Heliyon 10 (2024) e26759

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification and immunological role of cuproptosis in osteoporosis

Tongying Chen [a,1], Zhijie Gao [b,f,1], Yuedong Wang [c], Jiachun Huang [d], Shuhua Liu [a], Yanping Lin [d], Sai Fu [a], Lei Wan [d,e], Ying Li [d,g], Hongxing Huang [d,e,**], Zhihai Zhang [d,e,*]

[a] Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China
[b] The Second Clinical Medical School, Guangzhou University of Chinese Medicine, Guangzhou, China
[c] The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China
[d] The Third Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China
[e] Laboratory Affiliated to National Key Discipline of Orthopaedic and Traumatology of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China
[f] The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou, China
[g] Qifu Hospital Affiliated to Jinan University, Guangzhou, China

ARTICLE INFO

Keywords:
Osteoporosis
Cuproptosis
Immune infiltration
Mitochondria

ABSTRACT

Background: osteoporosis is a skeletal disorder disease features low bone mass and poor bone architecture, which predisposes to increased risk of fracture. Copper death is a newly recognized form of cell death caused by excess copper ions, which presumably involve in various disease. Accordingly, we intended to investigate the molecular clusters related to the cuproptosis in osteoporosis and to construct a predictive model.

Methods: we investigated the expression patterns of cuproptosis regulators and immune signatures in osteoporosis based on the GSE56815 dataset. Through analysis of 40 osteoporosis samples, we investigated molecular clustering on the basis of cuproptosis–related genes, together with the associated immune cell infiltration. The WGCNA algorithm was applied to detect cluster-specific differentially expressed genes. Afterwards, the optimum machine model was selected by calculating the performance of the support vector machine model, random forest model, eXtreme Gradient Boosting and generalized linear model. Nomogram, decision curve analysis, calibration curves, and the GSE7158 dataset was utilizing to confirm the prediction efficiency.

Results: Differences between osteoporotic and non-osteoporotic controls confirm poorly adjusted copper death-related genes and triggered immune responses. In osteoporosis, two clusters of molecules in connection with copper death proliferation were outlined. The assessed levels of immune infiltration showed prominent heterogeneity between the different clusters. Cluster 2 was characterized by a raised immune score accompanied with relatively high levels of immune infiltration. The functional analysis we performed showed a close relationship between the different immune responses and specific differentially expressed genes in cluster 2. The random forest machine model showed the optimum discriminatory performance due to relatively low

* Corresponding author. 261 Longxi Rd., Guangzhou, Guangdong, 510360, China.
** Corresponding author. The Third Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China.
E-mail addresses: hhx@gzucm.edu.cn (H. Huang), 13725295793@139.com (Z. Zhang).
[1] Co-first author:Tongying Chen, Zhijie Gao.

https://doi.org/10.1016/j.heliyon.2024.e26759
Received 5 May 2023; Received in revised form 11 February 2024; Accepted 20 February 2024
Available online 20 February 2024

residuals and root mean square errors. Finally, a random forest model based on 5 genes was built, showing acceptable performance in an external validation dataset (AUC = 0.750). Calibration curve, Nomogram, and decision curve analyses also evinced fidelity in predicting subtypes of osteoporosis.

Conclusion: Our study identifies the role of cuproptosis in OP and essentially illustrates the underlying molecular mechanisms that lead to OP heterogeneity.

1. Introduction

Osteoporosis is a condition in which bone mass decreases over time, leading to an increased rate of fractures. According to a universal epidemiological study on fractures causing by osteoporotic vertebral, more than 200 million people worldwide may have suffering from osteoporosis, which affects people's quality of life and places an additional burden on the economy [1]. According to the traditional view, the main contributing factors to osteoporosis are multifactorial imbalances in hormonal changes, calcium and vitamin D deficiency [2], oxidative stress, mitochondrial dysfunction, chronic inflammation and normal ageing. The literature is now beginning to support the inclusion of programmed cell death (PCD) as another important factor, with PCD (e.g. apoptosis, autophagy, iron toxicity, thermosis and necrosis) playing a key role in regulating bone metabolism by affecting osteoblast activity, suggesting that intervention in PCD by focusing on specific regulatory molecules could be effective in controlling osteoporosis [3]. Surprisingly, a recently identified type of controlled cell death brought about by excess copper ions (Cu^{2+}) is known as cupric cell death [[4,5]]. Copper is involved in the tricarboxylic acid cycle in mitochondria, during which it is present as cytochrome oxidase (COX) together with superoxide dismutase (SOD1). Therefore, Copper plays an essential role in many biological processes, whether it is redox homeostasis, iron conversion processes, oxidative phosphorylation of cells or cell growth. The form of copper-induced programmed cell death is very distinct from other cell death pathways such as apoptosis, iron toxicity and necrosis. Intracellular copper interacts with a number of proteins in mitochondria and acts on the lipid acylated component of the tricarboxylic acid (TCA) cycle. Cuproptosis can lead to cell death due to the accumulation of these Cubound lipid-acylated mitochondrial proteins - along with an induced reduction in Fe–S (iron-sulphur) clusters - causing proteotoxic stress [[6–9]]. Cuproptosis has been shown to affect a variety of cancers [10], such as breast [11], pancreatic [12], thyroid [13], leukemia [14], CRC [15], lung [16], prostate [17], and oral [18], with copper accumulation levels being detected at significantly different levels in serum and tumour tissue.

As cuproptosis is associated with mitochondria and oxidation, we believe that cuproptosis is crucial in osteoporosis, still, the association between osteoporosis and cuproptosis is still unknown. We therefore sought to clarify their association through bioinformatics.

2. Materials

2.1. Data gathering and preparation

We choose the two microarray datasets associated with the OP (GSE56815 and GSE7158) from the GEO database (GEO, www.ncbi.nlm.nih.gov/geo) utilizing the 'GEOquery' R package (version 2.60). The GSE56815 dataset (GPL96 platform) included mononuclear cell samples from 40 women not diagnosed with high BMD and 40 women diagnosed with osteoporosis would be further analysis. the GSE7158 dataset (GPL570 platform), including mononuclear cell samples from 14 subjects with no osteoporosis and 12 subjects who are suffering osteoporosis. The Robust Multiarray Average (RMA) method ("affy" R package, which version is 1.70.0) was applied to process and normalize these GEO datasets' raw gene expression profiles.

A measure of the level of infiltration of immune cells.

On the basis of the processed gene expression data, the relative amount of 22 immune cells in every sample was assessed with the help of the CIBERSORT algorithm and the LM22 signature matrix. CIBERSORT used Monte Carlo sampling to produce an inverse fold p-value for the sample respectively. p-values less than 0.05 were estimated as the exact immune cell fraction for each sample. The total of these 22 immune cells percentages in each sample was 1.

For the analysis of the presence between CRG and infiltrating immune cells.

The correlation coefficient between CRG expression and the relative proportion of immune cells allowed us to assess the association between CRG and OP-related immune cell characteristics. A p-value below 0.05 was regarded as a significant link due to the spearman correlation coefficient stipulation. Ultimately, the data were revealed using the "corrplot" R package (version 0.92).

2.2. Clustering of OP patients without supervision

Firstly, 4 CRGs in all were selected based on calculation. Underlying 4 CRG expression scores, we utilized unsupervised cluster analysis ("Consensus ClusterPlus" R package, which version is 2.60), performing the k-means algorithm to group the 40 OP samples into independent clusters and performed 1000 iterations. We chose an optimal number of subtypes k ($k = 2$) together with the optimal number of clusters was thoroughly estimated utilizing cumulative distribution function (CDF) curves, consensus matrices and consistent clustering scores (>0.9).

2.3. Analysis of gene set variation analysis (GSVA)

To elucidate the variability of enriched genomes between different CRGs clusters, GSVA enrichment was conducted with the R package "GSVA" (version 2.11). For additional GSVA analyses, MSigDBk can be accessed as "c2.cp.kegg.v7.4.symbols" and "c5.go.bp.v7.5.1.symbols". We downloaded the above files and utilized the R package limma (version 3.52.1) in order to get to distinguish between differentially expressed pathways and biological functions. We compared GSVA scores between clusters of CRGs. Cluster GSVA scores with significant variation had t-values greater than 2.

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

Co-expression modules were calculated using the R package for "WGCNA" (version 1,70.3). The data used for WGCNA analysis were then the top quater of genes with the greatest divergence, which ensures the accuracy of the results, then we constructed the weighted

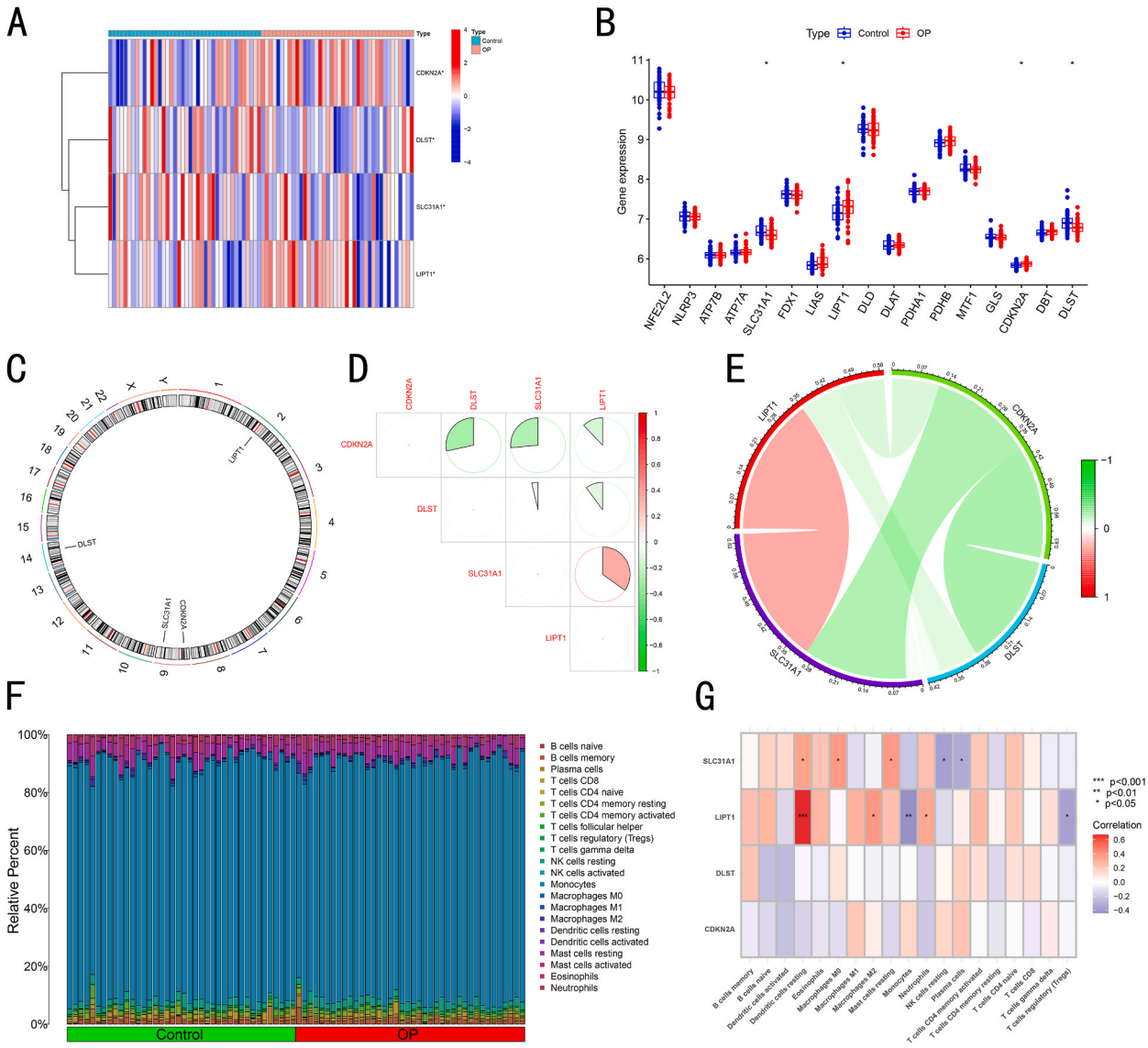

Fig. 1. Identification of CRGs dysregulated in OP.(A) Expression profiles of the 4 CRGs are shown in the heat map. (B) Boxplots are utilized to represent the expression of the 4 CRGs between OP and non-OP controls. *p < 0.05. (C) Positioning of the 4 CRGs on the chromosome. (D) Correlation scrutiny of the 4 differentially expressed CRGs. Green and red respectively own the meaning of positive and negative connections. Correlation coefficients are indicated by the occupation of the pie chart. (E) Gene bond system plot of the 4 differentially expressed CRGs. (F) Relative abundance of 22 infiltrated immunologic cells between OP and non-OP controls. (G) Correlation assays of 12 differentially expressed CRGs and inserts between immune cells. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

neighbour joining matrix and the topological overlap matrix (TOM). the TOM divergence metric (1-TOM) was utilized to identify modules on the basis of a hierarchical clustering tree algorithm. Each random colour represents a module. The module eigengene describes the global gene expression pattern of each module. The relationship between module and disease state is shown by module importance (MS). Gene Significance (GS) defines the association between a gene and a clinical phenotype.

2.4. Building a predictive model based on various machine learning techniques

The "caret" R package (version 6.0.91) to create machine studying models based on two distinct CRG clusters involves support vector machine models (SVM), random forest models (RF), eXtreme gradient boosting (XGB) and generalized linear models (GLM). RF is a set of independent decision trees used in ensemble machine learning methods for predictive classification or regression. By utilizing the SVM algorithm, A hyperplane can be created in the feature dimension that has the greatest margin for isolating positive and negative instances. A flexible way of assessing the association between naturally distributing causal data and classified or sequential stand-alone traits is via the use of GLM, the extension of the multinomial Regression Model. An ensemble of XGB boosted trees that are according to gradient upgrading allows careful comparison of sorting errors and model sophistication. Different clusters are assumed to respond to changes in the data and DEGs which are cluster-specific are selected as the explanatory variables. GSE7158 is randomly classified as the training set and GSE7158 as the validation set. the caret package adjusts the parameters of the model by lattice search

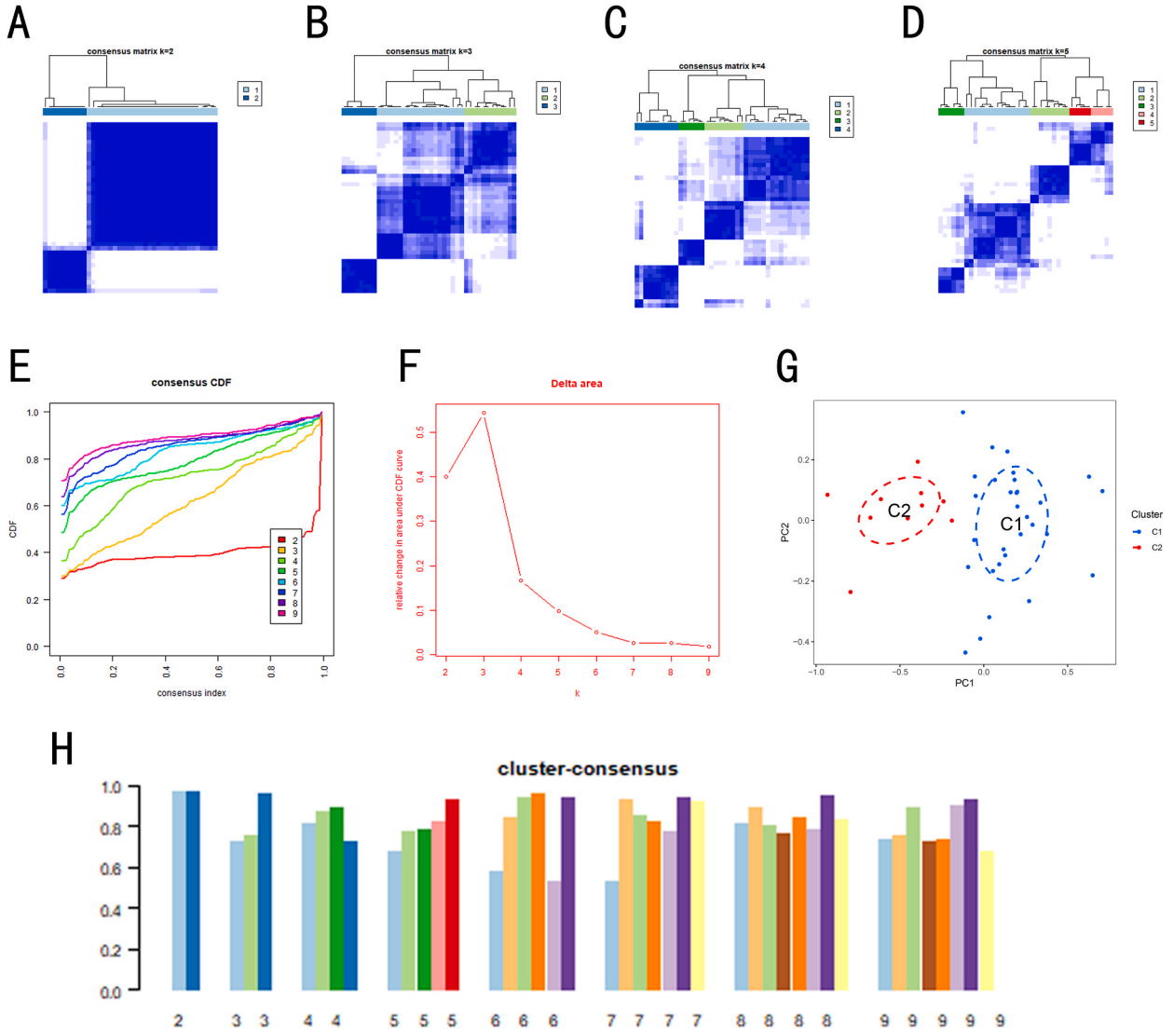

Fig. 2. Validation of molecular clusters connected with cuproptosis in OP. (A–D) Consensus clustering matrix when k = 2. (E–F) Representative cumulative distribution function (CDF) curves of the results. (E)the consensus index (F) CDF delta area curves. (G) t-SNE pictures the distribution of two subtypes. (H) the score of consensus clustering.

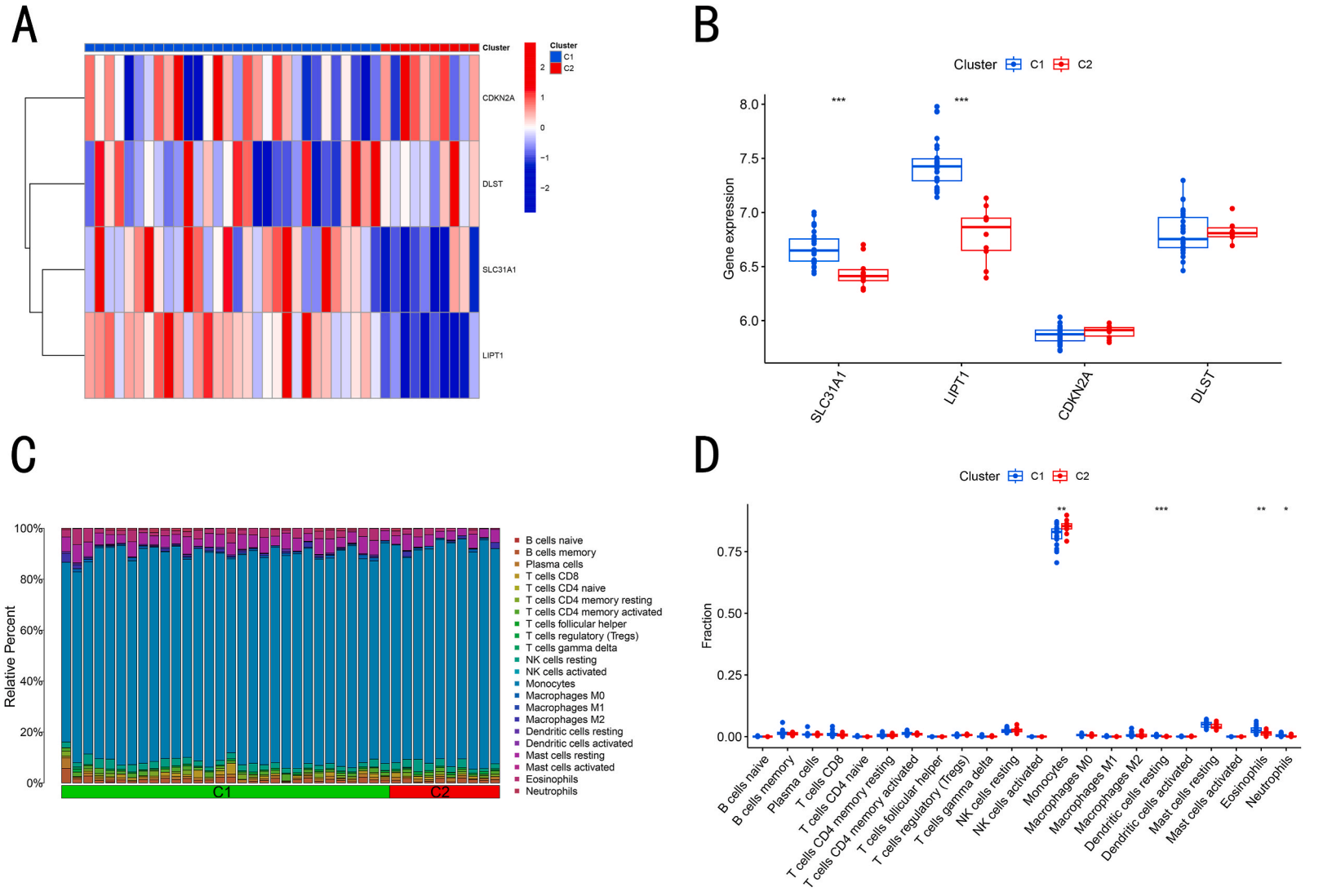

Fig. 3. Identification of the molecular and immunological traits between two clusters of cuproptosis. (A) Heat map showing the clinical features and expression patterns of the 4 CRGs between the two clusters of cuproptosis. (B) Boxplots showing the expression of the 4 CRGs between the two clusters of cuproptosis. ***p < 0.001, *p < 0.05. (C) Relative abundance of 22 endogenous immune cells between the two clusters of cuproptosis. (D) Boxplots showing the difference in immunologic infiltrating between the two clusters of cuproptosis. *p < 0.05, **p < 0.01 ***p < 0.001.

and the above computer learning model is proceeded with default parameters and evaluated by cross-validation of 5-fold. The software "DALEX" (version 2.4.0) was used to understand the distribution of residuals and the significance of features for the four aforementioned computer learning models. The R package "pROC" (version 1.18.0) was used to view the proportions under the ROC curve. As a

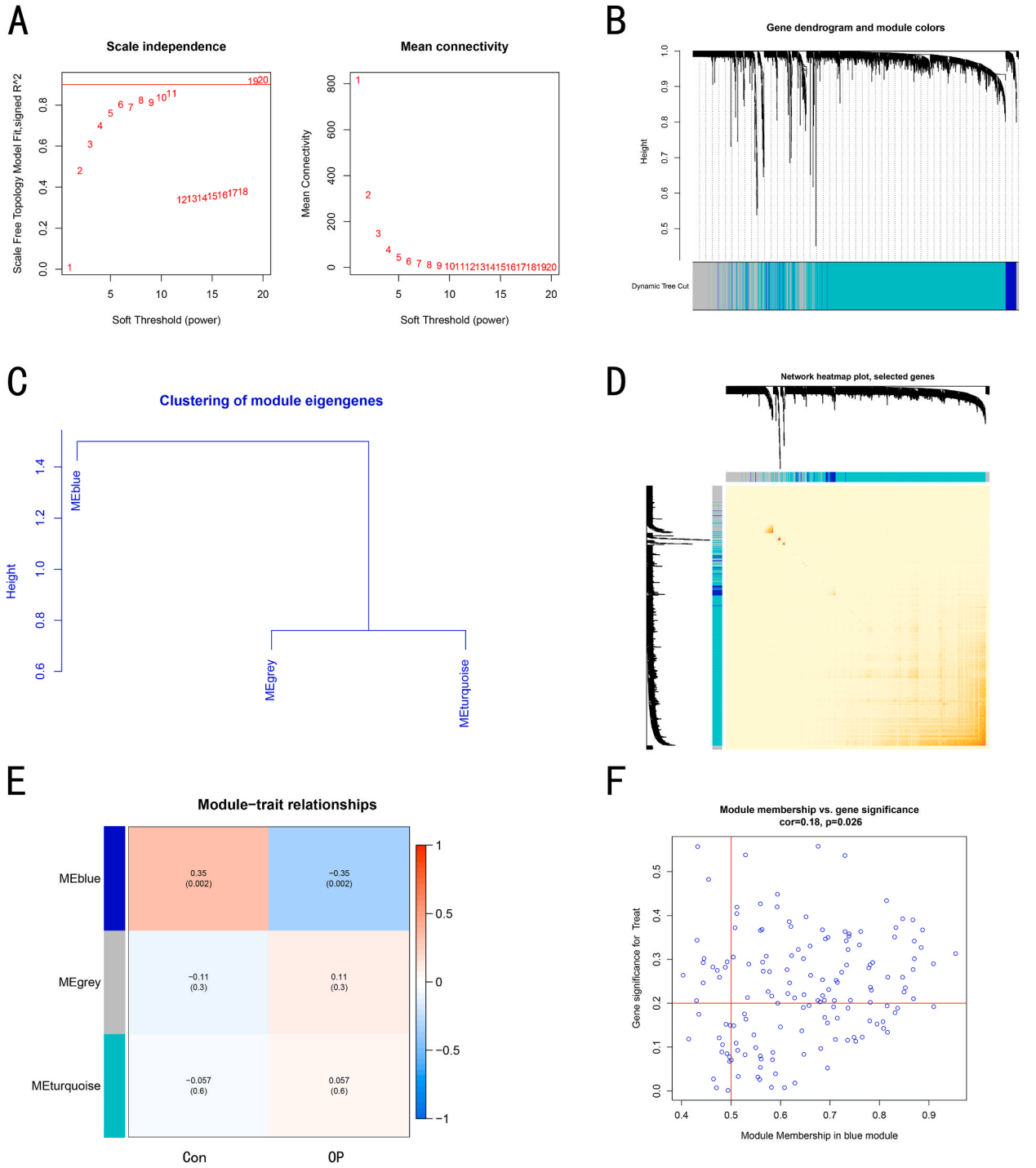

Fig. 4. Co-expression network of differentially expressed genes in OP. (A) Option of soft threshold power. (B) Clustered dendrograms of co-expression modules. Different colours represent different co-expression modules. (C) Clustered representation of module eigengenes. (D) Representative heat map of the correlation of the 3 modules. (E) Correlation analysis of module eigengenes and clinical status. Each row indicates a module; each column denotes a clinical state. (F) Scatter plots between module members in turquoise modules and the genetic significance of OPs. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

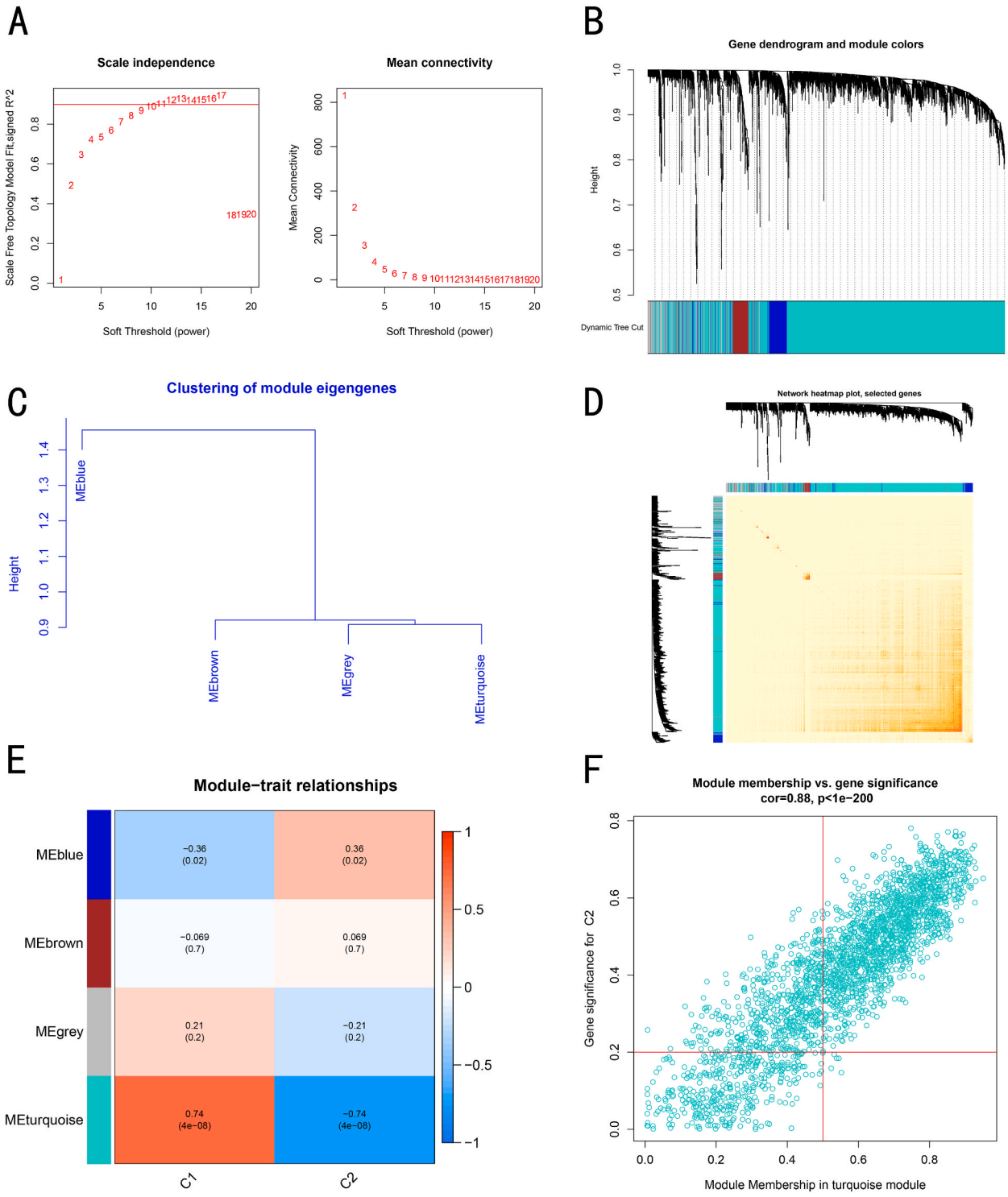

Fig. 5. Co-expression network of differentially expressed genes between two populations of cupula cells (A) Soft threshold selection. (B) Clustered dendrogram of coexpression modules. Different colours represent different co-expression modules. (C) Clustered representation of module eigengenes. (D) Representative heat map of correlations between the 11 modules. (E) Correlation analysis of module eigengenes and clinical status. Each row indicates a module, each column indicates a clinical status. (F) Scatter plots between module members in turquoise and cluster1 gene salience. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

result, we identified the best machine learning model, with the OP-related underlying predictive genes being the top 5 variables. Finally, ROC curve analysis was applied to the GSE7158 dataset to confirm the diagnostic value of the diagnosis model.

2.5. Establishing and testing a nomogram model

A nomogram model was constructed to estimate the prevalence of OP clusters (version 6.2.0) With the 'rms' R package. Each predictor is given a matching mark and the 'total score' is the result of adding the values of all the above predictors together. In order to calculate the forecasting power of the nomogram model, DCA and calibration curves were used.

3. Result

3.1. Dysfunctional copper death and activation of the immune response in patients with OP

To elucidate the biofunctions of cup mutation regulators in OP patients, we systematically estimated the manifestations of four CRGs in non-OP controls and OP through the GSE56815 dataset. 4 CRGs were selected as class II expressed genes. Among them, those with higher expression levels were CDKN2A and LIPT1, while SLC31A1 and DLST had lower expression levels in the OP than in the non-OP control (Fig. 1A–C). Thereafter, we conducted a correlation analysis between the two to investigate the function of the cupular apoptosis regulator during the development of OP. Surprisingly, SLC31A1 and LIPT1, expressed a strong cooperative relationship. Meanwhile, CDKN2A and DLST showed significant antagonistic effects. In addition, we found a significant association between CDKN2A and other related proteins. (Fig. 1D). The relationship between the two genes is inferred from the relationship network diagram (Fig. 1E).

To explicate the altered immune system between OP and non-OP, an immunoinfiltration analysis was performed predicated on the CIBERSORT method to demonstrate the distinction in the proportions of 22 cell types of infiltrated immune between OP and non-OP control patients. (Fig. 2F). The results of the study showed that Mast cells resting, M0 macrophages, M2 macrophages, activated dendritic cells, and neutrophils levels were higher in OP patients were infiltrated at higher levels in patients with OP (Fig. 1F), suggesting that the immune system is involved in the development of OP and plays an important role. Also, results from related studies showed a correlation between cuproptosis regulators and both dendritic cells resting and Monocytes (Fig. 1F). Based on these outcomes, CRG may have a profound impact in controlling the immunological infiltration status and associated molecules in patients with OP.

3.2. Recognition of cuproptosis clusters in OP

We conducted consensus clustering to classify 40 OP samples as stated by the expression profiles of four CRGs to elucidate the manifestation patterns associated with cuproptosis in ASD. The CDF curves fluctuated in the low range at consensus indices from 0.2 to 0.6, and the number of clusters was most stable when the k value was adjusted to 2 ($k = 2$). (Fig. 2A–E). The difference between the two CDF curves (k and k-1) is seen in the region under the CDF curve for $k = 2$ to 5 (Fig. 2F). In addition, the concordance score for each subtype was more than 0.9 only when $k = 2$. (Fig. 2H). We finally divided the 40 OP patients into two different clusters, Cluster1 ($n = 30$) and Cluster2 ($n = 10$). t-Distributed Random Neighbourhood Embedding (tSNE) analysis showed significant differences between these two clusters (Fig. 2G).

Distinction of cuproptosis regulators and immune percolation characteristics in cuproptosis clusters.

We comprehensively assessed the differences in expression of 4 CRGs between Cluster1 and Cluster2 to investigate the chemical variation between cluster groups. Different expression profiles of CRGs emerged between the two patterns of cuproptosis (Fig. 3A). high expression levels of LIPT1 and SLC31A1 were found in Cluster 1, while high expression of CDKN2A and DLST was found in Cluster 2 (Fig. 3B). In addition, the outcomes of the immune infiltration study suggests that there are some differences in the immune microenvironment between Cluster 1 and Cluster 2 of the cuproptosis mutations (Fig. 3C). Cluster 1 had a higher proportion of resting dendritic cells, eosinophils and neutrophils compared to Cluster 2 (Fig. 3C). Consistently, Cluster2 showed a higher number of monocytes (Fig. 3D), suggesting that Cluster2 may have a higher extent of immune infiltration.

Gene modules filtering and construction of co-expression network.

We utilized WGCNA technology to build a co-expression network and modules for normal and OP participants to determine the principal gene modules relevant to OP. We computed the expression variances for individual gene in GSE56815 then chose the top quater of these genes for additional inspection. Co-expressed gene modules were identified when the scale-free R^2 was equivalent to 0.9 and 9 as the value of soft power (Fig. 4A). In all 3 unique co-expression modules of various shades were collected with the dynamic cut method, while displaying the topological overlap matrix (TOM) heat map (Fig. 4B–D). These genes labeled in the 6 color modules were then consistently occupied for comparison and co-expression of neighboring module-clinical features (OP and Control). Last but not least, 316 genes in the turquoise module, had the highest correlation with OP (Fig. 4E). Furthermore, we recognized favorable correlations between the turquoise module and module-associated genes (Fig. 4F). Utilizing the technique of WGCNA, we also screened key gene modules that were firmly connected with the cupped cell population. $R^2 = 0.9$ was selected as the most correctly chosen soft threshold to build a scale-free network (Fig. 5A). In particular, four modules were found to be important and heat maps show the TOM of all genes associated with each module (Fig. 5B–D). Based on the investigation of the association between modules and clinical features (Cluster1 and Cluster2), there was a strong association between the turquoise modules (1560 genes) and OP clusters (Fig. 5E). Correlation studies showed a clear association between the genes of the Turquoise module and the selected modules (Fig. 5F).

3.3. Identification of cluster-specific DEGs and functional annotation

By supervising the junctions between the module-related genes of cuproptosis clusters and the module-related genes of OP and non-OP individuals, a total of 99 cluster-specific DEGs were discovered (Fig. 6A). The functional differences between the two clusters related to cluster-specific DEGs were further explored using the GSVA analysis. This findings showed that whereas the spliceosome, RNA degradation, Mismatch repair, Nucleotide excision repair, Propanoate metabolism, Peroxisome, Protein export, Terpenoid backbone biosynthesis, Proteasome, Basal transcription factors, Valine leucine and isoleucine degradation were increased in Cluster2, Metabolism of xenobiotics by cytochrome p450, taste transduction, Glycosphingolipid biosynthesis globo series, Dorso ventral axis formation, Adipocytokine_signaling pathway, Pathways in cancer, Glycosphingolipid biosynthesis lacto and neolacto series, Acute myeloid leukemia, Fc epsilon ri signaling pathway, Neurotrophin signaling pathway were reinforced in Cluster1 (Fig. 6B). Other than that, the findings of the functional enrichment analysis showed that Cluster1 was significantly connected to the control of Cluster of Actin Based Cell Projections, Asymmetric Neuroblast Division, Regulation Of Hormone Metabolic Process, Establishment Of Protein Localization To Mitochondrion, Neuroinflammatory Response, Regulation Of Tolerance Induction, Dopaminergic Neuron Differentiation. However, Microbody Membrane, RNA Cap Binding Complex, Vesicle Fusion with Golgi Apparatus, Ribosome Disassembly, Ribonucleoprotein Complex_Biogenesis, Ncrna Processing, Spliceosomal Complex Assembly, Amide Biosynthetic Process, Palmitoyltransferase Complex, Ribonucleoprotein Complex, RNA Splicing Via Transesterification Reactions were more abundant in Cluster2

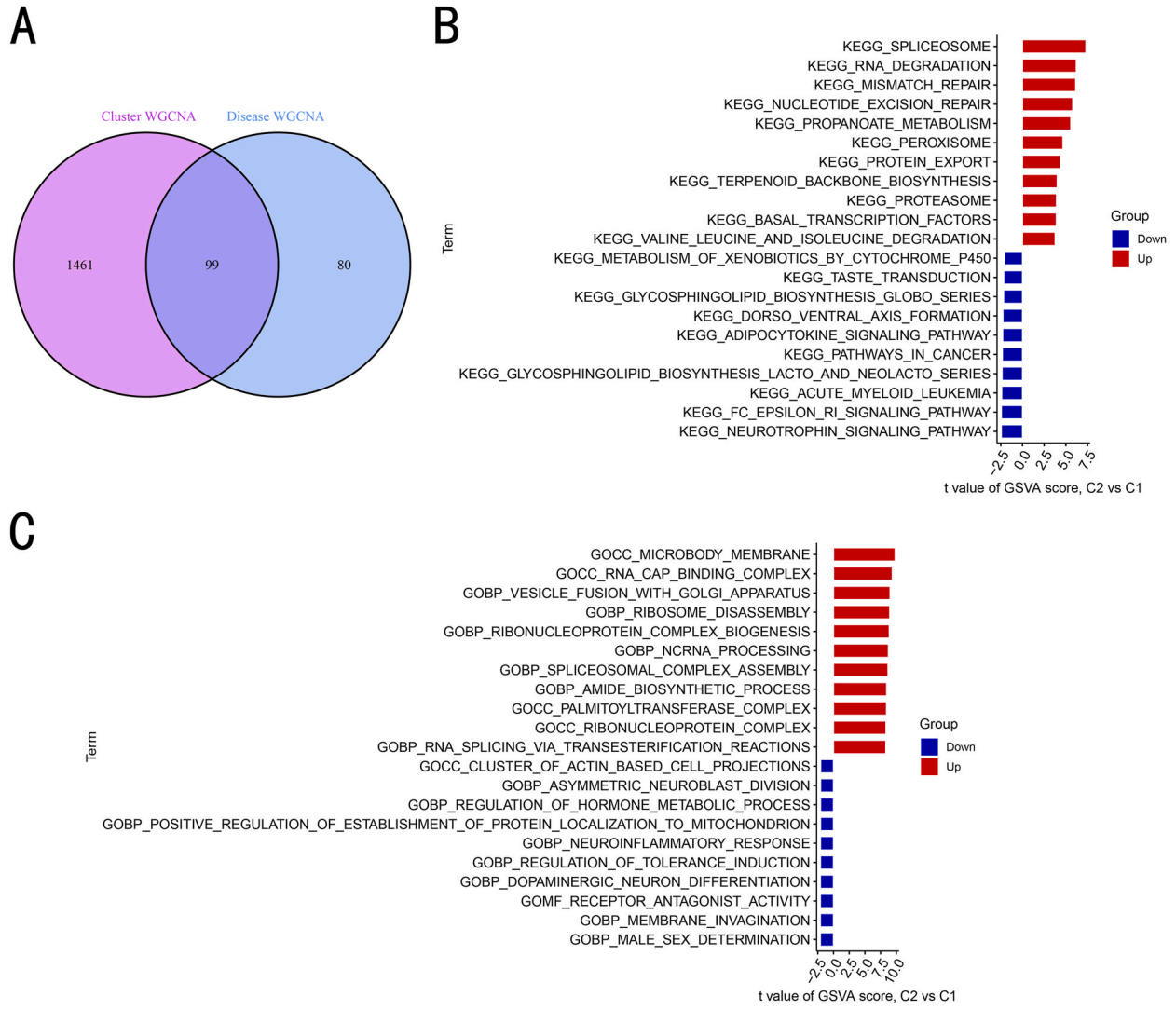

Fig. 6. Identification of biological traits and DEGs specific to each cluster between two cuproptosis clusters. (A) The points where the GSE56815 dataset's module-related genes and cuproptosis cluster-related genes intersect. (B) The GSVA method's t-values are used to rank differences in signature pathway activity between Cluster1 and Cluster2 samples. (C) Ranking by t-value of the GSVA method of biological function differences between Cluster1 and Cluster2 samples.

(Fig. 6C). As a result, we surmised that Cluster2 might be implicated in a range of immunological reactions.

3.4. Construction and assessment of machine learning models

We built four tested machine learning models [Support Vector Machine Model (SVM), Random Forest Model (RF), eXtreme Gradient Boosting (XGB) and Generalized Linear Model (GLM)]. Subtype-specific genes of high diagnostic value were further distinguished based on the expression profiles of 99 cluster-specific DEGs in the OP training cohort. The 'DALEX' software was used to visualise the distribution of residuals in the test set for each model and to interpret the 4 models. The machine learning model RF and GLM showed relatively reduced residuals (Fig. 7A and B). The top 10 significant characteristic variants of each model were then ranked with regard to their root mean square error (RMSE) (Fig. 7C). In the meantime, we calculated receiver operating characteristic (ROC) curves based on 5-fold cross-validation to evaluate the discriminatory performance of the four machine learning algorithms in the test set. The maximum area under the ROC curve (AUC) is shown by the RF machine learning models (GLM, AUC = 0.774; SVM, AUC = 0.694; RF, AUC = 0.750; XGB, AUC = 0.729, Fig. 7B). Collectively, these outcomes suggest that the RF model was the most efficient in distinguishing between patients in different clusters. The top five influences (EVA1B, RTCB, HEXB, SLC25A11 and TMEM126B) were then selected from the RF model as predictive gene but still needs further study.

We first created a nomogram to calculate the exposure to the cupped cell population in 40 OP patients to evaluate the RF model's

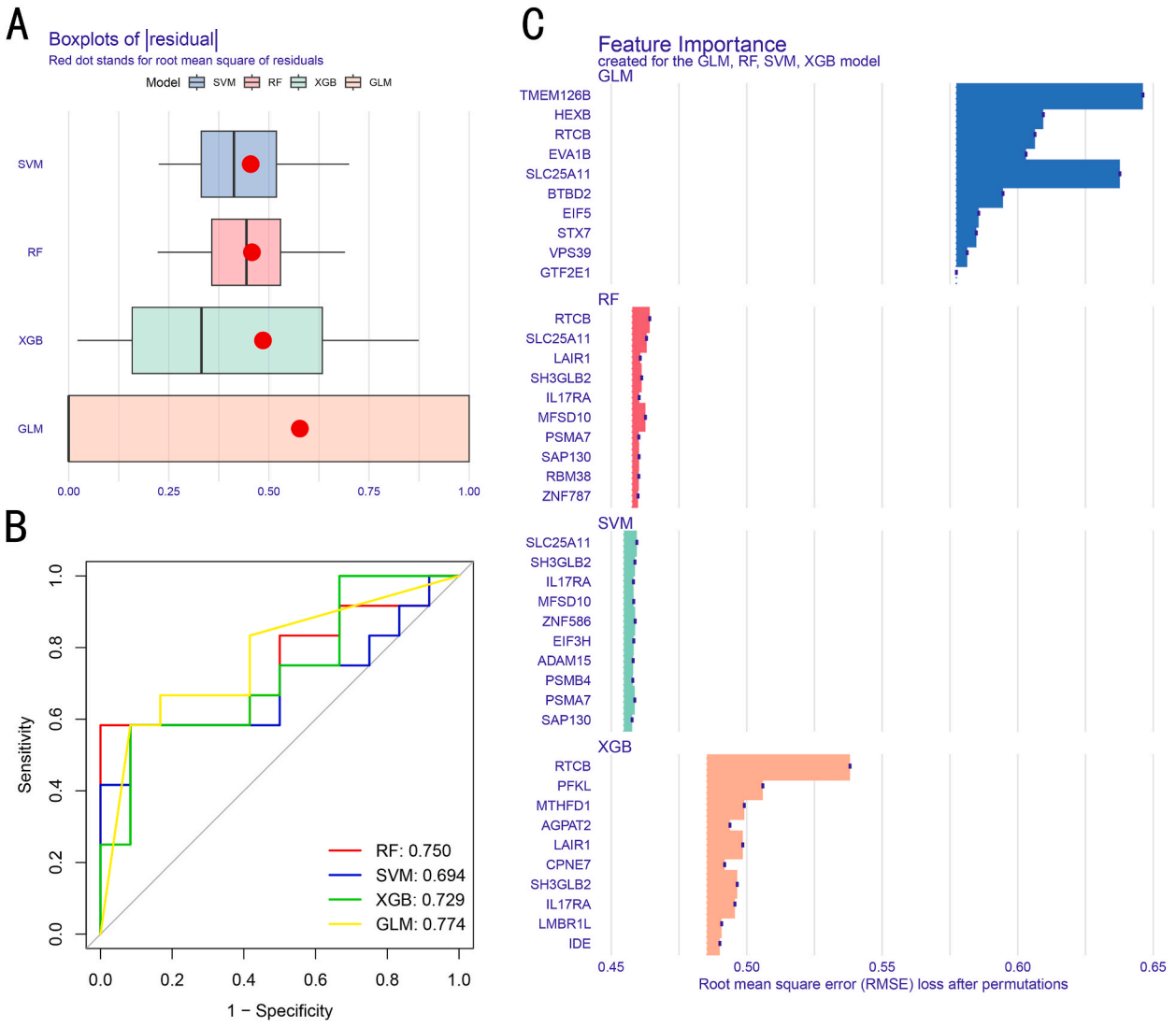

Fig. 7. Establishment and estimation of SVM, RF, XGB and GLM machine models. (A) The residuals for an individual computer learning model are shown by Boxplots. The Red spots denote the root mean square of the residuals (RMSE). (B) ROC interpretation of four computer learning models based upon 5-fold cross-validation in the testing cohort. (C) The distinguishing features in SVM, RF, XGB, and GLM machine models. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

predictive effect comprehensively (Fig. 8A). The prominent pathologic diagnostic value of the 5-gene prediction model was demonstrated by this result.

The predictive effect of the nomogram models was assessed using calibration curves and decision curve analysis (DCA). The calibration curves showed that the difference between the actual and expected risk associated with the OP cluster was very slight (Fig. 8B), and DCA showed that our nomograms were highly accurate and could potentially be used as a basis for making a clinical decision (Fig. 8C). We next confirmed our 5-gene forecasting model using two external mononuclear cell datasets, containing those from normal participants and OP patients. the AUC value for the GSE7158 dataset was 0.750 and the ROC curve showed that the 5-gene prediction model performed satisfactorily (Fig. 8D), indicating that our diagnosis model was equally effective in distinguishing OP from healthy individuals.

4. Discussion

Various factors are regarded as culprits for osteoporosis development nowadays, even though, oxidative metabolism in OP has advanced significantly over the past few decades, and the by-products of the energy generation reaction are mainly reactive oxygen species (ROS), which are mainly produced in mitochondria, and the accumulation of ROS levels is deleterious because they interact mainly with molecules including proteins, RNA, and DNA, which leads to the inhibition of the osteogenic lineage, thus, the reduction of ROS levels can be helpful by acting as signaling molecules that are essential for maintaining a balance between cell differentiation, self-renewability, and proliferation [[19–21]]. Therefore, mitochondrial malfunction has also been suggested to be a key factor in the etiology of bone illnesses as a loss in mitochondrial dynamic causes an imbalance in the process of bone remodeling [22]. What's more,

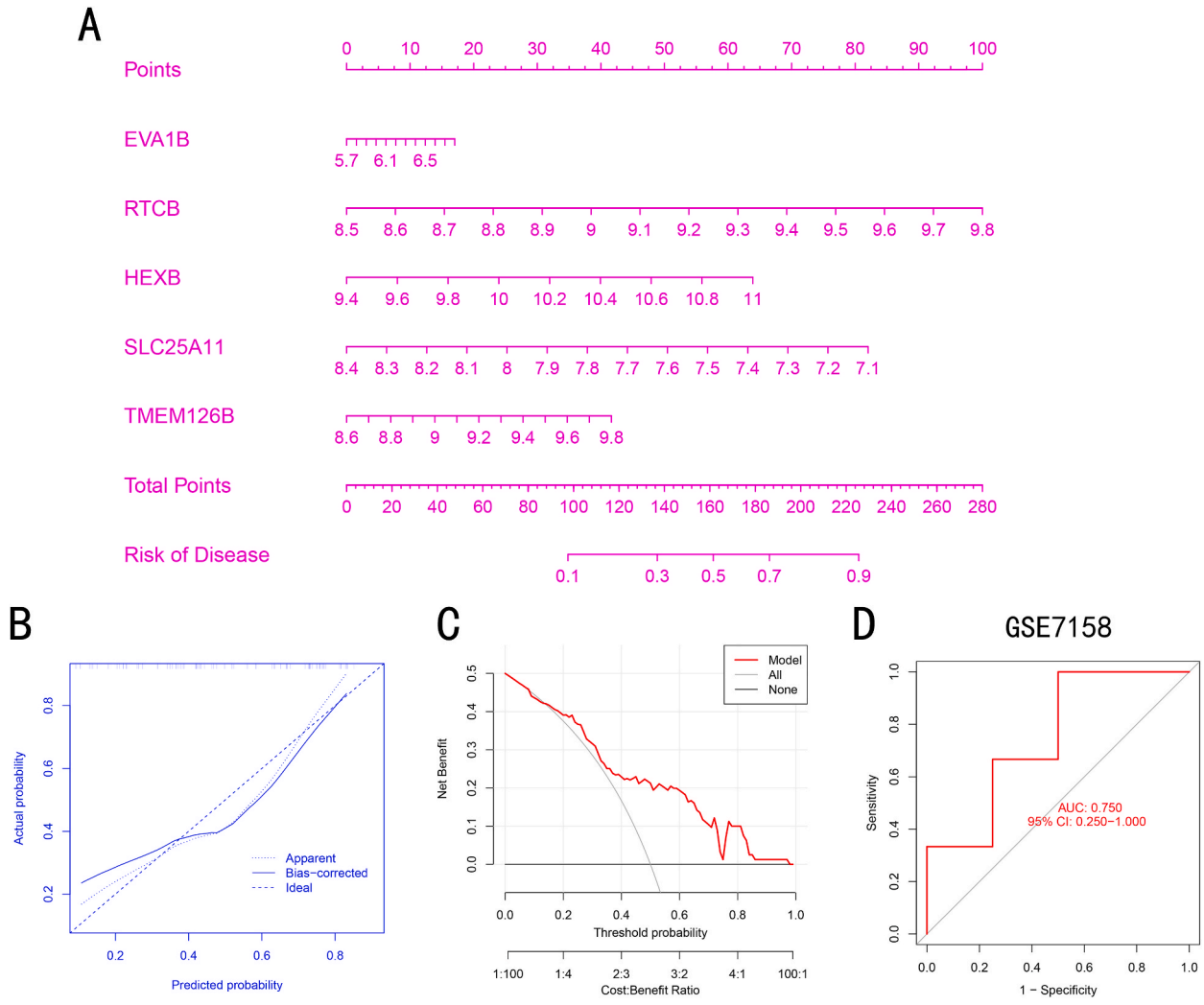

Fig. 8. Verification of the RF model centre on the 5-gene. (A) Structuring nomograms to predict the risk of the 5-gene accounting RF model for the OR cohort. (B,C) calibration curves construction (B) and DCA (C) to assess the predictability of the nomogram model. (D) a 5-fold cross-validated 5-gene RF-based model in GSE7158 for ROC analysis.

dysregulation of intracellular copper bioavailability can provoke oxidative stress and cytotoxicity.

According to the reports, mitochondrial stress elicits copper death by a process that results in excess accumulation of lipid acylated mitochondrial enzymes and the depletion of Fe–S cluster proteins [[23,24]]. Under normal conditions, Cu^{2+} entering the cell are converted to Cu^{+} by FDX1, which stimulates lipid acylation of mitochondrial proteins and increases the synthesis of important enzymes involved in the mitochondrial tricarboxylic acid (TCA) cycle (GCSH, DBT, DLAT and DLST). Additionally, FDX1 and the unsteadiness of proteins in the Fe–S cluster are strongly connected. Besides, crucial genes involved in the transport of Cu^{2+}, like SLC31A1 and ATP7B, may be crucial in controlling the occurrence of cuproptosis. Furthermore, it was discovered that cuproptosis-related harm might be reduced by inhibiting the activity of the mitochondrial pyruvate carrier as well as electron transport chains [25]. Since cuproptosis affects the TCA cycle of mitochondria, it can be predicted that cuproptosis is involved in osteoporosis as well. As mitochondria provide energy for osteoblasts, and the byproducts of oxidative phosphorylation, such as ROS, can be detrimental to them. Another reason may be that αKG, an important protein in the TCA cycle, ameliorates age-related osteoporosis [26].The precise mechanisms of cuproptosis in OP and its managing acts in different disorders have not been further studied. We hence tried to explicate the meticulous function of genes related to cuproptosis in the phenotyping of OP and the immunological microenvironment.

For the first time in our investigation, we thoroughly compared the expression of cuproptosis controllers in blood from healthy individuals and OP patients. More dysregulated CRGs were detected in OP patients than in healthy controls, pointing to a crucial function for CRGs in the development of OP. To better elucidate the connection between cup mutation modulators and OP, we then computed the relevance between CRGs. As shown by the interplay of CRGs present in OP patients, we found synergistic or antagonistic effects of various cuproptosis mutation modulators. Between control people and OP patients, there were differences in the number of immune cells present. Mast cells resting, M0 macrophages, M2 macrophages, activated dendritic cells, and neutrophils levels were higher in OP patients. In the meantime, anchored on the CRG expression landscape, we applied unsupervised clustering analysis to illustrate the diverse cupulocyte regulatory patterns in OP patients and found two independent cupulocyte-associated clusters. Cluster2 showed higher immunological scores and comparatively more immune infiltration. According to cluster-specific DEGs, Cluster1 was largely enriched in biological processes linked to adipocytokine, neuroblast, and mitochondrial location, while Cluster2 was distinguished by immune cell activation and ribosome. Growing evidence reinforces the role of bone marrow adipose tissue as an influential factor in the bone microenvironment. For example, Oxylipin-PPARg initiates adipocyte senescence, thus regulating bone deterioration [27]. Therefore, cuproptosis may affect osteoporosis through adipocytokine. Additionally, a novel neuropeptide called Cyp40 was found to play a critical role in osteoclastogenesis via the aryl hydrocarbon receptor (AhR)-Ras/Raf-*p*-Erk-NFATc1 pathway, improving bone quality in osteoporosis [28]. This suggests that cuproptosis may also influence osteoporosis through the nervous system.When considered as a whole, it would make sense to assume that Cluster 2 had more activated B cells and T cells to slow the course of OP and hence had a better prognosis.

In recent years, tools for predicting the prevalence of OP have been extended to machine learning models of demographic and imaging indicators [29]. These studies demonstrated that multifactor analyses have fewer errors and greater reliability of results than univariate analyses because they took into account the connection between variables. In the present study, we compared the predictive power of four different machine learning classifiers (SVM,RF, XGB and GLM) and built an RF-based prediction model accounting the expression profile of cluster-specific DEGs, which showed the highest forecasting validity in the test cohort (AUC = 0.9829), illustrating that RF-based computer learning in predicting subtypes of OP has adequate performance. To build a five-gene RF-based model, we then chose five key variables (EVA1B, RTCB, HEXB, SLC25A11 and TMEM126B). EVA1B known as the protein-coding gene, is an essential paralog of the EVA1A gene. It is a protein important for the endoplasmic reticulum and lysosomes that has a role in autophagy and apoptosis [[30–32]] As OP is associated with autophagy and apoptosis, there may be some correlation between OP and EVA1B, although the relationship between them haven't been elucidated yet. RTCB was initially discovered as a human tRNA splicing ligase complex's key subunit [[33,34]]. Later, it was shown that RTCB is also in charge of ligating the XBP1 mRNA exons following the intron's removal by IRE1 as part of the unfolded protein response (UPR)[[35–37]]. Although there still lack of directed research on the effect of RTCB on OP, relationship still can be found as tRNA splicing ligase and UPR regulates the homeostasis during OP. Hexosaminidase B (β-polypeptide) hexosaminidase (glycosylhydrolase) HEXB may play a role in promoting the degradation of glycosphingolipids on the surface of aged BMMSCs, plus affect lipid raft formation and influence signaling [38]. SLC25A11 is is a mitochondrial malate-α-ketoglutarate carrier family protein that is one of two retrotransporters of the malate-aspartate shuttle [39], which may be associated with OP though little evidence has been provided. TMEM126B, a transmembrane protein, is a component in complex I assembly. The mitochondrial complex I assembly complex is made up of evolutionary conserved signaling intermediate in toll pathway (ECSIT), acyl-CoA dehydrogenase family member 9 (ACDH), and complex I intermediate-associated protein (TMEM126B) [[40,41]]. The Q-module is easier to collect and incorporate into complex I of the mitochondrial respiratory chain thanks to this assembly complex and the translocase of inner mitochondrial membrane domain containing protein [42]. Thus, TMEM126B contributes to the 45 subunits that make up complex I having the right functional makeup [[43,44]]. Complex I inadequacies and a reduced ability for cellular respiration are linked to TMEM126B defects [[45,46]]. Chronic hypoxia will cause TMEM126B degradation and partial inactivation of complex I [47]. And the function of TMEM126B maybe involved in OP. New information on the diagnosis of OP is provided by the 5-gene-based RF model's ability to correctly predict OP in two external validation datasets (AUC = 0.8529 and 0.8333). More significantly, we used the EVA1B, RTCB, HEXB, SLC25A11, and TMEM126B to create a nomogram model for the diagnosis of OP subgroups. We discovered that this prediction model demonstrated astounding predictive efficacy, demonstrating its value for clinical applications.

5. Conclusion

Overall, our study unveiled an association between CRGs and infiltrated immune cells and proved significant immune heterogeneity between OP patients with significant clusters of cuproptosis degeneration. The 5-gene based RF model was chosen as the best computer learning model which can precisely evaluate the pathological results of OP subtypes and OP patients. Our study identifies the role of cuproptosis in OP and essentially illustrates the underlying molecular mechanisms that lead to OP heterogeneity. For future studies, it is recommended to collect additional parameters such as demographics or bone density and incorporate them into the prediction models. Additionally, a multicenter study design should be employed to validate and optimize the predictive models in different populations.

Data availability statement

The dataset underpinning the findings of this paper can be found on the GEO website with the following data accession identifiers. GSE56815 and GSE7158.

Conflict of interest

The authors state that the investigation was carried out without any commercial or financial relationship, which could be interpreted as a possible conflict of interest.

Additional information

No additional information is available for this paper.

CRediT authorship contribution statement

Tongying Chen: Writing – review & editing, Writing – original draft, Resources, Investigation, Data curation. **Zhijie Gao:** Writing – original draft, Validation, Methodology, Formal analysis. **Yuedong Wang:** Writing – review & editing, Validation, Supervision, Software, Investigation. **Jiachun Huang:** Visualization. **Shuhua Liu:** Visualization. **Yanping Lin:** Visualization. **Sai Fu:** Writing – review & editing. **Lei Wan:** Visualization, Conceptualization. **Ying Li:** Writing – review & editing. **Hongxing Huang:** Supervision, Funding acquisition, Conceptualization. **Zhihai Zhang:** Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Guangzhou Science and Technology Program (Project No.: 202102010124), National Natural Science Foundation of China (82174395, 82274551), Natural Science Foundation of Guangdong Province(2022A1515012067, 2022A1515011404)

References

[1] G. Ballane, J.A. Cauley, M.M. Luckey, G. El-Hajj Fuleihan, Worldwide prevalence and incidence of osteoporotic vertebral fractures, Osteoporos. Int. 28 (2017) 1531–1542.
[2] J.S. Kimball, J.P. Johnson, D.A. Carlson, Oxidative stress and osteoporosis, J Bone Joint Surg Am 103 (2021) 1451–1461.
[3] Z. Li, D. Li, R. Chen, S. Gao, Z. Xu, N. Li, Cell death regulation: a new way for natural products to treat osteoporosis, Pharmacol. Res. 187 (2023) 106635.
[4] S.R. Li, L.L. Bu, L. Cai, Cuproptosis: lipoylated TCA cycle proteins-mediated novel cell death pathway, Signal Transduct Target Ther 7 (2022) 158.
[5] A. Reshma, T. Tamilanban, V. Chitra, V. Subramaniyan, G. Gupta, N.K. Fuloria, M. Sekar, S. Fuloria, R. Sahu, J. Narayanan, S. Chakravarthy, S. Selvaraj, Anti-obesity effects of olivetol in adult zebrafish model induced by short-term high-fat diet, Sci. Rep. 13 (2023) 18449.
[6] L. Chen, J. Min, F. Wang, Copper homeostasis and cuproptosis in health and disease, Signal Transduct Target Ther 7 (2022) 378.
[7] V. Subramaniyan, N.S.A. Lubau, N. Mukerjee, V. Kumarasamy, Alcohol-induced liver injury in signalling pathways and curcumin's therapeutic potential, Toxicol Rep 11 (2023) 355–367.
[8] N.R. Rarokar, S.D. Saoji, N.V. Deole, M. Gaikwad, A. Pandey, C. Kamaraj, S.V. Chinni, V. Subramaniyan, G. Ramachawolran, S. Dharashivkar, Preparation and formula optimization of cephalexin loaded transferosomal gel by QbD to enhance the transdermal delivery: in vitro, ex vivo and in vivo study, J. Drug Deliv. Sci. Technol. 89 (2023) 104968.
[9] T. Santhoshkumar, R.K. Govindarajan, C. Kamaraj, N.S. Alharbi, K. Manimaran, D.H.Y. Yanto, V. Subramaniyan, K.-H. Baek, Biological synthesis of nickel nanoparticles using extracellular metabolites of Bacillus Sphaericus: characterization and vector-borne disease control applications, South Afr. J. Bot. 162 (2023) 481–494.
[10] X. Tong, R. Tang, M. Xiao, J. Xu, W. Wang, B. Zhang, J. Liu, X. Yu, S. Shi, Targeting cell death pathways for cancer therapy: recent developments in necroptosis, pyroptosis, ferroptosis, and cuproptosis research, J. Hematol. Oncol. 15 (2022) 174.
[11] V. Pavithra, T.G. Sathisha, K. Kasturi, D.S. Mallika, S.J. Amos, S. Ragunatha, Serum levels of metal ions in female patients with breast cancer, J. Clin. Diagn. Res. 9 (2015) 25–27.

[12] M.R. Lener, R.J. Scott, A. Wiechowska-Kozlowska, P. Serrano-Fernandez, P. Baszuk, K. Jaworska-Bieniek, G. Sukiennicki, W. Marciniak, M. Muszynska, J. Kladny, T. Gromowski, K. Kaczmarek, A. Jakubowska, J. Lubinski, Serum concentrations of selenium and copper in patients diagnosed with pancreatic cancer, Cancer Res Treat 48 (2016) 1056–1064.
[13] A.K. Baltaci, T.K. Dundar, F. Aksoy, R. Mogulkoc, Changes in the serum levels of trace elements before and after the operation in thyroid cancer patients, Biol. Trace Elem. Res. 175 (2017) 57–64.
[14] X.L. Zuo, J.M. Chen, X. Zhou, X.Z. Li, G.Y. Mei, Levels of selenium, zinc, copper, and antioxidant enzyme activity in patients with leukemia, Biol. Trace Elem. Res. 114 (2006) 41–53.
[15] L. Aubert, N. Nandagopal, Z. Steinhart, G. Lavoie, S. Nourreddine, J. Berman, M.K. Saba-El-Leil, D. Papadopoli, S. Lin, T. Hart, G. Macleod, I. Topisirovic, L. Gaboury, C.J. Fahrni, D. Schramek, S. Meloche, S. Angers, P.P. Roux, Copper bioavailability is a KRAS-specific vulnerability in colorectal cancer, Nat. Commun. 11 (2020) 3701.
[16] Y. Jin, C. Zhang, H. Xu, S. Xue, Y. Wang, Y. Hou, Y. Kong, Y. Xu, Combined effects of serum trace metals and polymorphisms of CYP1A1 or GSTM1 on non-small cell lung cancer: a hospital based case-control study in China, Cancer Epidemiol 35 (2011) 182–187.
[17] S.A.K. Saleh, H.M. Adly, A.A. Abdelkhaliq, A.M. Nassir, Serum levels of selenium, zinc, copper, manganese, and iron in prostate cancer patients, Curr. Urol. 14 (2020) 44–49.
[18] M. Baharvand, S. Manifar, R. Akkafan, H. Mortazavi, S. Sabour, Serum levels of ferritin, copper, and zinc in patients with oral cancer, Biomed. J. 37 (2014) 331–336.
[19] T. Zhou, Y. Yan, C. Zhao, Y. Xu, Q. Wang, N. Xu, Resveratrol improves osteogenic differentiation of senescent bone mesenchymal stem cells through inhibiting endogenous reactive oxygen species production via AMPK activation, Redox Rep. 24 (2019) 62–69.
[20] N. Mujafarkani, F.M.M. Ahamed, K.S. Babu, S. Debnath, A.A. Sayed, G.M. Albadrani, M.Q. Al-Ghadi, V. Kumarasamy, V. Subramaniyan, C. Kamaraj, M. M. Abdel-Daim, Unveiling a novel terpolymer-metal complex: a detailed exploration of synthesis, characterization, and its potential as an antimicrobial and antioxidant agent, Heliyon 9 (2023) e20459.
[21] R. Thapa, O. Afzal, G. Kumar, A.A. Bhat, W.H. Almalki, S.I. Alzarea, I. Kazmi, A.S.A. Altamimi, V. Subramaniyan, L. Thangavelu, Unveiling the connection: long-chain non-coding RNAs and critical signaling pathways in breast cancer, Pathol. Res. Pract. (2023) 154736.
[22] C.X. Zheng, B.D. Sui, X.Y. Qiu, C.H. Hu, Y. Jin, Mitochondrial regulation of stem cells in bone homeostasis, Trends Mol. Med. 26 (2020) 89–104.
[23] V. Oliveri, Selective targeting of cancer cells by copper ionophores: an overview, Front. Mol. Biosci. 9 (2022) 841814.
[24] Y. Wang, L. Zhang, F. Zhou, Cuproptosis: a new form of programmed cell death, Cell. Mol. Immunol. 19 (2022) 867–868.
[25] P.A. Cobine, D.C. Brady, Cuproptosis: cellular and molecular mechanisms underlying copper-induced cell death, Mol Cell 82 (2022) 1786–1787.
[26] Y. Wang, P. Deng, Y. Liu, Y. Wu, Y. Chen, Y. Guo, S. Zhang, X. Zheng, L. Zhou, W. Liu, Alpha-ketoglutarate ameliorates age-related osteoporosis via regulating histone methylations, Nat. Commun. 11 (2020) 5596.
[27] X. Liu, Y. Gu, S. Kumar, S. Amin, Q. Guo, J. Wang, C.-L. Fang, X. Cao, M. Wan, Oxylipin-PPARγ-initiated adipocyte senescence propagates secondary senescence in the bone marrow, Cell Metabol. 35 (2023) 667–684. e6.
[28] J. Li, B. Liu, H. Wu, S. Zhang, Z. Liang, S. Guo, H. Jiang, Y. Song, X. Lei, Y. Gao, Sensory nerves directly promote osteoclastogenesis by secreting peptidyl-prolyl cis-trans isomerase D (Cyp40), Bone Research 11 (2023) 64.
[29] J. Smets, E. Shevroja, T. Hugle, W.D. Leslie, D. Hans, Machine learning solutions for osteoporosis-A review, J. Bone Miner. Res. 36 (2021) 833–851.
[30] S. Zhao, H. Wang, EVA1A plays an important role by regulating autophagy in physiological and pathological processes, Int. J. Mol. Sci. 22 (2021).
[31] F.A. Rizwi, M. Abubakar, E.R. Puppala, A. Goyal, C.V. Bhadrawamy, V.G.M. Naidu, S. Roshan, B. Tazneem, W.H. Almalki, V. Subramaniyan, S. Rawat, G. Gupta, Janus kinase-signal transducer and activator of transcription inhibitors for the treatment and management of cancer, J. Environ. Pathol. Toxicol. Oncol. 42 (2023) 15–29.
[32] S. Mukherjee, S. Nag, N. Mukerjee, S. Maitra, R. Muthusamy, N.K. Fuloria, S. Fuloria, M.D. Adhikari, K. Anand, N. Thorat, V. Subramaniyan, S. Gorai, Unlocking exosome-based theragnostic signatures: deciphering secrets of ovarian cancer metastasis, ACS Omega 8 (2023) 36614–36627.
[33] Y. Akiyama, Y. Takenaka, T. Kasahara, T. Abe, Y. Tomioka, P. Ivanov, RTCB complex regulates stress-induced tRNA cleavage, Int. J. Mol. Sci. 23 (2022).
[34] J. Popow, M. Englert, S. Weitzer, A. Schleiffer, B. Mierzwa, K. Mechtler, S. Trowitzsch, C.L. Will, R. Lührmann, D. Söll, J. Martinez, HSPC117 is the essential subunit of a human tRNA splicing ligase complex, Science 331 (2011) 760–764.
[35] Y. Lu, F.X. Liang, X. Wang, A synthetic biology approach identifies the mammalian UPR RNA ligase RtcB, Mol Cell 55 (2014) 758–770.
[36] S.G. Kosmaczewski, T.J. Edwards, S.M. Han, M.J. Eckwahl, B.I. Meyer, S. Peach, J.R. Hesselberth, S.L. Wolin, M. Hammarlund, The RtcB RNA ligase is an essential component of the metazoan unfolded protein response, EMBO Rep. 15 (2014) 1278–1285.
[37] J. Jurkin, T. Henkel, A.F. Nielsen, M. Minnich, J. Popow, T. Kaufmann, K. Heindl, T. Hoffmann, M. Busslinger, J. Martinez, The mammalian tRNA ligase complex mediates splicing of XBP1 mRNA and controls antibody secretion in plasma cells, EMBO J. 33 (2014) 2922–2936.
[38] S.S. Jiang, C.H. Chen, K.Y. Tseng, F.Y. Tsai, M.J. Wang, I.S. Chang, J.L. Lin, S. Lin, Gene expression profiling suggests a pathological role of human bone marrow-derived mesenchymal stem cells in aging-related skeletal diseases, Aging (Albany NY) 3 (2011) 672–684.
[39] Q. Zhang, Y. Tang, S. Sun, Q. Xie, J. Yao, X. Wang, J. Qian, Z. Li, An extensive bioinformatics study on the role of mitochondrial solute carrier family 25 in PC and its mechanism behind affecting immune infiltration and tumor energy metabolism, J. Transl. Med. 20 (2022) 592.
[40] H. Heide, L. Bleier, M. Steger, J. Ackermann, S. Dröse, B. Schwamb, M. Zörnig, A.S. Reichert, I. Koch, I. Wittig, U. Brandt, Complexome profiling identifies TMEM126B as a component of the mitochondrial complex I assembly complex, Cell Metab 16 (2012) 538–549.
[41] S. Guerrero-Castillo, F. Baertling, D. Kownatzki, H.J. Wessels, S. Arnold, U. Brandt, L. Nijtmans, The assembly pathway of mitochondrial respiratory chain complex I, Cell Metab 25 (2017) 128–139.
[42] B. Andrews, J. Carroll, S. Ding, I.M. Fearnley, J.E. Walker, Assembly factors for the membrane arm of human complex I, Proc Natl Acad Sci U S A 110 (2013) 18934–18939.
[43] D.C. Fuhrmann, I. Wittig, B. Brune, TMEM126B deficiency reduces mitochondrial SDH oxidation by LPS, attenuating HIF-1alpha stabilization and IL-1 beta expression, Redox Biol. 20 (2019) 204–216.
[44] C. Wirth, U. Brandt, C. Hunte, V. Zickermann, Structure and function of mitochondrial complex I, Biochim. Biophys. Acta 1857 (2016) 902–914.
[45] L. Sanchez-Caballero, B. Ruzzenente, L. Bianchi, Z. Assouline, G. Barcia, M.D. Metodiev, M. Rio, B. Funalot, M.A. van den Brand, S. Guerrero-Castillo, J. P. Molenaar, D. Koolen, U. Brandt, R.J. Rodenburg, L.G. Nijtmans, A. Rotig, Mutations in complex I assembly factor TMEM126B result in muscle weakness and isolated complex I deficiency, Am. J. Hum. Genet. 99 (2016) 208–216.
[46] C.L. Alston, A.G. Compton, L.E. Formosa, V. Strecker, M. Olahova, T.B. Haack, J. Smet, K. Stouffs, P. Diakumis, E. Ciara, D. Cassiman, N. Romain, J.W. Yarham, L. He, B. De Paepe, A.V. Vanlander, S. Seneca, R.G. Feichtinger, R. Ploski, D. Rokicki, E. Pronicka, R.G. Haller, J.L. Van Hove, M. Bahlo, J.A. Mayr, R. Van Coster, H. Prokisch, I. Wittig, M.T. Ryan, D.R. Thorburn, R.W. Taylor, Biallelic mutations in TMEM126B cause severe complex I deficiency with a variable clinical phenotype, Am. J. Hum. Genet. 99 (2016) 217–227.
[47] D.C. Fuhrmann, I. Wittig, S. Drose, T. Schmid, N. Dehne, B. Brune, Degradation of the mitochondrial complex I assembly factor TMEM126B under chronic hypoxia, Cell. Mol. Life Sci. 75 (2018) 3051–3067.