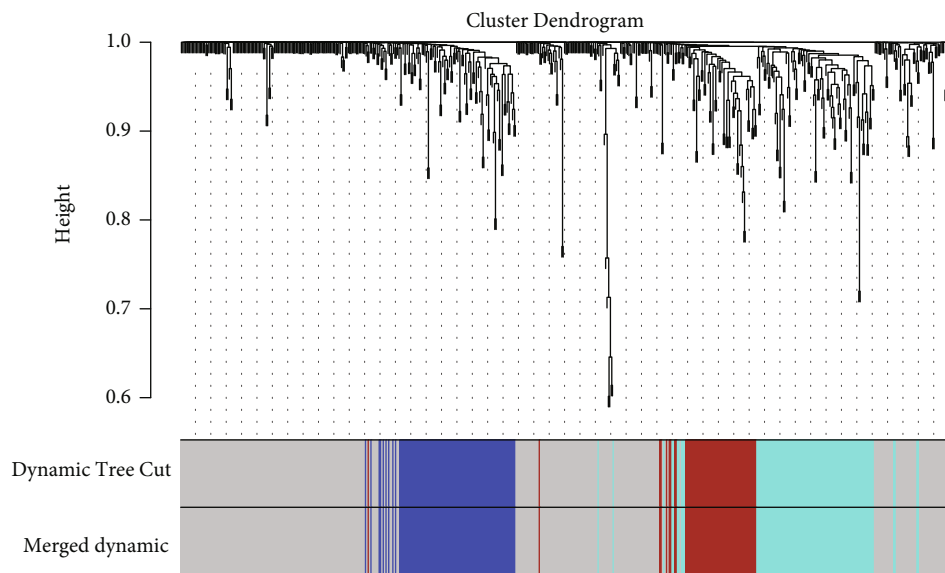


(a)

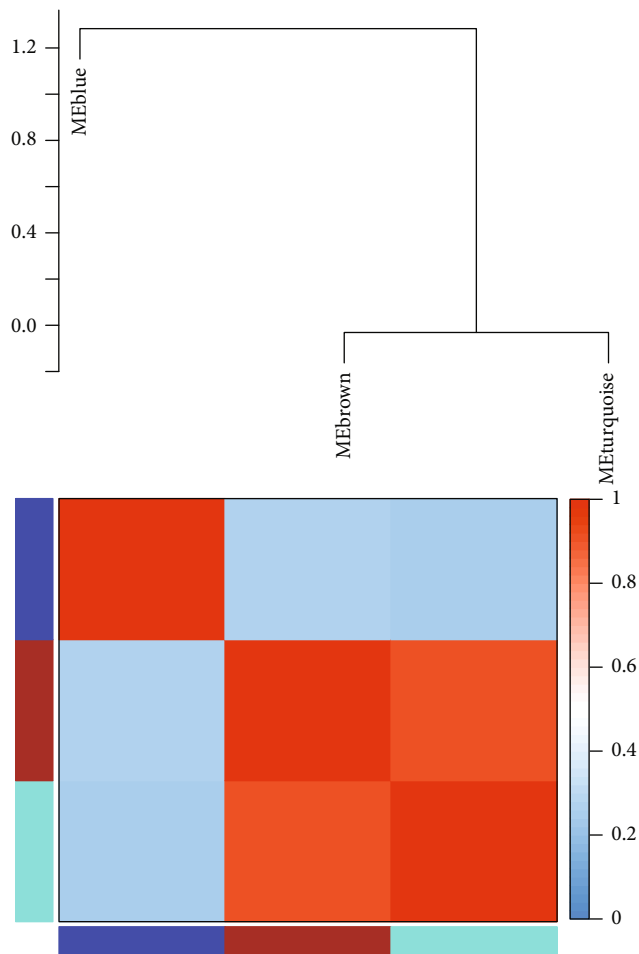


(b)

FIGURE 2: Continued.



(c)



(d)

FIGURE 2: Continued.

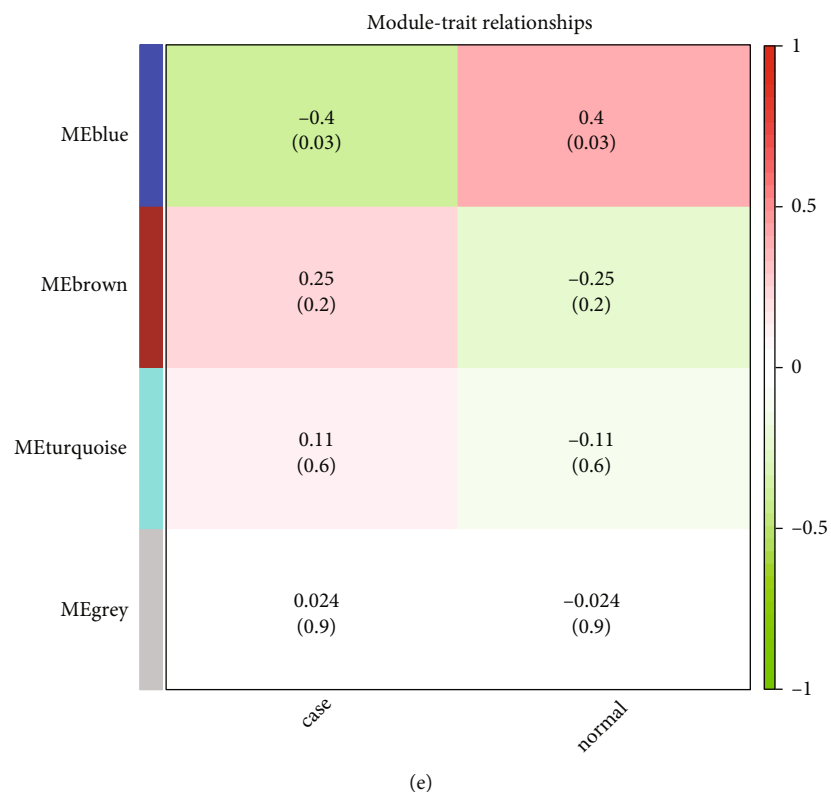


FIGURE 2: Identification of target genes by WGCNA. (a) Sample clustering analysis revealed that GSM3507251 sample was an outlier. (b) Soft threshold analysis suggested that gene associations were maximally consistent with the scale-free distribution and when $\beta = 26$. (c) Modules identified by merging modules with feature factors greater than 0.5 and setting the minimum number of genes in a module as 30. (d) Module correlations among the three identified modules. (e) Correlation between modules and SONFH.

2. Material and Methods

2.1. Data Collection and Preprocessing. The gene expression matrix of the peripheral serum for 30 SONFH and 10 control samples from China in the GSE123568 dataset, which was sequenced using the GPL15207 platform, was obtained from the GEO database.

2.2. Identification of DEGs in the Peripheral Serum of Patients with SONFH. “Limma” R package was used to screen DEGs between SONFH and control samples [16], and genes with $P < 0.01$ and $|\log_2^{FC}| > 1$ were regarded as DEGs.

2.3. Screening of Target Modules and Genes Based on WGCNA. To screen potential genes associated with SONFH, the expression matrix of DEGs was used to create a weighted gene coexpression network by using “WGCNA” R package [17]. First, clustering of all samples was performed to guarantee a reliable network. Second, we calculated the Pearson correlation coefficient between each pair of genes to evaluate the expression similarity of genes and acquire a correlation matrix. Moreover, we used the soft threshold function to convert the correlation matrix into a weighted neighborhood matrix. To ensure that gene correlations were maximally consistent with scale-free distribution, we used a soft connectivity algorithm to select the optimal soft threshold.

Subsequently, the neighborhood matrix was transformed into a topological overlap matrix (TOM). Furthermore, coexpression modules were obtained based on the criteria of dynamic tree cutting by setting the minimum number of genes in a module as 30. Finally, key modules were selected by correlation analysis, and genes in the key modules were considered as key genes.

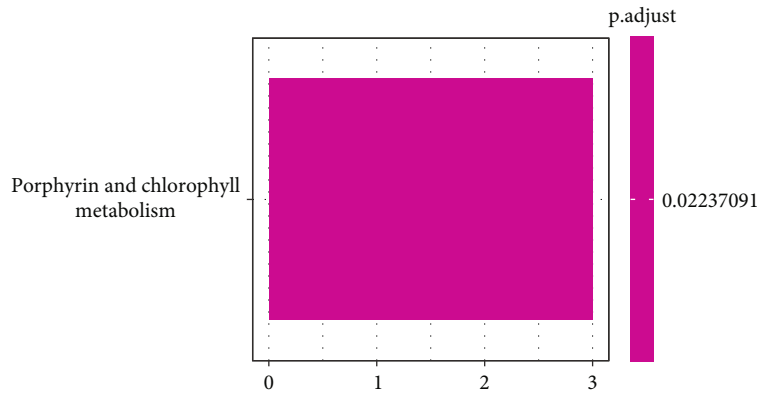
2.4. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Enrichment Analyses. Biological function enrichment of GO and KEGG analyses was performed by using the “clusterProfiler” R package [18]. GO enrichment analysis was performed to investigate the gene-related biological process (BP), molecular functions (MF), and cellular components (CC). KEGG enrichment analysis was conducted to explore gene-related signaling pathways. Statistical significance was set at adjusted P value < 0.05 .

2.5. Construction of a Protein-Protein Interaction (PPI) Network. A PPI network was constructed to investigate the protein interactions of genes through the Search Tool for the Retrieval of Interacting Genes (STRING, <https://string-db.org/>). Moreover, Cytoscape was selected to visualize the network, and the confidence score was set at 0.4.

2.6. Identification of Potential Biomarkers in the Peripheral Serum for SONFH Based on Machine Learning Algorithms.



(a)



(b)

FIGURE 3: Continued.

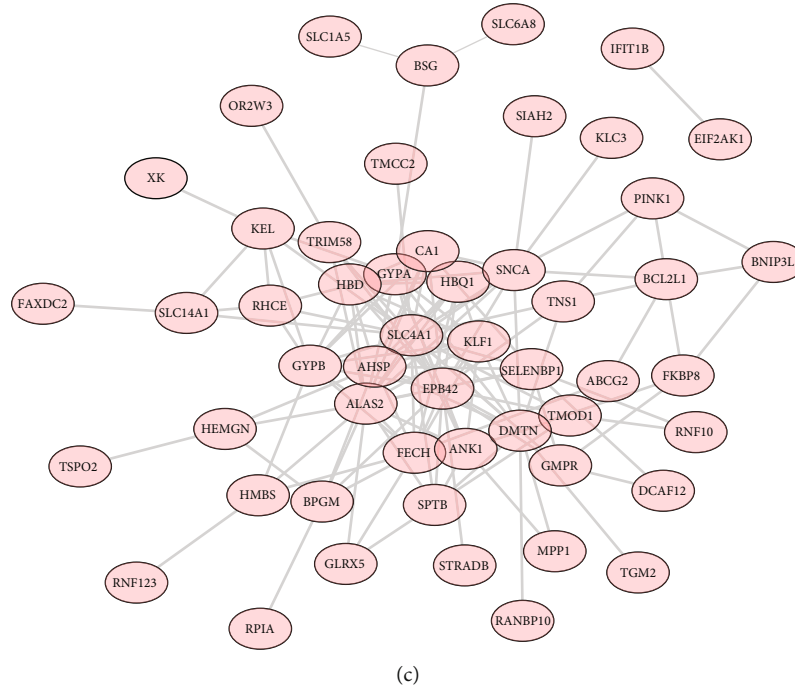


FIGURE 3: Functional enrichment analysis and PPI network: (a) GO results for the target genes; (b) KEGG results for the target genes; (c) PPI network of the target genes.

To begin with, the LASSO logistic regression algorithm [19] was performed to screen potential genes by using the “glmnet” R package [20], and receiver operating characteristic (ROC) analysis was selected to test the model reliability by calculating the area under the curve (AUC) value through the “pROC” R package [21]. Next, the SVM-RFE algorithm [22] was used to screen potential genes by using “e1071” R package [23]. In addition, the random forest (RF) algorithm [24] was conducted to screen potential genes by using the “randomForest” R package [25]. Similarly, the ROC curve was used to test the model reliability by using the “pROC” R package [21], and the top 10 genes based on %IncMSE ranking were regarded as the potential genes. Finally, overlapping genes among potential genes generated via LASSO, SVM-RFE, and RF algorithms were considered as potential biomarkers in the peripheral serum for SONFH.

2.7. Evaluation of the Expression Levels and Diagnostic Implications for Potential Biomarkers. Wilcoxon’s rank-sum test was used to analyze the expression levels of potential biomarkers, and ROC analysis was performed to evaluate whether potential biomarkers could differentiate SONFH samples from control samples by using the “pROC” R package [21].

2.8. Biological Functions and Regulating Mechanisms of Potential Biomarkers. Firstly, Gene Set Enrichment Analysis (GSEA) was performed using the “clusterProfiler” R package [18] to investigate the biological functions of potential biomarkers by the ordered gene expression matrix based on the Pearson correlation between each biomarker and other genes. Moreover, we also performed transcription

factor (TF) enrichment analysis for potential biomarkers using the CHIP-X Enrichment Analysis 3 (ChEA3) database (<https://amp.pharm.mssm.edu/chea3/>). We also constructed a competing endogenous RNA (ceRNA) network by predicting miRNAs and lncRNAs as potential biomarkers in the miRwalk and miRanda databases, separately.

3. Results

3.1. Identification of DEGs in the Peripheral Serum of Patients with SONFH. By setting the cut-off value as $P < 0.01$ and $|\log_2FC| > 1$, a total of 562 DEGs, including 318 upregulated and 244 downregulated genes, were identified in the peripheral serum of SONFH patients compared with control samples (Figures 1(a) and 1(b)).

3.2. Screening of Target Modules and Genes Based on WGCNA. To further screen genes related to SONFH, WGCNA was performed using 562 DEGs. As shown in Figure 2(a), clustering analysis of all samples showed that the GSM3507251 sample was poorly clustered. Therefore, this sample was excluded as an outlier in the WGCNA analysis. Next, the expression matrix of DEGs in the remaining 39 samples was used to construct a weighted gene coexpression network. The analysis of soft threshold selection revealed that gene associations were maximally consistent with the scale-free distribution and when $\beta = 26$ (scale free $R^2 = 0.85$, Figure 2(b)). Moreover, three coexpression modules were screened in the weighted gene coexpression network by merging modules with feature factors greater than 0.5 and setting the minimum number of genes in a module to 30 (Figure 2(c)). Furthermore, we also investigated the

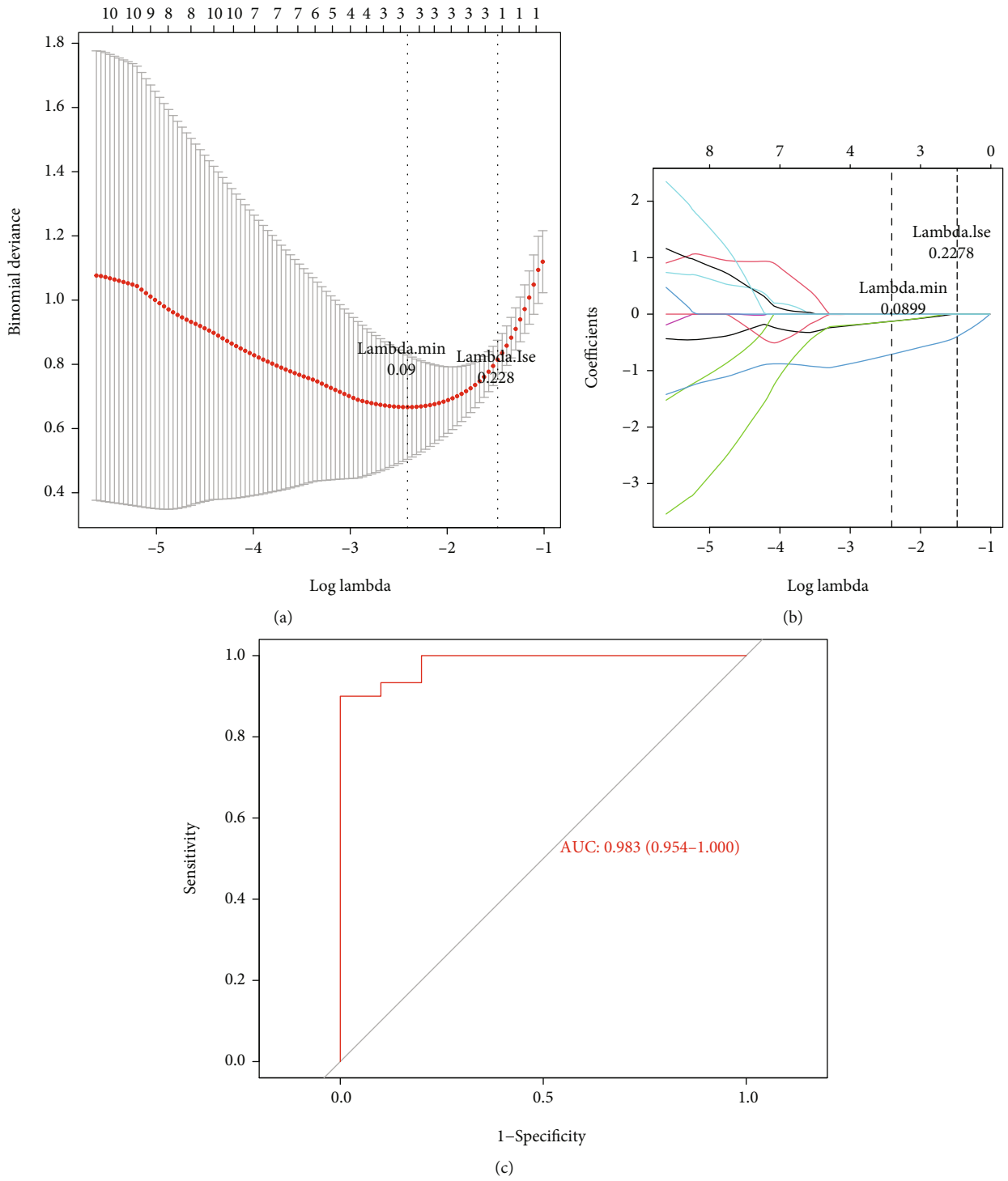


FIGURE 4: Continued.

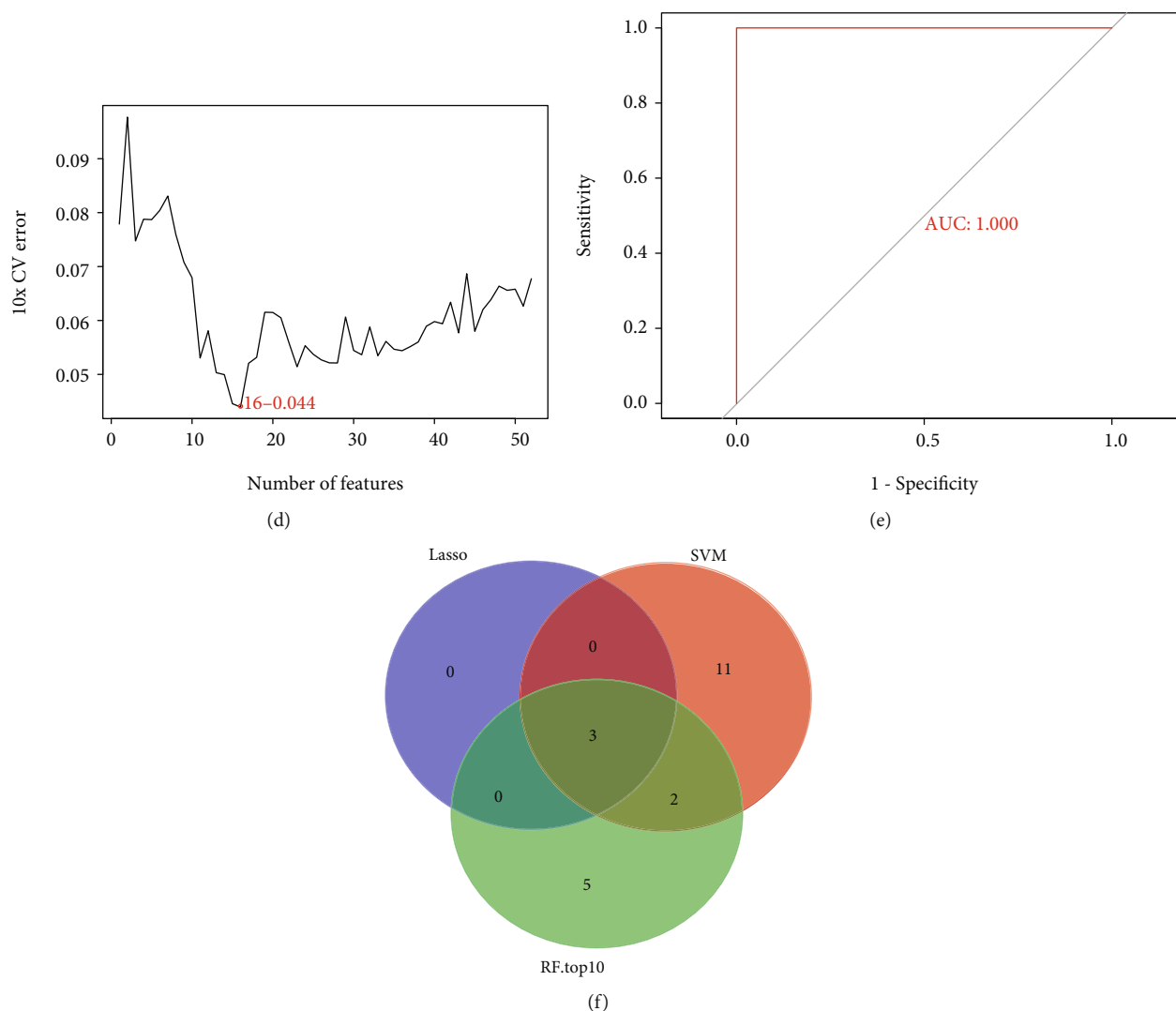


FIGURE 4: Identification of potential biomarkers for SONFH based on machine learning algorithms. (a) Log (Lambda) value of the three genes in LASSO model. (b) The most proper log (Lambda) value in LASSO model. (c) ROC curve of the LASSO model based on three genes. (d) The optimum error rate of SVM model based on 16 characteristic genes. (e) ROC curve of the RF module based on top 10 genes. (f) Venn diagram shows the overlapping genes in LASSO, SVM, and RF module.

module correlations and found that MEbrown and METurquoise presented stronger correlations than MEblue (Figure 2(d)). Finally, the MEblue module was selected as the target module because it had the highest correlation with SONFH (Figure 2(e)), and a total of 94 DEGs in the MEblue module were regarded as target genes.

3.3. GO and KEGG Enrichment Analyses. GO and KEGG enrichment analyses were performed to investigate the biological functions of the 94 target genes. GO analysis suggested that these target genes were mainly involved in erythrocyte differentiation, homeostasis, and development; myeloid cell homeostasis and development; and porphyrin metabolism-related BPs (Figure 3(a)). Moreover, these target genes were mainly associated with the cell cortex, cortical cytoskeleton, and mitochondrial and organelle outer membrane-related CCs (Figure 3(a)). Furthermore, KEGG enrichment analysis

suggested that these target genes were enriched only in the porphyrin- and chlorophyll metabolism-related signaling pathways (Figure 3(b)).

3.4. Construction of a PPI Network. To further investigate the protein interactions of the target genes, we constructed a PPI network. As shown in Figure 3(c), the PPI network included 52 nodes with 146 edges, and EBP42, ALAS2, FECH, TMOD1, ANK1, SLC4A1, HBQ1, GYPA, KLF1, and DMTN could affect more proteins. Thus, these 52 target genes were retained for subsequent analyses.

3.5. Identification of Potential Biomarkers in the Peripheral Serum for SONFH Based on Machine Learning Algorithms. To further identify the potential biomarkers in the peripheral serum of SONFH from 52 target genes, machine learning algorithms were selected and executed. Firstly,

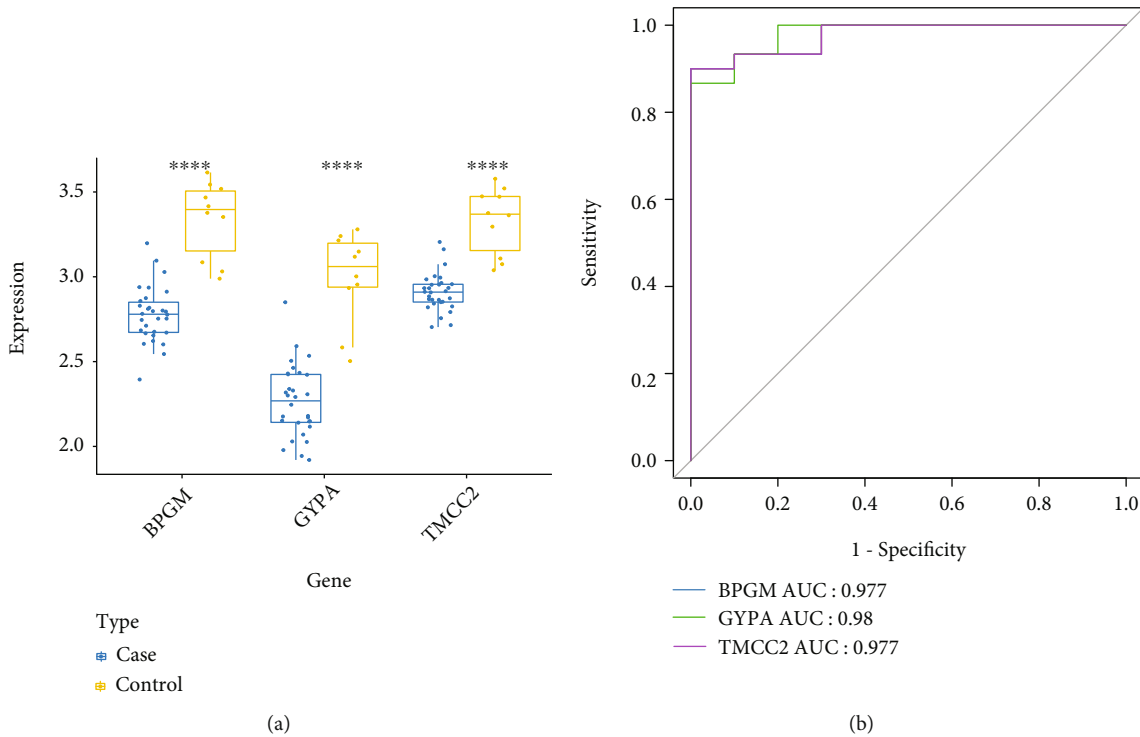


FIGURE 5: The expression levels and diagnostic implications of the potential biomarkers. (a) The expression levels of GYPA, TMCC2, and BPGM in SONFH patients and control samples. (b) ROC curves of GYPA, TMCC2, and BPGM.

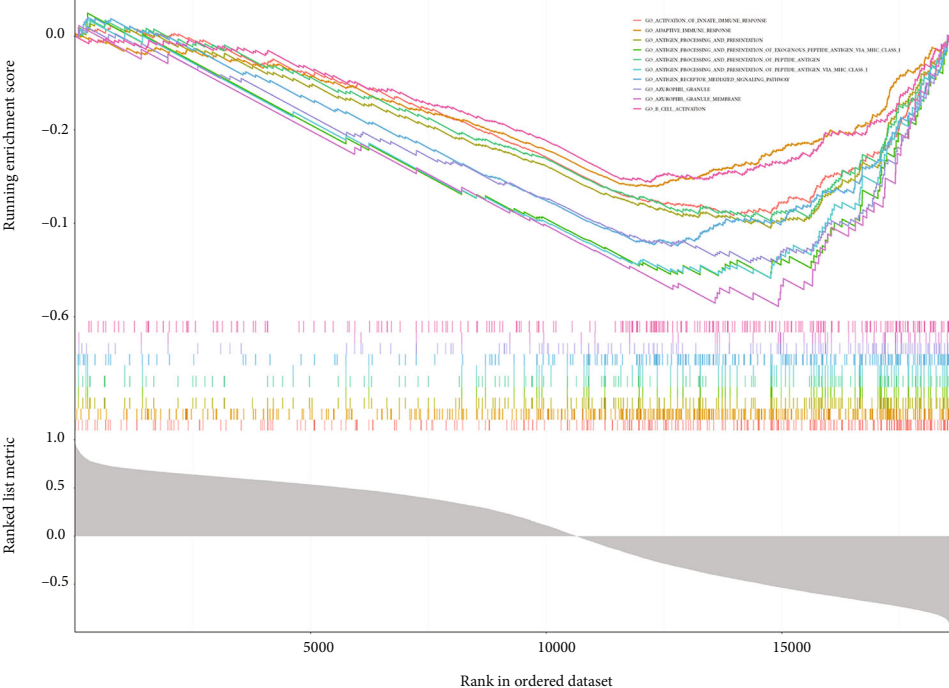
while constructing the LASSO model based on SONFH and control samples, λ analysis suggested that the model could accurately predict SONFH with $\lambda = 3$ (Figure 4(a)). Thus, GYPA, TMCC2, and BPGM were identified to build the LASSO module. We acquired the LASSO coefficient spectrum of the potential genes according to $\lambda = 3$ (Figure 4(b)). In addition, AUC analysis suggested that the LASSO module based on GYPA, TMCC2, and BPGM showed excellent performance (Figure 4(c)). On the other hand, SVM-RFE analysis revealed that the SVM model based on 16 characteristic genes showed an optimum error rate (0.044, Figure 4(d)). Thus, ANK1, BCL2L1, BPGM, BSG, FAXDC2, GYPA, HBQ1, HEMGN, IFIT1B, KEL, MPP1, SLC1A5, SPTB, TMCC2, TNS1, and TSPO2 were identified as the potential genes. At the same time, the RF algorithm identified the top 10 genes, including GYPA, RNF10, FECH, DMTN, FKBP8, BPGM, HEMGN, BNIP3L, and IFIT1B, from 52 target genes, and the RF module based on these 10 genes also showed good generalization performance (Figure 4(e)). Finally, three common potential genes, namely, GYPA, TMCC2, and BPGM, were regarded as the potential biomarkers in the peripheral serum of SONFH patients using the above three algorithms (Figure 4(f)).

3.6. Evaluation of the Expression Levels and Diagnostic Implications for the Potential Biomarkers. To further investigate the role of GYPA, TMCC2, and BPGM in SONFH, we first observed their expression levels in SONFH patients. Interestingly, we found that the expression of GYPA,

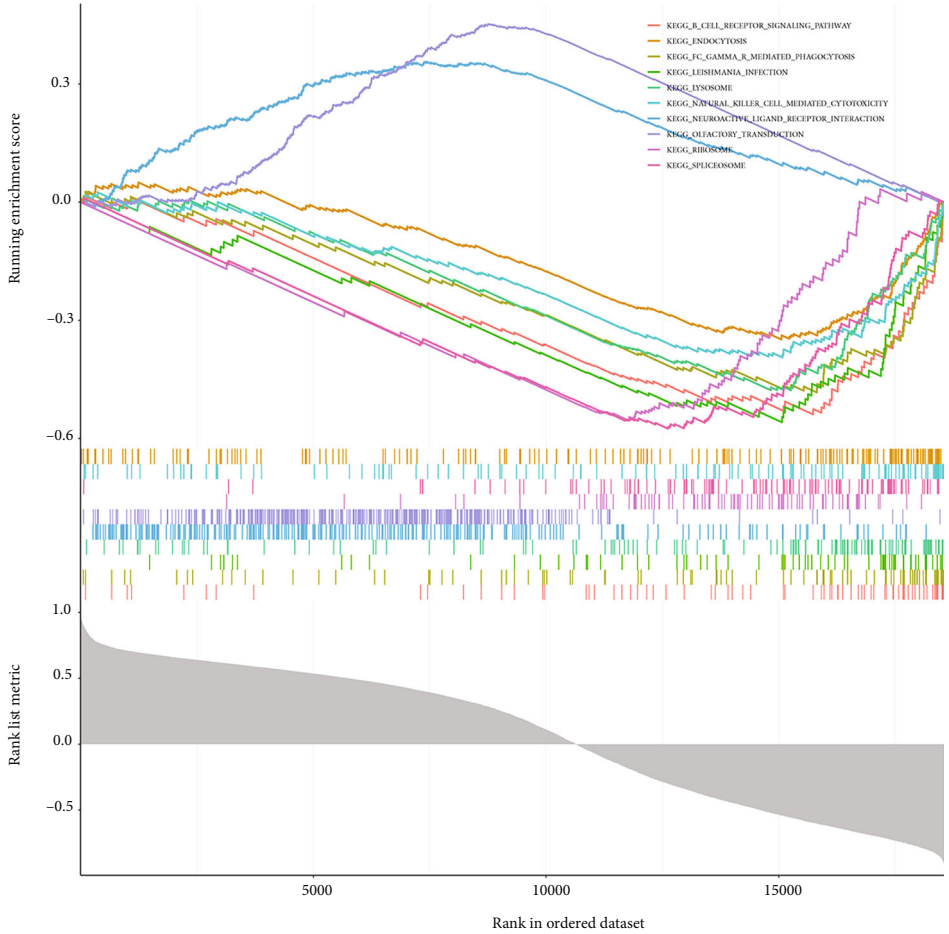
TMCC2, and BPGM was downregulated in SONFH patients compared with the control samples (Figure 5(a)). Moreover, ROC analyses suggested that GYPA, TMCC2, and BPGM might be used as diagnostic biomarkers in the peripheral serum of SONFH patients (Figure 5(b)).

3.7. Biological Functions and Regulating Mechanisms of the Potential Biomarkers. To further investigate the biological functions and regulatory mechanisms of GYPA, TMCC2, and BPGM, GSEA was performed based on their ordered gene expression matrix. As shown in Figure 6, GO and KEGG analyses revealed that GYPA, TMCC2, and BPGM are mainly involved in the B-cell receptor, endocytosis, FC gamma R-mediated D phagocytosis, T-cell, and natural killer cell-mediated cytotoxicity signaling pathways and are associated with activation of the innate immune response, adaptive immune response, antigen processing and presentation, antigen receptor signaling pathways mediated by antigen receptors, ATP metabolism, and B cell activation (Figure 6).

Moreover, to investigate the regulatory mechanisms of GYPA, TMCC2, and BPGM, we first predicted their TFs and constructed a TF-potential biomarker network (Figure 7(a)). Notably, GYPA might be regulated by ZBTB33, SOX9, ESR1, and BACH1; BPGM might be regulated by CTCFL, FOXO1, TCF7L2, USF2, RBPJ, and so on, and TMCC2 could be regulated by MAFK, JUN, SOX11, AR, and so on (Figure 7(a)). Furthermore, we also predicted the targeted miRNAs and lncRNAs of GYPA, TMCC2, and BPGM and

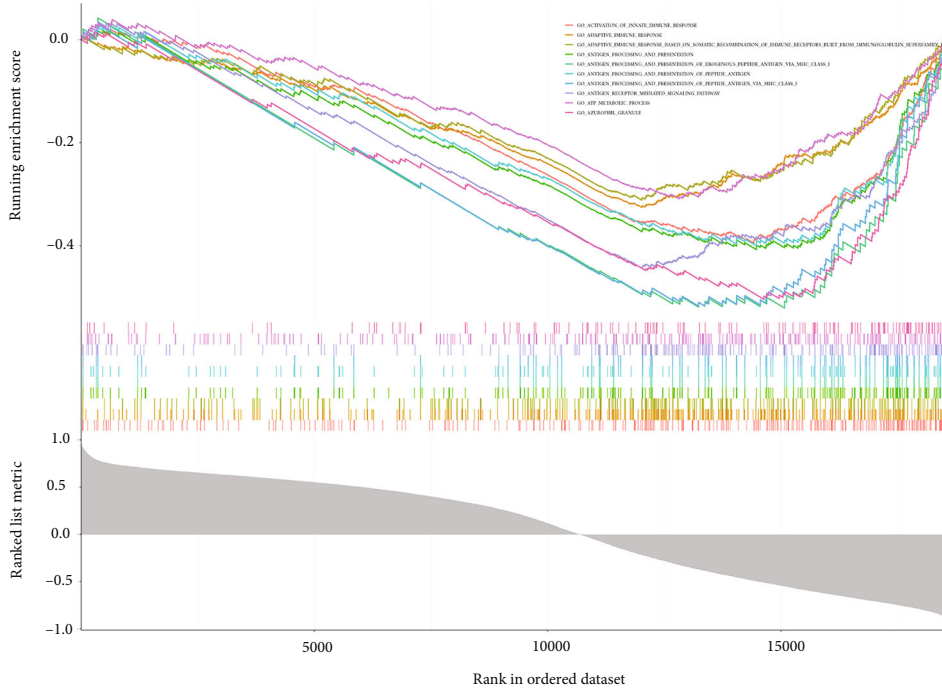


(a)

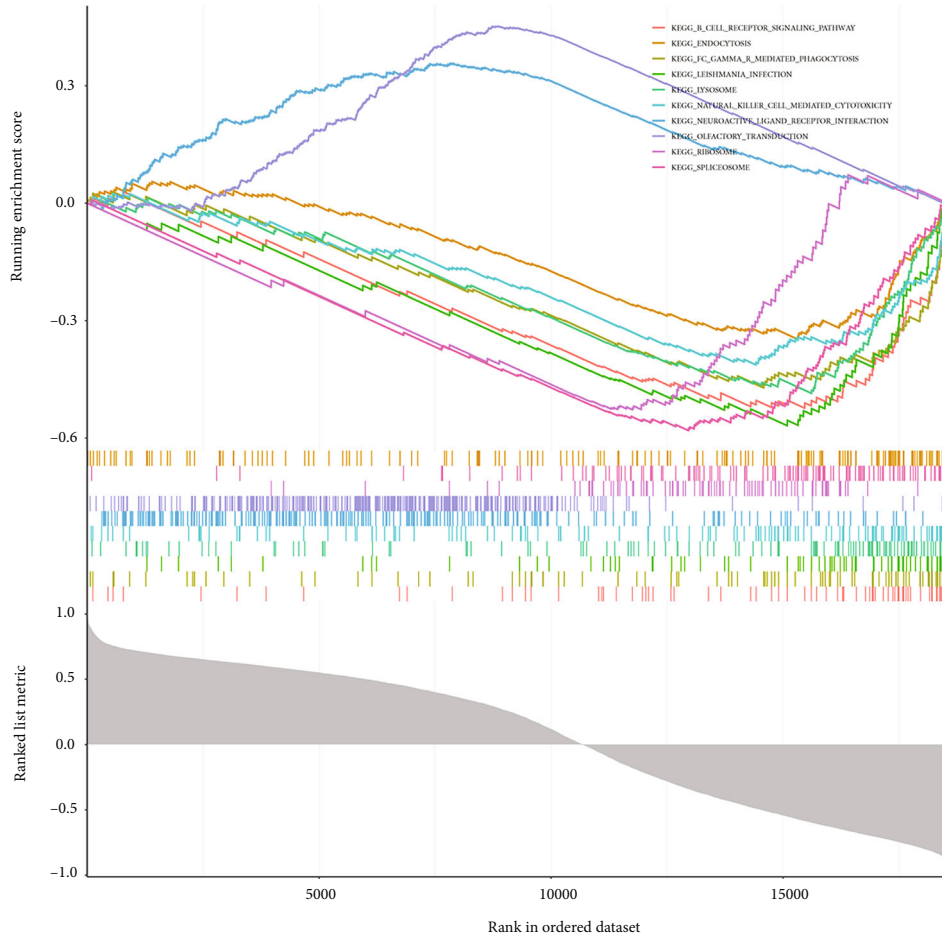


(b)

FIGURE 6: Continued.

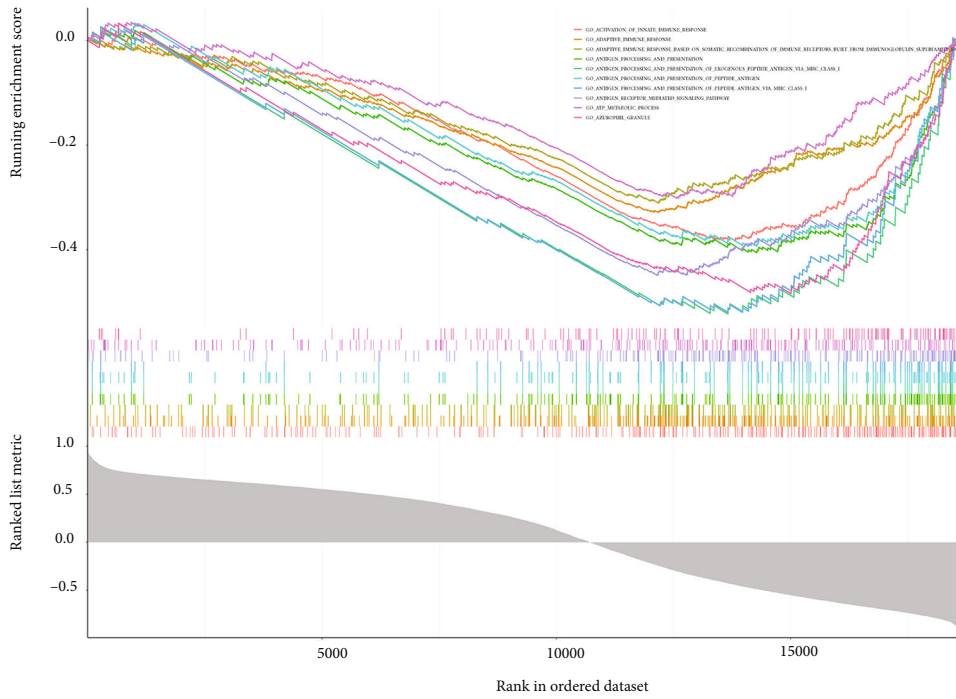


(c)

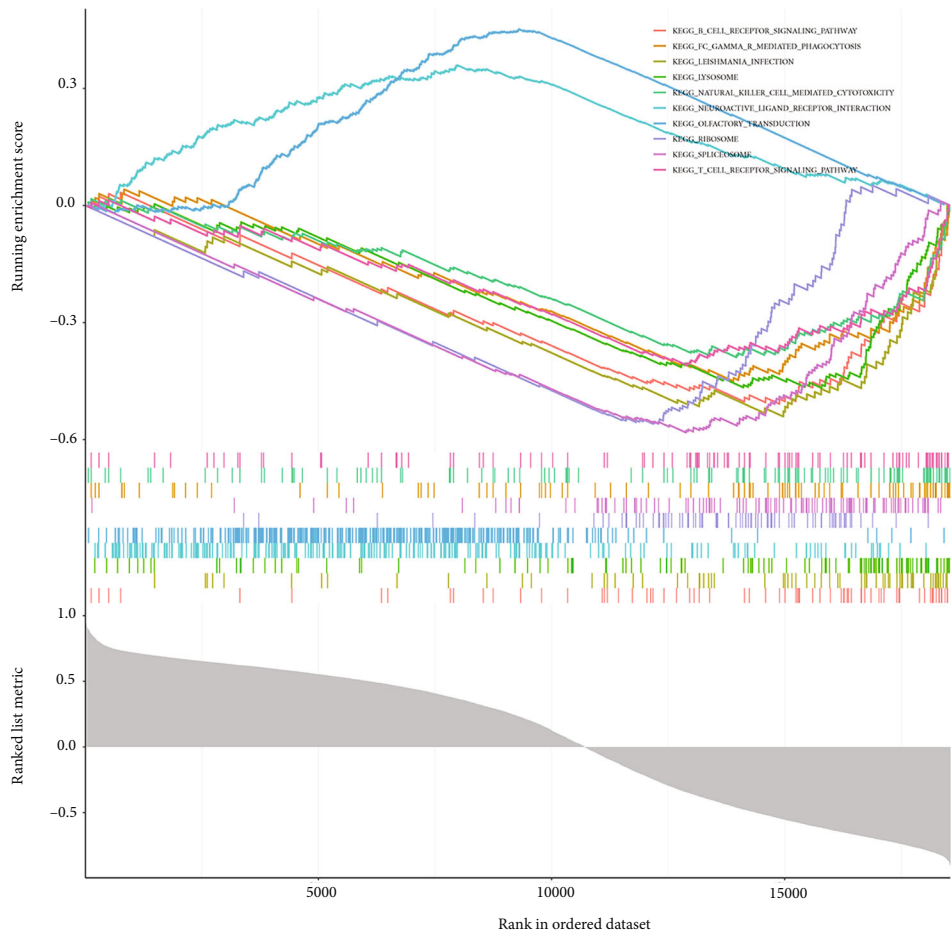


(d)

FIGURE 6: Continued.



(e)



(f)

FIGURE 6: GSEA of potential biomarkers: (a) GO results for BPGM; (b) KEGG results for BPGM; (c) GO results for GYPA; (d) KEGG results for GYPA; (e) GO results for TMCC2; (f) KEGG results for TMCC2.

constructed a ceRNA network (Figure 7(b)). Interestingly, only miRNAs were predicted for GYPA and BPGM, and the corresponding lncRNAs were not predicted (Figure 7(b)).

4. Discussion

SONFH is a complicated disease of the femoral head with complex pathogenesis, including genetic and environmental factors [26]. According to statistics, there are approximately 36,000 to 48,000 new SONFH patients in China each year [27, 28]. Currently, the clinical diagnosis of SONFH is sometimes full of significant difficulties due to the lack of frequent symptoms and valid diagnostic biomarkers, especially in the early stage of SONFH, and ultimately resulting in that most patients miss the best time for treatment and become difficult to reverse [29, 30]. Although total hip replacement is regarded as a proven treatment for femoral head necrosis, patients with SONFH are often too young to undergo total hip replacement. Thus, they still struggle with the psychological and economic burden of a revision surgery [31]. Therefore, identifying potential biomarkers is essential for the diagnosis of SONFH.

In the present study, we first identified 562 DEGs in the peripheral serum obtained from SONFH patients. Next, 94 genes of them were identified as target genes based on WGCNA. Interestingly, these 94 target genes were mostly related to erythrocyte differentiation, homeostasis, and development, myeloid cell homeostasis and development, and porphyrin metabolism-related BPs (Figure 3(a)). Thus, we speculated that these genes may play key roles in SONFH by regulating erythrocytes and myeloid cells. Recent studies have found that erythropoietin can protect rat models of SONFH by inhibiting the apoptosis of osteoblasts and osteocytes and increasing the expression of VEGF [32, 33]. In addition, increasing evidence has revealed that erythropoietin can promote bone repair in SONFH by regulating the hypoxia-inducible factor signaling pathway [34, 35]. Furthermore, erythropoietin also can prevent bone loss in mouse models of osteonecrosis of the femoral head by regulating osteogenesis, angiogenesis, and cell apoptosis [36]. On the other hand, we found that 94 target genes were enriched only in the porphyrin- and chlorophyll metabolism-related signaling pathways (Figure 3(b)). Clearly, metabolism is a key risk factor for nontraumatic SONFH [37, 38]. Therefore, our study may contribute to the understanding of the molecular mechanisms underlying SONFH.

Additionally, we further investigated the protein interactions among 94 target genes and screened 52 genes by constructing a PPI network. Finally, we identified GYPA, TMCC2, and BPGM as potential biomarkers using LASSO logistic regression, SVM-RFE, and RF algorithms. GYPA (Glycophorin A), a major sialoglycoprotein of the human erythrocyte membrane, has been found to be related to blood immunity by regulating the proliferation and differentiation of hematopoietic and immunocompetent cells in human marrow [39]. Thus, GYPA may play a critical role in SONFH by regulating blood immunity and red blood cells. TMCC2 (Transmembrane and Coiled-Coil Domain Family 2) has been revealed to be related to human erythroid

differentiation [40]. Moreover, TMCC2 can affect A β PP metabolism [41] and nitrogen metabolism and excretion [42]. Hence, TMCC2 may play a key role in SONFH by affecting erythroid differentiation and metabolism. BPGM (Bisphosphoglycerate Mutase), a small molecule found at high concentrations in red blood cells where it binds to and decreases the oxygen affinity of hemoglobin, is associated with erythrocyte metabolic reprogramming in chronic kidney disease [43]. More importantly, BPGM is related to erythrocytosis [44, 45] and neutrophilia [46]. Therefore, BPGM may play a decisive role in SONFH by affecting erythroid proliferation and immunity. Notably, no studies have reported the role of GYPA, TMCC2, and BPGM in SONFH. Thus, further investigations are necessary.

Finally, we also investigated the biological functions and regulatory mechanisms of GYPA, TMCC2, and BPGM. Interestingly, GSEA revealed that GYPA, TMCC2, and BPGM were mainly involved in immune-related BPs and signaling pathways, such as T- and B-cell receptor signaling pathways. Currently, an increasing number of studies have shown that the immune response has a significant impact on the occurrence and development of SONFH. For example, it has been suggested that macrophages and CD4⁺ T cells are associated with SONFH [47, 48]. Therefore, GYPA, TMCC2, and BPGM may play key roles in SONFH by regulating the immunity. Moreover, we also explored the regulatory mechanisms of GYPA, TMCC2, and BPGM and found that GYPA might be regulated by hsa-miR-3137, and BPGM might be regulated by hsa-miR-340-3p. However, to the best of our knowledge, their regulatory mechanisms are rarely studied. Thus, further studies are required in the future.

5. Conclusion

In summary, 562 DEGs were screened between the peripheral serum of SONFH patients and control samples. Moreover, GYPA, TMCC2, and BPGM were identified as potential biomarkers in the peripheral serum of SONFH patients based on WGCNA and machine learning algorithms. Furthermore, we found that GYPA might be regulated by hsa-miR-3137 and that BPGM might be regulated by hsa-miR-340-3p. Therefore, our study may contribute to the understanding of SONFH and may help in improving the diagnosis of SONFH. However, further studies are needed to investigate the roles of GYPA, TMCC2, and BPGM.

Data Availability

The GSE123568 dataset was obtained from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>).

Conflicts of Interest

The authors declare they have no competing interests.

Authors' Contributions

JYL designed and conceived this research. JZ and CH acquired the data, performed the data analysis, and wrote

the draft. ZHL, SR, ZLS, and KCH helped in the data analysis. WGX, GYH, and CH modified the manuscript. All authors reviewed and approved the manuscript for publication. Jian Zhang and Chi Huang contributed equally to this work.

Acknowledgments

The work was supported by the National Science Foundation of China (81971828).

References

- [1] M. A. Mont, J. J. Cherian, R. J. Sierra, L. C. Jones, and J. R. Lieberman, "Nontraumatic osteonecrosis of the femoral head: where do we stand today? A ten-year update," *The Journal of Bone and Joint Surgery. American Volume*, vol. 97, no. 19, pp. 1604–1627, 2015.
- [2] J. R. Lieberman, D. J. Berry, M. A. Mont et al., "Osteonecrosis of the hip: management in the 21st century," *Instructional Course Lectures*, vol. 52, pp. 337–355, 2003.
- [3] Covid-19 National Incident Room Surveillance Team, "COVID-19 Australia: epidemiology report 34: reporting period ending 31 January 2021," *Communicable diseases intelligence*, vol. 45, 2021.
- [4] J. Moya-Angeler, A. L. Gianakos, J. C. Villa, A. Ni, and J. M. Lane, "Current concepts on osteonecrosis of the femoral head," *World Journal of Orthopedics*, vol. 6, no. 8, pp. 590–601, 2015.
- [5] K. Issa, J. J. Jauregui, M. McElroy, S. Banerjee, B. H. Kapadia, and M. A. Mont, "Unnecessary magnetic resonance imaging of hips: an economic burden to patients and the healthcare system," *The Journal of Arthroplasty*, vol. 29, no. 10, pp. 1911–1914, 2014.
- [6] H. Y. Zhu, Y. C. Gao, Y. Wang, and C. Q. Zhang, "Circulating exosome levels in the diagnosis of steroid-induced osteonecrosis of the femoral head," *Bone & Joint Research*, vol. 5, no. 6, pp. 276–279, 2016.
- [7] S. K. Tripathy, T. Goyal, and R. K. Sen, "Management of femoral head osteonecrosis: current concepts," *Indian journal of orthopaedics*, vol. 49, pp. 28–45, 2015.
- [8] X. Huang, S. Liu, L. Wu, M. Jiang, and Y. Hou, "High throughput single cell RNA sequencing, bioinformatics analysis and applications," *Advances in Experimental Medicine and Biology*, vol. 1068, pp. 33–43, 2018.
- [9] T. Wang, X. Zheng, R. Li et al., "Integrated bioinformatic analysis reveals YWHAB as a novel diagnostic biomarker for idiopathic pulmonary arterial hypertension," *Journal of Cellular Physiology*, vol. 234, no. 5, pp. 6449–6462, 2019.
- [10] Y. Yao, S. Xie, and F. Wang, "Identification of key genes and pathways in chronic rhinosinusitis with nasal polyps using bioinformatics analysis," *American Journal of Otolaryngology*, vol. 40, pp. 191–196, 2019.
- [11] G. Xue, L. Hua, N. Zhou, and J. Li, "Characteristics of immune cell infiltration and associated diagnostic biomarkers in ulcerative colitis: results from bioinformatics analysis," *Bioengineered*, vol. 12, no. 1, pp. 252–265, 2021.
- [12] Y. Chen, R. Liao, Y. Yao, Q. Wang, and L. Fu, "Machine learning to identify immune-related biomarkers of rheumatoid arthritis based on WGCNA network," *Clinical rheumatology*, 2021.
- [13] Y. X. Zhu, J. Q. Huang, Y. Y. Ming, Z. Zhuang, and H. Xia, "Screening of key biomarkers of tendinopathy based on bioinformatics and machine learning algorithms," *PLoS One*, vol. 16, no. 10, article e0259475, 2021.
- [14] R. Sun, S. Li, K. Zhao, M. Diao, and L. Li, "Identification of ten core hub genes as potential biomarkers and treatment target for hepatoblastoma," *Frontiers in Oncology*, vol. 11, article 591507, 2021.
- [15] K. Chai, J. Liang, X. Zhang et al., "Application of machine learning and weighted gene co-expression network algorithm to explore the hub genes in the aging brain," *Frontiers in Aging Neuroscience*, vol. 13, article 707165, 2021.
- [16] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, p. e47, 2015.
- [17] P. Langfelder and S. Horvath, "WGCNA: an R package for weighted correlation network analysis," *BMC Bioinformatics*, vol. 9, no. 1, p. 559, 2008.
- [18] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.
- [19] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, pp. 267–288, 1996.
- [20] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.
- [21] X. Robin, N. Turck, A. Hainard et al., "pROC: an open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinformatics*, vol. 12, p. 77, 2011.
- [22] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [23] M.-L. Huang, Y.-H. Hung, W. M. Lee, R. K. Li, and B.-R. Jiang, "SVM-RFE based feature selection and Taguchi parameters optimization for multiclass SVM classifier," *Scientific World Journal*, vol. 2014, article 795624, pp. 1–10, 2014.
- [24] L. Breiman, *Machine Learning*, Springer, 2001.
- [25] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News*, vol. 23, 2002.
- [26] C. Chang, A. Greenspan, and M. E. Gershwin, "The pathogenesis, diagnosis and clinical manifestations of steroid-induced osteonecrosis," *Journal of Autoimmunity*, vol. 110, article 102460, 2020.
- [27] Y. Song, Z. du, M. Ren et al., "Association of gene variants of transcription factors PPAR γ , RUNX2, Osterix genes and COL2A1, IGFBP3 genes with the development of osteonecrosis of the femoral head in Chinese population," *Bone*, vol. 101, pp. 104–112, 2017.
- [28] Y. Tian, F. An, J. Wang et al., "MMP2 and MMP10 polymorphisms are related to steroid-induced osteonecrosis of the femoral head among Chinese Han population," *Biomed Research International*, vol. 2019, Article ID 8298193, 2019.
- [29] W. Fu, B. Liu, B. Wang, and D. Zhao, "Early diagnosis and treatment of steroid-induced osteonecrosis of the femoral head," *International Orthopaedics*, vol. 43, pp. 1083–1087, 2019.
- [30] Y. Chen, C. Zeng, H. Zeng et al., "Comparative serum proteome expression of the steroid-induced femoral head osteonecrosis in adults," *Experimental and Therapeutic Medicine*, vol. 9, no. 1, pp. 77–83, 2015.

- [31] J. T. Best, "Revision total hip and total knee arthroplasty," *Orthopaedic Nursing*, vol. 24, no. 3, pp. 174–179, 2005.
- [32] L. Y. Jiang, X. Yu, and Q. J. Pang, "Research in the precaution of recombinant human erythropoietin to steroid-induced osteonecrosis of the rat femoral head," *The Journal of International Medical Research*, vol. 45, no. 4, pp. 1324–1331, 2017.
- [33] S. Chen, J. Li, H. Peng, J. Zhou, and H. Fang, "Administration of erythropoietin exerts protective effects against glucocorticoid-induced osteonecrosis of the femoral head in rats," *International Journal of Molecular Medicine*, vol. 33, no. 4, pp. 840–848, 2014.
- [34] D. Li, Q. Hu, G. Tan, X. Xie, Z. Yang, and P. Kang, "Erythropoietin enhances bone repair effects via the hypoxia-inducible factor signal pathway in glucocorticoid-induced osteonecrosis of the femoral head," *The American Journal of the Medical Sciences*, vol. 355, no. 6, pp. 597–606, 2018.
- [35] D. Li, X. Xie, Z. Yang, C. Wang, Z. Wei, and P. Kang, "Enhanced bone defect repairing effects in glucocorticoid-induced osteonecrosis of the femoral head using a porous nano-lithium-hydroxyapatite/gelatin microsphere/erythropoietin composite scaffold," *Biomaterials Science*, vol. 6, no. 3, pp. 519–537, 2018.
- [36] T. Xu, H. Jin, Y. Lao et al., "Administration of erythropoietin prevents bone loss in osteonecrosis of the femoral head in mice," *Molecular Medicine Reports*, vol. 16, no. 6, pp. 8755–8762, 2017.
- [37] T. Kuroda, N. Tanabe, A. Wakamatsu et al., "High triglyceride is a risk factor for silent osteonecrosis of the femoral head in systemic lupus erythematosus," *Clinical Rheumatology*, vol. 34, no. 12, pp. 2071–2077, 2015.
- [38] Y. Zhang, R. Sun, L. Zhang, L. Feng, and Y. Liu, "Effect of blood biochemical factors on nontraumatic necrosis of the femoral head: logistic regression analysis," *Der Orthopäde*, vol. 46, no. 9, pp. 737–743, 2017.
- [39] S. V. Sennikov, T. V. Injelevskaya, S. V. Krysov et al., "Production of hemo- and immunoregulatory cytokines by erythroblast antigen+ and glycophorin A+ cells from human bone marrow," *BMC Cell Biology*, vol. 5, no. 1, p. 39, 2004.
- [40] L. S. Ludwig, C. A. Lareau, E. L. Bao et al., "Transcriptional states and chromatin accessibility underlying human erythropoiesis," *Cell Reports*, vol. 27, no. 11, pp. 3228–3240.e7, 2019.
- [41] P. C. Hopkins, R. Sáinz-Fuertes, and S. Lovestone, "The impact of a novel apolipoprotein E and amyloid- β protein precursor-interacting protein on the production of amyloid- β ," *Journal of Alzheimer's Disease*, vol. 26, no. 2, pp. 239–253, 2011.
- [42] H. Honerlagen, H. Reyer, M. Oster et al., "Identification of genomic regions influencing N-metabolism and N-excretion in lactating Holstein-Friesians," *Frontiers in Genetics*, vol. 12, article 699550, 2021.
- [43] T. Xie, C. Chen, Z. Peng et al., "Erythrocyte metabolic reprogramming by sphingosine 1-phosphate in chronic kidney disease and therapies," *Circulation Research*, vol. 127, no. 3, pp. 360–375, 2020.
- [44] N. Petousi, R. R. Copley, T. R. Lappin et al., "Erythrocytosis associated with a novel missense mutation in the BPGM gene," *Haematologica*, vol. 99, no. 10, pp. e201–e204, 2014.
- [45] I. Lazana, A. Mohamedali, F. Smith, H. de Lavallade, D. McLornan, and K. Raj, "Uniparental disomy (UPD) of a novel bisphosphoglycerate mutase (BPGM) mutation leading to erythrocytosis," *British Journal of Haematology*, vol. 192, no. 1, pp. 220–223, 2021.
- [46] W. J. Hong and J. Gotlib, "Hereditary erythrocytosis, thrombocytosis and neutrophilia," *Best Practice & Research. Clinical Haematology*, vol. 27, no. 2, pp. 95–106, 2014.
- [47] D. Kamal, R. O. Trăistaru, C. K. Kamal, D. O. Alexandru, D. A. Ion, and D. C. Grecu, "Macrophage response in patients diagnosed with aseptic necrosis of the femoral head presenting different risk factors," *Romanian Journal of Morphology and Embryology*, vol. 56, pp. 163–168, 2015.
- [48] B. J. Rabquer, G. J. Tan, P. J. Shaheen, G. K. Haines III, A. G. Urquhart, and A. E. Koch, "Synovial inflammation in patients with osteonecrosis of the femoral head," *Clinical and Translational Science*, vol. 2, no. 4, pp. 273–278, 2009.