Dovepress
Taylor & Francis Group

ORIGINAL RESEARCH

Unveiling Cuproptosis-Driven Molecular Clusters and Immune Dysregulation in Ankylosing Spondylitis

Bowen Wei[1,2,]*, Siwei Wang[1,2,]*, Suiran Li[1,2,]*, Qingxiang Gu[3], Qingyun Yue 🄳[4], Zhuo Tang[1,2], Jiamin Zhang[1,2], Wei Liu 🄳[1,2]

[1]Department of Rheumatism and Immunity, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, People's Republic of China; [2]National Clinical Research Center for Chinese Medicine Acupuncture and Moxibustion, Tianjin, People's Republic of China; [3]Center of Preventive Treatment of Disease, Cangzhou Hospital of Integrated Traditional Chinese and Western of Hebei Province, Cangzhou, Hebei, People's Republic of China; [4]School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, People's Republic of China

*These authors contributed equally to this work

Correspondence: Wei Liu, Email fengshiliuwei@163.com

Background: Ankylosing spondylitis (AS) is a chronic autoimmune disease characterized by inflammation of the sacroiliac joints and spine. Cuproptosis is a newly recognized copper-induced cell death mechanism. Our study explored the novel role of cuproptosis-related genes (CRGs) in AS, focusing on immune cell infiltration and molecular clustering.

Methods: By analyzing the peripheral blood gene expression datasets obtained from GSE73754, GSE25101, and GSE11886, we identified the expression patterns of cellular factors and immune infiltration cell related to cuproptosis. Subsequently, we employed weighted gene co-expression network analysis (WGCNA) to identify differentially expressed genes (DEGs) within each cluster and utilized the "GSVA" and "GSEABase" software packages to examine variations in gene sets enriched across various CRG clusters. Finally, we selected the best-performing machine learning model to predict genes associated with AS. Datasets (GSE25101 and GSE73754) and ELISA to assess the expression levels of the five genes and their corresponding proteins.

Results: Seven cuproptosis-related DEGs and four immune cell types were identified, revealing significant immune heterogeneity in the immune cell infiltration between the two cuproptosis-related molecular clusters in AS. The eXtreme Gradient Boosting (XGB) model showed the highest predictive accuracy, achieving an area under the receiver operating characteristic curve (AUC) of 0.725, and 5-gene prediction models were established. It showed satisfactory performance in the GSE25101 dataset (AUC = 0.812). According to the blood serum samples of AS patients and controls, PELI1 had a higher expression level (AUC = 0.703, $p = 0.07$), while ICAM2 and RANGAP1 had lower expression levels (AUC = 0.724, 0.745, and $p = 0.011$, 0.000, respectively) in AS patients.

Conclusion: We explored the correlation of cuproptosis in AS, and developed the optimal machine learning model to identify high-risk genes associated with AS. We also explored the pathogenesis and treatment strategies of AS, targeting *PELI1, ICAM2*, and *RANGAP1*.

Keywords: ankylosing spondylitis, cuproptosis, immune cell infiltration, machine learning model

Introduction

Ankylosing spondylitis (AS) is a chronic autoimmune disease marked by persistent inflammation,[1] with the latest statistics indicating a global prevalence rate of approximately 0.1–1.4%.[2] It is mainly marked by inflammation in the sacroiliac joints or spine, resulting in structural damage, abnormal bone remodeling, and spinal rigidity. Clinically, its main manifestations include arthritis, enthesitis, and other peripheral symptoms, and is also linked to conditions such as uveitis, psoriasis, and inflammatory bowel disease.[3] Long-term disease progression can result in characteristic postural changes, leading to a significant reduction in the patient's quality of life and placing a considerable financial burden on them.[4]

Currently, the exact pathogenesis of AS is not fully understood. It is generally believed to be influenced by genetics, infections, immune responses, and environmental factors.[5] Genetic factors are crucial in determining susceptibility to AS, and about 90% of the susceptibility can be attributed to it.[6] Since Brewerton and Schlosstein found that human leukocyte antigen B27 gene (*HLA-B27*) was highly correlated with AS, *HLA-B27* was the most closely related known gene with AS.[7,8] Although more and more AS susceptibility genes have been found, such as Endoplasmic reticulum aminopeptidase 1 (*ERAP1*), Interleukin-23 receptor (*IL-23R*), Anthrax toxin receptor 2 (*ANTXR2*), Interleukin-1 receptor type 2 (*IL-1R2*), etc,[6] at present, the pathogenic gene loci have not been identified and there is a lack of specific biomarkers.

Studies have reported a link between the occurrence of AS and both apoptosis and necrosis.[9,10] And cuproptosis is the copper (Cu) induced cell death mode discovered by Tsvetkov P et al in 2022 and named for the first time.[11] Cu, a vital trace mineral for the human body, is a necessary co-factor for many enzymes related to basic cell functions.[12] Cu is transported to specific subcellular compartments, including the mitochondria, trans-Golgi network, and nucleus, via various protein carriers, such as cytochrome c oxidase 17 (COX17), copper chaperone for superoxide dismutase (CCS) and antioxidant 1 (Atox1). And then participate in the synthesis of biological compounds such as mitochondrial energy metabolism, redox homeostasis, and neurotransmitter metabolism.[13,14] At the same time, the dysregulation of storage and utilization of Cu can induce cytotoxicity and oxidative stress, leading to the occurrence of various diseases.[12] Cuproptosis represents a unique form of cell death that relies on mitochondrial respiration. Tsvetkov P et al used Cu ionophores, such as elesclomol (ES), to explore the mechanism of Cu ion cytotoxicity. A study revealed that cells with a high dependence on mitochondrial respiration exhibit increased sensitivity to ES treatment, and cuproptosis serves a critical function in the tricarboxylic acid (TCA) cycle.[11] Research has shown that serum Cu levels are significantly higher in patients with AS, and the extent of this increase is positively correlated with disease severity.[15,16] Therefore, we hypothesized that excessive Cu accumulation may exert cytotoxic effects and regulate the expression of related genes, affecting the immune function of AS patients. However, the relevant mechanism of cuproptosis in AS is still unclear and needs further exploration.

Despite progress in identifying genetic predispositions (eg, *HLA-B27*), specific biomarkers for AS remain limited. This study aimed to bridge this gap by investigating cuproptosis-related genes (CRGs), which represent an underexplored pathway in AS. Identifying CRGs not only deepens understanding of AS pathogenesis but may also lead to more precise diagnostic tools or therapeutic strategies.

In this work, we used the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) to conduct a comprehensive analysis of the differential expression and immune profiles of CRGs between AS patients and controls, with the goal of further investigating the potential pathogenic mechanisms underlying AS. This analysis focused on the main cuproptosis-related differentially expressed genes (CuDEGs). Then we divided 76 AS patients into two clusters according to the characteristics of cuproptosis and examined the variations in immune cells between the clusters to identify specific patterns or correlations. Next, weighted gene co-expression network analysis (WGCNA) was applied to identify DEGs of specific clusters. This process involves not only the analysis of gene expression but also the study of immune cell differences in different clusters to determine the association between AS and cuproptosis. In addition, a variety of machine learning algorithms were employed to construct predictive models, with the aim of identifying distinct molecular clusters. Meanwhile, to verify the connection between cuproptosis and AS, the relationship between the 5 hub genes and AS was further confirmed via the enzyme-linked immunosorbent assay (ELISA).

Materials and Methods
Data Acquisition and Preprocessing
Using "ankylosing spondylitis" as the keyword, gene expression data from the GEO database, encompassing both AS patients and healthy individuals, was retrieved for analysis. The selected datasets included GSE73754 (GPL10558 platform, 52 AS patients and 20 controls), GSE25101 (GPL6947 platform, 16 AS patients and 16 controls), and GSE11886 (GPL570 platform, 8 AS patients and 9 controls). During the analysis, the three datasets were merged with the "sva" package for further analysis. Among them, GSE25101 and GSE73754 datasets were also independently used for validation. Prior to analysis, all data were transformed into log2 format. Based on the findings of Tsvetkov et al,

a total of 19 CRGs were identified, including *NFE2L2, NLRP3, ATP7B, ATP7A, SLC31A1, FDX1, LIAS, LIPT1, LIPT2, DLD, DLAT, PDHA1, PDHB, MTF1, GLS, CDKN2A, DBT, GCSH* and *DLST*.[11,17–20]

Immune Cell Infiltration Assessment

In this study, the "e1071" and "preprocesscore" packages were used to create CIBERSORT files. The CIBERSORT algorithm (https://cibersortx.stanford.edu/) was applied via the LM22 signature matrix to estimate the relative abundance of 22 immune cell types within each sample. Monte Carlo sampling was employed within CIBERSORT to calculate the inverse multiple product *p*-value for each sample, with those showing $p < 0.05$ considered valid and reliable for immune cell composition analysis.[21] Heatmaps were generated using the "pheatmap" package, and boxplots were created with the "reshape2" and "ggpubr" packages, to clearly show the data distribution and statistical characteristics.

Correlation Analysis Between CRGs and Infiltrating Immune Cells (IICs)

This study examined the correlation between CRG expression levels and the proportional distribution of IICs. When the correlation co-efficient was $p < 0.05$, it was considered that the correlation was significant. The findings were ultimately depicted using the "reshape2", "tidyverse", and "ggplot2" packages.

Unsupervised Clustering and Principal Component Analysis (PCA) of AS

Unsupervised clustering was conducted based on the expression profiles of 19 CRGs, using the ConsensusClusterPlus software ("ConsensusClusterPlus" package, version 2.60) with 1000 iterations.[22] By applying the k-means algorithm, 76 AS samples were classified into different clusters. The optimal number of clusters was determined by analyzing the cumulative distribution function (CDF) curve, consensus matrix, and achieving a cluster-consensus score above 0.9, setting the maximum number of subtypes between k = 2-9. PCA was then conducted to highlight the differences among the identified subtypes.

WGCNA

The co-expression module was delineated using the "WGCNA" package (version 1.70.3).[23] To enhance result accuracy, subsequent WGCNA concentrated on the top 25% of genes with the highest variability. An optimal soft threshold (soft power) was selected to construct a weighted adjacency matrix, which was then transformed into a topological overlap matrix (TOM). With a minimum module size of 100, modules were derived using the TOM dissimilarity measure (1-TOM) via hierarchical clustering. Each module was randomly assigned a unique color. The module eigengene represented the overall expression pattern within each module, illustrating the connection between module members (MM) and disease states. The term "gene significance" was used to describe the association between AS genes and clinical phenotypes.

Gene Set Variation Analysis (GSVA) Analysis

The "GSVA" and "GSEABase" packages were utilized for differential pathway analysis to reveal the differences in enriched gene sets across various clusters.

Constructing Predictive Models Using Various Machine Learning Techniques

To identify key genes closely related to disease and clustering, the "VennDiagram" program was used to determine the consensus genes in these gene sets. Next, data from the various CRG clusters were integrated, and predictive models were built using random forests (RF), support vector machines (SVM), generalized linear models (GLMs), and eXtreme Gradient Boosting (XGB) models. The machine learning model, developed using the "caret" package (version 6.0.91), aims to identify key predictors of clustering characteristics associated with AS. As a comprehensive machine learning method, RF performs classification or regression prediction using multiple independent decision trees.[24] SVM create a hyperplane with the maximum margin in the feature space to effectively separate positive and negative examples.[25] GLM, an extension of multiple linear regression, are well-suited for assessing the relationship between normally distributed dependent variables and categorical or continuous independent variables, offering increased flexibility.[26]

XGB is an ensemble tree method built on gradient boosting, enabling precise comparison between classification errors and model complexity.[27] The "DALEX" package (version 2.4.0) was utilized to compare the four machine learning models in terms of feature importance and residual distribution. In the "pROC" software (version 1.18.0), we visualized the area under the receiver operating characteristic (ROC) curve (AUC). Based on this, the optimal machine learning model was identified, and the top five most significant variables were selected as the predictive hub genes associated with AS. Lastly, ROC curve analysis was conducted on the GSE25101 dataset to evaluate the diagnostic accuracy of the model.

Construction and Validation of the Column Line Plot Model

To assess the risk of AS clustering, we constructed a column line plot model using the "rms" package (version 6.2.0). Each predictor was assigned a score, with the "total score" representing the cumulative value of these predictions. The model's predictive accuracy was evaluated through calibration curves and decision curve analysis (DCA).

Independent Validation Analysis

To verify the prediction model, two datasets (GSE25101 and GSE73754) were used, and the ROC curve was generated using the "pROC" package to differentiate AS from non-AS, while the accuracy of five hub genes was confirmed. The visualization of ROC curve was visualized in the "pROC" package.

Clinical Specimens

From June 2023 to February 2024, a total of 31 AS patients and 30 healthy individuals were recruited from the First Teaching Hospital of Tianjin University of Traditional Chinese Medicine (Tianjin, China). The inclusion criteria for AS patients were based on the 2009 classification criteria for axial spondyloarthritis established by the International Society for Spondyloarthritis;[28] Between the ages of 18 and 65, regardless of gender; medication use was restricted to DMARDs and NSAIDs. The clinical trial received approval from the Ethics Committee of the First Teaching Hospital of Tianjin University of Traditional Chinese Medicine (TYLL2023[Y]005). All participants gave formal, informed, and written consent for serum collection. After blood samples were centrifuged at 3000 rpm for 10 minutes at 4°C, serum was extracted and stored at − 80°C. The assay was completed within one week following the collection of all samples.

ELISA

ELISA kits were purchased from SHANGHAI TONGWEI BIOTECHNOLOGY CO., LTD (Shanghai, China). SRPK1 (TW-7211), PELI1 (TW-7562), ICAM2 (TW-7062), RANGAP1 (TW-7235), and BAZ1A (TW-72150) were used in the experiment. All tests were performed following the manufacturer's provided instructions.

Statistical Analysis

All statistical analyses were performed using SPSS 20.0 software, with a significance threshold set at $p < 0.05$. Categorical variables were expressed as percentages, while quantitative data were presented as mean ± standard deviation. For comparisons between two groups, an independent sample t-test was used when the data met the assumptions of normality and homogeneity of variance. If these assumptions were violated, the nonparametric Mann–Whitney U-test was applied instead.

Results

Dysregulation of CRGs and Immune Cell Infiltration in AS

The detailed design flow of the study was depicted in Figure 1. The results of merging the dataset before and after interference removal and batch effect correction were shown in Figure 2A and B. In this dataset, we evaluated the expression levels of 19 CRGs in AS patients and controls, of which 7 CuDEGs showed differential expression. The levels of *NFE2L2, SLC31A1, MTF1,* and *DBT* were higher in AS patients, while *LIPT1, PDHA1,* and *PDHB* were elevated expressed in the controls (Figure 3A and B). Through the correlation analysis of these DEGs, we found the
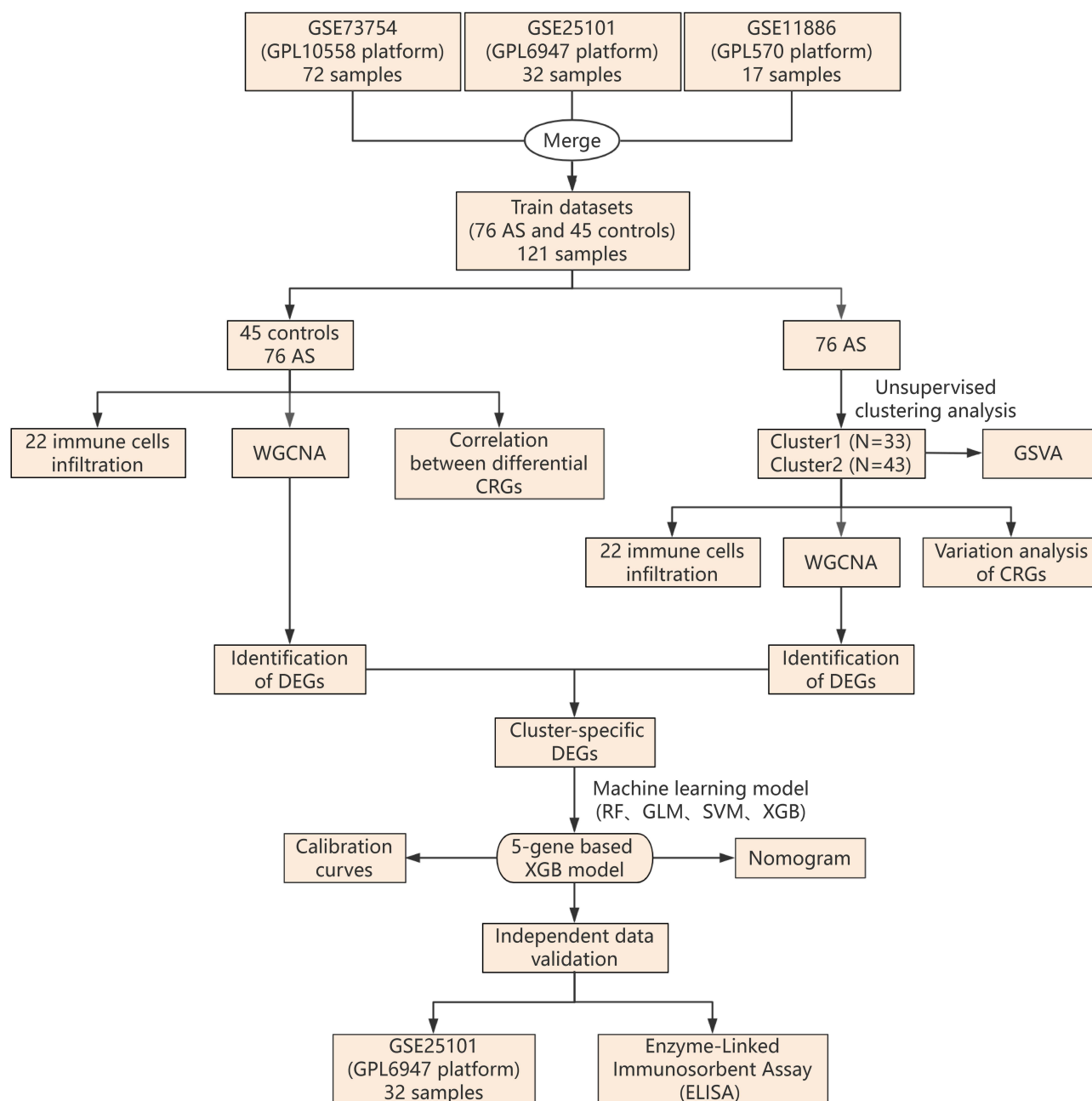
Figure 1 Study flow chart.

synergy between *PDHB, DBT*, and *PDHA1*, while *MTF1* showed significant antagonism with *PDHA1* and *PDHB* (Figure 3C). The relationships among these DEGs were illustrated in the gene correlation map (Figure 3D).

Immune Cell Infiltration in AS

The CIBERSORT algorithm identified differences in the proportions of 22 immune cell types between AS patients and controls (Figure 4A). The analysis revealed significant variations in the proportions of CD8 T cells, γ/δ T cells, and other immune cells, and activated NK cells were reduced, whereas the proportion of neutrophils was elevated ($p < 0.05$) in AS patients compared to controls (Figure 4B). Additionally, correlation analysis demonstrated that the 7 CuDEGs were linked to IICs (Figure 4C). These findings suggested that the CuDEGs may play a pivotal role in regulating molecular processes and immune cell infiltration in AS.
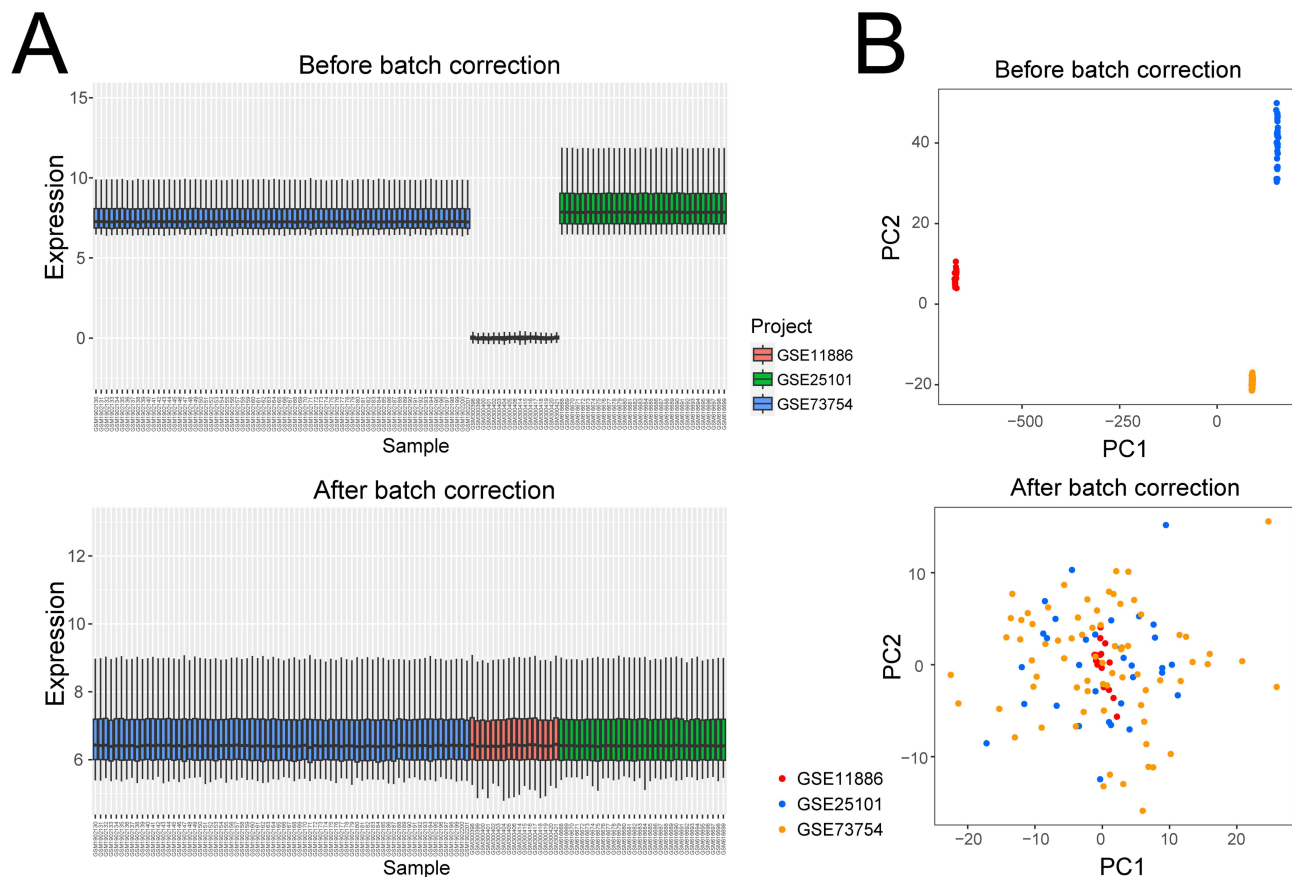
Figure 2 Batch correction results. (**A**) Expression levels before and after merging. (**B**) PCA results before and after merging.

Identification of Cuproptosis-Related Molecular Clusters in AS

At k = 2, the clustering results were determined to be the most stable and reliable (Figure 5A), as indicated by the CDF curve, which exhibited minimal fluctuation (Figure 5B), and a consistency score greater than 0.9 (Figure 5C). PCA demonstrated clear distinctions between the two identified clusters (Figure 5D).

Immune Cell Infiltration Characteristics of the Cuproptosis-Related Molecular Cluster in AS

The differential expression analysis of CRGs between Cluster1 and Cluster2 showed that *SLC31A1* and *MTF1* were significantly upregulated in Cluster1, while *LIPT1, PDHA1, PDHB*, and *DBT* had higher expression levels in Cluster2 (Figure 6A and B). Furthermore, immune cell infiltration analysis revealed that naive B cells, CD8 T cells, and activated CD4 memory T cells were more abundant in Cluster1, whereas monocytes and neutrophils were more prevalent in Cluster2 (Figure 6C and D).

Gene Module Screening and Co-Expression Network Construction

The WGCNA algorithm was employed to build gene modules and co-expression networks for AS and controls, with the objective of identifying key gene modules associated with AS. The results showed that using a soft threshold of 11 and achieving a scale-free R^2 of 0.9, co-expressed gene modules were identified (Figure 7A). The dynamic cutting algorithm generated three co-expression modules, each represented by distinct colors, and displayed the TOM heatmap (Figure 7B–D). The genes within these modules were then analyzed to assess the correlation between AS and the 264 genes in the gray module, which co-expressed with genes in the controls (Figure 7E). Additionally, the gray module showed a positive correlation with module-related genes (Figure 7F).
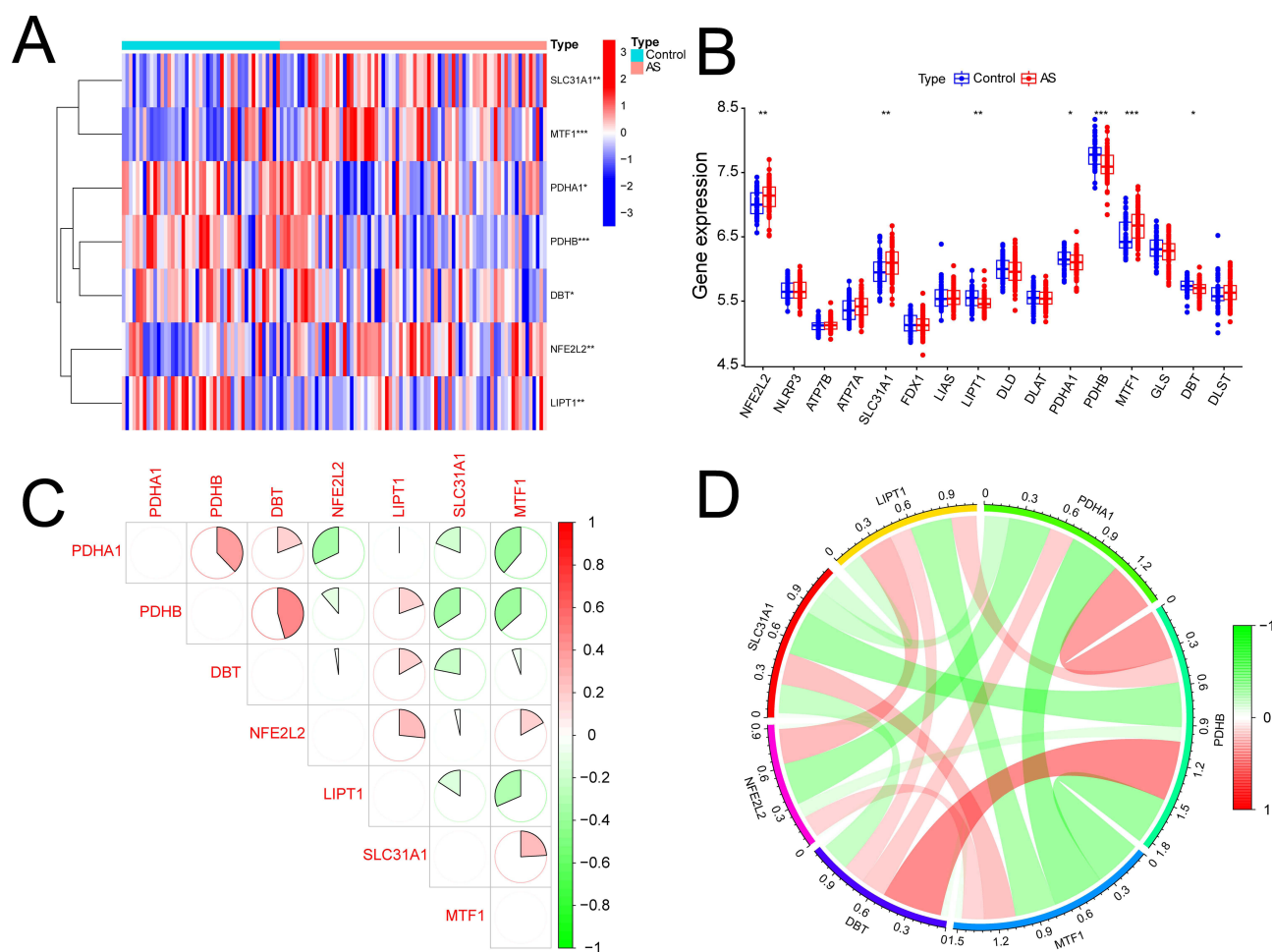
Figure 3 Identification of dysregulated CRGs in AS. (**A**) Heatmap showing the expression profiles of 7 CuDEGs. (**B**) Boxplots depicting the differential expression of CRGs between AS patients and controls. (**C**) Correlation analysis of the 7 CuDEGs, with red representing positive correlations, green representing negative correlations, and the correlation coefficient indicated by the size of the pie charts. (**D**) Gene interaction map of the 7 CuDEGs, where red denotes positive correlations, green denotes negative correlations, and the correlation coefficient was represented by the thickness of the connecting lines.
Notes: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

To further identify key gene modules linked to AS, we applied the WGCNA algorithm to construct gene modules and co-expression networks for the two molecular clusters related to cuproptosis. Using a soft threshold of 12 and achieving a scale-free R^2 of 0.9 (Figure 8A), we generated co-expression modules. The dynamic cutting algorithm identified four distinct co-expression modules, each represented by a unique color, with the TOM heatmap displayed (Figure 8B–D). By analyzing the gene co-expression patterns between the two clusters, we found that the turquoise module, comprising 1052 genes, had the strongest association with the cuproptosis-related molecular clusters (Figure 8E). Furthermore, a significant correlation was observed between the turquoise module and the module-related genes (Figure 8F).

Cluster-Specific DEGs

By intersecting the cuproptosis-related molecular clusters with DEGs in AS, 190 genes were identified (Figure 9A) and further analyzed. Additionally, GSVA was carried out to examine the functional distinctions of specific DEGs across the two clusters. The functional enrichment analysis results revealed that Cluster1 was upregulated in the cytosolic DNA sensing pathway, the calcium ion signaling pathway and the T cell receptor signaling pathway (Figure 9B), and was upregulated in stem cell division, glucan biosynthesis, and negative feedback regulation of Th17 immune response (Figure 9C).
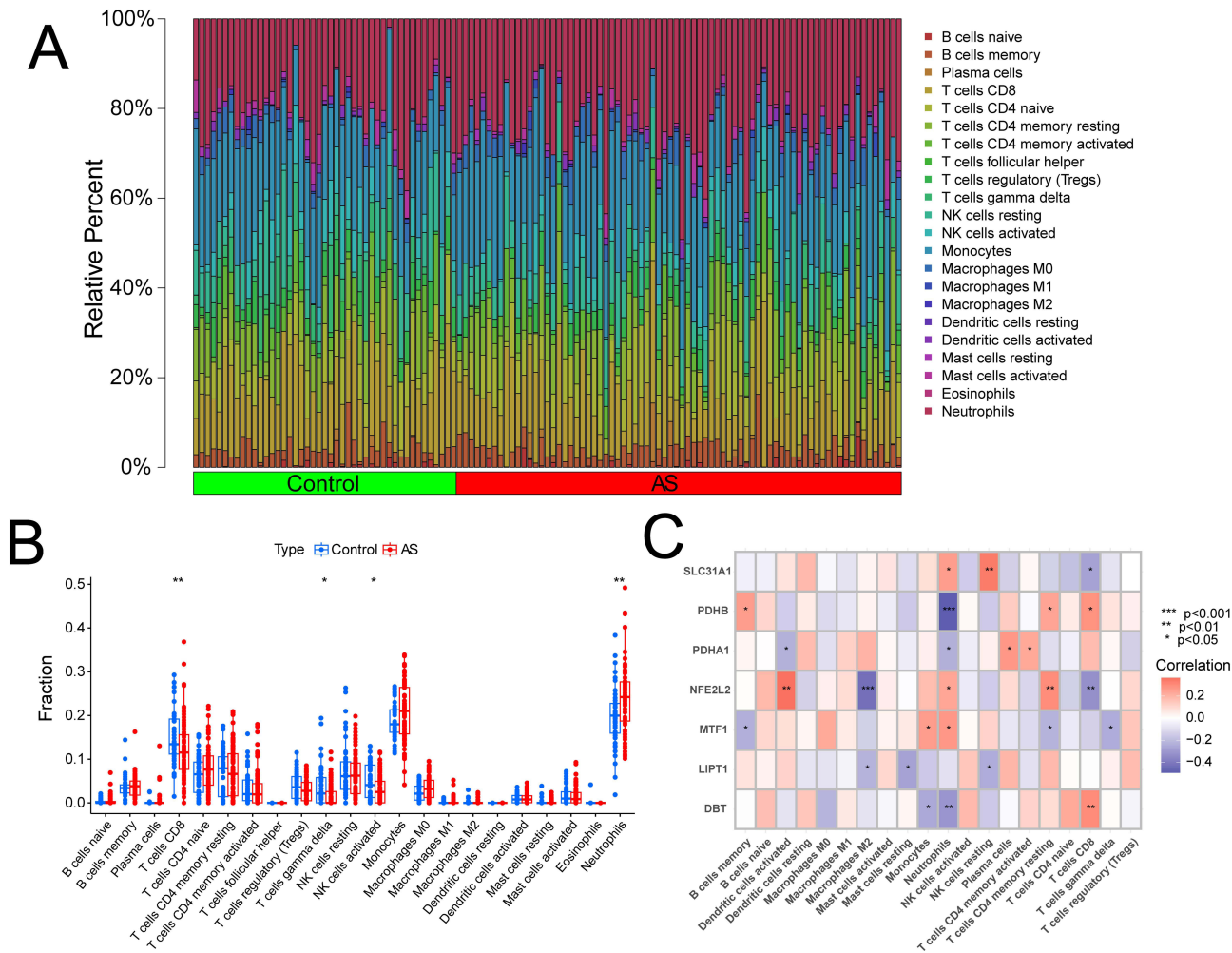
Figure 4 Overview of IICs in AS. (**A**) Relative proportions of 22 IICs in AS patients compared to controls. (**B**) Boxplots illustrating differences in IICs between AS patients and controls. (**C**) Heatmap depicting the correlation between IICs and the 7 CuDEGs, with red representing positive correlations and blue representing negative correlations.

Notes: *p < 0.05, **p < 0.01, ***p < 0.001.

Construction of Machine Learning Models

We built four machine learning models—SVM, RF, XGB, and GLM—to identify genes specific to each cluster with diagnostic significance. Among these models, XGB, RF, and SVM exhibited the lowest residual values (Figure 10A and B). To evaluate the predictive accuracy of each model, five-fold cross-validation was performed to estimate the AUC. The XGB model achieved the highest AUC (0.752) (Figure 10C). Feature importance maps for each model were also generated (Figure 10D). Based on residual values and ROC analysis, the XGB model was chosen as the optimal model for distinguishing between clusters in AS. In conclusion, five hub genes—*SRPK1, PELI1, ICAM2, RANGAP1*, and *BAZ1A*—were identified from the XGB model for further analysis.

A nomogram was developed to evaluate the risk associated with cuproptosis-related factors in 76 AS patients, enabling the assessment of the predictive accuracy of the XGB machine learning model (Figure 11A). The accuracy of column line plot model prediction was verified by calibration curve and DCA. The calibration curve results revealed a moderate difference between the predicted risk and the actual cluster risk in AS (Figure 11B). DCA findings further supported the high accuracy of the nomogram (Figure 11C), indicating its potential as a useful tool for informing treatment decisions. Moreover, the 5-gene prediction model was validated using data from whole blood samples. According to the ROC results, the AUC of GSE25101 was 0.812 (Figure 11D and E), indicating that it is equally effective in distinguishing AS patients from non-AS patients with the five genes in GSE73754. The AUC values of
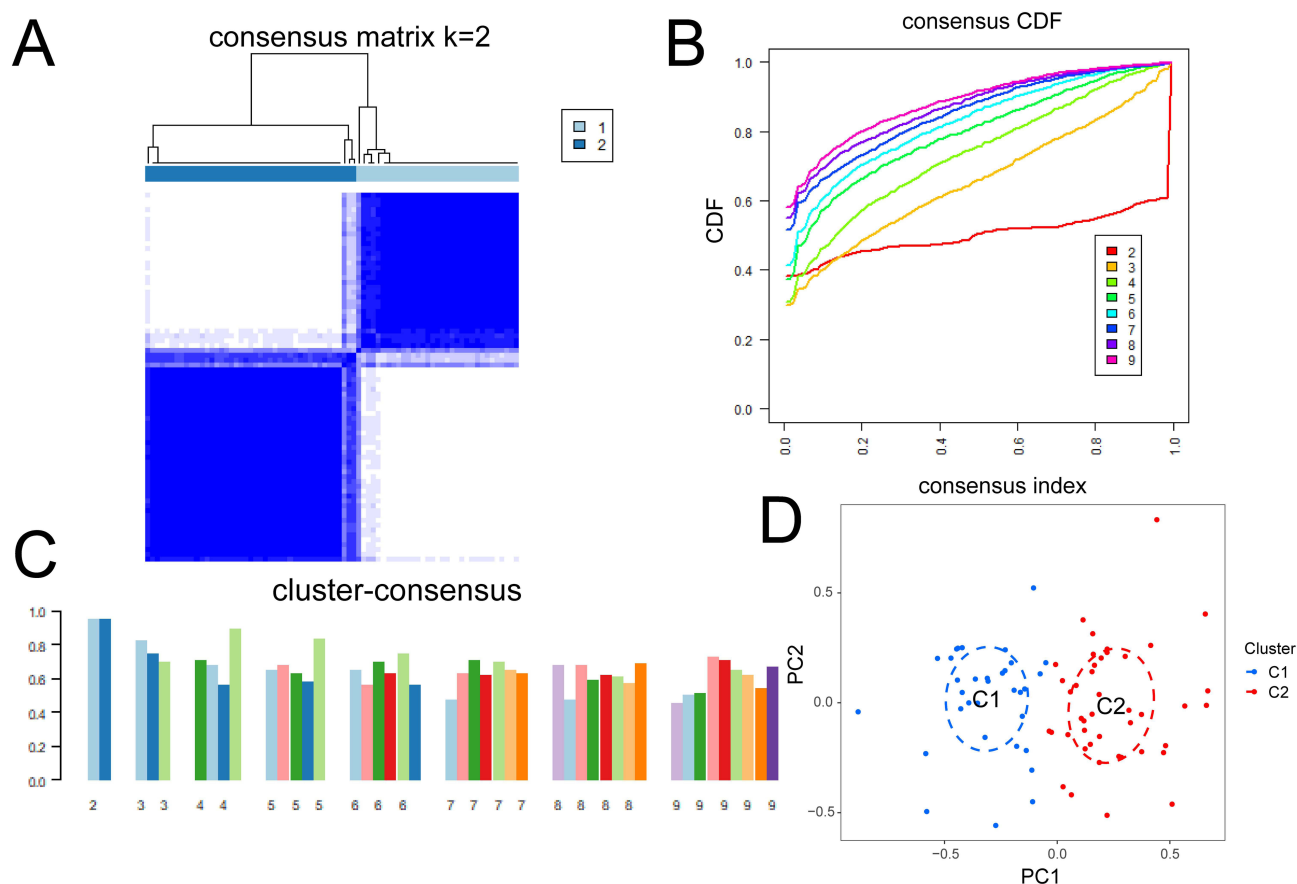
Figure 5 Identification of cuproptosis-related molecular clusters in AS. (**A**) Consensus matrix at k = 2. (**B**) CDF curve. (**C**) Consensus clustering score. (**D**) PCA plot showing the distribution of the two identified clusters.

SRPK1, PELI1, ICAM2, RANGAP1, and *BAZ1A* were 0.793, 0.703, 0.724, 0.745, and 0.886, respectively. Subsequently, the levels of the five proteins in the serum of AS patients and controls were validated using an independent dataset and further confirmed via ELISA.

Validation of the Concentrations of the 5 Genes in the Serum of AS Patients

Table 1 provides a comparison of the clinical characteristics between AS patients and controls. AS patients exhibited elevated levels of CRP and ESR compared to controls. Additionally, the protein expression levels of the five hub genes were confirmed using serum samples. The findings indicated that compared with the controls, the levels of PELI1 in the serum of AS patients were significantly increased (552.19 ± 328.41 vs 412.79 ± 161.56, $p = 0.007$) (Figure 12B), while ICAM2 and RANGAP1 were significantly decreased (201.03 ± 43.25 vs 255.96 ± 99.24, $p = 0.011$; 204.31 ± 40.29 vs 204.31 ± 40.29, $p = 0.000$) (Figure 12C and D). But SRPK1 and BAZ1A were not statistically significant (Figure 12A and E).

Discussion

AS is a chronic inflammatory disorder with a subtle onset, primarily affecting the spine and sacroiliac joints. In its later stages, the disease can cause irreversible structural damage and significantly impair mobility. At present, there is still no clear etiology and pathogenesis, and no obvious breakthrough in treatment.[2] Therefore, it is crucial to discover novel and effective biomarkers for both the diagnosis and treatment of AS, investigate characteristic genes associated with AS, and discover potential therapeutic targets. This will offer new insights for early diagnosis and intervention in AS. Cuproptosis, a newly identified cell death mechanism, has been associated with a range of diseases.[14] Relevant literature
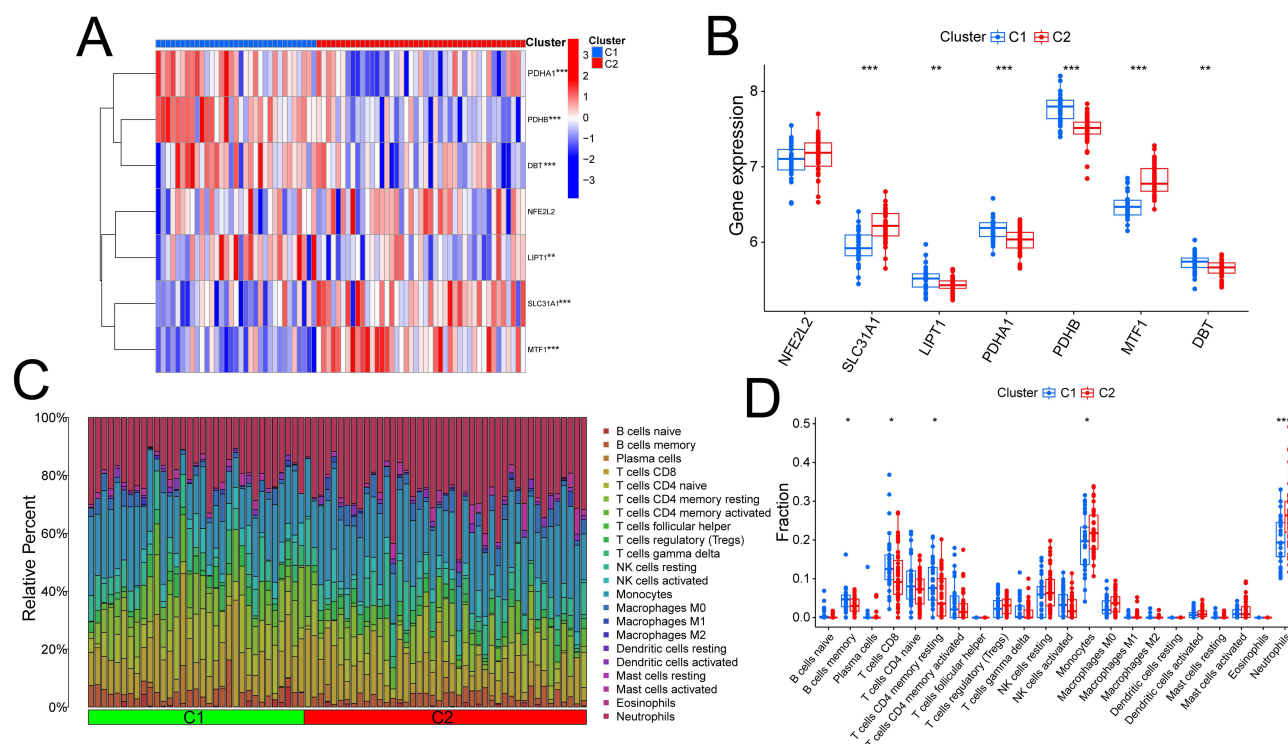
Figure 6 Characterization of molecular and immune signature infiltrates in the two clusters. (**A**) Heatmap displaying the expression levels of 7 CuDEGs. (**B**) Boxplots displaying the expression levels of the 7 CuDEGs. (**C**) Relative proportions of the 22 IICs. (**D**) Boxplots show the infiltration of 22 IICs.
Notes: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

shows that the level of Cu in the serum of AS patients increases, which is closely related to acute inflammation.[15,16] This indicates that AS may have a certain correlation with cuproptosis. However, its underlying mechanism remains unclear and requires further investigation. In AS patients, CD8[+] T lymphocytes are preferentially drawn to inflamed joints, leading to a reduced level of CD8[+] T cells in the blood and an elevated level in the synovium.[29] Another study showed that the disruption of the TNF signaling pathway in CD8[+] T cells plays a key role in the development of AS.[30] Gamma delta T (γ/δ T) cells, a specific subset of CD3[+] T cells, were initially discovered and studied in the context of autoimmune rheumatism. Despite being a highly conserved group of T cells, they are critical in various aspects of immunobiology.[31] A study[32] showed that in patients with AS, γ/δ T cells produce IL-17 at a rate five times higher than other cells, indicating that they may play an important role in the pathogenesis of AS. *HLA-B27* positivity is the most prominent genetic risk factor for AS patients,[33] and *HLA-B27* combined with *KIR3DL1* can initiate inhibitory signals in NK cells.[34] Evidence on the changes in NK cell proportions in the peripheral blood of AS patients is conflicting. While two studies have reported an increase in NK cell numbers in the peripheral blood of AS patients,[35,36] the opposite results were obtained in other studies.[37,38]

This study sought to investigate AS-related CuDEGs and their involvement in the physiological and pathological processes of the disease. To identify potential therapeutic targets for AS, we analyzed the CuDEGs between AS patients and controls. From the expression profiles of 19 CRGs, we found that the expressions of *NFE2L2, SLC31A1, MTF1, LIPT1, PDHA1, PDHB*, and *DBT* had statistical significance, as CuDEGs with diagnostic significance for AS. *NFE2L2* is a transcriptional activator that acts by inhibiting NF-κB signaling pathway and inhibiting pro-inflammatory cytokines to suppress inflammation, induce the expression of related proteins to cope with oxidative stress and regulate antioxidant defense.[39] *SLC31A1*, as a Cu transporter,[21] is responsible for most of the Cu uptake in cells. *MTF1* can promote the production of inflammatory factors while suppressing T cell activation and is a hub gene be of paramount importance in joint destruction and inflammation.[40] *LIPT1* supports mitochondrial respiration, TCA cycle and promotes fatty acid synthesis, a vital element in glucose and glutamine metabolism.[41] *PDHA1* regulates mitochondrial signaling, oxidative phosphorylation, cellular respiration and electron transport activities, pyruvate metabolism, carbohydrate metabolism,
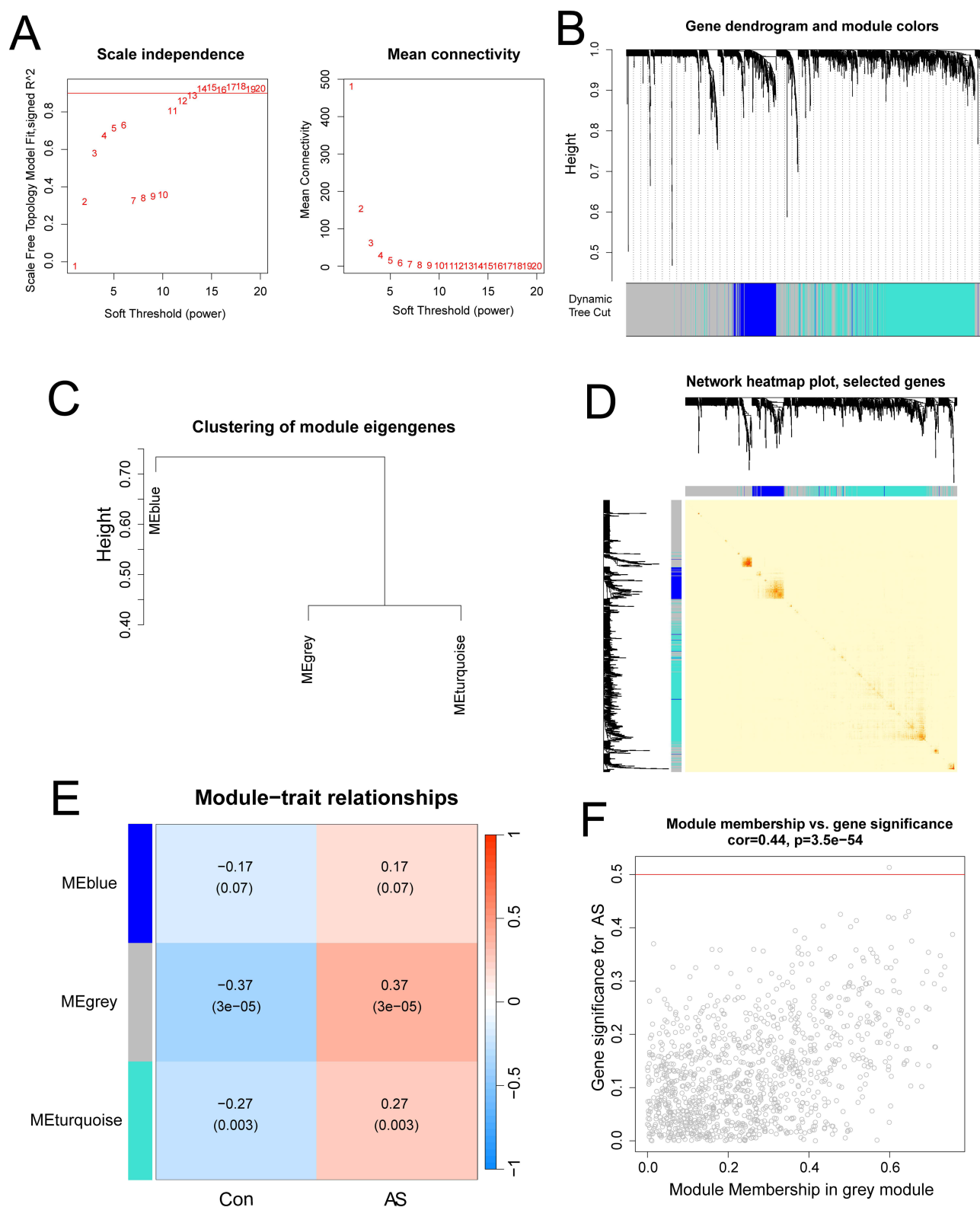
Figure 7 Co-expression network of AS-related DEGs. (**A**) Determination of soft threshold. (**B**) Gene dendrogram within the co-expression module, with each module represented by a distinct color. (**C**) Cluster diagram of characteristic genes in the module. (**D**) Heatmap showing correlations between modules. (**E**) Correlation analysis between module-related genes and clinical characteristics. (**F**) Scatter plot illustrating the correlation between gene significance for AS and genes in the gray module.
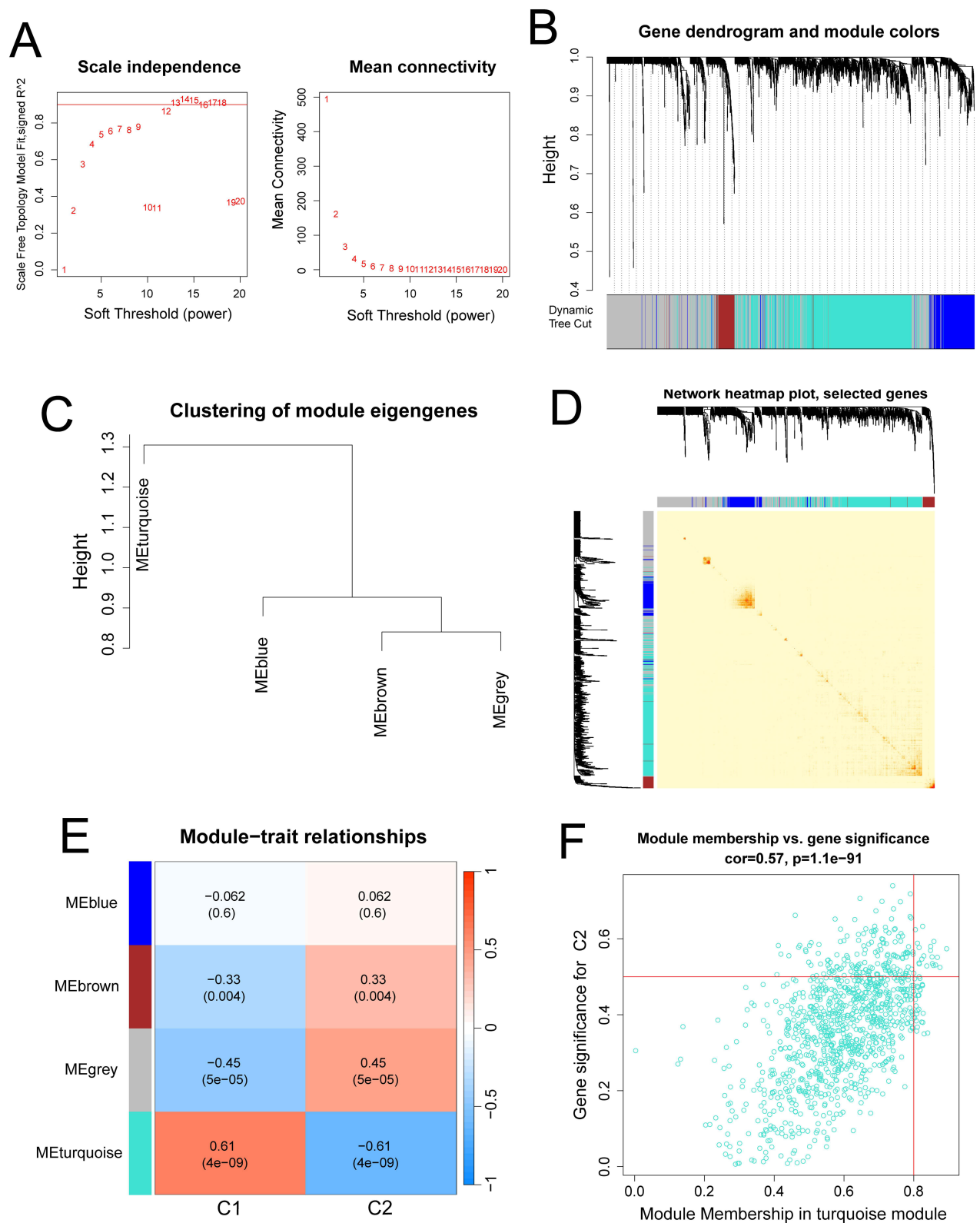
Figure 8 Co-expression network of DEGs in the clusters. (**A**) Selection of soft threshold. (**B**) The gene dendrogram of the co-expression module, with each module represented by a different color. (**C**) Cluster diagram of characteristic genes in the module. (**D**) Heatmap of correlations between modules. (**E**) Correlation analysis between module-related genes and clinical characteristics. (**F**) Scatter plot showing the correlation between turquoise module genes and the significance of Cluster2 genes.
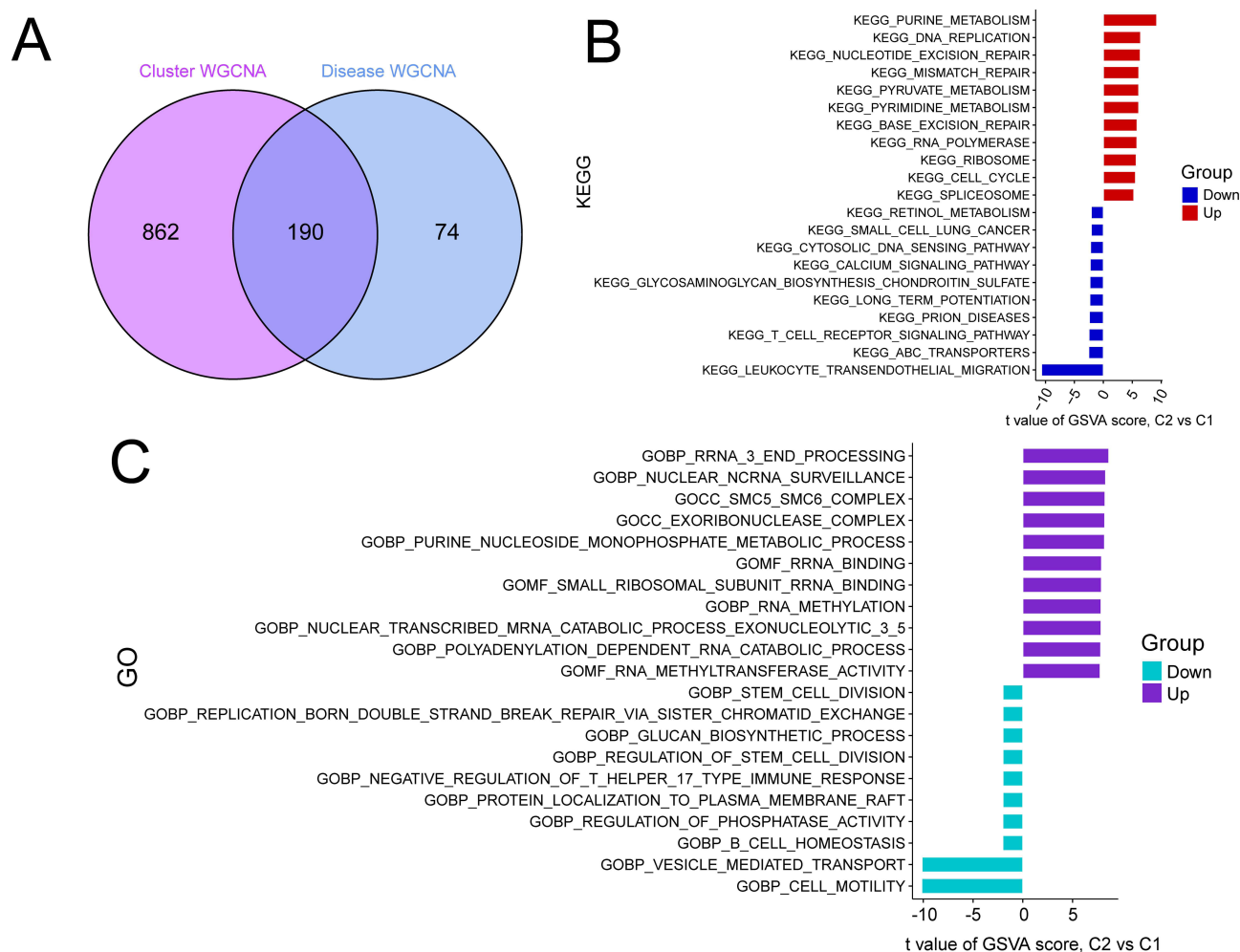
Figure 9 The cluster-specific DEGs and enrichment analysis. (**A**) The intersection of cuproptosis-related molecular clusters and DEGs in AS patients versus controls. (**B** and **C**) KEGG and GO enrichment of DEGs between the two clusters.

and citric acid cycle.[42] *PDHB* is a nuclear-encoded pyruvate dehydrogenase that facilitates the conversion of pyruvate to acetyl-CoA and aids in protecting skeletal muscle by inhibiting the FoxP1-Arih2 axis.[43,44] *DBT* can promote oxidative stress and it boosts NF-κB activity and stimulates the production of inflammatory cytokines such as TNF-α and IL-6.[45]

Through unsupervised clustering analysis and PCA, we identified two molecular clusters associated with cuproptosis to classify AS. Utilizing the specific DEGs to each cluster, we developed four machine learning models (RF, SVM, XGB, and GLM) to assess their predictive performance. XGB model has the highest prediction performance and constructed 5-gene prediction models of *SRPK1, PELI1, ICAM2, RANGAP1*, and *BAZ1A*. The GSE73754 dataset and ELISA were utilized to verify the protein expression levels of *PELI1, ICAM2*, and *RANGAP1*, and a possible mechanism diagram of its action was drawn (Figure 13), this further demonstrated that the prediction model possesses high accuracy and significant clinical applicability.

SRSF protein kinase 1 *(SRPK1)* is an enzyme that phosphorylates serine/arginine domain rich splicing factors. It is involved in numerous cellular processes, such as cell cycle regulation, innate immune responses, chromatin remodeling, and the regulation of several mRNA processing pathways, in addition to modulating inflammation through interactions with various transcription factors and signaling pathways.[46,47] In addition, studies have found that *SRPK1* can regulate the expression of AKT3 at the transcriptional or translational level to mediate the inflammatory response.[48] E3 ubiquitin *PELI1* is an E3 ubiquitin ligase that is highly expressed in T lymphocytes. It functions as a key negative regulator, suppressing T cell activation and autoimmunity in vivo by facilitating K48-linked ubiquitination and degradation of
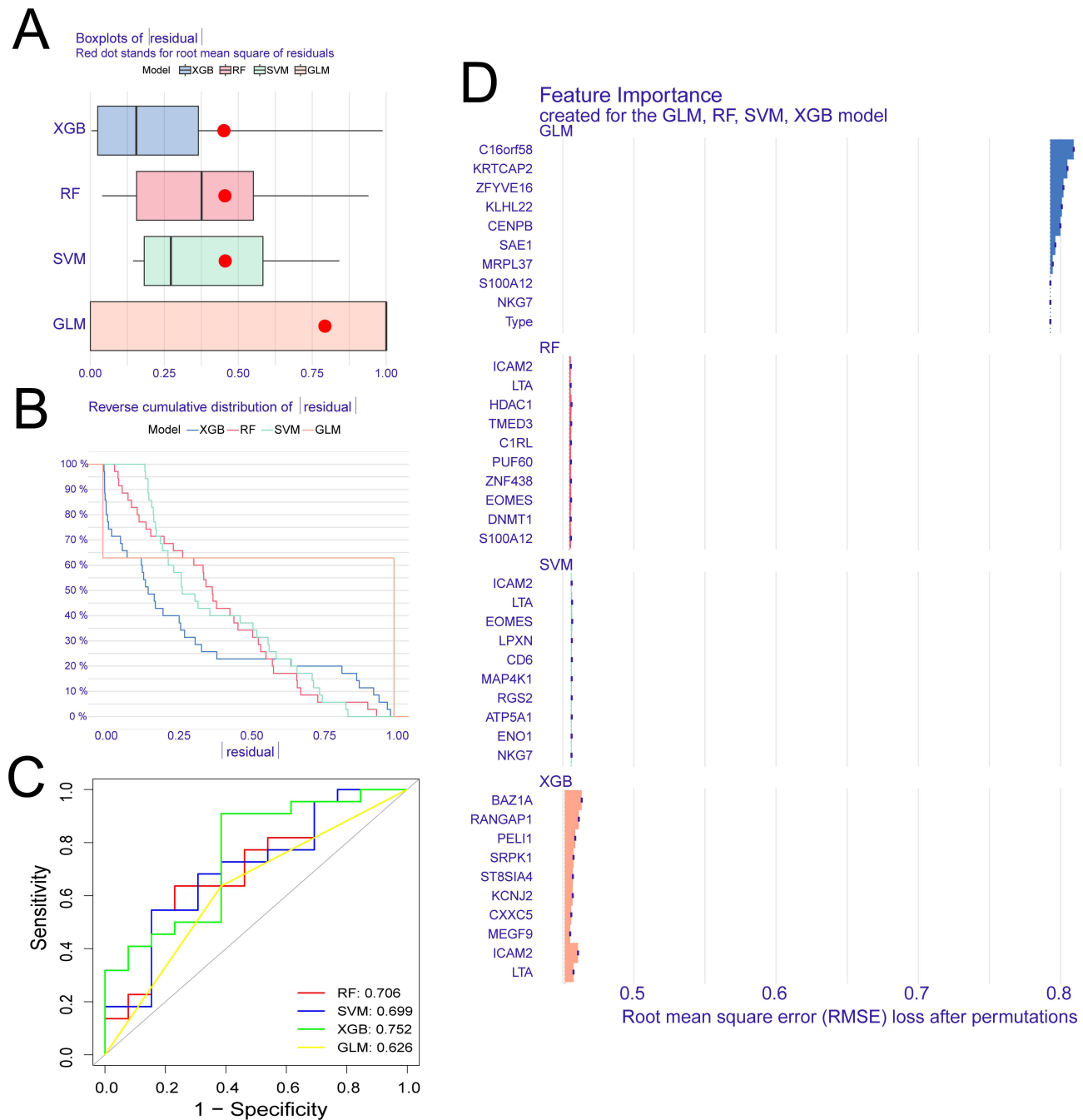
Figure 10 Construction of four machine learning models. (**A**) Boxplots showing the residuals of the four models, with red dots indicating the root mean square of the residuals. (**B**) Cumulative distribution curve of the residuals. (**C**) ROC analysis of the machine learning model performed with five-fold cross-validation. (**D**) Histogram displaying the key features of the four machine learning models.

c-Rel., and regulates various inflammatory diseases.[34] Research has found that miR-155 is differentially expressed in spinal joint diseases such as Graves' disease, RA, and AS, demonstrating strong diagnostic performance. Serum miR-155 may affect the occurrence and development of ankylosing spondylitis by regulating *SRPK1*.[49,50] In addition, studies have shown that miR-155 may also regulate the potential role of *PELI1* in Tfh cells, influencing the onset and progression of AS.[51] Intercellular adhesion molecule proteins act as ligands for the leukocyte adhesion protein LFA-1 (integrin α-L/β-2). *ICAM2* may participate in lymphocyte recirculation by inhibiting LFA-1-dependent cell adhesion, contributing to antigen-specific immune responses, NK cell-mediated clearance, lymphocyte recirculation, and other essential adhesion
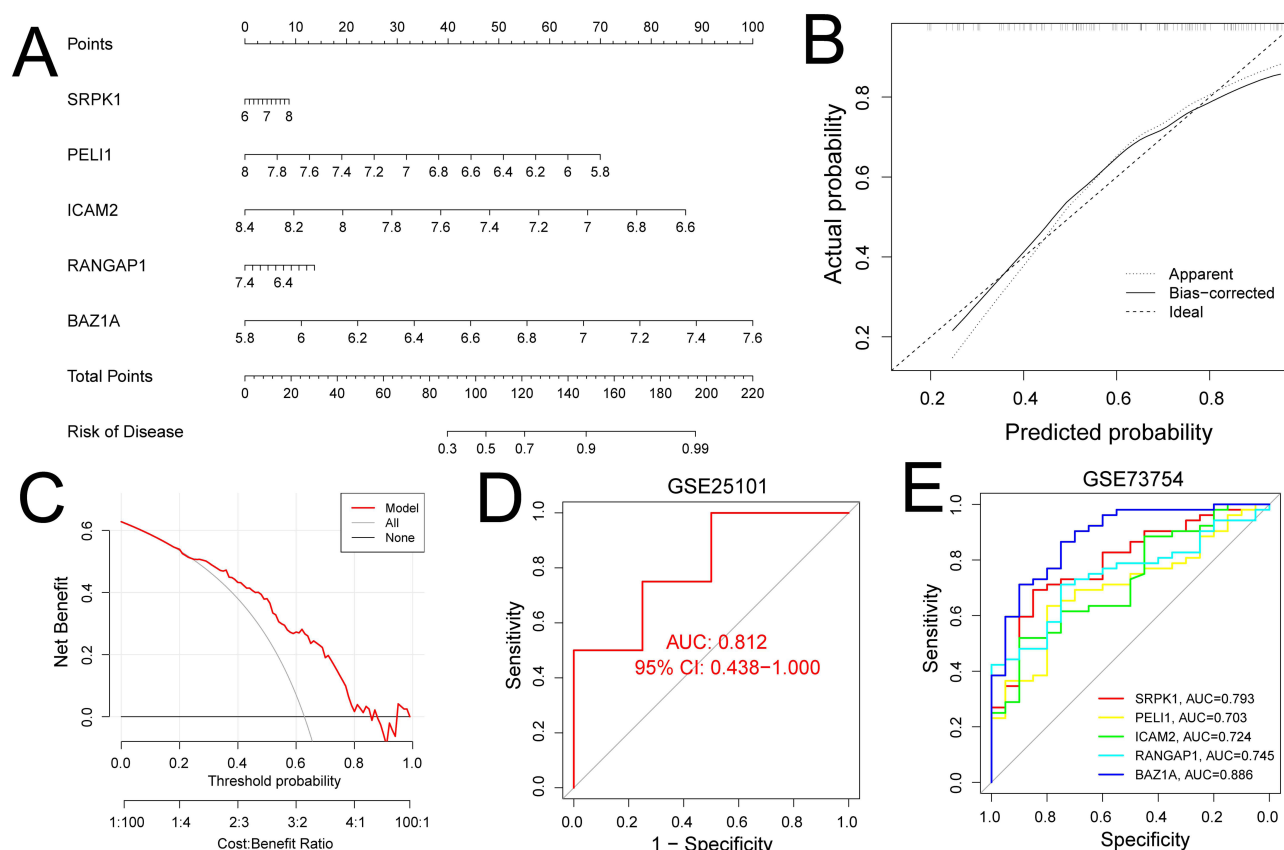
Figure 11 Validation of the XGB model based on 5 genes. (**A**) Nomogram predicting the likelihood of AS cluster involvement. (**B**) Calibration curve. (**C**) DCA results. (**D**) ROC plot for validation using the GSE25101 dataset. (**E**) ROC analysis of the 5 genes (*SRPK1, PELI1, ICAM2, RANGAP1,* and *BAZ1A*) in the GSE73754 dataset.

interactions required for immune surveillance and response. *ICAM2* is overexpressed in platelets of AS patients, which may trigger the inflammatory cascade and promote the development of AS.[35] RanGTPase activating protein 1 (*RANGAP1*) is a critical component of the nuclear pore complex, responsible for regulating the nuclear transport of

Table 1 Comparison of the Clinical Characteristics Between AS Patients and Controls

Characteristics	Control (n = 30)	AS (n = 31)	*p*-value
Age, years	36.90 ± 6.21	34.74 ± 9.90	0.311
Male, n (%)	21 (70.00%)	24 (77.42%)	0.510
RBC (×10^{12}/L)	4.71 ± 0.45	4.80 ± 0.49	0.461
WBC (×10^9/L)	6.96 ± 1.15	6.53 ± 1.28	0.116
PLT (×10^9/L)	234.30 ± 47.68	263.42 ± 70.24	0.251
HGB (g/L)	133.83 ± 10.12	138.29 ± 20.08	0.141
ALT (U/L)	20.00 ± 4.57	27.57 ± 28.35	0.444
AST (U/L)	19.64 ± 5.24	19.98 ± 10.10	0.530
Urea (mmol/L)	4.53 ± 1.58	4.99 ± 1.29	0.214
CR (μmol/L)	66.33 ± 12.22	68.73 ± 13.64	0.363
CRP (mg/L)	2.36 ± 2.14	5.41 ± 5.32	0.011*
ESR (mm/h)	3.95 ± 1.15	15.32 ± 10.25	0.000***
RF (IU/mL)	9.94 ± 1.16	10.61 ± 2.17	0.151
HLA-B27, n (%)	NA	19 (61.29%)	NA

Notes: *p < 0.05; ***p < 0.001.
Abbreviations: RBC, red blood cell; WBC, white blood cell; PLT, platelet; HGB, hemoglobin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CR, creatinine; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; RF, rheumatoid factor; *HLA-B27*, human leukocyte antigen-B27; NA, not applicable.
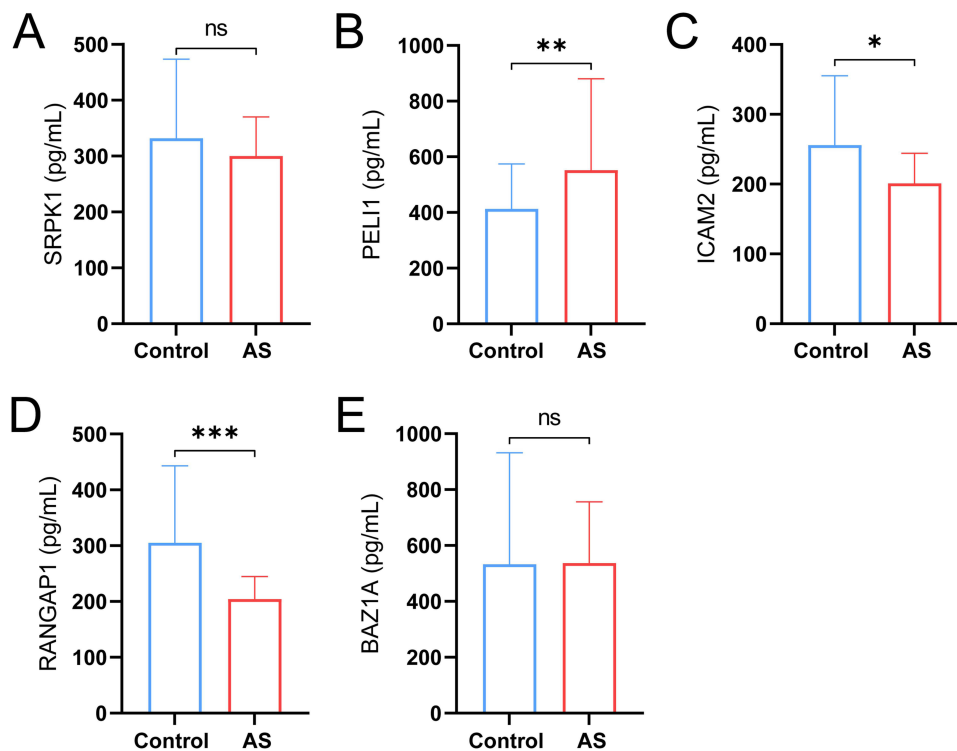
Figure 12 Protein expression levels of the five genes in serum samples. (**A** and **E**) Serum levels of SRPK1 and BAZ1A did not show significant differences between AS patients and controls. (**B**) PELI1 levels were notably higher in the serum of AS patients compared to controls. (**C** and **D**) ICAM2 and RANGAP1 levels were significantly lower in AS patients compared to controls.

Notes: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$; ns, not significant.

intracellular proteins,[36] and is divided into non-sumoylated *RANGAP1*, which is mainly localized in the cytoplasm, and SUMO1 bound *RANGAP1*, which resides in NPC.[37] Studies have demonstrated that tumor necrosis factor receptor-associated factor 6 (TRAF6), which mediates several essential signaling pathways and is involved in immune surveillance and inflammation, can enter the nucleus via the nuclear pore complex with the involvement of *RANGAP1*. This implies that *RANGAP1* may indirectly contribute to immune-related inflammation.[52] The Bromodomain adjacent to zinc finger domain protein 1A (*BAZ1A*) encodes a subunit of the adenosine triphosphate (ATP)-dependent chromatin assembly factor (ACF), which is a component of a chromatin remodeling complex involved in DNA damage response and repair.[38]

The AUC of the dataset validated by the XGB model based on the prediction model of five genes is 0.812, which also provides a new idea for the diagnosis of AS. The nomogram model subtypes of *SRPK1* (AUC = 0.793), *PELI1* (AUC = 0.703), *ICAM2* (AUC = 0.724), *RANGAP1* (AUC = 0.745), and *BAZ1A* (AUC = 0.886) were used to diagnose AS. Validation of the concentrations of ICAM2 and RANGAP1 in AS patients showed that their serum levels were significantly lower, whereas PELI1 levels were significantly higher. SRPK1 and BAZ1A did not show significant differences in the ELISA results. The research may still need more samples for validation mining. Based on the above, the results of XGB model using five genes were of great significance to evaluating AS clusters and exploring new ideas for diagnosis and treatment of AS.

Certainly, at present, the pathogenesis of AS is not completely clear. After preliminary data mining, the clinical sample size of our study may not be enough, resulting in certain limitations of the study. Despite the in-depth bioinformatics analysis, further experiments are required to investigate the regulatory mechanisms of AS pathogenesis and the expression of CRGs. Furthermore, additional AS samples are required to enhance the accuracy of cuproptosis-related molecular clustering. Lastly, clinical samples must be evaluated to validate the accuracy of the predictive models.
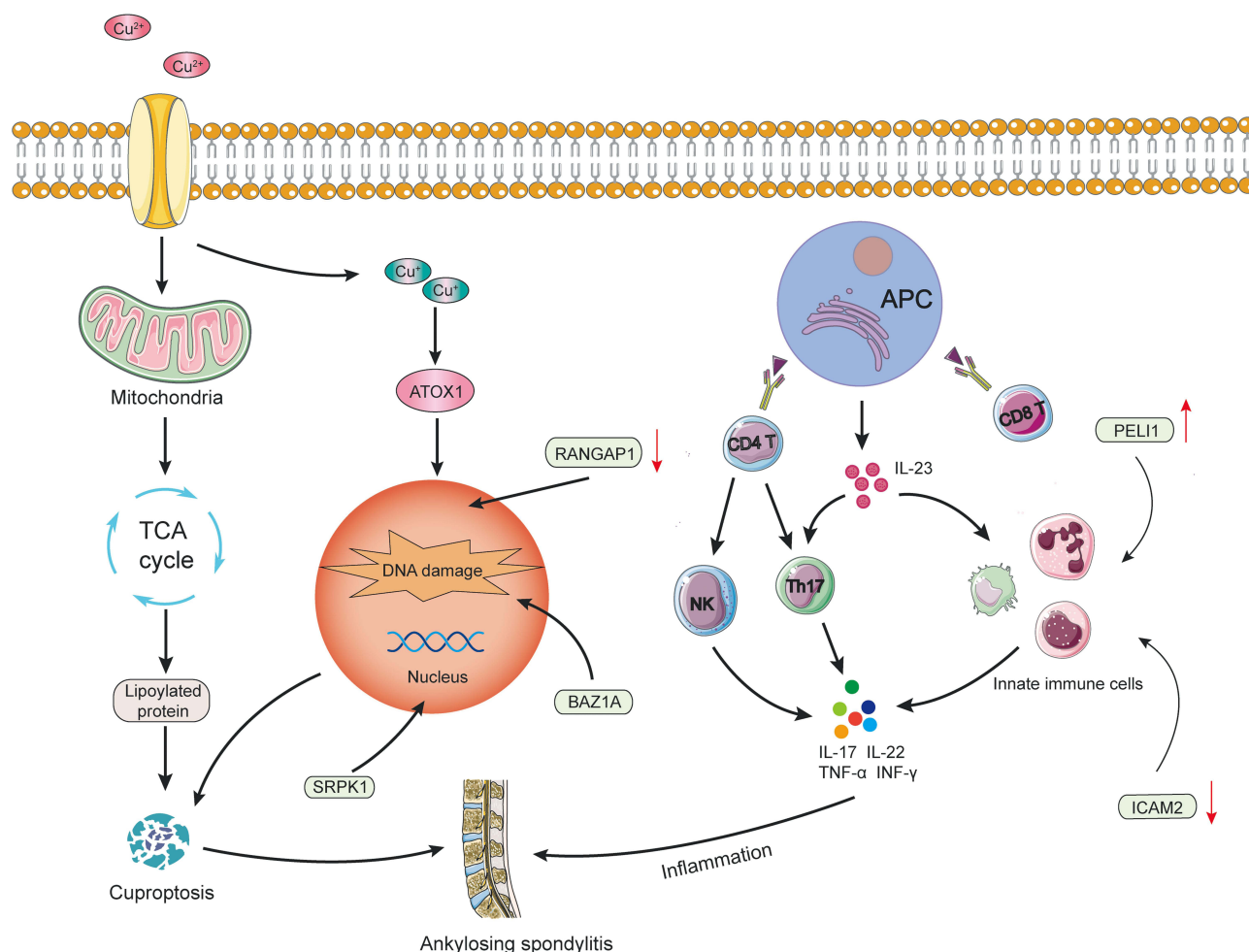
Figure 13 The propose mechanisms of the 5 genes in AS.

Conclusion

In summary, this study investigated the significance of cuproptosis in AS through comprehensive bioinformatics analysis, revealing a connection between AS, CRGs, and immune cell infiltration. The best model for predicting AS was constructed based on five genes using the XGB model, and high-risk related genes were screened. Our current study laid the foundation for further exploring the pathogenesis of AS and determined the prognostic biological indicators of cuproptosis associated with AS.

Data Sharing Statement

The datasets generated and/or analyzed in the study can be found in the Gene Expression Omnibus datasets (GEO, https://www.ncbi.nlm.nih.gov/geo/, GSE73754, GSE25101, and GSE11886).

Ethics Approval and Informed Consent

The study was conducted following the Declaration of Helsinki and approved by the Ethics Committee of the First Teaching Hospital of Tianjin University of traditional Chinese medicine (TYLL2023[Y]005). The participants consented in writing to take part in this research.

Consent for Publication

This article was published with the written consent of the patient.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

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.

References

1. Kumar S, Doss RSA, Rebekah G, et al. Prevalence of HLA-B*27 subtypes in the Tamil population of India with ankylosing spondylitis and its correlation with clinical features. *Hum Immunol*. 2021;82:404–408. doi:10.1016/j.humimm.2021.03.001
2. Navarro-Compán V, Sepriano A, El-Zorkany B, van der Heijde D. Axial spondyloarthritis. *Ann Rheumatic Dis*. 2021;80:1511–1521. doi:10.1136/annrheumdis-2021-221035
3. Sieper J, Poddubnyy D. Axial spondyloarthritis. *Lancet*. 2017;390:73–84. doi:10.1016/s0140-6736(16)31591-4
4. Xiong Y, Cai M, Xu Y, et al. Joint together: the etiology and pathogenesis of ankylosing spondylitis. *Front Immunol*. 2022;13:996103. doi:10.3389/fimmu.2022.996103
5. Wang R, Ward MM. Epidemiology of axial spondyloarthritis: an update. *Curr Opin Rheumatol*. 2018;30:137–143. doi:10.1097/bor.0000000000000475
6. Tam LS, Gu J, Yu D. Pathogenesis of ankylosing spondylitis. *Nat Rev Rheumatol*. 2010;6:399–405. doi:10.1038/nrrheum.2010.79
7. Arévalo M, Gratacós Masmitjà J, Moreno M, et al. Influence of HLA-B27 on the ankylosing spondylitis phenotype: results from the REGISPONSER database. *Arthritis Res Therapy*. 2018;20:221. doi:10.1186/s13075-018-1724-7
8. Brewerton DA, Hart FD, Nicholls A, Caffrey M, James DC, Sturrock RD. Ankylosing spondylitis and HL-A 27. *Lancet*. 1973;1:904–907. doi:10.1016/s0140-6736(73)91360-3
9. Ma C, Wen B, Zhang Q, et al. Emodin induces apoptosis and autophagy of fibroblasts obtained from patient with ankylosing spondylitis. *Drug Des Devel Ther*. 2019;13:601–609. doi:10.2147/dddt.S182087
10. Ni WJ, Leng XM. Down-regulated miR-495 can target programmed cell death 10 in ankylosing spondylitis. *Mol Med*. 2020;26:50. doi:10.1186/s10020-020-00157-3
11. Tsvetkov P, Coy S, Petrova B, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science*. 2022;375:1254–1261. doi:10.1126/science.abf0529
12. Ge EJ, Bush AI, Casini A, et al. Connecting copper and cancer: from transition metal signalling to metalloplasia. *Nat Rev Cancer*. 2022;22:102–113. doi:10.1038/s41568-021-00417-2
13. Chen J, Jiang Y, Shi H, Peng Y, Fan X, Li C. The molecular mechanisms of copper metabolism and its roles in human diseases. *Pflugers Arch*. 2020;472:1415–1429. doi:10.1007/s00424-020-02412-2
14. Chen L, Min J, Wang F. Copper homeostasis and cuproptosis in health and disease. *Signal Transduct Target Ther*. 2022;7:378. doi:10.1038/s41392-022-01229-y
15. Aiginger P, Kolarz G, Willvonseder R. Copper in ankylosing spondylitis and rheumatoid arthritis. *Scand J Rheumatol*. 1978;7:75–78. doi:10.3109/03009747809098838
16. Jayson MI, Davis P, Whicher JT, Walters G. Serum copper and caeruloplasmin in ankylosing spondylitis, systemic sclerosis, and morphea. *Ann Rheumatic Dis*. 1975;35:443–445. doi:10.1136/ard.35.5.443
17. Bian Z, Fan R, Xie L. A novel cuproptosis-related prognostic gene signature and validation of differential expression in clear cell renal cell carcinoma. *Genes*. 2022;13:851. doi:10.3390/genes13050851
18. Kahlson MA, Dixon SJ. Copper-induced cell death. *Science*. 2022;375:1231–1232. doi:10.1126/science.abo3959
19. Tang D, Chen X, Kroemer G. Cuproptosis: a copper-triggered modality of mitochondrial cell death. *Cell Res*. 2022;32:417–418. doi:10.1038/s41422-022-00653-7
20. Wang Y, Zhang L, Zhou F. Cuproptosis: a new form of programmed cell death. *Cell Mol Immunol*. 2022;19:867–868. doi:10.1038/s41423-022-00866-1
21. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*. 2015;12:453–457. doi:10.1038/nmeth.3337

22. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26:1572–1573. doi:10.1093/bioinformatics/btq170

23. 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

24. Rigatti SJ. Random Forest. *J Insur Med*. 2017;47:31–39. doi:10.17849/insm-47-01-31-39.1

25. Gold C, Sollich P. Model selection for support vector machine classification. *Neurocomputing*. 2003;55:221–249. doi:10.1016/S0925-2312(03)00375-8

26. Dennerlein S, Rehling P. Human mitochondrial COX1 assembly into cytochrome c oxidase at a glance. *J Cell Sci*. 2015;128(5):833–837. doi:10.1242/jcs.161729

27. Chen T, He T, Benesty M. xgboost: extreme gradient boosting. 2016.

28. Sieper J, Rudwaleit M, Baraliakos X, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheumatic Dis*. 2009;68(Suppl 2):ii1–44. doi:10.1136/ard.2008.104018

29. Gracey E, Yao Y, Qaiyum Z, Lim M, Tang M, Inman RD. Altered cytotoxicity profile of CD8+ T cells in ankylosing spondylitis. *Arthritis Rheumatol*. 2020;72:428–434. doi:10.1002/art.41129

30. Braun J. Discontinuing tumour necrosis factor inhibitors in non-radiographic axial spondyloarthritis. *Lancet*. 2018;392:98–100. doi:10.1016/s0140-6736(18)31440-5

31. Bank I. The role of gamma delta T cells in autoimmune rheumatic diseases. *Cells*. 2020;9:462. doi:10.3390/cells9020462

32. Pedersen SJ, Maksymowych WP. The pathogenesis of ankylosing spondylitis: an update. *Curr Rheumatol Rep*. 2019;21:58. doi:10.1007/s11926-019-0856-3

33. Mathieu A, Cauli A, Fiorillo MT, Sorrentino R. HLA-B27 and ankylosing spondylitis geographic distribution as the result of a genetic selection induced by malaria endemic? A review supporting the hypothesis. *Autoimmunity Rev*. 2008;7:398–403. doi:10.1016/j.autrev.2008.03.013

34. Peruzzi M, Wagtmann N, Long EO. A p70 killer cell inhibitory receptor specific for several HLA-B allotypes discriminates among peptides bound to HLA-B*2705. *J Exp Med*. 1996;184:1585–1590. doi:10.1084/jem.184.4.1585

35. Azuz-Lieberman N, Markel G, Mizrahi S, et al. The involvement of NK cells in ankylosing spondylitis. *Int Immunol*. 2005;17:837–845. doi:10.1093/intimm/dxh270

36. Snaith HA, Armstrong CG, Guo Y, Kaiser K, Cohen PT. Deficiency of protein phosphatase 2A uncouples the nuclear and centrosome cycles and prevents attachment of microtubules to the kinetochore in Drosophila microtubule star (mts) embryos. *J Cell Sci*. 1996;109(Pt 13):3001–3012. doi:10.1242/jcs.109.13.3001

37. Ren C, Li M, Zheng Y, et al. Single-cell RNA-seq reveals altered NK cell subsets and reduced levels of cytotoxic molecules in patients with ankylosing spondylitis. *J Cell Mol Med*. 2022;26:1071–1082. doi:10.1111/jcmm.17159

38. Sun L, Xi S, He G, et al. Two to tango: dialogue between adaptive and innate immunity in type 1 diabetes. *J Diabetes Res*. 2020;2020:4106518. doi:10.1155/2020/4106518

39. Ma Q. Role of nrf2 in oxidative stress and toxicity. *Annu Rev Pharmacol Toxicol*. 2013;53:401–426. doi:10.1146/annurev-pharmtox-011112-140320

40. Tsuchiya H, Ota M, Sumitomo S, et al. Parsing multiomics landscape of activated synovial fibroblasts highlights drug targets linked to genetic risk of rheumatoid arthritis. *Ann Rheumatic Dis*. 2021;80:440–450. doi:10.1136/annrheumdis-2020-218189

41. Ni M, Solmonson A, Pan C, et al. Functional assessment of lipoyltransferase-1 deficiency in cells, mice, and humans. *Cell Rep*. 2019;27:1376–1386.e1376. doi:10.1016/j.celrep.2019.04.005

42. Deng L, Jiang A, Zeng H, Peng X, Song L. Comprehensive analyses of PDHA1 that serves as a predictive biomarker for immunotherapy response in cancer. *Front Pharmacol*. 2022;13:947372. doi:10.3389/fphar.2022.947372

43. Jiang X, Ji S, Yuan F, et al. Pyruvate dehydrogenase B regulates myogenic differentiation via the FoxP1-Arih2 axis. *J Cachexia Sarcopenia Muscle*. 2023;14:606–621. doi:10.1002/jcsm.13166

44. Wang H, Yang Z, He X, et al. Cuproptosis related gene PDHB is identified as a biomarker inversely associated with the progression of clear cell renal cell carcinoma. *BMC Cancer*. 2023;23:804. doi:10.1186/s12885-023-11324-0

45. Chantong B, Kratschmar DV, Lister A, Odermatt A. Dibutyltin promotes oxidative stress and increases inflammatory mediators in BV-2 microglia cells. *Toxicol Lett*. 2014;230:177–187. doi:10.1016/j.toxlet.2014.03.001

46. Manley JL, Krainer AR. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev*. 2010;24:1073–1074. doi:10.1101/gad.1934910

47. Zhou Z, Fu XD. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma*. 2013;122:191–207. doi:10.1007/s00412-013-0407-z

48. Yao Y, Wang H, Xi X, Sun W, Ge J, Li P. miR-150 and SRPK1 regulate AKT3 expression to participate in LPS-induced inflammatory response. *Innate Immun*. 2021;27:343–350. doi:10.1177/17534259211018800

49. Liu L, Hu X. Predictive values of circulating miR-146a and miR-155 for disease activity and clinical response to TNF-α blocking therapy in patients with ankylosing spondylitis. *Int J Rheum Dis*. 2024;27:e15004. doi:10.1111/1756-185x.15004

50. Qian BP, Ji ML, Qiu Y, et al. Identification of serum miR-146a and miR-155 as novel noninvasive complementary biomarkers for ankylosing spondylitis. *Spine*. 2016;41:735–742. doi:10.1097/brs.0000000000001339

51. Liu WH, Kang SG, Huang Z, et al. A miR-155-Peli1-c-Rel pathway controls the generation and function of T follicular helper cells. *J Exp Med*. 2016;213:1901–1919. doi:10.1084/jem.20160204

52. Pham LV, Zhou HJ, Lin-Lee YC, et al. Nuclear tumor necrosis factor receptor-associated factor 6 in lymphoid cells negatively regulates c-Myb-mediated transactivation through small ubiquitin-related modifier-1 modification. *J Biol Chem*. 2008;283:5081–5089. doi:10.1074/jbc.M706307200