Nature and Science of Sleep

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Using Bioinformatics and Machine Learning to Predict the Genetic Characteristics of Ferroptosis-Cuproptosis-Related Genes Associated with Sleep Deprivation

Liang Wang[1,2,*], Shuo Wang[3,*], Chujiao Tian[3], Tao Zou[3], Yunshan Zhao[2], Shaodan Li[3], Minghui Yang[3], Ningli Chai[1]

[1]Department of Gastroenterology, First Medical Center, Chinese PLA General Hospital, Beijing, 100853, People's Republic of China; [2]Health Medicine Department, the 955th Hospital of the Army, Changdu, Tibet, 854000, People's Republic of China; [3]Department of TCM, Sixth Medical Center, Chinese PLA General Hospital, Beijing, 100048, People's Republic of China

*These authors contributed equally to this work

Correspondence: Shaodan Li, Department of TCM, Sixth Medical Center, Chinese PLA General Hospital, Beijing, 100048, People's Republic of China, Email lsd301@126.com; Ningli Chai, Department of Gastroenterology, First Medical Center, Chinese PLA General Hospital, Beijing, 100853, People's Republic of China, Email chainingli@vip.163.com

Purpose: Sleep deprivation (SD), a common sleep disease in clinic, has certain risks, and its pathogenesis is still unclear. This study aimed to identify ferroptosis-cuproptosis-related genes (FCRGs) associated with SD through bioinformatics and machine learning, thus elucidating their biological significance and clinical value.

Methods: SD-DEGs were obtained from GEO. We intersected key WGCNA module genes of DE-FCRGs with SD-DEGs to obtain SD-DE-FCRGs. GO and KEGG analyses were performed. Machine learning was used to screen SD-DE-FCRGs, and filtered genes were intersected to obtain SD characteristic genes. ROC curves were used to evaluate the accuracy of SD characteristic genes. CIBERSORT was used to analyze the correlation between SD-DE-FCRGs and immune cells. We constructed a ceRNA network of SD-DE-FCRGs and used DGIbd to predict gene drug targets.

Results: The 156 DEGs were identified from GSE98566. Five SD-DE-FCRGs from DE- FCRGs and SD-DEGs were analyzed via WGCNA, and enrichment analysis involved mainly ribosome regulation, mitochondrial pathways, and neurodegenerative diseases. Machine learning was used to obtain Four SD-DE-FCRGs (IKZF1, JCHAIN, MGST3, and UQCR11), and these gene analyses accurately evaluated the distribution model (AUC=0.793). Immune infiltration revealed that SD hub genes were correlated with most immune cells. Unsupervised cluster analysis revealed significant differential expression of immune-related genes between two subtypes. GSVA and GSEA revealed that enriched biological functions included oxidative phosphorylation, ribonucleic acid, metabolic diseases, activation of oxidative phosphorylation, and other pathways. Four SD-DE-FCRGs associated with 29 miRNAs were identified via the construction of a ceRNA network. The important target lenalidomide of IKZF1 was predicted.

Conclusion: We first used bioinformatics and machine learning to screen four SD-DE-FCRGs. These genes may affect the involvement of infiltrating immune cells in pathogenesis of SD by regulating FCRGs. We predicted that lenalidomide may target IKZF1 from SD-DE-FCRGs.

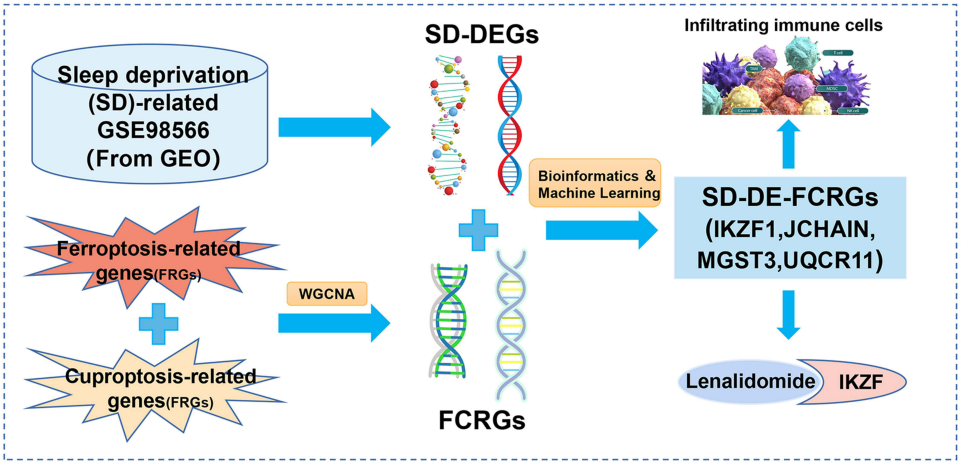
Keywords: sleep deprivation, ferroptosis, cuproptosis, bioinformatics, machine learning, immune infiltration

Introduction

Adequate sleep is an important support for human physical and mental health and social activities.[1] With the development of the global social economy and changes in human lifestyle, people's sleep time has significantly shortened. As a common clinical disease of the nervous system, sleep deprivation (SD) has become a very common social phenomenon.

Received: 13 May 2024
Accepted: 12 September 2024
Published: 24 September 2024

Graphical Abstract



Recent studies have shown that sleeping less than 7 hours per night increases the risk of death by 12%-35%.[2] Long-term SD can harm human health and increase the risk of developing other diseases.[3]

In recent years, with the development and popularization of high-throughput sequencing technology, disease research models based on data mining technologies such as bioinformatics and machine learning have gradually matured, providing important methods and means for mining potential mechanisms, potential biomarkers and therapeutic targets of various diseases.[4] Sleep-related bioinformatics research is an emerging field worthy of in-depth study. Sleep studies have shown that obstructive sleep apnea, as a sleep disorder, may have the same bioinformatics basis as several sleep phenotypes in terms of etiology and causal pathways.[5] Exploring the genetic relationship between sleep function and sleep deprivation from a bioinformatics perspective is highly beneficial for evidence-based and personalized sleep deprivation treatment in the future.[6] Ferroptosis and cuproptosis, new modes of cell death that have been discovered one after another and are different from the traditional apoptosis pathway, have been confirmed to play important roles in the pathophysiological processes of cancer, the nervous system and other diseases, mostly involving immunity, cognition and inflammation.[7–9] Previous studies have shown that SD can lead to neuronal ferroptosis and affect hippocampal function and that copper transport signaling is a potential therapeutic target for sleep disorder-related myocardial injury and even ischemic myocardial injury.[10,11] At present, the pathogenesis of SD is not clear, and no relevant literature has analyzed the role of ferroptosis combined with cuproptosis in SD via data mining methods.

In this study, we integrate previously published transcriptome data related to SD, propose the use of bioinformatics and machine learning, identify the relevant hub genes involved in SD and their relationships with immune cell infiltration through ferroptosis combined with the cuproptosis pathway, screen key genes, and mine potential drugs that can treat SD. This study provides a reference for further exploration of the key target genes of SD-differentially expressed-ferroptosis-cuproptosis-related genes (SD-DE-FCRGs) and screening potential drugs for the treatment of SD. The specific design scheme and process of this study are shown in Figure 1.

Materials and Methods

Data Acquisition

The original SD dataset GSE98566 (annotation platform: GPL6244) was obtained from the GEO database (https://www.ncbi.nlm.nih.gov/GEO/). Ferroptosis-related genes (FRGs) were downloaded from the FerrDb database (http://www.zhounan.org/ferrdb/) and retrieved from previous publications. Cuproptosis-related genes (CRGs) were obtained from the literature and materials databases.[12–15]

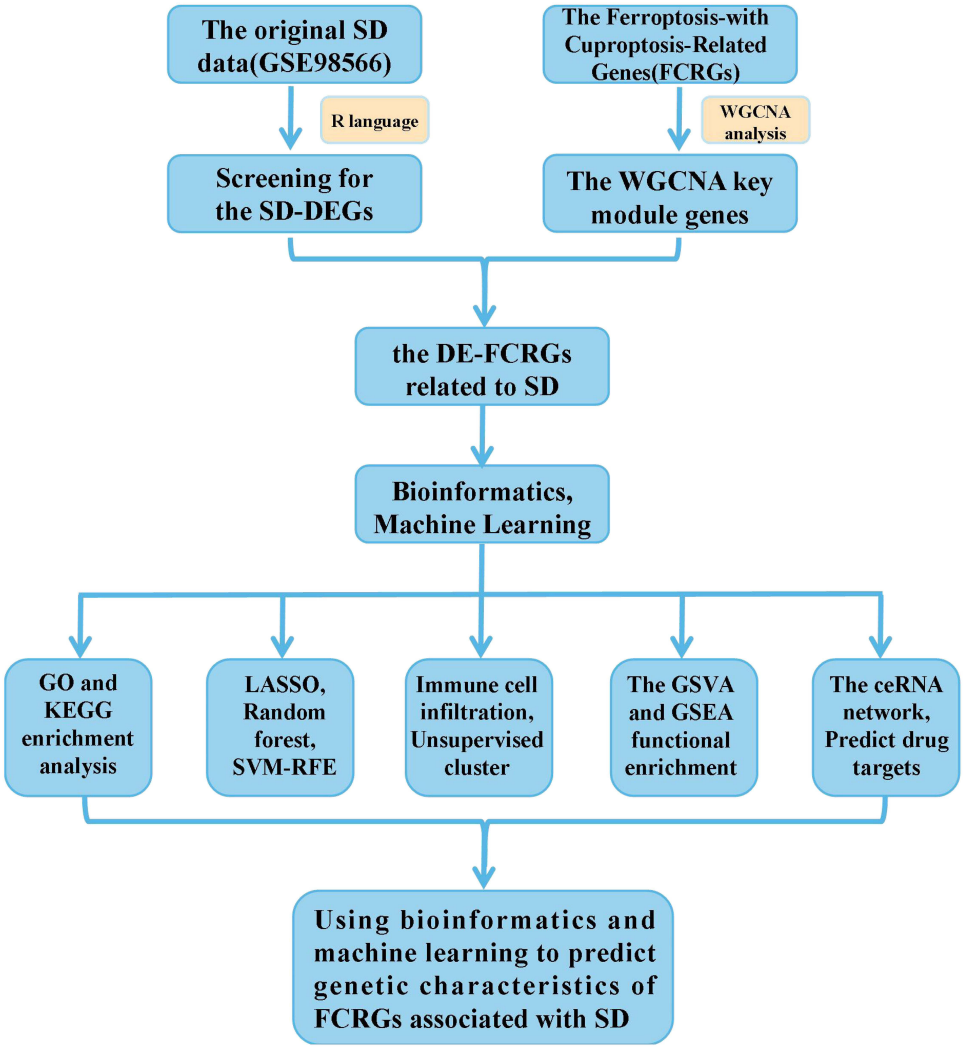


Figure 1 Flowchart of the overall study design.

Identification of SD-DE-FCRGs

We used the R package to process, count, and analyze expression files. We downloaded a series of matrix files for the dataset from GEO. The R package "limma" was used to normalize the data and locate differentially expressed genes (DEGs), with $P<0.05$ and $\log2 |FC|>0.3$ considered significant DEGs.[16] The "ggplot2" package was used to construct a volcano map of the DEGs. In addition, the mRNA expression data obtained via log2 transformation were arranged into a heatmap via the "pheatmap" R package. The "GSVA" package in R was applied to compute the FCRGs scores of each patient with SDs, which could serve as an indicator of FCRGs activity. The "wilcox.test" function in R was then employed to compare the discrepancy in the FCRGs scores between different clusters.

Weighted Gene Coexpression Network Analysis (WGCNA) to Determine the Core SD-DE-FCRGs

On the basis of the results of WGCNA, we obtained gene modules related to ferroptosis and cuproptosis expression and explored the correlations between gene networks and ferroptosis-cuproptosis, as well as the core SD-DE-FCRGs in the network.[17] We used the GSVA method to quantify the scores of ferroptosis genes and analyzed the differences between the ferroptosis scores and ferroptosis scores via the Wilcoxon test.

Functional Enrichment Analysis of the Core SD-DE-FCRGs

Gene Ontology (GO) analysis is based on the analysis of the cellular component (CC), biological process (BP), and molecular function (MF) of candidate genes. The Kyoto Encyclopedia of Genes and Genomes (KEGG) mainly generates pathway maps on the basis of the relationship network formed by gene sequences, genome information, metabolism, disease, etc. To analyze the enrichment of gene biological functions and pathways, the "clusterProfiler" package was used for enrichment analysis of brown module genes, and the "org.db" package and "ggplot2" package were used for data and information visualization. GO and KEGG enrichment pathways with p and q values less than 0.05 were considered significant categories.

Machine Learning Feature Gene Selection

Three machine learning algorithms (LASSO, random forest, and SVM-RFE) were used to screen feature.[18,19] Compared with regression analysis, LASSO is a dimensionality reduction method that demonstrates superiority in evaluating high-dimensional data. LASSO analysis searches for the smallest classification error λ to determine variables and is mainly used for screening feature variables and constructing the best classification model. LASSO analysis was conducted via the glmnet software package via 10-fold cross validation and the turning/privacy parameters. Recursive feature elimination (RFE) of the random forest algorithm is a supervised machine learning method that was used to sort SD-DE-FCRGs. The prediction performance was estimated through 10-fold cross-validation, and the relative importance of the top 5 genes was identified as the characteristic gene. SVM-RFE was superior to linear discriminant analysis and mean squared error in selecting relevant features and removing redundant features. SVM-RFE was applied to feature selection through tenfold cross validation. The genes associated with these genes are considered hub genes in SD diagnosis.

ROC analysis was used to estimate diagnostic efficacy by combining hub genes to establish nomograms.[20] The calibration curve is used to evaluate the accuracy of the column chart. The clinical practicality of the column charts was evaluated through decision curve analysis.

Immune Cell Infiltration

We used ssGSEA to analyze the infiltration level of immune cells on the basis of the expression profiles of 28 immune-related features. The estimated proportions of 28 immune cell subtypes in each sample were subsequently visualized via the "forcats" and "tidyHetmap" packages. The "coreplot" package was used to analyze the interaction between immune cells and further analyze the impact of disease-immune cell interactions. The "vioplot" package was used to plot the relative content of immune cells between different subtypes.[21]

Unsupervised Clustering of SD-DE-FCRGs

On the basis of the expression levels of the 5 SD-DE-FCRGs, unsupervised cluster analysis was performed via the R package "Consensus Cluster Plus" to classify the SD samples into different clusters.[22] The consensus matrix graph, consensus cumulative distribution function (CDF) graph, relative change in area under the CDF curve and tracking graph are used to find the optimal number of clusters. Principal component analysis (PCA) was used to determine the differences in HF progression and immune-related gene expression between the two subtypes. The PCA diagram was depicted via the ggplot2 package.

Gene Set Variation Analysis (GSVA) and Gene Set Enrichment Analysis (GSEA)

We obtained the "c2. cp. kegg. Hs. symbols" file from the MSigDB online database (http://software.broadinstitute.org/gsea/msigdb) and conducted GSVA enrichment analysis via R packages to evaluate the activity and biological function of different pathways among different subtypes.[23] We used the "vioplot" package to plot the relative content of immune cells among the different subtypes.

The implementation of GSEA aims to elucidate the biological significance of characteristic genes in terms of functionality. The "c2. cp. kegg. v7.4. symbols. gmt". The gene set from MSigDB was used as a reference set.[24,25] To obtain standardized enrichment scores for each analysis, a gene set arrangement was conducted once. FDR<0.05 was considered to indicate significant enrichment.

Building a ceRNA Network and Predicting Target Drugs

We used the multiMiR R package as a tool for predicting miRNA binding sites to estimate the miRNAs with which SD-DE-FCRGs may interact. We subsequently constructed a ceRNA network of mRNAs via the visualization function of Cytoscape 3.8.0. Finally, we used the DGIbd database (https://dgidb.org/) to predict the target drugs of the SD-DE-FCRGs.[26]

Statistical Analysis

All the statistical analyses were conducted via R software (version 4.1.2). Receiver operating characteristic analysis was used to measure the extent to which the model predicted outcomes. Immune infiltrating cells were compared via the Wilcoxon test. Visualization was performed via the "ggplot2", "pheatmap", and "forestplot" packages. The Wilcoxon test was used to analyze the differences between the two groups. All the statistical tests conducted were bilateral, and values of $p<0.05$ were considered statistically significant.

Results

Data Preprocessing and Identification of DEGs

The original SD dataset GSE98566 was collected from the GEO database and included 92 SD samples and 71 normal samples. We used a filter with a |log2-fold change (FC)| > 0.3 and a p value < 0.05 in the analysis process of the SD dataset GSE98566. A total of 156 DEGs were identified in this analysis, including 138 highly expressed up-regulated genes (CA1, CLEC12A, CLEC1B, etc). and 18 highly expressed down-regulated genes (DEFA1B, DEFA4, EIF1AY, etc)., as shown in Figure 2A. The visualization of the DEGs in the form of volcano maps and heatmaps is shown in Figure 2B. These 728 FRGs were downloaded from the FerrDb database (http://www.zhounan.org/ ferrdb/) and retrieved from previous publications. These 13 CRGs were obtained from the literature and materials databases. A differential expression box diagram of the ferroptosis scores and cuproptosis scores in the SD group and normal group is shown in Figure 2C.

The Construction of WGCNA

A coexpression network of all genes in the dataset was constructed via the WGCNA-R package, and hierarchical clustering was performed (Figure 3A), with a soft threshold set to 7 (Figure 3B). The weighted adjacency matrix was transformed into a TOM to estimate the network connectivity, and the hierarchical clustering method was used to construct the clustering tree structure of the TOM matrix (Figure 3C). The different branches of the clustering tree represent different gene modules, and different colors represent different modules. On the basis of the weighted correlation coefficient of genes, genes are classified according to their expression pattern, and genes with similar patterns are classified into one module. Genes are divided into multiple modules through their gene expression patterns. The

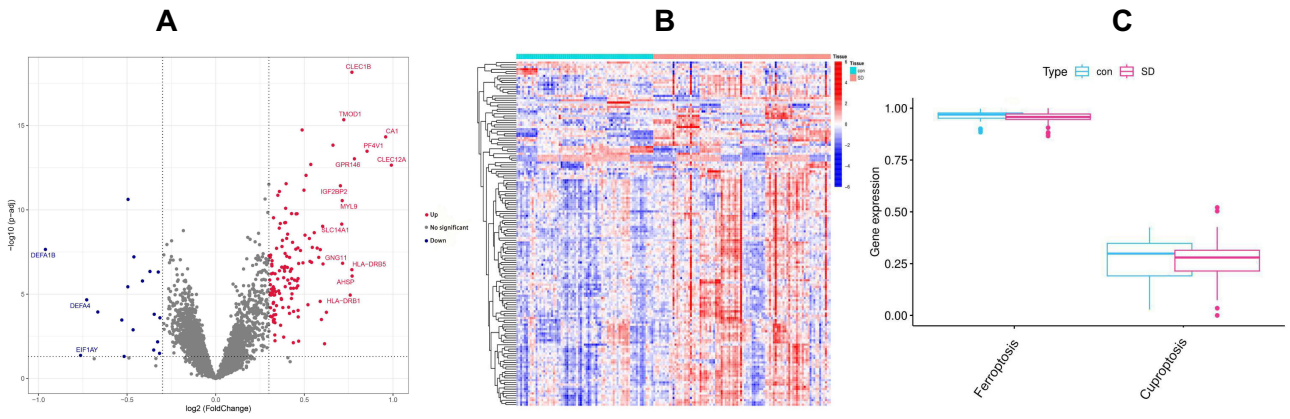


Figure 2 The Identification of SD-DEGs and DE-FCRGs. (**A**) Volcano plot of SD-DEGs. Note: The red dots indicate 138 highly expressed up-regulated genes (CA1, CLEC12A, CLEC1B, etc.), the gray dots indicate no significance expression of differential genes, and the blue dots indicate 18 highly expressed down-regulated genes (DEFA1B, DEFA4, EIF1AY, etc). (**B**) Clustered heatmap of SD-related FCRGs expression levels. (**C**) Box diagram of the DE-FCRGs scores in the SD group and normal group.

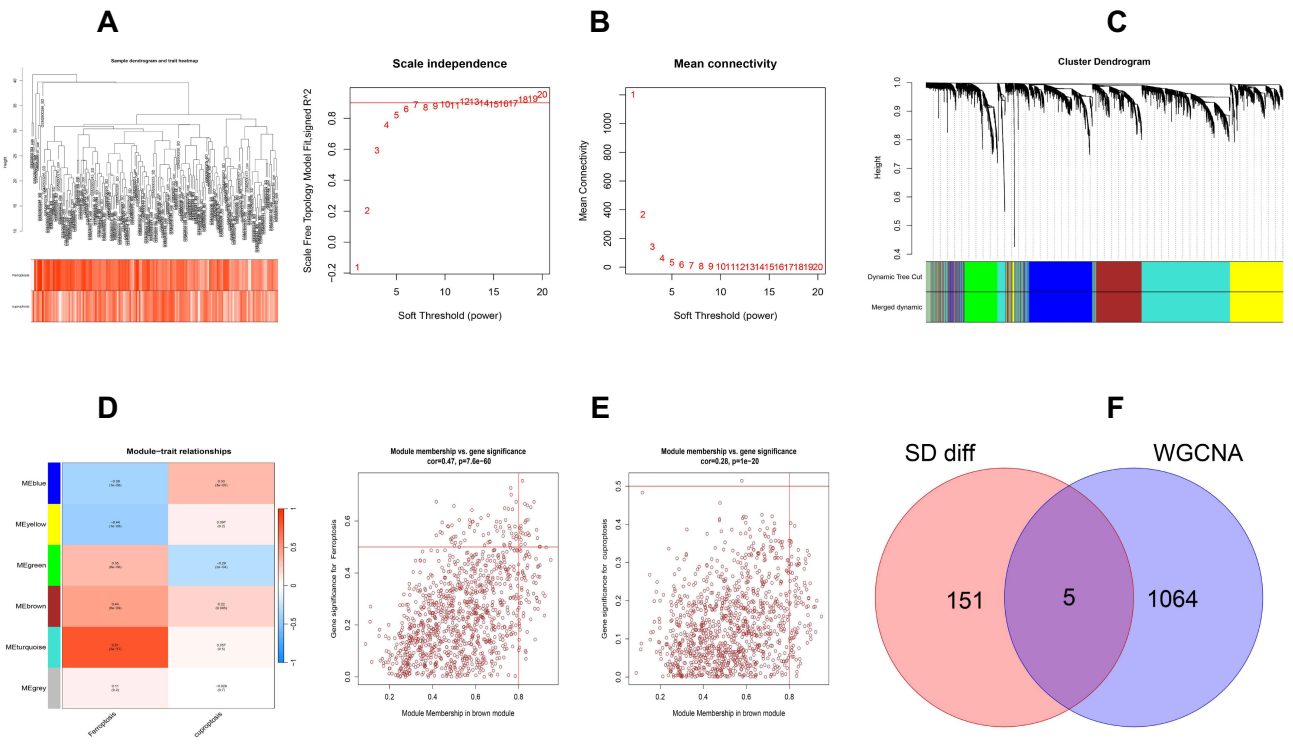


Figure 3 The construction of the WGCNA graph. (**A**) Dendrogram and trait heatmap formed by hierarchical clustering of samples. (**B**) Building a gene relationship network. (**C**) Constructing gene modules through hierarchical clustering. (**D**) DE-FCRG quantification module feature association. (**E**) Identification of the brown module as most relevant to DE-FCRGs. (**F**) Intersection of SD-DEGs and WGCNA key module genes (SD-DE-FCRGs).

associations among the six modules evaluated revealed that brown modules were most closely associated with ferroptosis and cuproptosis and contained 1069 genes (Figure 3D and E). The Venn online tool was used to identify five SD-DE-FCRGs (IKZF1, JCHAIN, MGST3, UQCR11, and XCL1) from the intersection of the SD-DEGs and key module genes identified via WGCNA (Figure 3F).

GO and KEGG Enrichment Pathway Analysis

We investigated the main functions of the SD-DE-FCRGs through GO and KEGG enrichment pathway analyses. GO enrichment analysis included BP, CC, and MF, where BP included mRNA processing, histone modification, RNA splicing, RNA splicing, transesterification reactions with bulged adenosine as the nucleophile, mRNA splicing via the spliceosome, RNA splicing via transesterification reactions, regulation of mRNA metabolic processes, Golgi vesicle transport, regulation of DNA metabolic processes, cytosolic transport, etc. Moreover, we also demonstrated CC enrichment of ribosomes, ribosomal subunits, mitochondrial protein-containing complexes, spliceosomal complexes, respirasomes, mitochondrial respirasomes, catalytic step 2 spliceosomes, U2-type spliceosomal complexes, cytosolic ribosomes, and large ribosomal subunits. The MF terms included structural constituent of ribosome, transcription coregulator activity, oxidoreduction-driven active transmembrane transporter activity, RNA polymerase II-specific DNA-binding transcription factor binding, DNA-binding transcription factor binding, histone binding, transcription coactivator activity, primary active transmembrane transporter activity, electron transfer activity, and NADH dehydrogenase (ubiquinone) activity (Figure 4A and B). According to the KEGG analysis, the DEGs were enriched in signaling pathways such as amyotrophic lateral sclerosis, pathways related to neurodegeneration-multiple diseases, Huntington's disease, Alzheimer's disease, Parkinson's disease, thermogenesis, prion disease, ribosomes, chemical carcinogenesis-reactive oxygen species, and spliceosomes (Figure 4C and D).

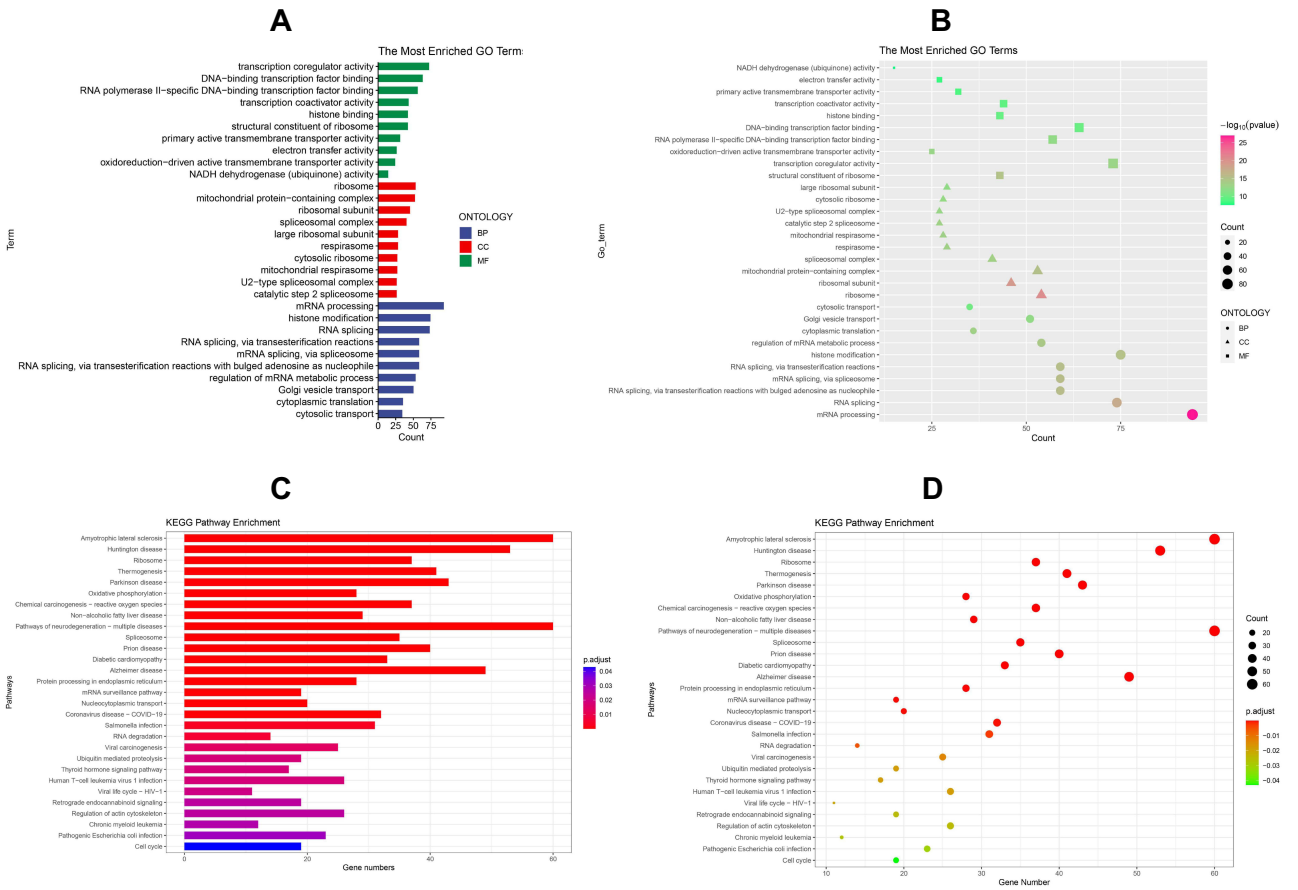


Figure 4 The GO and KEGG analysis of SD-DE-FCRGs. (**A** and **B**) The most enriched GO analysis terms, including BP, CC, and MF. (**C** and **D**) Enriched terms in the KEGG pathway analysis.

Selection of Feature Genes in Machine Learning

As a technical means of implementing artificial intelligence (AI), machine learning can automatically analyze existing data to obtain rules or models and use guidelines to predict unknown data. Using the LASSO regression algorithm, five genes (IKZF1, JCHAIN, MGST3, UQCR11, XCL1) extracted from SD and DE-FCRGs were identified as potential diagnostic biomarkers (Figure 5A and B). The RF algorithm identified 5 genes with diagnostic value (Figure 5C and D). Through the SVM-REF algorithm, five genes extracted from these potential targets were used as candidate biomarkers (Figure 5E and F). The characteristic gene Wayne maps obtained from the three algorithms were subsequently overlaid to obtain four diagnostic biomarkers, IKZF1, JCHAIN, MGST3, and UQCR11 (Figure 5G).

ROC Curve

The area under the ROC curve is called the AUC and is currently considered the standard method for evaluating the accuracy of predictive distribution models. The receiver operating characteristic curves of the four disease characteristic genes obtained from machine learning methods were drawn, and the AUC values of the four disease characteristic genes were greater than 0.5. Finally, the AUC value of the detection model was 0.793, indicating that IKZF1, JCHAIN, MGST3, and UQCR11 were SD characteristic genes with a certain accuracy, as shown in Figure 6A–D.

Evaluation of Immune Cell Infiltration

To further confirm the correlation between the expression of SD-DE-FCRGs and 24 types of immune cells, the CIBERSORT algorithm was used to analyze the proportion of infiltrating immune cells, and a map of 24 immune cells in the SD sample was drawn (Figure 7A).

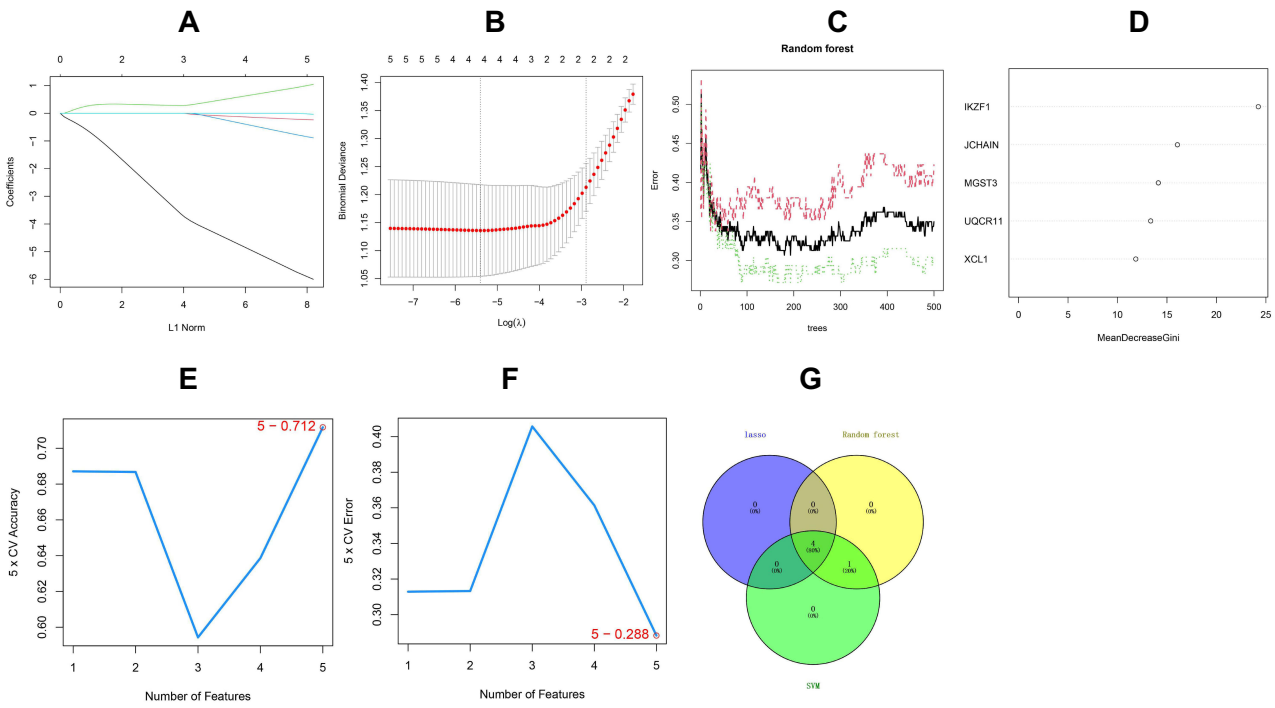


Figure 5 Three machine learning algorithms for screening SD-DE-FCRGs characteristic genes. LASSO algorithm for screening candidate genes. LASSO coefficient profiles, (**C-D**) Construction of the random forest model, (**E** and **F**) SVM-RFE algorithm. Notes: The vertical axis in the figure represents the size of the error, whereas the horizontal axis represents the number of genes.(**G**) The Wayne map of SD-DE-FCRGs was obtained via the intersection of the LASSO, SVM-RFE and RF algorithms.

Correlation analysis between four types of SD-DE-FCRGs (IKZF1, JCHAIN, MGST3, and UQCR11) and 24 immune cells revealed that the majority of immune cells were significantly correlated with SD-DE-FCRGs (Figure 7B). IKZF1 expression was positively correlated with T helper cells, TILs, and MHC class I. Neutrophils, CD8+ T cells, and inflammation-promoting macrophages were positively correlated; IKZF1 expression was negatively correlated with NK cells, the type II IFN response, T-cell coinhibition, and pDCs (Figure 7C). JCHAI expression was positively correlated with APC cosimulation, cytological activity, the type II IFN response, and Th1 cells. JCHAI expression and neutrophil, T helper cell, HLA, and CD8+ T cell, macrophage, and type I IFN responses were negatively correlated (Figure 7D). MGST3 expression was positively correlated with type II IFN response, NK cells, cytosolic activity, Th1 cells, T-cell coinhibition, APC cosimulation, checkpoints, pDCs, T-cell cosimulation, and Th2 cells. MGST3 expression and neutrophil, T helper cell, macrophage, Treg, HLA, and TIL counts were negatively correlated (Figure 7E). UQCR11 expression was positively correlated with NK cells, the type II IFN response, APC cosimulation, T-cell coinhibition, parainflammation, and the type I IFN response. UQCR11 expression was negatively correlated with T helper cells, neutrophils, TILs, macrophages, inflammation-promoting cells, MHC class I cells, CD8+ T cells, CCRs, HLA cells, and B cells (Figure 7F).

The correlations among the 24 immune cell scores were analyzed on the basis of the expression levels of IKZF1, JCHAIN, MGST3, and UQCR11. Different immune cells, especially NK cells, Th1 cells, APC Co simulations, B cells, checkpoints, T-cell Co simulations, macrophages, CCRs, type II IFN responses, TILs, T helper cells, and MHC class I cells, exhibited different effects in the two groups (Figure 7G). These results further support that the levels of IKZF1, JCHAIN, MGST3, and UQCR11 may affect the immune activity of immune cells.

Unsupervised Clustering of SD-DE-FCRGs

A consensus matrix graph, consensus CDF graph, relative change in the area under the CDF curve, and tracking graph were used to find the optimal number of clusters (Figure 8A–C). PCA was conducted, indicating a good degree of clustering between the SD and normal groups (Figure 8D). The Wilcoxon test was used to determine the differences in the expression of immune-related genes between the two subtypes of SD (Figure 8E).

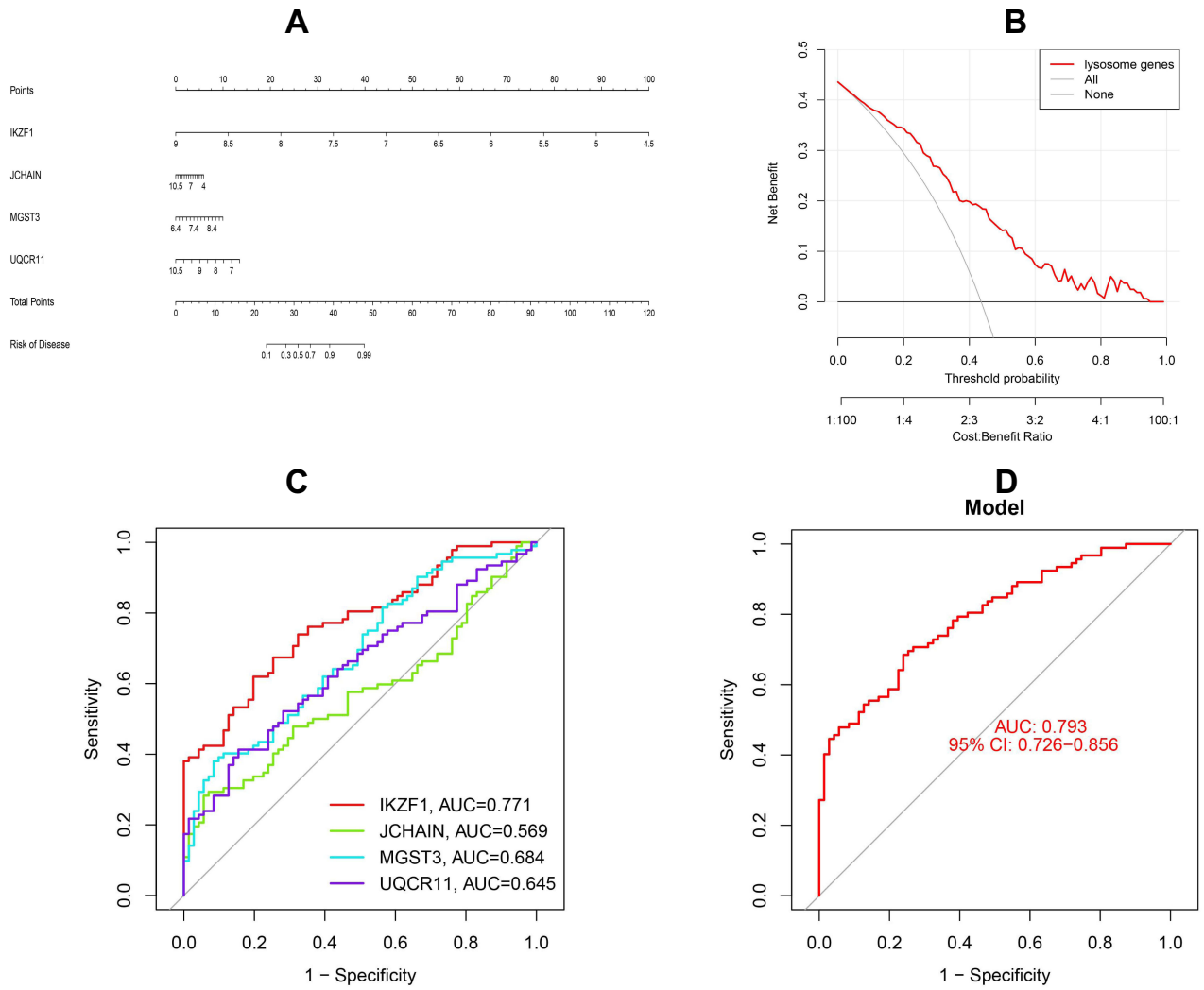


Figure 6 The receiver operating characteristic curves of four SD-DE-FCRGs. (**A**) Nomogram for predicting SD risk on the basis of signature genes. (**B**) Decision curve showing the clinical value of the nomogram. (**C**) Time ROC curve of the SD-DE-FCRGs risk score. (**D**) ROC curve for the validation set.

GSVA and GSEA Functional Enrichment

The R package GSVA was used for enrichment analysis. The significantly enriched pathways and biological functions of group A and group B in the SD patients were associated mainly with oxidative phosphorylation, melanogenesis, various metabolic and degenerative diseases and other signaling pathways, revealing a possible molecular mechanism underlying the difference between the two groups of SD patients (Figure 9A).

Single-gene enrichment analysis revealed that the influence of four key molecules on SD involves multiple signaling pathways, such as IKZF1, JCHAIN, MGST3, and UQCR11, and the activation of related signaling pathways, such as oxidative phosphorylation, ribosome, porphyrin and chlorophyll metabolism, and neural degenerative disease, with a significant correlation (Figure 9B–E).

Building the ceRNA Network and Predicting Drugs

We found that four characteristic genes were associated with 29 miRNAs by constructing a ceRNA network of mRNAs. The red hexagons in the figure represent SD-DE-FCRGs; the green ovals represent the miRNA map (Figure 10A and B). Moreover, important targets, such as lenalidomide, for IKZF1 gene drugs were predicted (Figure 10C and D).

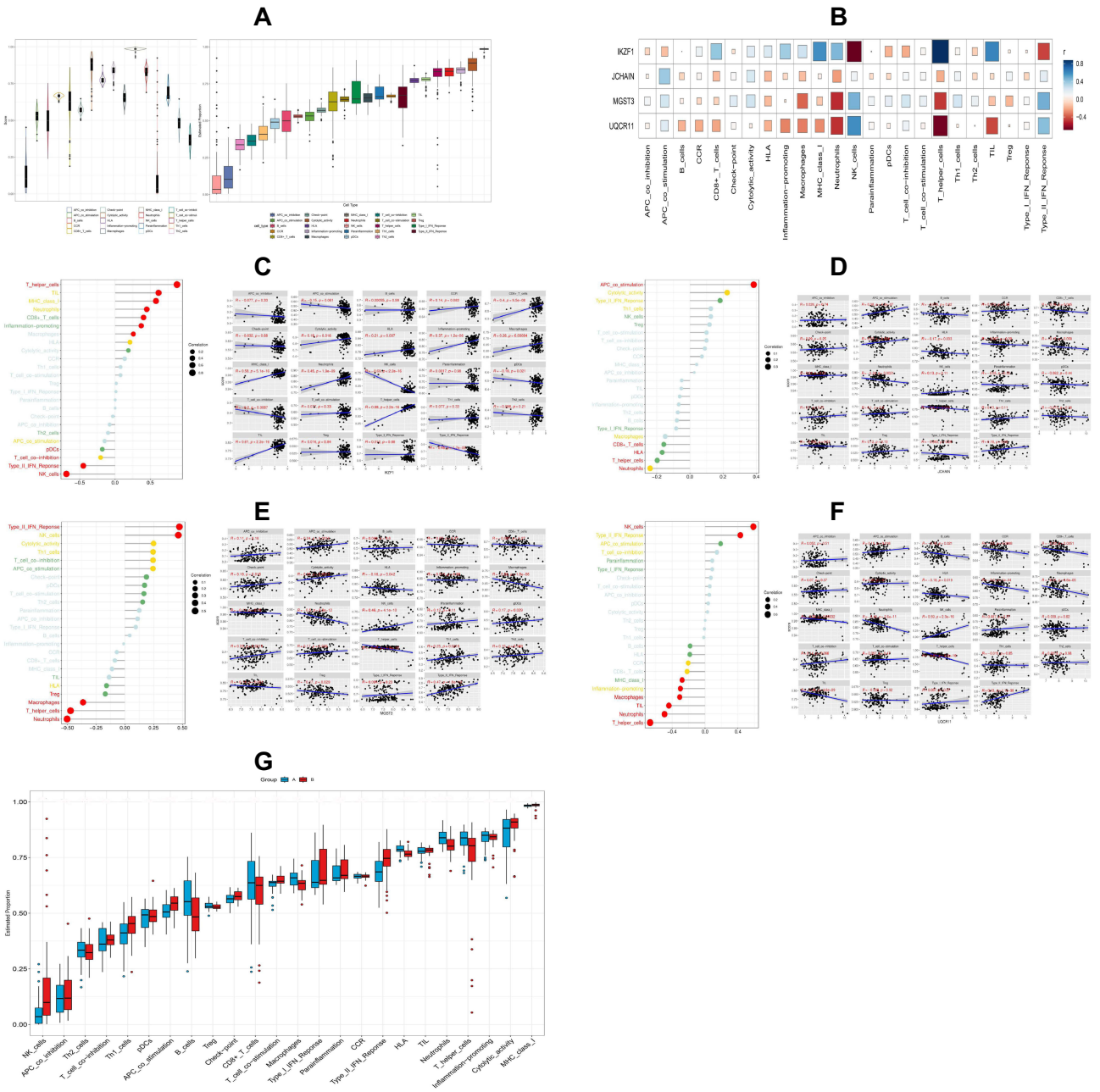


Figure 7 Correlations between key SD-DE-FCRGs and immune cells. (**A**) Box plot of the scores and estimated proportions of different immune cell infiltrates in SD patients. (**B**) Correlations between the hub genes and immune cells. (**C**–**F**) Correlations between IKZF1, JCHAIN, MGST3, and UQCR11 genes and immune cells. Notes: The red, yellow and green dots indicate P<0.05, and gray dots indicate P≥0.05. (**G**) Correlations between two sets of risk scores and 24 different immune cell subtypes in patients with SD.

Discussion

It is generally believed that SD specifically refers to a decrease in sleep time and frequent staying up late for various subjective reasons. Owing to changes in modern lifestyles, overall sleep time in society has sharply decreased, resulting in SD affecting the work efficiency and quality of life of millions of people worldwide and gradually leading to affective disorders such as mental and psychological disorders.[27] The widespread application of bioinformatics analysis and high-throughput sequencing technology, as well as the gradual maturity of machine learning in bioinformatics applications, provides important artificial intelligence methods and means for further data mining of potential mechanisms,

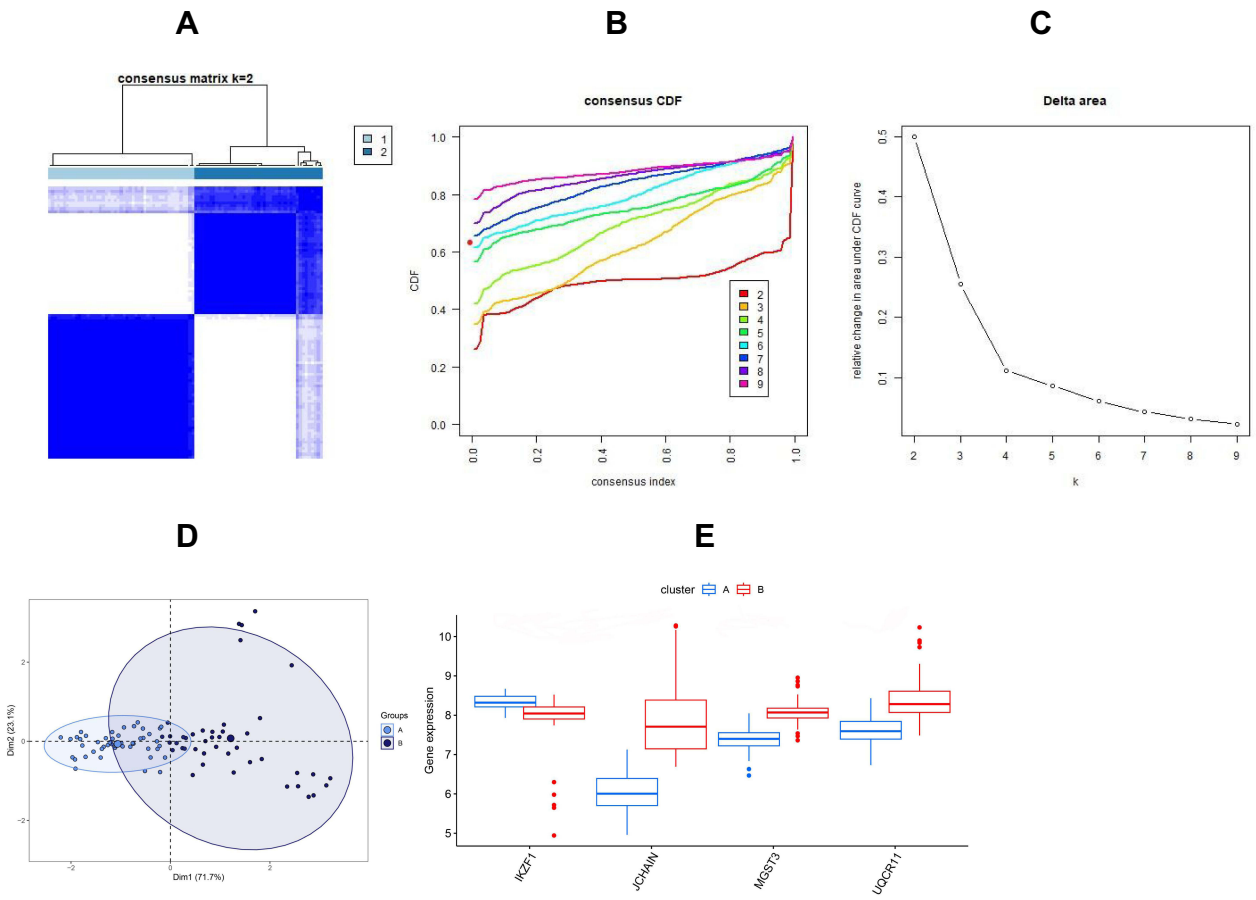


Figure 8 Construction of two ferroptosis–cuproptosis subtypes in SD. (**A**) Consensus clustering matrix when k=2. (**B** and **C**) Relative alterations in the area under the CDF curve. (**D**) PCA plot showing the distribution of the two subclusters. (**E**) Box plot showing the differential expression of SD-DE-FCRGs between subtypes.

biomarkers, and therapeutic targets of various diseases.[28] Scientists such as Yin Yen Lee have integrated multiple whole-genome datasets and applied machine learning to identify genes and pathways related to sleep regulation.[29]

In recent years, ferroptosis and cuproptosis[15] have been confirmed to be highly correlated with diseases and iron- or copper-dependent programmed new cell death pathways and have been found to affect the occurrence and development of many diseases, such as nervous system diseases and tumors.[15,30,31] Research has shown[32] that the disruption of iron homeostasis or even ferroptosis in the brain can significantly damage the oxidative metabolism of neurons, exerting significant effects on synaptic plasticity, myelin formation, neurotransmitter synthesis, and energy metabolism.[33] Research has suggested that copper can bind to various proteins or enzymes and participate in regulating multiple physiological processes, such as energy metabolism, mitochondrial respiration, and antioxidant activity.[12,34] These effects are all related to the disruption of the physiological mechanisms of neurological diseases.

However, as ferroptosis is a common manifestation of clinical neurological diseases, there have been no reports on the mechanism of ferroptosis combined with cuproptosis in SD. Therefore, the purpose of this study was to explore biomarkers related to SD and ferroptosis and further analyze their relationships with immune cell infiltration, signaling pathways, drug sensitivity, etc., providing new directions for clinical research on SD.

In this study, we identified 155 DEGs between the SD and control groups on the basis of the GSE98566 dataset. WGCNA of the intersection of FCRGs and SD-DEGs yielded 5 SD-DE-FCRGs. The GO and KEGG enrichment analyses involved mainly ribosome regulation, mitochondrial pathways, oxidative phosphorylation, and neurodegenerative diseases. We applied three types of machine learning to ultimately obtain four feature genes (IKZF1, JCHAIN, MGST3, and UQCR11) as potential biomarkers for SD diagnosis and validated them via receiver operating characteristic

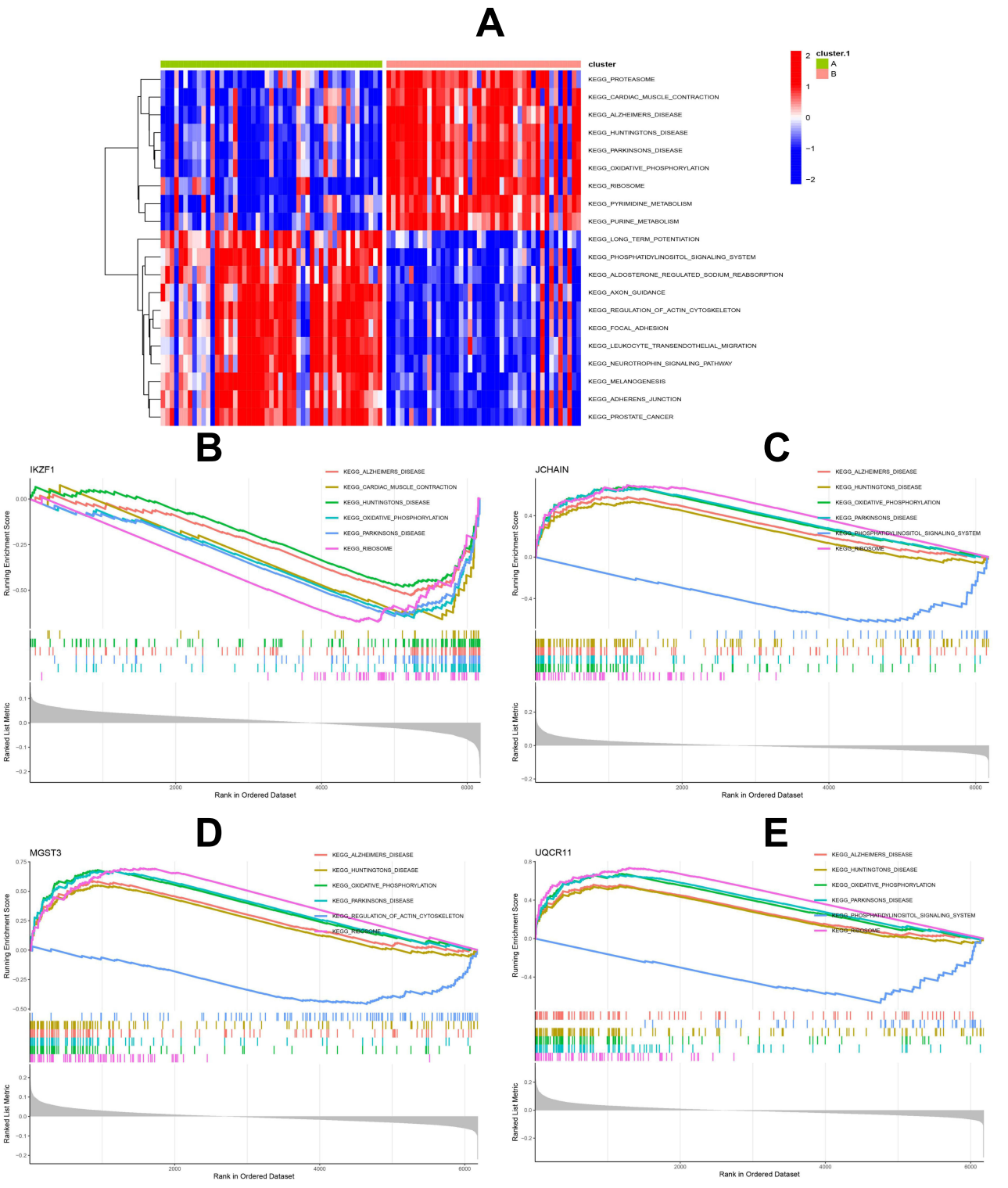


Figure 9 The GSVA and GSEA functional enrichment. (**A**) GSVA KEGG enrichment; (**B**) GSEA map of the IKZF1 gene; (**C**) GSEA map of the JCHAIN gene; (**D**) GSEA map of the MGST3 gene; (**E**) GSEA map of the UQCR11 gene.

Notes: The upper half of the graph represents the score of each gene in pathway enrichment, the lower half represents the RANK value expression of each gene, the left half represents high expression, and the right half represents low expression.

(ROC) analysis. Importantly, immune infiltration analysis revealed that the majority of immune cells were high in SD patients and that their hub genes were correlated with most immune cells. There were significant differences in the expression of 12 immune cell types and in the expression of four characteristic immune-related genes between the two

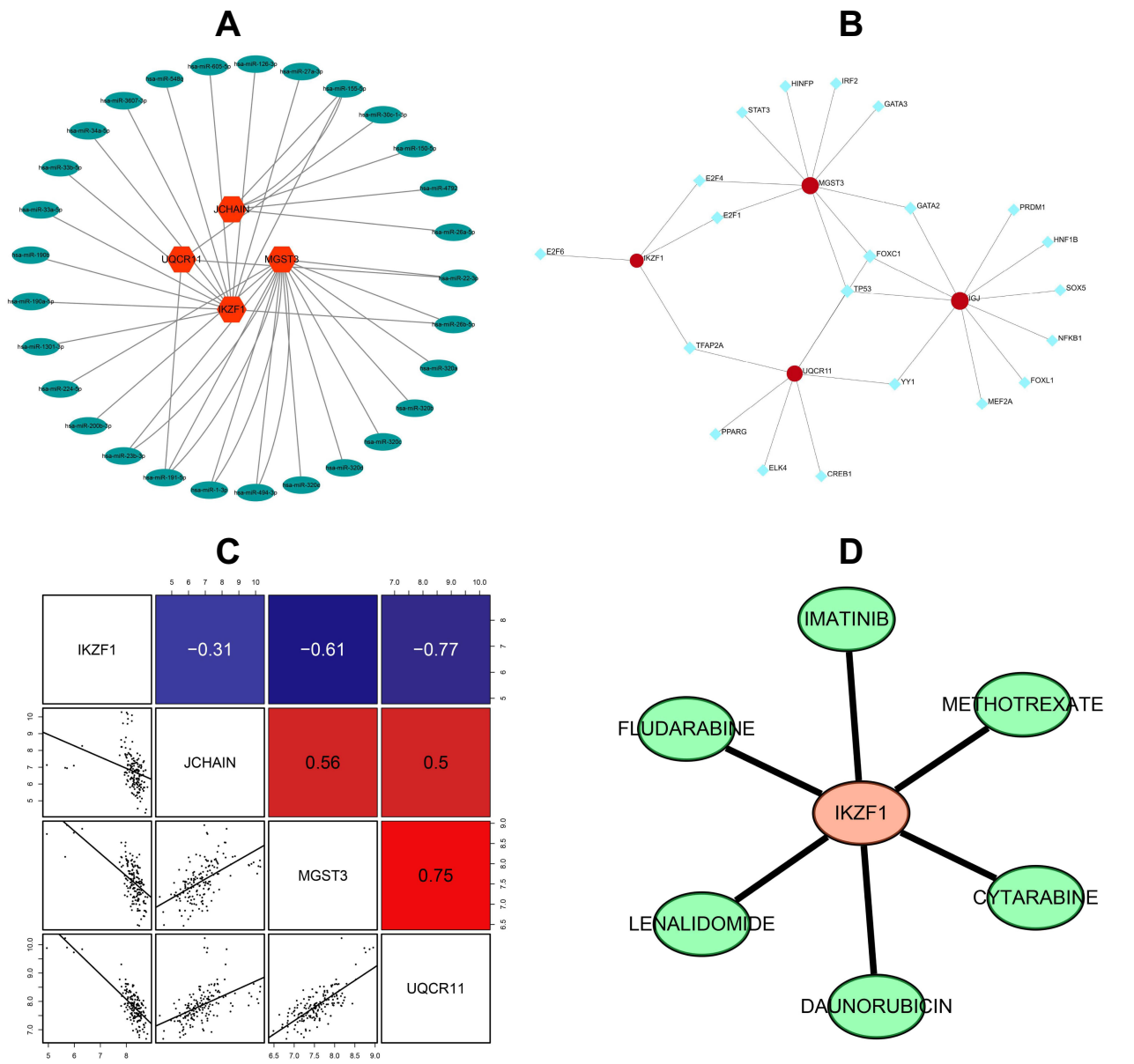


Figure 10 Building the ceRNA network and predicting drugs of four SD-DE-FCRGs. (**A**) SD-DE-FCRGs predict miRNAs. (**B**) CeRNA network map of SD-DE-FCRGs. Notes: The circles represent SD-DE-FCRGs; the blue diamonds represent lncRNAs. (**C**) Drug targets for predicting SD-DE-FCRGs. (**D**) Prediction of targets for IKZF1 gene drugs.

subtypes. IKZF1 was upregulated in T helper cells and downregulated in NK cells. JCHAIN was upregulated in APC costimulation and downregulated in neutrophils. MGST3 was upregulated in type II IFN response and NK cells and downregulated in T helper cells. UQCR11 was upregulated in NK cells and the type II IFN response and downregulated in T helper cells and neutrophils. GSEA revealed that IKZF1, JCHAIN, MGST3, and UQCR11 were related to the activation of oxidative phosphorylation, ribosomes, porphyrin and chlorophyll metabolism, neurodegenerative diseases, and other related signaling pathways, and there were significant correlations. We found that 4 characteristic genes were associated with 29 miRNAs by constructing a ceRNA network of mRNAs. We predicted important targets of IKZF1 gene drugs, such as lenalidomide, and conducted correlation analysis. These results indicate that IKZF1 is highly correlated with genes involved in the MAPK signaling pathway, cytotoxic particle movement and exocytosis, and NK cell activation. In addition, GSEA and KEGG analysis revealed that the DE-FCRGs between low- and high-IKZF1-expressing samples in GSE59867 are closely related to the immune response and NK cell-mediated inflammatory

regulation. These findings suggest that the biological function of IKZF1 in DE FCRGs may be related to the immune response related to SD and is especially closely related to the immune response of NK cells.

IKZF1 belongs to the zinc finger DNA-binding protein family of transcription factors, is involved in multiple signaling pathways and regulatory mechanisms and is particularly crucial for immune cell development and homeostasis.[35] IKZF1 is involved in the direct regulation of the expression of IL-2, a key and differentially regulated factor in many T-cell gene programs in T cells.[36,37] B lymphoid transcription factors such as IKZF1 are crucial for early B-cell development.[38] Research has shown that IKZF1 forces chronic energy deprivation, leading to constitutive activation of the energy pressure sensor AMPK, whereas the dominant negative mutant of IKZF1 alleviates glucose and energy limitations.[39] SD affects the body's energy metabolism. PET studies have shown that, as the main user of glucose, the brain glucose utilization rate seems to decrease after SD.[40] In addition, IKZF1/3 regulates a positive feedback loop involving interferon regulatory factor 4 (IRF4) and c-Myc, thereby maintaining the abnormal proliferation of MM cells.

The JCHAIN gene is expressed to synthesize J-chain immunoglobulin (j-chain), a cysteine-rich peptide chain synthesized by plasma cells, and is a B-cell immune response-related gene. The main function of the j-chain is to connect multiple Ig monomers into polymers.[41] The JCHAIN gene is expressed during the differentiation stage of B lymphocytes and is a marker or prognostic factor for many diseases. The MGMT3 gene is an important enzyme-coding gene in the human genome that has important biological functions in cells.[42] The MGST3 enzymes belong to the glutathione S-transferase family and have antioxidant effects. They protect cells from oxidative stress by reacting with glutathione and lipid peroxides. Research has shown that the antioxidant effect of the MGST3 enzyme is crucial for cell survival and functional maintenance, is closely related to the occurrence and development of various diseases, and has important value in personalized drug therapy. Research has shown that MGST3 can regulate the size of the hippocampus in the mammalian brain, which is associated with specific genes related to neurological diseases.[43] The UQCR11 gene encodes the smallest known component of the ubiquinone cytochrome c reductase complex that can form a part of the mitochondrial respiratory chain.[44] This gene is needed for energy metabolism in the oxidative phosphorylation process and is also the main source of ROS, which are closely related to mitochondrial energy metabolism.[45–47] Research has shown that UQCR11 plays a role in the biological process of mitochondrial dysfunction in neurological diseases.[48]

Lenalidomide, an analog of the immune regulatory drug thalidomide[49,] has various cellular effects, such as immune regulation and antiangiogenic, anti-inflammatory, and antiproliferation effects.[50] In this study, we found through the DGIbd database that the IKZF1 gene is an important target for lenalidomide. Research has shown that lenalidomide has a protective effect on many inflammation-related diseases. Lenalidomide increases the ubiquitination of the Ikaros family zinc finger transcription factor IKZF and its members and specifically reduces protein levels.[51] Lenalidomide was previously considered a "magical drug" that can provide safe and good sleep.[52] Currently, lenalidomide, a new type of immunoregulatory and anti-inflammatory drug, is expected to be helpful for people with related diseases and sleep disorders caused by immune inflammation.[53]

Our study has increased the understanding of bioinformatics and machine learning related to ferroptosis combined with necroptosis during SD. However, this study has certain limitations. First, the proposed machine learning model was constructed and validated via a public database with a single source, and additional prospective real-world data are needed to confirm its clinical significance. Second, the regulatory mechanism of FCRGs and their functions in SD immune infiltration are not yet clear, and more in vitro and in vivo experiments are needed to assist in verification. Finally, further research is needed to confirm the accuracy and stability of the machine learning model and determine whether it can be used to predict the feasibility and safety of SD treatment drugs in clinical practice.

Conclusion

In summary, we first used bioinformatics and three machine learning methods to screen four feature genes related to SD-DE-FCRGs (IKZF1, JCHAIN, MGST3, and UQCR11) and found that they may affect the involvement of infiltrating immune cells in the pathogenesis of SD by regulating FCRGs. We predicted that lenalidomide may be a target of the core feature gene IKZF1 of SD-DE-FCRGs. However, this study still has certain limitations, and the dataset used was relatively small. In the future, in vitro and in vivo experiments are needed to verify the regulatory mechanism of SD-DE-

FCRGs, predict drug safety, etc., to further clarify their pathogenesis and provide a reference for subsequent SD experimental research.

Data Share Statement

Both the SD-related database and the aggregated analysis data in this study are available for public download by qualified researchers. The SD dataset GSE98566 (annotation platform: GPL6244) was obtained from the GEO database (https://www.ncbi.nlm.nih.gov/GEO/). The FRGs were obtained from the FerrDb database (http://www.zhounan.org/ferrdb/). Data from the current study can be obtained from the corresponding author (Shaodan Li, Email: lsd301@126.com) upon reasonable request.

Ethics Approval and Consent to Participate

The present study was conducted in strict accordance with the Declaration of Helsinki and its later amendments or comparable ethical standards. Our research project was approved by the Medical Ethics Committee of Chinese PLA General Hospital.

Acknowledgments

We are grateful to the China Postdoctoral Science Foundation, the National Natural Science Foundation of China and the special plan for the cultivation and improvement of TCM service capabilities.

Author Contributions

All authors made a significant contribution to the work reported, whether 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; agreed on the journal to which the article has been submitted; and agreed to be accountable for all aspects of the work.

Funding

This study was supported by the 75th batch of general support from the China Postdoctoral Science Foundation, National Natural Science Foundation of China (82174283, 82474485), the special plan for the cultivation and improvement of TCM service capabilities (2021ZY009, 2021ZY048), and the Basic and Clinical Research on Prevention and Treatment of Geriatric Diseases with Integrated Traditional Chinese and Western Medicine-National Grant (111-gb).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Li Z-H, Cheng L, Wen C, et al. Activation of CNR1/PI3K/AKT pathway by tanshinone IIA protects hippocampal neurons and ameliorates sleep deprivation-induced cognitive dysfunction in rats. *Front Pharmacol*. 2022;13:823732. doi:10.3389/fphar.2022.823732
2. Svensson T, Saito E, Svensson AK, et al. Association of sleep duration with all- and major-cause mortality among adults in Japan, China, Singapore, and Korea. *JAMA Network Open*. 2021;4(9):e2122837. doi:10.1001/jamanetworkopen.2021.22837
3. Sabia S, Dugravot A, Léger D, et al. Association of sleep duration at age 50, 60, and 70 years with risk of multimorbidity in the UK: 25-year follow-up of the Whitehall II cohort study. *PLoS Med*. 2022;19(10):e1004109. doi:10.1371/journal.pmed.1004109
4. Chen X, Li Q, Zhang Z, et al. Identification of potential diagnostic biomarkers from circulating cells during the course of sleep deprivation-related myocardial infarction based on bioinformatics analyses. *Front Cardiovasc Med*. 2022;9:843426. doi:10.3389/fcvm.2022.843426
5. Sofer T, Kurniansyah N, Murray M, et al. Genome-wide association study of obstructive sleep apnea in the Million Veteran Program uncovers genetic heterogeneity by sex. *EBioMedicine*. 2023;90:104536. doi:10.1016/j.ebiom.2023.104536
6. Pandi-Perumal SR, Saravanan KM, Paul S, et al. Waking up the sleep field: an overview on the implications of genetics and bioinformatics of sleep. *Mol Biotechnol*. 2024;66(5):919–931. doi:10.1007/s12033-023-01009-1
7. Stockwell BR. A powerful cell-protection system prevents cell death by ferroptosis. *Nature*. 2019;575(7784):597–598. doi:10.1038/d41586-019-03145-8
8. Kahlson MA, Dixon SJ. Copper-induced cell death. *Science*. 2022;375(6586):1231–1232. doi:10.1126/science.abo3959
9. Fuding W. *FerroptosisM*. Beijing: Science Press; 2022.

10. Chen N, Guo L, Wang L, et al. Sleep fragmentation exacerbates myocardial ischemia–reperfusion injury by promoting copper overload in cardiomyocytes. *Nat Commun*. 2024;15(1):3834. doi:10.1038/s41467-024-48227-y
11. Yuan M, Wang F, Sun T, et al. Vitamin B6 alleviates chronic sleep deprivation-induced hippocampal ferroptosis through CBS/GSH/GPX4 pathway. *Biomed Pharmacother*. 2024;174:116547. doi:10.1016/j.biopha.2024.116547
12. Chen L, Min J, Wang F. Copper homeostasis and cuproptosis in health and disease. *Signal Transduct Target Ther*. 2022;7(1):378. doi:10.1038/s41392-022-01229-y
13. Huang J, Xu Z, Yuan Z, et al. Identification of a cuproptosis-related lncRNA signature to predict the prognosis and immune landscape of head and neck squamous cell carcinoma. *Front Oncol*. 2022;12:983956. doi:10.3389/fonc.2022.983956
14. Ren X, Pan C, Pan Z, et al. Knowledge mapping of copper-induced cell death: a bibliometric study from 2012 to 2022. *Medicine (Baltimore)*. 2022;101(45):e31133. doi:10.1097/MD.0000000000031133
15. Tsvetkov P, Coy S, Petrova B, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science*. 2022;375(6586):1254–1261. doi:10.1126/science.abf0529
16. Wang T, Jiang X, Lu Y, et al. Identification and integration analysis of a novel prognostic signature associated with cuproptosis-related ferroptosis genes and relevant lncRNA regulatory axis in lung adenocarcinoma. *Aging (Albany NY)*. 2023;15(5):1543–1563. doi:10.18632/aging.204561
17. Liu Z, Ma H, Lai Z. The role of ferroptosis and cuproptosis in curcumin against hepatocellular carcinoma. *Molecules*. 2023;28(4):1623. doi:10.3390/molecules28041623
18. Huang YY, Bao TY, Huang XQ, et al. Machine learning algorithm to construct cuproptosis- and immune-related prognosis prediction model for colon cancer. *World J Gastrointest Oncol*. 2023;15(3):372–388. doi:10.4251/wjgo.v15.i3.372
19. Liu L, Liang L, Yang C, et al. Machine learning-based solution reveals cuproptosis features in inflammatory bowel disease. *Front Immunol*. 2023;14:1136991. doi:10.3389/fimmu.2023.1136991
20. Liu R, Liu Y, Zhang F, et al. A cuproptosis random forest cox score model-based evaluation of prognosis, mutation characterization, immune infiltration, and drug sensitivity in hepatocellular carcinoma. *Front Immunol*. 2023;14:1146411. doi:10.3389/fimmu.2023.1146411
21. Zhang H, Nagai J, Hao L, et al. Identification of key genes and immunological features associated with copper metabolism in parkinson's disease by bioinformatics analysis. *Mol Neurobiol*. 2023;61(2):799–811. doi:10.1007/s12035-023-03565-8
22. Chen H, Zhang J, Sun X, et al. Mitophagy-mediated molecular subtypes depict the hallmarks of the tumor metabolism and guide precision chemotherapy in pancreatic adenocarcinoma. *Front Cell Dev Biol*. 2022;10:901207. doi:10.3389/fcell.2022.901207
23. Ferreira MR, Santos GA, Biagi CA, et al. GSVA score reveals molecular signatures from transcriptomes for biomaterials comparison. *J Biomed Mater Res A*. 2021;109(6):1004–1014. doi:10.1002/jbm.a.37090
24. Molendijk J, Yip R, Parker BL. urPTMdb/TeaProt: upstream and downstream proteomics analysis. *J Proteome Res*. 2023;22(2):302–310. doi:10.1021/acs.jproteome.2c00048
25. Su J, Peng J, Wang L, et al. Identification of endoplasmic reticulum stress-related biomarkers of diabetes nephropathy based on bioinformatics and machine learning. *Front Endocrinol (Lausanne)*. 2023;14:1206154. doi:10.3389/fendo.2023.1206154
26. He Y, Zou C, Cai Z. Construction and comprehensive analysis of the ceRNA network to reveal key genes for benign tracheal stenosis. *Front Genet*. 2022;13:891741. doi:10.3389/fgene.2022.891741
27. Troynikov O, Watson CG, Nawaz N. Sleep environments and sleep physiology: a review. *J Therm Biol*. 2018;78:192–203. doi:10.1016/j.jtherbio.2018.09.012
28. Kumar N, Narayan Das N, Gupta D, et al. Efficient automated disease diagnosis using machine learning models. *J Healthc Eng*. 2021;2021:9983652. doi:10.1155/2021/9983652
29. Lee YY, Endale M, Wu G, et al. Integration of genome-scale data identifies candidate sleep regulators. *Sleep*. 2023;46(2):zsac279. doi:10.1093/sleep/zsac279
30. Dixon SJ, Lemberg KM, Lamprecht MR, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell*. 2012;149(5):1060–1072. doi:10.1016/j.cell.2012.03.042
31. Tang D, Chen X, Kang R, et al. Ferroptosis: molecular mechanisms and health implications. *Cell Res*. 2021;31(2):107–125. doi:10.1038/s41422-020-00441-1
32. Ferreira A, Neves P, Gozzelino R. Multilevel Impacts of Iron in the Brain: the Cross Talk between Neurophysiological Mechanisms, Cognition, and Social Behavior. *Pharmaceuticals (Basel)*. 2019;12(3):126. doi:10.3390/ph12030126
33. Wang X, Wang Z, Cao J, et al. Melatonin alleviates acute sleep deprivation-induced memory loss in mice by suppressing hippocampal ferroptosis. *Front Pharmacol*. 2021;12:708645. doi:10.3389/fphar.2021.708645
34. An Y, Li S, Huang X, et al. The role of copper homeostasis in brain disease. *Int J Mol Sci*. 2022;23(22):13850. doi:10.3390/ijms232213850
35. Leng D, Yang Z, Sun H, et al. Comprehensive analysis of tumor microenvironment reveals prognostic ceRNA network related to immune infiltration in sarcoma. *Clin Cancer Res*. 2023;19(3986):4001–4010. doi:10.1158/1078-0432.CCR-22-3396
36. Huang J, Zhou Q. CD8+T-cell-related gene biomarkers in macular edema of diabetic retinopathy. *Front Endocrinol*. 2022;13:907396. doi:10.3389/fendo.2022.907396
37. Ye MT, Wang Y, Zuo Z, et al. Integrated clinical genotype-phenotype characteristics of early T-cell precursor acute lymphoblastic leukemia. *Cancer*. 2023;129(1):49–59. doi:10.1002/cncr.34515
38. Lopes BA, Meyer C, Bouzada H, et al. The recombinome of IKZF1 deletions in B-cell precursor ALL. *Leukemia*. 2023;37(8):1727–1731. doi:10.1038/s41375-023-01935-8
39. Chan LN, Chen Z, Braas D, et al. Metabolic gatekeeper function of B-lymphoid transcription factorsJ. *Nature*. 2017;542(7642):479–483. doi:10.1038/nature21076
40. Morselli L, Leproult R, Balbo M, et al. Role of sleep duration in the regulation of glucose metabolism and appetite. *Best Pract Res Clin Endocrinol Metab*. 2010;24(5):687–702. doi:10.1016/j.beem.2010.07.005
41. King LB, Corley RB. Characterization of a presecretory phase in B-cell differentiation. *Proc Natl Acad Sci U S A*. 1989;86(8):2814–2818. doi:10.1073/pnas.86.8.2814
42. Kelner MJ, Diccianni MB, Yu AL, et al. Absence of MGST1 mRNA and protein expression in human neuroblastoma cell lines and primary tissue. *Free Radic Biol Med*. 2014;69:167–171. doi:10.1016/j.freeradbiomed.2014.01.021

43. González-Castillo C, Ortuño-Sahagún D, Guzmán-Brambila C, et al. The absence of pleiotrophin modulates gene expression in the hippocampus in vivo and in cerebellar granule cells in vitro. *Mol Cell Neurosci*. 2016;75:113–121. doi:10.1016/j.mcn.2016.07.004
44. Wang X, Huang K, Yang F, et al. Association between structural brain features and gene expression by weighted gene coexpression network analysis in conversion from MCI to AD. *Behav Brain Res*. 2021;410:113330. doi:10.1016/j.bbr.2021.113330
45. Achreja A, Yu T, Mittal A, et al. Metabolic collateral lethal target identification reveals MTHFD2 paralog dependency in ovarian cancer. *Nat Metab*. 2022;4(9):1119–1137. doi:10.1038/s42255-022-00636-3
46. Han N, Luo R, Liu J, et al. Transcriptomic and proteomic analysis reveals mechanisms of patulin-induced cell toxicity in human embryonic kidney cells. *Toxins*. 2020;12(11):681. doi:10.3390/toxins12110681
47. Feng Z, Hom ME, Bearrood TE, et al. Targeting colorectal cancer with small-molecule inhibitors of ALDH1B1. *Nat Chem Biol*. 2022;18 (10):1065–1075. doi:10.1038/s41589-022-01048-w
48. Liu W, Li M, Zhang W, et al. Leveraging functional annotation to identify genes associated with complex diseases. *PLoS Comput Biol*. 2020;16 (11):e1008315. doi:10.1371/journal.pcbi.1008315
49. Osada N, Kikuchi J, Iha H, et al. c-FOS is an integral component of the IKZF1 transactivator complex and mediates lenalidomide resistance in multiple myeloma. *Clin Transl Med*. 2023;13(8):e1364. doi:10.1002/ctm2.1364
50. Liu A, Li S, Donnenberg V, et al. Immunomodulatory drugs downregulate IKZF1 leading to expansion of hematopoietic progenitors with concomitant block of megakaryocytic maturation. *Hematologica*. 2018;103(10):1688–1697. doi:10.3324/hematol.2018.188227
51. Krönke J, Hurst SN, Ebert BL. Lenalidomide induces degradation of IKZF1 and IKZF3. *Oncoimmunology*. 2014;3(7):e941742. doi:10.4161/21624011.2014.941742
52. Kumar N, Sharma U, Singh C, et al. Thalidomide: chemistry, therapeutic potential and oxidative stress induced teratogenicity. *Curr Top Med Chem*. 2012;12(13):1436–1455. doi:10.2174/156802612801784407
53. Asher C, Furnish T. Lenalidomide and thalidomide in the treatment of chronic pain. *Expert Opin Drug Saf*. 2013;12(3):367–374. doi:10.1517/14740338.2013.775242