



OPEN ACCESS

EDITED BY

Chenyu Sun,
AMITA Health, United States

REVIEWED BY

Wenzhao Wang,
Shandong University, China
Chen Xinxi,
Jilin University, China

*CORRESPONDENCE

Ming Lu

✉ dlmlm@2008.sina.com

Xueling Qu

✉ quxuelingmm@sina.com

Shouyu Wang

✉ wangshouyu666@126.com

Junwei Zong

✉ aweizone@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 19 March 2023

ACCEPTED 21 June 2023

PUBLISHED 14 August 2023

CITATION

Wang X, Meng L, Zhang J, Zhao Z, Zou L, Jia Z, Han X, Zhao L, Song M, Zong J, Wang S, Qu X and Lu M (2023)

Identification of ferroptosis-related molecular clusters and genes for diabetic osteoporosis based on the machine learning.

Front. Endocrinol. 14:1189513.

doi: 10.3389/fendo.2023.1189513

COPYRIGHT

© 2023 Wang, Meng, Zhang, Zhao, Zou, Jia, Han, Zhao, Song, Zong, Wang, Qu and Lu.

This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of ferroptosis-related molecular clusters and genes for diabetic osteoporosis based on the machine learning

Xingkai Wang^{1,2†}, Lei Meng^{3†}, Juwei Zhang^{4†}, Zitong Zhao⁵, Linxuan Zou², Zhuqiang Jia^{6,7}, Xin Han^{7,8}, Lin Zhao⁹, Mingzhi Song², Junwei Zong^{2*}, Shouyu Wang^{2*}, Xueling Qu^{10*} and Ming Lu^{1*}

¹Department of Trauma and Tissue Repair Surgery, Dalian Municipal Central Hospital, Dalian, China,

²Department of Orthopaedic Surgery, The First Affiliated Hospital of Dalian Medical University, Dalian, China,

³Department of Surgery, The First Affiliated Hospital of Nanhua Medical University, Hengyang, China, ⁴Health Inspection and Quarantine, College of Medical Laboratory, Dalian Medical University, Dalian, China, ⁵International Department, Beijing No.80 High School, Beijing, China,

⁶Department of Surgery, The First Affiliated Hospital of Dalian Medical University, Dalian, China,

⁷Department of Surgery, Naqu People's Hospital, Tibet, China, ⁸Department of Orthopaedic Surgery, The Second Affiliated Hospital of Dalian Medical University, Dalian, China,

⁹Department of Quality Management, Dalian Municipal Central Hospital, Dalian, China, ¹⁰Changjianglu Pelvic Floor Repair Center, Dalian Women and Children's Medical Group, Dalian, China

Background: Diabetic osteoporosis exhibits heterogeneity at the molecular level. Ferroptosis, a controlled form of cell death brought on by a buildup of lipid peroxidation, contributes to the onset and development of several illnesses. The aim was to explore the molecular subtypes associated with ferroptosis in diabetic osteoporosis at the molecular level and to further elucidate the potential molecular mechanisms.

Methods: Integrating the CTD, GeneCards, FerrDb databases, and the microarray data of GSE35958, we identified ferroptosis-related genes (FRGs) associated with diabetic osteoporosis. We applied unsupervised cluster analysis to divide the 42 osteoporosis samples from the GSE56814 microarray data into different subclusters based on FRGs. Subsequently, FRGs associated with two ferroptosis subclusters were obtained by combining database genes, module-related genes of WGCNA, and differentially expressed genes (DEGs). Eventually, the key genes from FRGs associated with diabetic osteoporosis were identified using the least absolute shrinkage and selection operator (LASSO), Boruta, support vector machine recursive feature elimination (SVM RFE), and extreme gradient boosting (XGBoost) machine learning algorithms. Based on ROC curves of external datasets (GSE56815), the model's efficiency was examined.

Results: We identified 15 differentially expressed FRGs associated with diabetic osteoporosis. In osteoporosis, two distinct molecular clusters related to ferroptosis were found. The expression results and GSVA analysis indicated that 15 FRGs exhibited significantly different biological functions and pathway activities in the two ferroptosis subclusters. Therefore, we further identified 17 FRGs associated with diabetic osteoporosis between the two subclusters. The results of the comprehensive analysis of 17 FRGs demonstrated that these genes were heterogeneous and had a specific interaction between the two subclusters.

Ultimately, the prediction model had a strong foundation and excellent AUC values (0.84 for LASSO, 0.84 for SVM RFE, 0.82 for Boruta, and 0.81 for XGBoost). IDH1 is a common gene to all four algorithms thus being identified as a key gene with a high AUC value (AUC = 0.698).

Conclusions: As a ferroptosis regulator, IDH1 is able to distinguish between distinct molecular subtypes of diabetic osteoporosis, which may offer fresh perspectives on the pathogenesis of the disease's clinical symptoms and prognostic heterogeneity.

KEYWORDS

diabetic osteoporosis, ferroptosis, molecular clusters, machine learning, prediction model

1 Introduction

According to the World Health Organization, diabetes prevalence rates have been rising over the past few decades. Currently, more than 463 million people worldwide have diabetes, and by 2040, that figure is projected to double (1). Not only is diabetes dangerous in its own right, but its numerous complications also have a severe impact on patients. Diabetic osteoporosis, as a chronic complication of the skeletal system, can lead to major pain and skeletal deformities, as well as high levels of disability and mortality, making the treatment and rehabilitation of diabetic patients more difficult, not only affecting the quality of life but also increasing the financial burden (2, 3). Unfortunately, less attention has been paid to osteoporosis than to the complications traditionally associated with diabetes such as macrovascular disease and microvascular disease. Consequently, further exploration of the pathogenesis and prevention of diabetic osteoporosis is warranted.

In contrast to apoptosis, necroptosis, and autophagy, ferroptosis is a unique kind of iron-dependent programmed cell death that is characterized by an excessive buildup of reactive oxygen species and lipid peroxidation (4). Ferroptosis has been shown to be associated with the pathophysiological mechanisms underlying several illnesses and contribute to the onset and progression of such diseases. A growing amount of research is now confirming that ferroptosis may be a valuable research direction for the prevention and treatment of osteoporosis (5–7). Nevertheless, the molecular subtypes and processes relating to ferroptosis in diabetic osteoporosis need to be elucidated. Thus, comprehending the connection between FRGs and the development of diabetes osteoporosis and locating molecular clusters based on FRGs helps to understand the molecular heterogeneity of diabetic osteoporosis.

In the current work, we thoroughly identified and in-depth investigated FRGs that differed in expression between healthy people and patients with diabetic osteoporosis. 42 osteoporosis patients were separated into two ferroptosis-associated subclusters based on 15 distinct ferroptosis gene expression profiles linked to diabetic osteoporosis, and the molecular interactions between the

two clusters were further investigated. In addition, utilizing multiple machine learning methods, a robust prediction model was established and validated by employing an external dataset. The receiver operating characteristic (ROC) curve results demonstrated that FRGs could be a prospective predictor of diabetic osteoporosis subtypes. Figure 1 depicts the research flowchart for this study.

2 Materials and methods

2.1 Data source and processing

Through applying the R package of GEOquery, three microarray datasets (GSE35958, GSE56814, and GSE56815) related to osteoporosis were downloaded from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>) (8). The detailed sample information for the three datasets is presented in Table 1. Taking $\text{adj. } P\text{-Value} < 0.05$ and $|\log_2\text{fold change (FC)}| \geq 1$ as the cutoff value, the GSE35958 dataset was selected to be used for differential gene expression analysis, which is based on the limma package in the R software (9). We then screen diabetes-related genes from two disease databases, including Comparative Toxicogenomics Database (CTD) (<http://ctdbase.org/>) (10), and GeneCards (<https://www.genecards.org/>) (11). In addition, ferroptosis-related genes (FRGs) were obtained from the FerrDb database (<http://www.zhounan.org/ferrdb/>) (12). Finally, 15 FRGs associated with diabetic osteoporosis were obtained as the differentially expressed genes (DEGs) between control and osteoporosis samples, and a comprehensive analysis was performed.

2.2 Unsupervised clustering of FRGs associated with diabetic osteoporosis

To further divide the osteoporosis into different subclusters based on the expression level of the 15 FRGs associated with diabetic osteoporosis, an unsupervised cluster analysis was

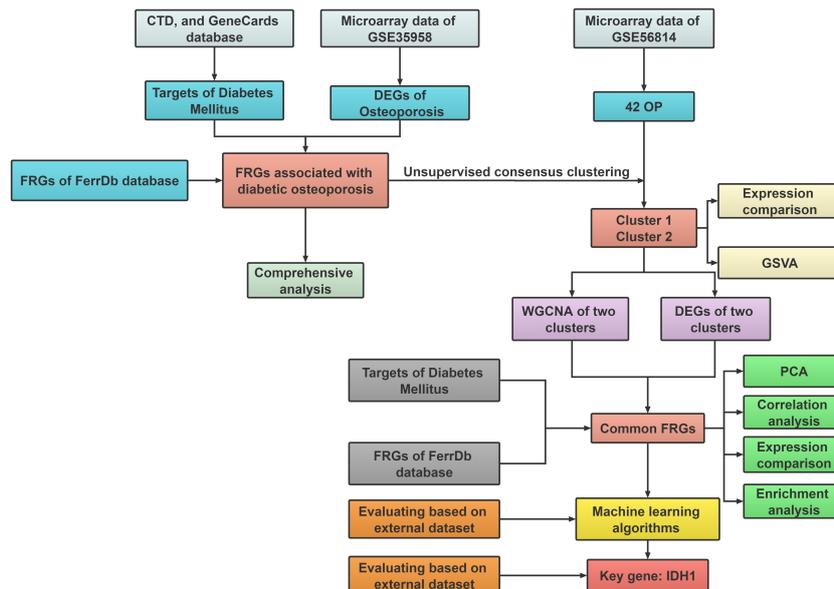


FIGURE 1 The flow chart of the study design and analysis: DEGs analysis and subcluster identification; Comprehensive analysis; Machine learning and validation.

TABLE 1 The detailed sample information for the three datasets.

Dataset	Platform	Count	Osteoporosis	Control
GSE35958	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	9	5	4
GSE56814	GPL5175 [HuEx-1_0-st] Affymetrix Human Exon 1.0 ST Array [transcript (gene) version]	82	42	40
GSE56815	GPL96 [HG-U133A] Affymetrix Human Genome U133A Array	80	40	40

performed to distinguish the 42 osteoporosis samples in the GSE56814 dataset into distinct clusters by applying the R package of ConsensusClusterPlus (13). The ideal cluster number was confirmed using the CDF (cumulative distribution function) and the area under the CDF curve.

2.3 Expression and gene set variation analysis

We observed the expression of 15 FRGs associated with diabetic osteoporosis in different subclusters by drawing the heatmap and violin plot. We used GSEA, an unsupervised, non-parametric algorithm (14), to assess the corresponding biological characteristics and pathway activity of different ferroptosis subclusters of osteoporosis. The “c2.cp.kegg.v7.4.symbols.gmt” and “c5.go.v2022.1.Hs.symbols.gmt” gene sets were utilized for the GSEA enrichment analysis, and biological functions and pathways were judged to be substantially enriched when the adjusted p-value was less than 0.05.

2.4 Weighted gene co-expression network analysis

We applied the R package of WGCNA to construct a co-expression network of all genes in 42 osteoporosis samples (15). Firstly, the gene expression matrix is loaded in the R software to check for missing values and identify outliers. Secondly, we construct a scale-free network to select a soft threshold value, which is considered to be the parameter cutoff value for the construction of the adjacency matrix. A network relationship is usually defined by an adjacency matrix. Thirdly, we transformed the adjacency matrix into a topology matrix allowing similarities between genes to be represented at the expression and network topology levels. To discover gene co-expression modules, block module function, and module division analyses were lastly carried out. The association between each module and osteoporosis was determined, and the most pertinent modules were sorted based on the Pearson correlation analysis results. The genes in these modules were thought to be osteoporosis-related module genes.

2.5 Identifying DEGs of ferroptosis subclusters

The R package of limma in Bioconductor was used to identify DEGs by comparing the expression values between different ferroptosis subclusters. The criteria were $\text{adj. P-Value} < 0.05$ and $|\log_2\text{fold change (FC)}| \geq 1$, and genes fulfilling this condition were determined to be DEGs. The asymptotic volcano map and heatmap displaying the DEGs were created using the ggplot2 and pheatmap package, respectively.

2.6 Comprehensive analysis of FRGs associated with diabetic osteoporosis of ferroptosis subclusters

By integrating diabetes-related genes from the CTD and GeneCards databases, FRGs from the FerrDb database, DEGs, and osteoporosis-related module genes between ferroptosis subclusters, altogether 17 FRGs associated with diabetic osteoporosis of ferroptosis subclusters were obtained. We then evaluate the efficiency of 17 genes using principal component analysis (PCA). The ggcorrplot package and the ggplot2 package were applied for the correlation analysis of 17 genes and expression analysis in different subclusters. Enrichment analysis of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis were executed using Metascape (<https://metascape.org/gp/index.html>), which is a public database of gene annotation and analysis resources (16).

2.7 Robust prediction model created with a variety of machine learning techniques

We applied the R packages “glmnet,” “caret,” “Boruta” and “XGBoost” to build a machine learning model (17). The entire dataset of 17 genes was subjected to the least absolute shrinkage and selection operator (LASSO), Boruta, Support Vector Machine Recursive Feature Elimination (SVM RFE), and extreme gradient boosting (XGBoost) analyses to identify key genes that belong to four predictive models. The 40 osteoporosis samples in GSE56815 were similarly divided into two different subclusters by 15 FRGs associated with diabetic osteoporosis and used as an external dataset to validate the efficiency of the four predictive models and key genes by the ROC curves.

3 Results

3.1 Identification of FRGs specifically expressed in diabetic osteoporosis

The expression profile data of 4 control samples and 5 osteoporosis samples in GSE35958 were normalized using the “limma” package (Figures 2A, B). Based on the $\text{adj. P-Value} < 0.05$ and $|\log_2\text{fold change (FC)}| \geq 1$, a total of 1102 DEGs, including

677 up-regulated and 425 down-regulated genes, were identified by differential analysis (Figure 2C). We obtained 3,8253 and 1,4818 diabetes-related genes from the CTD database and the GeneCards database, respectively. The 259 ferroptosis-related genes (FRGs) were obtained from the FerrDb database. We overlapped the DEGs, diabetes-related genes, and ferroptosis-related genes, 15 overlapped genes were obtained, namely FRGs associated with diabetic osteoporosis, which was shown by the Venn diagram (Figure 2D). The details of these genes were shown in Supplementary Table S1. We can observe that the 15 FRGs associated with diabetic osteoporosis are significantly differentially expressed between the osteoporotic samples and the control samples by drawing a heatmap (Figure 2E). To clarify the relationship between the 15 FRGs, Spearman correlation analysis was employed (Figure 2F). In addition, the localization of the 15 FRGs on the chromosome is shown in the loop graph (Figure 2G).

3.2 Identification of ferroptosis subclusters in osteoporosis

We explored the ferroptosis subclusters in osteoporosis using unsupervised clustering to analyze the expression of 15 FRGs associated with diabetic osteoporosis in 42 osteoporotic samples. The number of subtypes is most stable when $k = 2$ of the consensus matrix, representing the two well-defined clusters (Figure 3A). As shown in Figure 3B, the CDF curve for $k = 2$ has minimal fluctuations in the consistency index range of 0-1.0. The CDF diagram showed the relative change in area for variable values of k (Figure 3C). Principal component analysis (PCA) further supported the finding that the two clusters differed considerably (Figure 3D).

3.3 Differences between ferroptosis subclusters

To better understand the distinctions between the two ferroptosis subclusters, we analyzed the expression differences and variations in the pathway and biological activity of the 15 FRGs in the two subclusters. The expression of 15 FRGs between the two subclusters was apparently distinguished from that in the control and osteoporotic samples (Figure 4A). Cluster1 exhibited higher expression levels of FBXW7, G6PD, MAPK3, PML PGD, SLC1A5, SQSTM1, TP53, and YWHAE, while Cluster1 exhibited higher expression levels of ALOX5, BAP1, BRD4, CDKN1A, EGFR and NNMT (Figure 4B). Biological function results from GSVA analysis showed that nuclear protein containing complex, response to wounding, and wound healing were downregulated in Cluster1, while neurogenesis, cellular response to biotic stimulus, and regulation of telomerase activity were upregulated in Cluster2 (Figure 4C). In addition, the enrichment pathway of Cluster1 is mainly upregulated, such as in colorectal cancer, thyroid cancer, and small cell lung cancer, while Cluster2 is mainly associated with the downregulation pathway, like basal cell carcinoma, huntingtons disease, and amyotrophic lateral sclerosis als (Figure 4D). The

details of the GSVA analysis were shown in [Supplementary Table S2](#). These findings showed that there are notable differences between ferroptosis clusters of osteoporosis patients in terms of the expression, enriched pathways, and biological roles of 15 FRGs. For various subclusters of ferroptosis, specific therapeutic approaches are needed.

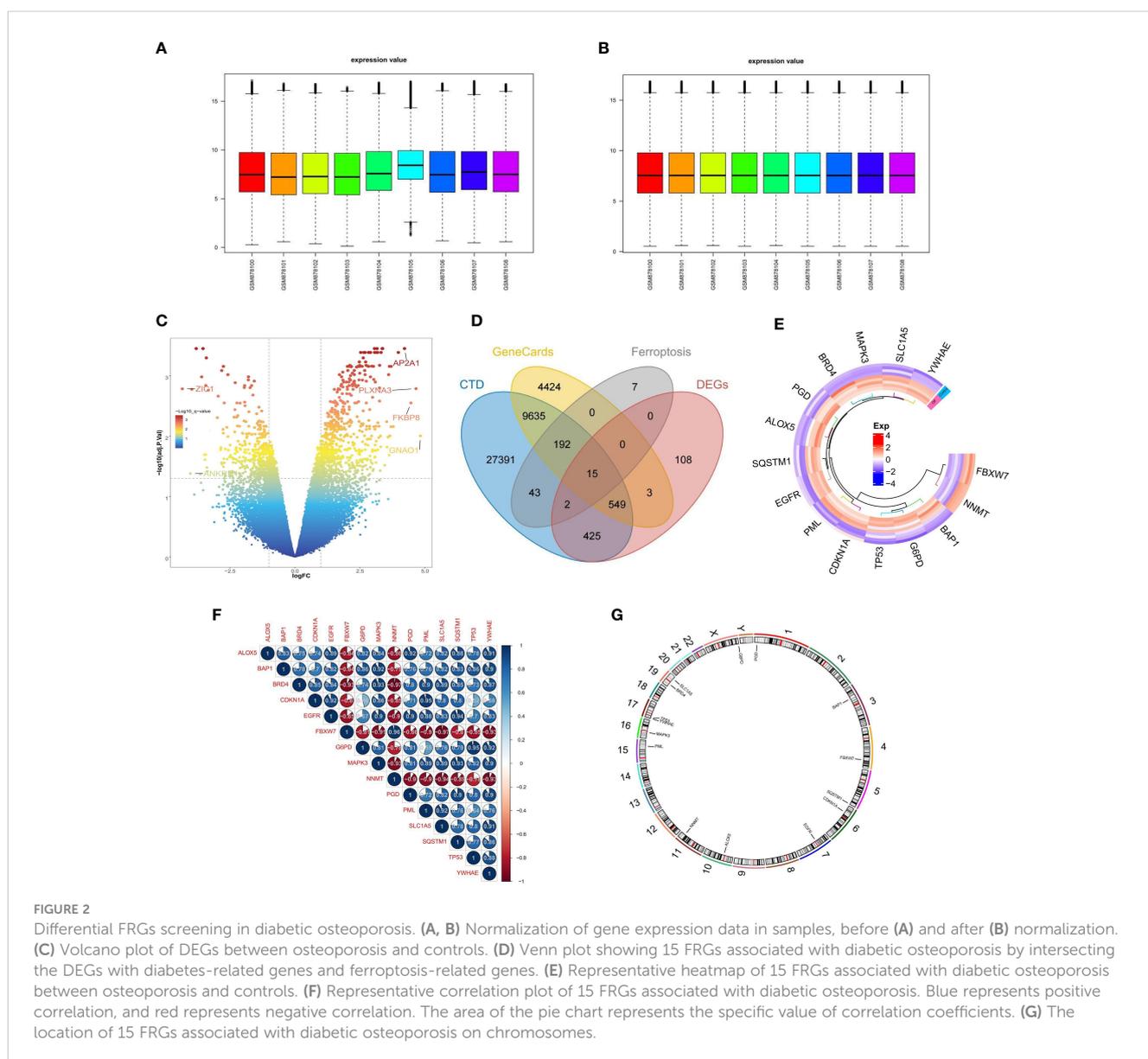
3.4 Differential genes analysis between ferroptosis clusters

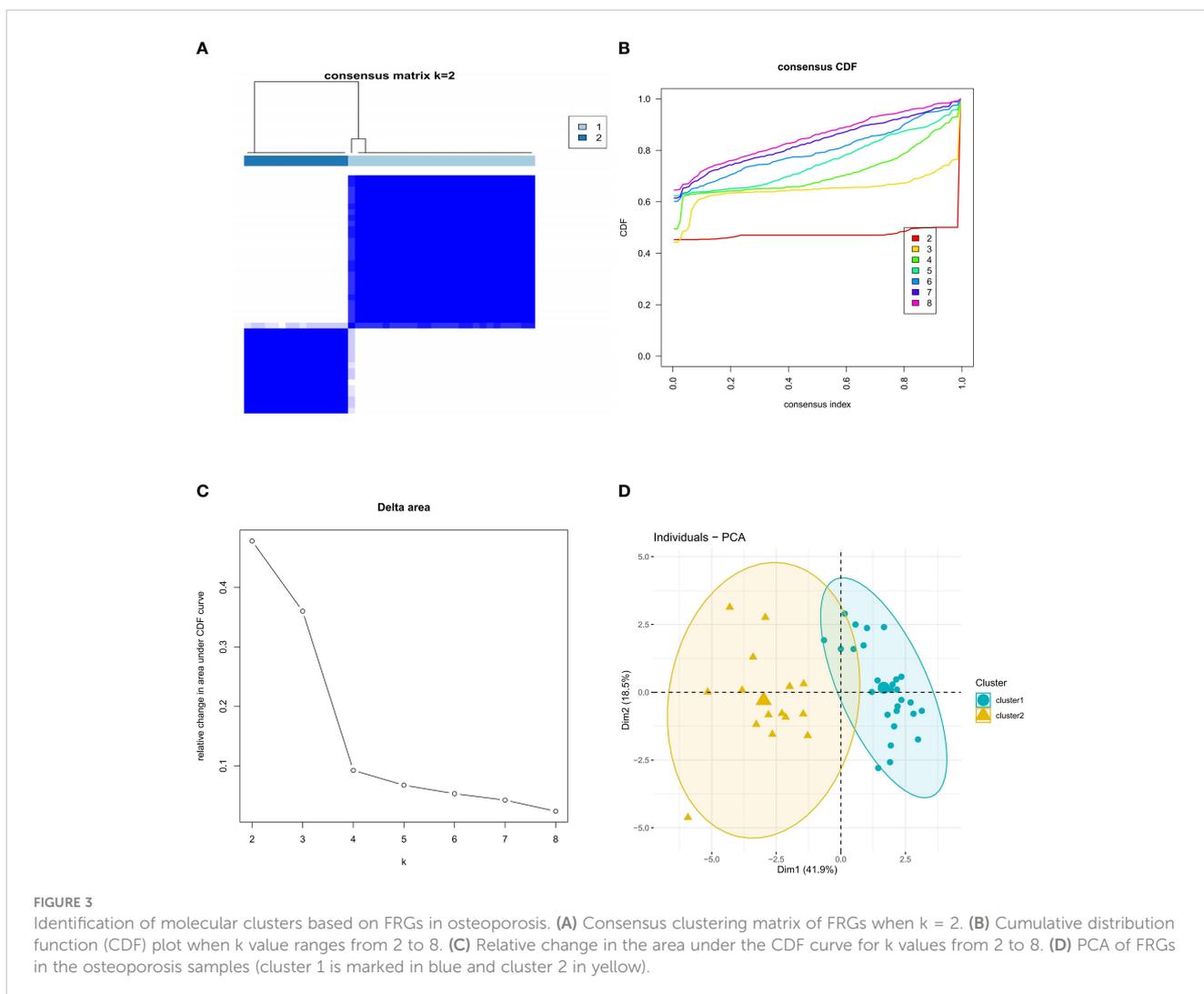
The possible modules with the strongest connections to the ferroptosis subclusters were constructed by the WGCNA algorithm, based on the gene expression profiles. As shown in [Figure 5A](#), the sample with the serial number GSM1369716 was excluded as an outlier. The ideal soft threshold for maintaining a network with scale-free topology was determined to be 6 ($R2 = 0.85$) ([Figure 5B](#)). Based on correlation clustering, the 15 signature modules were

categorized and given various color labels ([Figure 5C](#)). The blue module (4,591 genes) had the strongest connection with Cluster1 ($R = -0.89$) and Cluster2 ($R = -0.89$) among these modules ([Figure 5D](#)). We observed a significant correlation between the blue module and the module-related genes ($cor = 0.91$) ([Figure 5E](#)). Subsequently, we identified the DEGs of the ferroptosis subclusters using $adj. P\text{-Value} < 0.05$ and $|\log_2\text{fold change (FC)}| \geq 1$ as the cutoff value. 1,376 DEGs in total were found; 265 of them showed up-regulation, whereas 1,111 showed down-regulation ([Figures 5F–H](#)).

3.5 Comprehensive analysis of FRGs associated with diabetic osteoporosis of ferroptosis subclusters

We obtained a total of 17 FRGs associated with diabetic osteoporosis of ferroptosis subclusters, by combining genes from





the database and dataset (Figure 6A). The details of genes were shown in Supplementary Table S3. The PCA results revealed 17 FRGs associated with diabetic osteoporosis which effectively distinguish between the two ferroptosis subclusters (Figure 6B). In addition, the relationship network diagram of the 17 FRGs associated with diabetic osteoporosis showed a significant positive correlation with each other, contributing to a comprehensive analysis of the interrelationship among the genes (Figure 6C). Meanwhile, all genes except BNIP3 were highly expressed in cluster1, as shown in Figure 6D. GO and KEGG enrichment analyses were carried out to further investigate the probable biological function and pathway activity of the 17 FRGs relevant to diabetic osteoporosis. The significant results from GO enrichment analysis revealed that 17 FRGs were primarily associated with cellular response to external stimulus, response to oxidative stress, and neuron apoptotic process (Figure 6E). Moreover, 17 FRGs were mainly involved in various classical signaling pathways based on the KEGG enrichment analyses, including Lipid and atherosclerosis, Mitophagy animal, and

Endocrine resistance (Figure 6F). Supplementary Table S4 displayed the findings of the GO and KEGG analyses in detail.

3.6 Construction of prediction model and identification of key gene

Based on the whole dataset, we used four established machine learning methods (LASSO, SVM RFE, Boruta, and XGBoost) to find important genes from 17 FRGs associated with diabetic osteoporosis. These algorithms yielded 4, 7, 16, and 6 genes, respectively (Figures 7A–E). We then verified the efficiency of the four machine learning algorithms by ROC curves utilizing GSE56815 as an external dataset. All four algorithms had high area under curve (AUC) values that were more than 0.8 and we considered the results of the prediction models to be reliable (Figure 7F). IDH1 is a common gene belonging to all four algorithms (Figure 7G). The details of genes were shown in Supplementary Table S5. Considering the accuracy of the identified genes, we plotted ROC curves with the external dataset GSE56815,

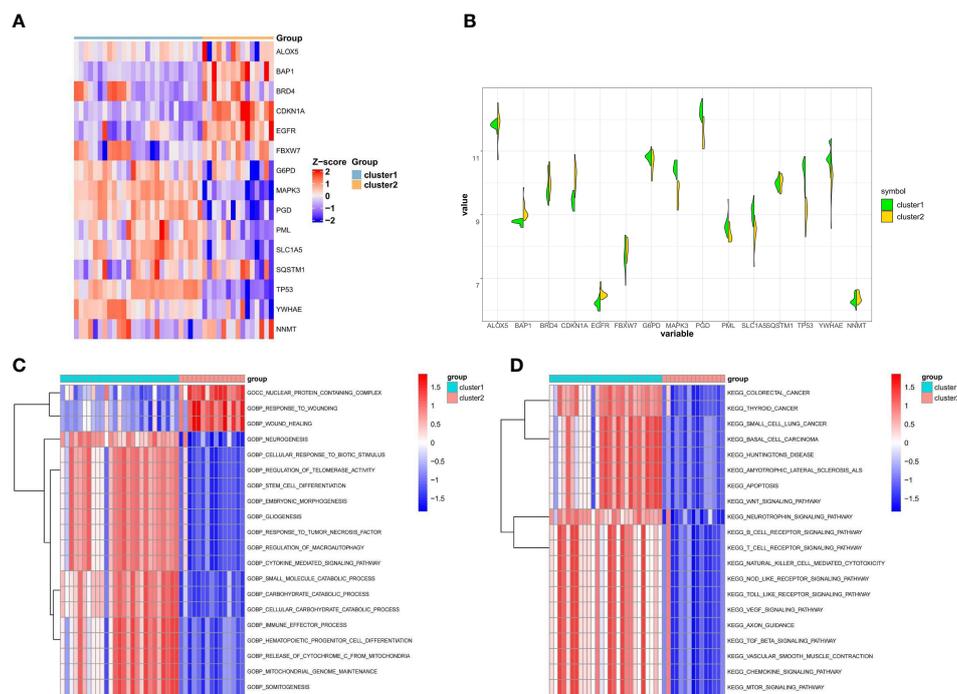


FIGURE 4 Difference analysis between the two ferroptosis clusters. **(A)** Representative heatmap of 15 FRGs between the two ferroptosis clusters. **(B)** Violin plot of 15 FRGs expression between the two ferroptosis clusters. **(C, D)** GSEA enrichment analysis in different ferroptosis clusters showing biological functions **(C)** and significantly activated pathways **(D)**.

which showed a high predictive efficiency (Figure 7H). Ultimately, we identified IDH1 from 17 FRGs as a prospective indicator of diabetic osteoporosis subtypes through four highly efficient prediction models.

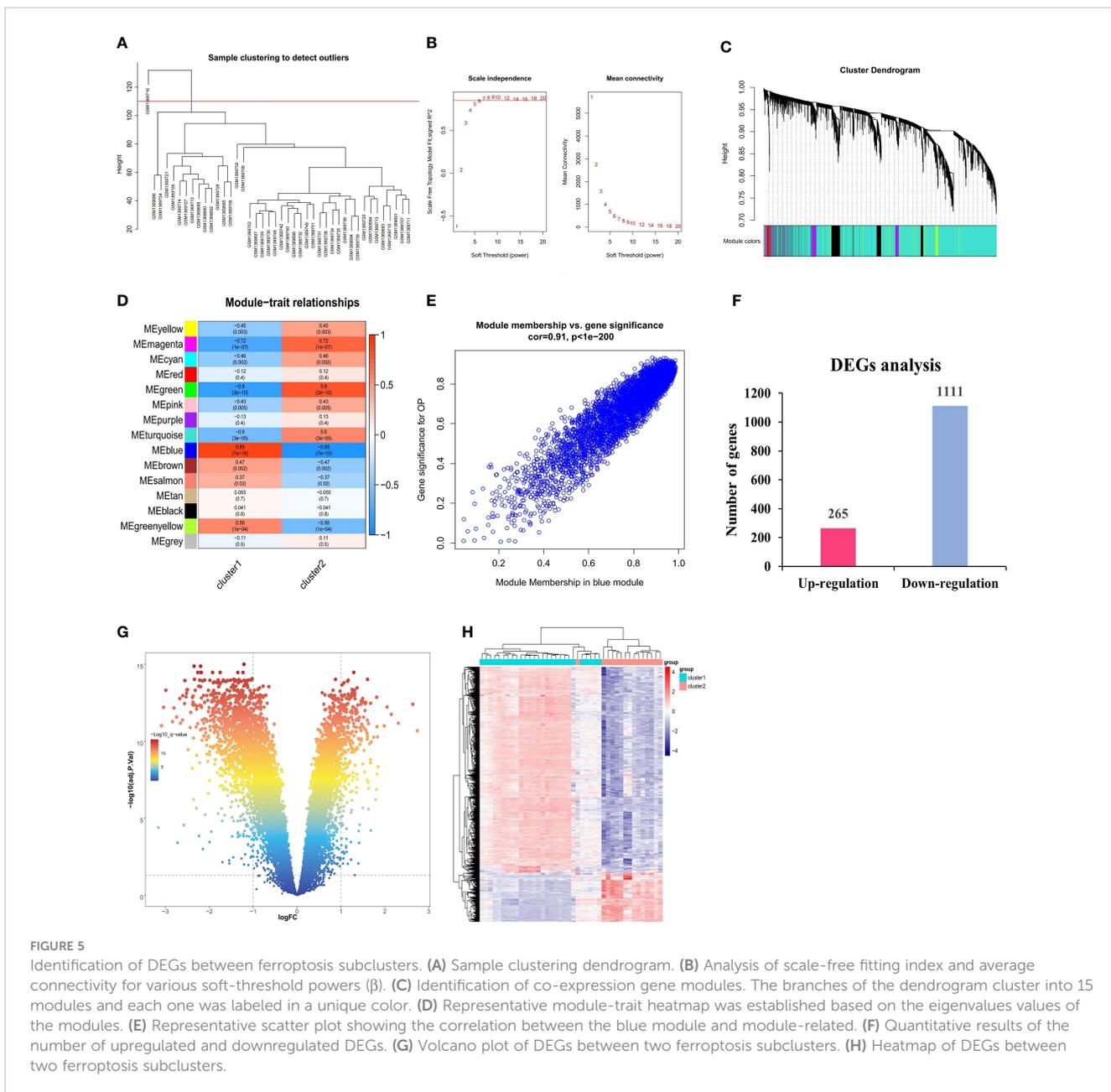
4 Discussion

Diabetic osteoporosis, a chronic complication of diabetes in the skeletal system with a high risk of fracture, is not only a medical problem but also a critical social problem, making early detection and timely intervention extremely significant (18). The role of FRGs in the underlying pathogenesis of diabetic osteoporosis has attracted strong research interest from scholars both nationally and internationally. In addition, studies have confirmed evidence of ferroptosis in the bone tissue of diabetic osteoporotic rats, and inhibitors of ferroptosis may improve osteoporosis symptoms (19). However, the molecular mechanisms of FRGs in different molecular subtypes of diabetic osteoporosis were not widely reported. The molecular heterogeneity of diabetic osteoporosis is made more understandable by the identification of molecular clusters based on FRGs.

In the current study, we synthesized database data and gene expression profiling data for the analysis of FRGs associated with diabetic osteoporosis. In total, we identified 15 FRGs that can distinguish diabetic osteoporosis from normal samples,

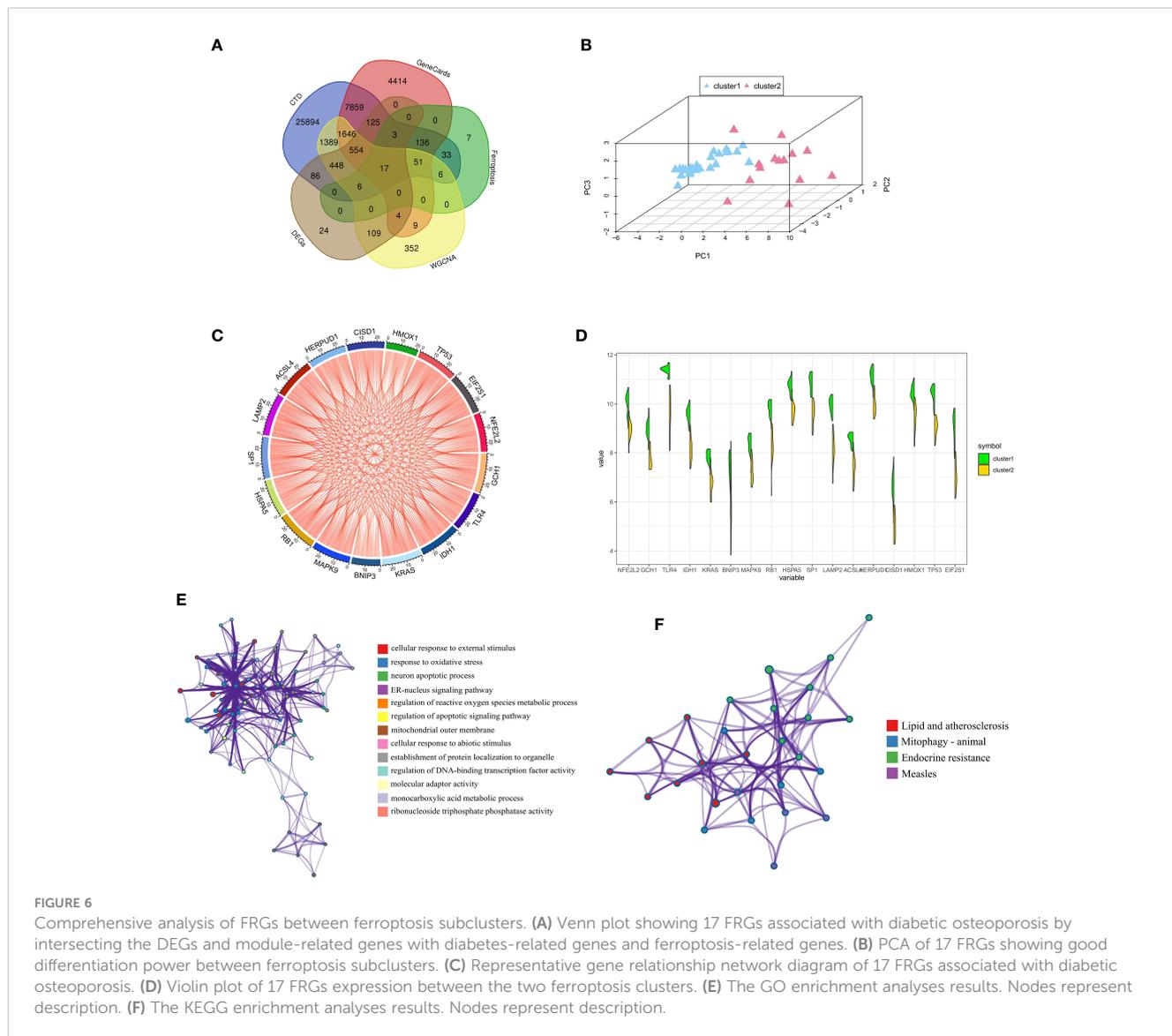
suggesting that ferroptosis may exert an overarching role in the pathological development of diabetic osteoporosis. Subsequently, correlation and expression analysis between ferroptosis regulators elucidated the complex relationship between diabetic osteoporosis and normal individuals. Furthermore, the expression of 15 FRGs associated with diabetic osteoporosis allowed for the classification of 42 osteoporotic samples into two different categories. Expression and enrichment analysis of the 15 FRGs exhibited significant differences between the two molecular clusters, suggesting that ferroptosis could be a prospective indicator of osteoporotic subtypes. According to GSEA enrichment analysis, cluster1 was predominantly enriched in the cellular response to biotic stimulus, regulation of telomerase activity, and stem cell differentiation. Cluster2 was prominently associated with nuclear protein containing the complex, response to wounding, and wound healing. Ferroptosis as a heterogeneous regulator of diabetic osteoporosis, we combined WGCNA and differential expression analysis to identify 17 FRGs that could effectively distinguish between the two subtypes. Ultimately, we identified IDH1 from 17 FRGs as a prospective indicator of diabetic osteoporosis subtypes, using multiple machine learning algorithms and validated by uniting external datasets.

IDH1, a gene located on chromosome 2q34, encodes the IDH1 protein consisting of 414 amino acids situated in the



cytoplasm and peroxisome (20). IDH1 is a homodimer that forms an active NADP⁺ binding site among the structural domains (21). NADP⁺ dependent IDH1 catalyzes the oxidative decarboxylation of isocitrate to produce alpha-ketoglutarate, which leaves DNA and histones in a demethylated state (22). In previous studies, mutations in IDH1 were found in association with a variety of malignancies and rare cases, such as low-grade diffuse gliomas, periosteal cartilage tumours, cholangiocarcinoma, acute myeloid leukaemia, and Ollier's disease and Maffucci's syndrome (23). In malignancies, IDH1 mutations promote the accumulation of lipid reactive oxygen species (ROS) by reducing the protein levels of glutathione peroxidase 4 (GPX4), which subsequently leads to ferroptosis (24). This phenomenon indicates that tumour-derived IDH1

mutations sensitise cells to ferroptosis. IDH1 primarily exerts metabolic effects *in vivo*, IDH1 ectopic expression suppresses the production of brown adipocytes as a novel therapeutic target against obesity and related metabolic diseases such as type II diabetes and may represent a therapeutic target for the treatment of metabolic diseases (25). Besides metabolic functions, IDH1 also regulates gene expression through epigenetic modifications of histones. However, it has never been studied whether IDH1 affects ferroptosis and thus worsens the progression of diabetic osteoporosis. Diabetes induces fat accumulation causing atherosclerosis and narrowing of the vascular lumen, resulting in inadequate blood supply to the bone (26). With a poor vascular system, the blood supply to the bone tissue is compromised which will not work properly and structural



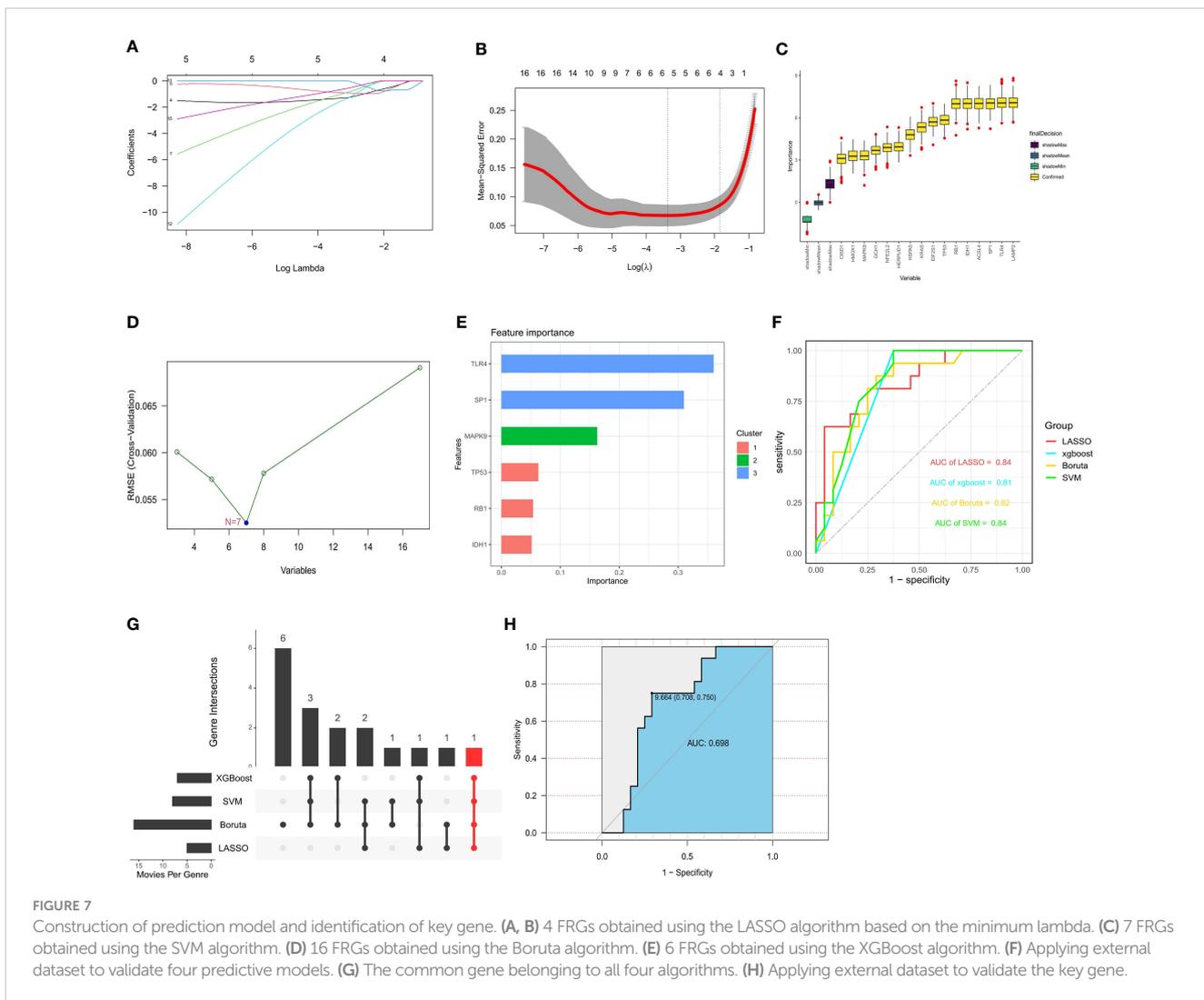
abnormalities, such as microcracks, may occur. However, it has been demonstrated that IDH1 inhibition regulates the progression of atherosclerosis by improving macrophage viability and apoptosis, and may alleviate atherosclerosis by activating NRF2 to ameliorate ox-LDL-induced ferroptosis in macrophages (27). Furthermore, IDH1 production of NADPH in the cytoplasm reduces intracellular oxidative stress, which as a complication of diabetes can also lead to osteoporosis (28). In this study, IDH1 was identified combined with machine learning prediction models and expressed significantly different in the analysis of diabetic osteoporosis subtypes. Moreover, plotting the ROC curve with the external data set as the validation set resulted in a high AUC value (0.698). The above results sufficiently demonstrate that IDH1 has a high probability of acting as a regulator of ferroptosis in differentiating the heterogeneity of diabetic osteoporosis subtypes, which provides a strong basis for further studies.

There are certain research limitations that need to be made clear. We lacked a sufficiently large number of diabetic osteoporosis

samples to assess the accuracy of ferroptosis genes in machine learning model predictions and to validate the stability of subpopulations. More prognostic information needs to be gathered in further studies to assess the prognostic value of ferroptosis in distinct subtypes of diabetic osteoporosis. Additionally, IDH1 needs further literature support and experimental validation as a regulator of iron death which distinguishes diabetic osteoporosis subtypes.

5 Conclusion

In conclusion, we found two clusters of ferroptosis in diabetic osteoporosis and confirmed the distinctive characteristics of each. Four distinct machine learning prediction models (LASSO, XGBoost, Boruta, and SVM) based on 17 ferroptosis genes discovered ferroptosis regulators capable of distinguishing diabetic osteoporosis subtypes. Ultimately, as a ferroptosis regulator validated by the external datasets, IDH1 has the



capacity to precisely distinguish molecular subtypes of diabetic osteoporosis, which may provide novel insights into the pathophysiology of the clinical symptoms and prognostic heterogeneity in diabetic osteoporosis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

XW, LM and JunZ collected the original data, finished the analysis and drafted the initial version. ZZ, LZ, ZJ, XH, LZ and MS helped revise the manuscript. ML and XQ offered constructive comments on experimental studies. ML, XQ, SW and JunZ provided the funding. ML designed and conceptualized the study, put forward many constructive comments for the final version and supervised the

study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (82074426, 82104864, 82204822), Natural Science Foundation of Liaoning Province (2021-BS-215, 2022-MS-25, 2023-MS-13), Liaoning Revitalization Talents Program (XLYC1802014), Liaoning Key Research and Development Planning Project (2017226015), Basic Research Projects of Liaoning Provincial Department of Education (LJKMZ20221286), Natural Science Foundation of Tibet Autonomous Region and Regional Science (XZ202301ZR0030G) and Technology Project of Naqu City.

Acknowledgments

We thank all the participants who supported our study. In particular, thanks to the GEO database, CTD database, GeneCards database, FerrDb database, and R software for the analytical data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract* (2022) 183:109119. doi: 10.1016/j.diabres.2021.109119
- Abdulameer SA, Sulaiman SAS, Hassali MAA, Subramaniam K, SAHIB MN. Osteoporosis and type 2 diabetes mellitus: what do we know, and what we can do? *Patient Preference Adherence* (2012) 6:435–48. doi: 10.2147/PPA.S32745
- Vilaca T, Schini M, Harnan S, Sutton A, Poku E, Allen IE, et al. The risk of hip and non-vertebral fractures in type 1 and type 2 diabetes: a systematic review and meta-analysis update. *Bone* (2020) 137:115457. doi: 10.1016/j.bone.2020.115457
- Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* (2012) 149:1060–72. doi: 10.1016/j.cell.2012.03.042
- Ma H, Wang X, Zhang W, Li H, Zhao W, Sun J, et al. Melatonin suppresses ferroptosis induced by high glucose via activation of the Nrf2/HO-1 signaling pathway in type 2 diabetic osteoporosis. *Oxid Med Cell Longevity* (2020) 2020:9067610. doi: 10.1155/2020/9067610
- Gao Z, Chen Z, Xiong Z, Liu X. Ferroptosis - a new target of osteoporosis. *Exp Gerontology* (2022) 165:111836. doi: 10.1016/j.exger.2022.111836
- Lin Y, Shen X, Ke Y, Lan C, Chen X, Liang B, et al. Activation of osteoblast ferroptosis via the METTL3/ASK1-p38 signaling pathway in high glucose and high fat (HGHF)-induced diabetic bone loss. *FASEB J Off Publ Fed Am Societies For Exp Biol* (2022) 36:e22147. doi: 10.1096/fj.202101610R
- 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* (2013) 41:D991–5. doi: 10.1093/nar/gks1193
- 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:e47. doi: 10.1093/nar/gkv007
- Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wiegiers J, Wiegiers TC, et al. Comparative toxicogenomics database (CTD): update 2021. *Nucleic Acids Res* (2021) 49:D1138–43. doi: 10.1093/nar/gkaa891
- Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, et al. GeneCards version 3: the human gene integrator. *Database J Biol Database Curation* (2010) 2010:baq020. doi: 10.1093/database/baq020
- Zhou N, Bao J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database J Biol Database Curation* (2020) 2020:baaa021. doi: 10.1093/database/baaa021
- Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinf (Oxford England)* (2010) 26:1572–3. doi: 10.1093/bioinformatics/btq170
- Ferreira MR, Santos GA, Biagi CA, Silva Junior WA, Zambuzzi WF. GSA score reveals molecular signatures from transcriptomes for biomaterials comparison. *J Biomed Materials Res Part A* (2021) 109:1004–14. doi: 10.1002/jbm.a.37090

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2023.1189513/full#supplementary-material>

SUPPLEMENTARY TABLE 1

The details of database genes, DEGs and intersecting genes.

SUPPLEMENTARY TABLE 2

The details of GSEA analysis.

SUPPLEMENTARY TABLE 3

The details of database genes, DEGs, module-related genes of WGCNA and intersecting genes.

SUPPLEMENTARY TABLE 4

The detailed analysis results of GO and KEGG.

SUPPLEMENTARY TABLE 5

The details of the genes obtained by four machine learning algorithms.

- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* (2019) 10:1523. doi: 10.1038/s41467-019-09234-6
- Huang D, Zheng S, Liu Z, Zhu K, Zhi H, Ma G. Machine learning revealed ferroptosis features and a novel ferroptosis-based classification for diagnosis in acute myocardial infarction. *Front In Genet* (2022) 13:813438. doi: 10.3389/fgene.2022.813438
- Schacter GI, Leslie WD. Diabetes and osteoporosis: part II, clinical management. *Endocrinol Metab Clinics North America* (2021) 50:287–97. doi: 10.1016/j.jec.2021.03.006
- Wang X, MA H, Sun J, Zheng T, Zhao P, Li H, et al. Mitochondrial ferritin deficiency promotes osteoblastic ferroptosis via mitophagy in type 2 diabetic osteoporosis. *Biol Trace Element Res* (2022) 200:298–307. doi: 10.1007/s12011-021-02627-z
- Shechter I, Dai P, Huo L, Guan G. IDH1 gene transcription is sterol regulated and activated by SREBP-1a and SREBP-2 in human hepatoma HepG2 cells: evidence that IDH1 may regulate lipogenesis in hepatic cells. *J Lipid Res* (2003) 44:2169–80. doi: 10.1194/jlr.M300285-JLR200
- Xu X, Zhao J, Xu Z, Peng B, Huang Q, Arnold E, et al. Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J Biol Chem* (2004) 279:33946–57. doi: 10.1074/jbc.M404298200
- Geisbrecht BV, Gould SJ. The human PICD gene encodes a cytoplasmic and peroxisomal NADP(+)-dependent isocitrate dehydrogenase. *J Biol Chem* (1999) 274:30527–33. doi: 10.1074/jbc.274.43.30527
- Bruce-Brand C, Govender D. Gene of the month: IDH1. *J Clin Pathol* (2020) 73:611–5. doi: 10.1136/jclinpath-2020-206813
- Wang T-X, Liang J-Y, Zhang C, Xiong Y, Guan K-L, Yuan H-X. The oncometabolite 2-hydroxyglutarate produced by mutant IDH1 sensitizes cells to ferroptosis. *Cell Death Dis* (2019) 10:755. doi: 10.1038/s41419-019-1984-4
- Kang HS, Lee JH, Oh K-J, Lee EW, Han BS, Park K-Y, et al. IDH1-dependent α -KG regulates brown fat differentiation and function by modulating histone methylation. *Metabolism: Clin Exp* (2020) 105:154173. doi: 10.1016/j.metabol.2020.154173
- Kolluru GK, Bir SC, Kevil CG. Endothelial dysfunction and diabetes: effects on angiogenesis, vascular remodeling, and wound healing. *Int J Vasc Med* (2012) 2012:918267. doi: 10.1155/2012/918267
- Li B, Wang C, Lu P, Ji Y, Wang X, Liu C, et al. IDH1 promotes foam cell formation by aggravating macrophage ferroptosis. *Biology* (2022) 11:1392. doi: 10.3390/biology11101392
- Molenaar RJ, Radvovjevitch T, Maciejewski JP, Van Noorden CJF, Bleeker FE. The driver and passenger effects of isocitrate dehydrogenase 1 and 2 mutations in oncogenesis and survival prolongation. *Biochim Et Biophys Acta* (2014) 1846:326–41. doi: 10.1016/j.bbcan.2014.05.004