

Identification of Ferroptosis-Related Genes in Type 2 Diabetes Mellitus Based on the Machine Learning

Sen Wang¹, yongpan lu¹, Yixin Zhang¹, Yuli Zhao¹, Huimin Guo¹, and Li Feng¹

¹Shandong Provincial Qianfoshan Hospital

March 07, 2024

Abstract

Background: Type 2 diabetes mellitus (T2DM), which has a high incidence and several harmful consequences, poses a severe danger to human health. More research is being done on ferroptosis' function in T2DM. This study uses a bioinformatics technique to look for new diagnostic T2DM biomarkers associated with ferroptosis. **Methods:** In order to identify ferroptosis-related genes (DEGs) that are differently expressed between T2DM patients and healthy individuals, we first obtained T2DM sequencing data and ferroptosis-related genes (FRGs) from the Gene Expression Omnibus (GEO) database and FerrDb database. Then, drug-gene interaction networks and ceRNA networks linked to the marker genes were built after marker genes were filtered by two machine learning algorithms (LASSO and SVM-RFE algorithms). Finally, to confirm the expression of marker genes, the GSE76895 dataset was utilized. The protein expression of some marker genes between T2DM and non-diabetic tissues was also examined by Western Blotting, Immunohistochemistry (IHC) and Immunofluorescence (IF), respectively. **Results:** We obtained 58 DEGs associated with ferroptosis. GO and KEGG enrichment analysis showed that these DGEs were significantly enriched in hypoxia and ferroptosis. Subsequently, eight marker genes (SCD, CD44, HIF1A, BCAT2, MTF1, HILPDA, NR1D2 and MYCN) were screened by LASSO and SVM- RFE machine learning algorithms, and a model was constructed based on these eight genes. These newly discovered marker genes may be linked to alterations in the immune microenvironment in T2DM patients. In addition, based on these 8 genes, we obtained 48 drugs and a complex ceRNA network map. Finally, Western Blotting, IHC and IF results of clinical samples further confirmed the results of public databases. **Conclusions:** The diagnosis and etiology of T2DM can be greatly aided by eight ferroptosis-related genes, opening up novel therapeutic avenues.

Identification of Ferroptosis-Related Genes in Type 2 Diabetes Mellitus Based on the Machine Learning

Sen Wang¹⁺, Yongpan Lu²⁺, Yixin Zhang¹, Yuli Zhao¹, Huimin Guo¹, and Li Feng ^{1*}

¹Department of Medical Ultrasound, Shandong First Medical University, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qian Foshan Hospital, Shandong Medicine and Health Key Laboratory of Abdominal Medical Imaging, No.16766, Jingshi Road, Jinan 250014, Shandong Province, China.

²The First Clinical Medical College, Shandong University of Traditional Chinese Medicine, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qian Foshan Hospital, Jingshi Road, Jinan 250014, Shandong Province, China.

*Correspondence:

Li Feng, M.D, ph.D.

ORCID: 0000-0003-3284-9247

Affiliation: Department of Medical Ultrasound, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qian Foshan Hospital.

Postal address: No.16766, Jingshi Road, Jinan 250014, Shandong Province, China.

Fax number:0531-82969995

Email: qfsfengli@163.com

+These authors have contributed equally to this work

Identification of Ferroptosis-Related Genes in Type 2 Diabetes Mellitus Based on the Machine Learning

Sen Wang¹⁺, Yongpan Lu²⁺, Yixin Zhang¹, Yuli Zhao¹, Huimin Guo¹, and Li Feng^{1*}

¹Department of Medical Ultrasound, Shandong First Medical University, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qian Foshan Hospital, Shandong Medicine and Health Key Laboratory of Abdominal Medical Imaging, No.16766, Jingshi Road, Jinan 250014, Shandong Province, China.

²The First Clinical Medical College, Shandong University of Traditional Chinese Medicine, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qian Foshan Hospital, Jingshi Road, Jinan 250014, Shandong Province, China.

* Correspondence: qfsfengli@163.com

+These authors have contributed equally to this work

Abstract :

Background : Type 2 diabetes mellitus (T2DM), which has a high incidence and several harmful consequences, poses a severe danger to human health. More research is being done on ferroptosis' function in T2DM. This study uses a bioinformatics technique to look for new diagnostic T2DM biomarkers associated with ferroptosis. **Methods :** In order to identify ferroptosis-related genes (DEGs) that are differently expressed between T2DM patients and healthy individuals, we first obtained T2DM sequencing data and ferroptosis-related genes (FRGs) from the Gene Expression Omnibus (GEO) database and FerrDb database. Then, drug-gene interaction networks and ceRNA networks linked to the marker genes were built after marker genes were filtered by two machine learning algorithms (LASSO and SVM-RFE algorithms). Finally, to confirm the expression of marker genes, the GSE76895 dataset was utilized. The protein expression of some marker genes between T2DM and non-diabetic tissues was also examined by Western Blotting, Immunohistochemistry (IHC) and Immunofluorescence (IF), respectively. **Results :** We obtained 58 DEGs associated with ferroptosis. GO and KEGG enrichment analysis showed that these DGEs were significantly enriched in hypoxia and ferroptosis. Subsequently, eight marker genes (SCD, CD44, HIF1A, BCAT2, MTF1, HILPDA, NR1D2 and MYCN) were screened by LASSO and SVM- RFE machine learning algorithms, and a model was constructed based on these eight genes. This model also has a high diagnostic power, and three ferroptosis genes (HIF1A, HILPDA and SCD) are strongly associated with hypoxia and lipid metabolism in T2DM. These newly discovered marker genes may be linked to alterations in the immune microenvironment in T2DM patients. In addition, based on these 8 genes, we obtained 48 drugs and a complex ceRNA network map. Finally, Western Blotting, IHC and IF results of clinical samples further confirmed the results of public databases. **Conclusions :** The diagnosis and etiology of T2DM can be greatly aided by eight ferroptosis-related genes, opening up novel therapeutic avenues.

Keywords : type 2 diabetes mellitus, bioinformatics, ferroptosis, diagnostic, machine learning, gene expression omnibus

1 INTRODUCTION

Diabetes mellitus (DM) is a chronic metabolic disease characterized by high blood glucose levels caused by insulin resistance in peripheral tissues or insufficient pancreatic insulin secretion (Yang *et al.* 2019). Type 1 diabetes mellitus (T1DM) and type 2 diabetes mellitus (T2DM) are the two primary subtypes of diabetes, with T2DM accounting for nearly 90% of all cases (Weyer *et al.* 1999; Gheibi *et al.* 2020). As living standards

rise, the population ages, and there is a global pandemic of nonalcoholic fatty liver disease (NAFLD), the prevalence of T2DM will rapidly rise worldwide (Marin-Penalver *et al.* 2016; Younossi *et al.* 2019). According to estimates, the prevalence of T2DM will be 9.3% (463 million) in the world in 2019, 10.2% (578 million) in 2030, and 10.9% (700 million) in 2045 (Saeedi *et al.* 2019). Additionally, diabetes has a wide range of intricate side effects, such as macrovascular conditions like coronary heart disease, peripheral artery disease, and stroke, as well as microvascular conditions like diabetic nephropathy, peripheral neuropathy, and retinopathy (Viigimaa *et al.* 2020; Tomic *et al.* 2022). Some patients don't know they have T2DM until they start experiencing major side effects. Therefore, finding potential biomarkers and understanding the molecular causes of T2DM is essential for early detection and the avoidance of its consequences.

Ferroptosis was first hypothesized in 2012 and is a type of iron-dependent controlled cell death that is accompanied by an aberrant buildup of lipid reactive oxygen species (L-ROS) (Dixon *et al.* 2012; Hadian and Stockwell 2020). The majority of earlier research on ferroptosis focused on malignancies (El Hout *et al.* 2018). According to an increasing body of research, ferroptosis also plays a significant part in the onset of non-neoplastic disorders such as Parkinson's (Cong *et al.* 2019), Alzheimer's disease (Do Van *et al.* 2016), pulmonary fibrosis (Tsubouchi *et al.* 2019), brain damage (Tsubouchi *et al.* 2019), etc. Additionally, elevated ferritin levels have been seen in people with T2DM and gestational diabetes, indicating a link between excess iron storage and the onset of T2DM (Liu *et al.* 2020; Gautam *et al.* 2021). Erastin, a ferroptosis inducer, influences the development and operation of human pancreatic islet-like cell clusters by specifically inhibiting the Xc-cystine/glutamate antiporter necessary for GSH biosynthesis (Linkermann *et al.* 2014; Li and Leung 2020). In vitro erastin treatment of human islet β -cells resulted in considerably lower glucose-stimulated insulin secretion (GSIS) capability, based on research. On the other hand, GSIS damage was prevented by pretreatment with a ferroptosis inhibitor, such as Fer-1 or DFO (Bruni *et al.* 2018). According to recent research, various T2DM medications on the market can prevent ferroptosis. For instance, ACSL4 is a crucial element of ferroptosis, and rosiglitazone is the most potent ACSL4 inhibitor (Kim *et al.* 2001; Yuan *et al.* 2016). However, numerous genes involved in ferroptosis in T2DM have not yet been discovered, necessitating more research on these genes.

The mechanism of ferroptosis in the pathogenic phase of T2DM and associated consequences is the major focus of current studies on ferroptosis and T2DM. This study sought to add to previous research and establish ferroptosis as a therapeutic target for T2DM by examining the genetic relationship between genes relevant to ferroptosis and T2DM. It also used Cytoscape software to create a drug-gene interaction map. Then, to investigate the possible regulatory effects of miRNAs and lncRNAs on ferroptosis-related marker genes in T2DM, we built a competitive endogenous RNA (ceRNA) regulatory network of the marker genes. Our findings offer a fresh viewpoint for the clinical diagnosis and management of T2DM and may be useful in clarifying the possible contribution of the ferroptosis process to the pathogenesis of T2DM.

2 MATERIALS AND METHODS

2.1 Data acquisition

The Gene Expression Omnibus (GEO) database, accessible at <https://www.ncbi.nlm.nih.gov/geo/>, provided the gene expression information used in this analysis for samples with T2DM and normal tissue. The RNA sequences of 68 T2DM and 62 normal samples are included in the collection GSE78721. This dataset served as a training set for the main body of this research's study. The expression of the marker genes was validated using the GSE76895 dataset, which included samples from 32 normal and 36 T2DM patients. Additionally, FerrDb (<http://www.zhounan.org/ferrdb/>) was utilized to obtain the FRGs ($n = 358$) used in this investigation. These FRGs included the three categories of driver, suppression, and marker FRGs.

2.2 Screening Ferroptosis-Related Differentially Expressed Genes

We examined differentially expressed genes (DEGs) between T2DM samples and normal samples using the R software's limma package. Significant genes were those with $p < 0.05$. The heatmap was then used to illustrate the overlap of DEGs and ferroptosis-related genes.

2.3 Ferroptosis-Related DEGs: Functional Enrichment Analysis

The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of ferroptosis-related DEGs were performed in R using the clusterProfiler package(Wu *et al.* 2021). The GO analysis covered three categories: biological process (BP), cellular component (CC), and molecular function (MF), which was useful in exploring biological functions(Ashburner *et al.* 2000). The KEGG analysis was utilized to investigate probable biological activities, illnesses, substances, and drugs(Kanehisa and Goto 2000).

2.4 Identification of optimal diagnostic gene biomarkers for T2DM

The glmnet package was used to minimize the data dimensions by employing the least absolute shrinkage and selection operator (LASSO) method (Friedman *et al.* 2010; Yang *et al.*2019). The ferroptosis-related DEGs found in T2DM patients and normal samples were then cross-validated in the LASSO logistic regression approach to look for disease hallmark genes. Simultaneously, a support vector machine-recursive feature elimination (SVM-RFE) model was created using SVM software, and the average misjudgment rates of their 10-fold cross-validations were compared(Wu *et al.* 2022). Finally, the outputs of the LASSO and SVM-RFE algorithms were intersected to select the best T2DM biomarkers, which were depicted by a Venn diagram. The logistic regression models based on these genes were created using the glmnet package of the R programming language. Researchers might assess the diagnostic effectiveness of the logistic regression models by computing receiver operating characteristic (ROC) curves and calculating the area under the curve (AUC). ROC curves were also used to assess a gene’s ability to distinguish between samples with and without T2DM.

2.5 Single-gene Gene Set Enrichment Analysis (GSEA)

To further explore the related pathways of the eight marker genes, we performed GSEA analysis to identify pathways enriched in T2DM patients and “c2.cp.kegg.v7.0.symbols.gmt” from the MSigDB database was adopted as the reference dataset. Patients were divided into high and low expression groups according to the expression levels of the eight marker genes. Annotated gene sets were used to distinguish subtypes by the identified differentially expressed genes. We computed the consistency P-value for each gene set, and P-values less than 0.05 were considered significantly enriched. Subsequently, significantly enriched gene sets were sorted according to their correlations from high to bottom.

2.6 Single-gene enrichment analysis using Gene Set Variation Analysis (GSVA)

The enrichment of transcriptomic gene sets may be determined using the non-parametric, unsupervised GSVA approach. In order to assess the biological activities of the samples, GSVA first translates gene-level changes into pathway-level changes by rating the sets of genes(Hanzelmann *et al.* 2013). We used the KEGG pathway set as the background gene set for this analysis. GSVA evaluation of every marker gene. The possible biological function alterations of various samples were assessed at the same time as the GSVA score difference between samples from the high- and low-expression groups of the marker gene were analyzed using the limma software.

2.7 Immune infiltration analysis

Immune cell infiltrations were calculated using the bioinformatics method CIBERSORT (<https://cibersortx.stanford.edu/>), which was used to quantify the relative proportions of the 22 infiltrating immune cell types in the GSE78721 dataset. The total of all examined immune cell type fractions for each sample was 1. Using violin plots created using the vioplot program, the differences in immune cell infiltration between T2DM patients and controls were shown. Furthermore, immune cells and gene expression were analyzed using Spearman correlation in this study.

2.8 Establishment of a Nomogram

Using the rms package, marker genes were included to create a nomogram. In addition, we assessed the accuracy of this nomogram by calibration curves, decision curves, and clinical impact curves.

2.9 Drug-gene interaction

Screening for drugs that modulate marker genes Using Drug gene interaction database (DGIdb). This final medicine list only contained medications that were DrugBank-sourced and authorized by the Food and Drug Administration.

2.10 Construction of ceRNA network

To predict the binding of marker genes to miRNA, we used three different programs (miRanda, miRDB, and TargetScan), which all regarded marker genes to be miRNA target genes. The spongeScan database provided us with the targeting connection between miRNA and lncRNA. Then, using the Cytoscape program, we built the lncRNA-miRNA-mRNA regulatory network.

2.11 Patients and tissue samples

Diabetic ulcers and normal skin tissues were prospectively collected from six patients that were enrolled in Shandong Provincial Hospital of Traditional Chinese Medicine, Jinan, China. Tissue after ulcer debridement surgery in 3 diabetic patients and after circumcision surgery in 3 non-diabetic patients. Half of the above tissue was divided into two parts and rapidly frozen in a 1.5 ml snap cap tube in liquid nitrogen and stored at - 80 °C for subsequent molecular analysis. This study was conducted in accordance with the Declaration of Helsinki. All the experiments were approved by the ethics committee of Affiliated Hospital of Shandong University of Traditional Chinese Medicine & Shandong Provincial Hospital of Traditional Chinese Medicine (Approval Number: AF/SC-08/02.0) and were performed in accordance with the guidelines and regulations.

The other half of the above tissue was fixed in 10% paraformaldehyde fixative, embedded in paraffin and sectioned at 3µm for histochemical analyses which were performed by haematoxylin eosin (H&E) staining and immunohistochemical (IHC) staining. Deparaffinization, rehydration, antigen retrieval, endogenous peroxidase blocking, and goat serum (#SP-9001; ZSGB-BIO, Beijing, China) blocking of paraffin sections. Next, the wound skin tissue sections were incubated with primary antibody against CD44 (Proteintech; 60224-1-Ig) and MYCN (Abcam; ab16898) at a dilution of 1:200 at 4°C overnight. On the second day, all sections were incubated with biotin-labelled goat anti mouse IgG polymer for 15 min at room temperature and incubated with horseradish enzyme-labelled streptavidin working solution for 15min at room temperature. Finally, the slides were counterstained with diaminobenzidine (DAB; #ZLI-9018, ZSGB-BIO) and haematoxylin (CAS. 517-28-2; Beijing Solarbio Science & Technology). For H&E staining, after deparaffinization and rehydration, slides were stained with haematoxylin and eosin (CAS. 17372-87-1; Beijing Solarbio Science & Technology) according to the manufacturer's instructions. All the sections were then dehydrated, cleared, and sealed. The images were observed and captured using an Olympus IX73 microscope (Olympus, Tokyo, Japan). Using ImageJ software to calculate the positive rate.

2.12 Western blot analysis

Total protein was extracted from tissue sample using RIPA lysis buffer (Beyotime Biotechnology) supplemented with Thermo Scientific Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific). Protease inhibitors and phosphatase inhibitors (1:100) were added during protein extraction (MedChemExpress), and a Pierce BCA Protein Analysis Kit (Thermo Fisher Scientific) was used to measure protein concentrations. Protein samples were separated by 10% SDS-PAGE and transferred to PVDF membranes. The membranes were blocked in 5% skim milk and incubated with the respective primary antibodies over-night at 4 °C. The samples were incubated with horseradish per-oxidase-conjugated secondary antibodies (1:5000 dilution; Cell Signaling Technology) for 1 h at room temperature, and an iBright FL1500 imaging system (Invitrogen) and Super Signal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific, Invitrogen) were used to detect and analyze protein expression levels. Antibody information was provided for anti-CD44 (Proteintech; 60224-1-Ig), anti-MYCN (Abcam; ab16898), and rabbit anti-GAPDH (Cell Signaling Technology; WB: 1/1000).

2.13 IHC and Immunofluorescence (IF) staining

We performed IHC and IF staining of tissues. IHC of tissues was performed as described previously. IF staining, paraffin sections were dewaxed to water, and the sections were sequentially placed in environmentally friendly dewaxing solution for 10 min - anhydrous ethanol for 5 min - anhydrous ethanol for 5 min - anhydrous ethanol for 5 min - distilled water. The antigen repair was then carried out, and the repair was completed by natural cooling. The slides were placed in PBS (PH7.4) and washed 3 times with shaking on a decolorization shaker for 5 min each time. the sections were slightly shaken dry and then closed with a histochemical pen by drawing circles around the tissue and adding 3% BSA dropwise for 30 min. the sections were incubated overnight at 4degC in a wet box after adding the prepared primary antibody dropwise. The slides were placed in PBS (pH 7.4) and washed 3 times with shaking on a decolorized shaker for 5 min each time. The slides were washed three times in PBS (pH 7.4) on a decolorization shaker for 5 min each time. observed and recorded using a Nikon Eclipse Ti2 confocal microscope (Nikon Instruments (Shanghai) Co., Ltd., Shanghai, China). The following primary antibodies were used: anti-CD44 (Proteintech; 60224-1-Ig), anti-MYCN (Abcam; ab16898).

2.14 Statistical analysis

R software 4.2.1 was used for the statistical analysis. The link between 58 ferroptosis-related DEGs was discovered using Pearson correlation analysis. The data for all experiments were shown as means \pm standard deviation (SD) of three biological replicates, and all data analyses were performed using GraphPad Prism version 9.0.0 (GraphPad Software, San Diego, CA). Statistical analysis between groups was performed using Student's t test to determine significance. A statistically significant P value is less than 0.05.

3 RESULTS

3.1 Identification of Ferroptosis-Related DEGs Between T2DM and Control

The GSE78721 dataset discovered 58 FRGs that showed differential expression between T2DM and normal samples, including 20 down-regulated and 38 up-regulated genes (Table 1). The ferroptosis-related DEGs' standardized expression was displayed in the clustering heatmap of Figure 1A. Figure 1B depicts the interaction between 58 genes that are associated with ferroptosis in diverse ways. The majority of these genes have a high degree of correlation with one another.

3.2 Enrichment Analysis of Ferroptosis-Related DEGs

GO and KEGG enrichment analyses were done to further study the biological activities and pathways of these ferroptosis-related DEGs, as shown in Figures 2A and 2B, respectively. The important GO-BP categories were primarily connected with hypoxia, such as stress hypoxia, reduced oxygen levels, and oxygen levels. The ferroptosis-related DEGs were highly enriched in the transcription regulator complex and the basal portion of the cell in GO-CC analysis. The pathways enriched by GO-MF were mostly related to transcription factors. The top 20 enriched pathways, according to KEGG analysis, were mostly involved in viral hepatitis, growth and thyroid hormone production, autophagy-related processes, and ferroptosis. Surprisingly, ferroptosis-related DEGs were clearly enriched in cancer-related signatures such as non-small cell lung cancer, renal cell carcinoma, chemical carcinogenesis-receptor activation, viral carcinogenesis, central carbon metabolism in cancer, and so on. These findings suggest that hypoxia and transcription factors may play a crucial role in the development of T2DM, while also providing some novel paths for the link between T2DM and cancer.

3.3 8 Ferroptosis-Related DEGs were identified as diagnostic genes for T2DM

The LASSO and SVM- RFE were used to screen the significant ferroptosis-related DEGs to distinguish T2DM from normal people in GSE78721. In the LASSO logistic regression algorithm, we selected the 15 genes at the time of minimum cross-validation error (Figure 3A, B). The best diagnostic genes for T2DM were ultimately determined to be 18 genes (highest precision =0.731, minimum RMSE =0.269) after using the SVM-RFE method to filter 58 ferroptosis-related DEGs (Figure 3C, D). The 8 genes (SCD, CD44, HIF1A, BCAT2, MTF1, HILPDA, NR1D2, MYCN) that overlapped these two algorithms were selected (Figure 3E).

The AUC for all genes was larger than 0.6 when we drew ROC curves for these eight biomarkers that were identified by both machine learning systems (Figure 3F). We also create a logistic regression model using the R glmnet package based on these eight biomarkers. According to our ROC curve results, the 8 marker gene-based logistic regression model offered more sensitivity and precision than the independent marker genes for discriminating T2DM samples from normal samples, with an AUC of 0.832 (95% CI 0.760-0.896). (Figure 3G). Additionally, we display in Figure 4 the interactions and expression of the eight marker genes in the GSE78721 dataset.

3.4 Various Pathways Associated with Marker Genes

We carried out a single-gene GSEA-KEGG pathway analysis to further investigate the distinct signaling pathways connected to the marker genes. The first six routes for each marker gene are displayed (Figures 5A-H). After a thorough analysis, we discovered that these eight marker genes were primarily enriched in the pathways for lysosome, cell cycle, ribosome, peroxisome, ubiquitin-mediated proteolysis, fatty acid metabolism, and various disease pathways (including those for Parkinson’s disease and Huntington’s disease). In addition, we discovered that the marker genes were enriched in the chemokine signaling route, B cell receptor signaling pathway, olfactory transduction, and neuroactive ligand-receptor interaction.

3.5 GSVA

Then, we analyzed the differentially activated pathways between the high- and low-expression groups according to the expression level of each marker gene in combination with GSVA results (see Supplemental1, Supplemental Content, which illustrates the GSVA results of these marker genes). We found that the high expression of CD44 and HIF1A in the disease has a high similarity (Figures 6), probably through inducing T2DM by activating LINOLEIC ACID METABOLISM and NEUROACTIVE LIGAND RECEPTOR INTERACTION. In addition, they were associated with TASTE and OLFACTORY TRANSDUCTION. The low expression of HILPDA and SCD in T2DM also has a high similarity, and they were mainly related to the synthesis, metabolism and degradation of substances such as BIOSYNTHESIS OF UNSATURATED FATTY ACIDS, TERPENOID BACKBONE BIOSYNTHESIS, FATTY ACID METABOLISM, BETA ALANINE METABOLISM, PYRUVATE METABOLISM, SELENOAMINO ACID METABOLISM, VALINE LEUCINE AND ISOLEUCINE DEGRADATION, etc. In addition, low expression of SCD and high expression of MTF1 were associated with SYSTEMIC LUPUS ERYTHEMATOSUS, PARKINSONS DISEASE, and PATHOGENIC ESCHERICHIA COLIINFECTION. MYCN low expression in the disease was only involved in ALPHA LINOLENIC ACID METABOLISM. High expression of NR1D2 has enriched in OLFACTORY TRANSDUCTION and NEUROACTIVE LIGAND RECEPTOR INTERACTION. It is worth noting that BCAT2, CD44 and HIF1A were directly related to the pathway MATURITY ONSET DIABETES OF THE YOUNG.

3.6 Immune landscape analysis

We investigated immune cell infiltration between T2DM patients and normal samples using CIBERSORT. The percentage of immune cells from 62 normal and 68 T2DM samples was shown in Figure 7A. The interaction of the immune cells revealed that resting mast cells and active mast cells had the most pronounced negative correlation with $r = 0.51$, while T cells CD4 naive and B cells naive had the most significant positive connection with $r = 0.83$. (Figure 7B). Figure 7C demonstrates that whereas dendritic cells activated are more prevalent in T2DM patients, dendritic cells resting are less common in T2DM patients than in normal samples.

3.7 Eight Biomarkers and Infiltrating Immune Cells Correlation Analysis

As illustrated in Figure 8, the correlation analysis revealed a significant link between these 8 marker genes and immune cells. We discovered that CD44 strongly correlated with the activation of NK cells and B cells’ memories, respectively. HIF1A had a substantial negative correlation with both activated NK cells and CD8 T cells. NR1D2, MYCN, MTF1, and HILPDA all had a negative correlation with dendritic cell activation. Macrophages M0, mast cells that are activated, monocytes, and neutrophils all have a positive

correlation with MTF1. Of course, marker genes related to neutrophils also include HILPDA and HIF1A. This data implies that these newly discovered marker genes may be linked to alterations in the immune microenvironment in T2DM patients.

3.8 Construction and assessment of the nomogram for patients with T2DM

A nomogram was created as a diagnostic tool for T2DM by including marker genes (Figure 9A). Each marker gene in the nomogram was assigned a score, and the overall score was obtained by adding the scores of all marker genes. Higher overall scores increased the chance of acquiring T2DM, while lower total scores were linked to lower risks of developing T2DM. The nomogram's great accuracy was shown by the calibration and decision curve (Figures 9B and 9C). The nomogram also maintained excellent accuracy in identifying high-risk T2DM patients, as seen in the clinical impact curves (Figure 9D).

3.9 Drug-gene interaction

We searched through the DGIdb database for drugs that might affect the marker genes. The Cytoscape software-visualized results were displayed in (Figure 10). We had queried 48 medicines targeting marker genes, including 30 for HIF1A, 9 for MYCN, 5 for SCD, 3 for CD44 and NTF1 targeted 1 drug. Unfortunately, the medications connected to BCAT2, HILPDA, and NR1D2 were not predicted by us. In addition, we also searched the structural formulae of the above 48 drugs using the DrugBank database. 32 drug structures in all were found. A total of 16 drug structures were retrieved from 30 HIF1A-targeted drugs (see Supplementaly2, Supplemental Content, which illustrates the structures of 32 drug). Among them, PX-478 and NITROGLYCERIN have known inhibitors of HIF1A. A total of 8 drug structures were retrieved from the 9 targeted drugs of MYCN. The corresponding drug structures were derived for all five targets of SCD and three targets of CD44.

3.10 Marker gene-based ceRNA networks

We built a ceRNA network based on these 8 marker genes to highlight the interaction between lncRNA, miRNA, and mRNA, as shown in Figure 11. The network includes 671 nodes (7 marker genes, 349 miRNAs and 315 lncRNAs). BCAT2 had 18 miRNAs associated with it, CD4 had 69 miRNAs connected with it, HIF1A had 62 miRNAs associated with it, MTF had 115 miRNAs linked with it, MYCN had 96 miRNAs associated with it, NR1D2 had 85 miRNAs associated with it, and SCD had 70 miRNAs related with it. We discovered that 19 lncRNAs could control the expression of SCD and MTF1 through competitive binding to hsa-miR-149-3p, 16 lncRNAs control the expression of SCD, MTF1, NR1D2 and HIF1A through competitive binding to hsa-miR-149-3p, 15 lncRNAs control the expression of NR1D2, MTF1, and MYCN through competitive binding to hsa-miR-515-5p, 15 lncRNAs control the expression of SCD, MTF1, and CD44 through competitive binding to hsa-miR-18a-3p, 8 lncRNAs control the expression of MYCN, NR1D2, and CD44 through competitive binding to hsa-miR-590-3p, 14 lncRNAs control the expression of MTF1 through competitive binding to hsa-miR-129- 5p, 20 lncRNAs control the expression of NR1D2 through competitive binding to hsa-miR-766-3p.

3.11 Expression of the marker gene in the validation set

We verified the differential expression of these eight marker genes and the precision of the logistic regression model using the GSE76895 dataset. We also discovered that BCAT2, HILPDA, MYCN, and CD44 had differential expressions, and the patterns of these expression were consistent with those in the GSE78721 dataset (Figure 12A). Among these, CD44 expression ($p = 0.0013$) was greater in T2DM patients than in controls, but BCAT2 expression ($p = 0.043$), HILPDA expression ($p = 0.044$), and MYCN expression ($p = 4e-06$) were lower in T2DM patients. The ROC curves for the marker genes in the validation set are shown in Figure 12B. With an AUC of 0.872 (95% CI 0.782-0.948), the logistic regression model based on 8 marker genes continued to be extremely precise and specific in the validation group (Figure 12C).

3.12 Protein expression of CD44 and MYCN in T2DM and non-diabetic tissues

We performed experimental validation by clinical tissues, and the experimental results were consistent with

our results by bioinformatics analysis, CD44 was indeed highly expressed in diabetic tissues and low in non-diabetic tissues; MYCN was low in diabetic tissues and high in non-diabetic tissues. As shown in Figure 13A-B, Western blot showed a trend of high expression of CD44 and low expression of MYCN in diabetic tissues. In Figure 13C-F, IHC also showed a trend of high expression of CD44 and low expression of MYCN in diabetic tissues. Also, we performed IF validation, In Figure 13G-L, the expression trends of CD44 and MYCN in diabetic tissues were consistent with Western blot and IHC. Our experimental results are extremely suggestive for the prevention and development of diabetes, and future in vivo and in vitro experimental interventions can be performed to further explore the role of CD44 and MYCN in the development and genetic aspects of diabetes.

4 Discussion

T2DM is an increasingly common metabolic disease and poses a significant public health burden (Carbone *et al.* 2019; Magliano *et al.* 2021). In recent years, the diagnosis and treatment of T2DM have been increasingly studied, but the prognosis for patients with T2DM remains poor due to the limited understanding of the pathogenesis of the disease and the numerous complications associated with drug therapy (Ali *et al.* 2022). Many studies now suggest that ferroptosis plays a substantial role in T2DM and its consequences, and that it is a risk factor for T2DM development. However, the precise process remains unknown (Sha *et al.* 2021; Li *et al.* 2022). In this study, we screened eight differential genes associated with ferroptosis by two machine learning algorithms (LASSO and SVM) and finally constructed a diagnostic model based on these eight differential genes (SCD, CD44, HIF1A, BCAT2, MTF1, HILPDA, NR1D2 and MYCN). In the training cohort, the model had high predictive ability (AUC=0.832). It also demonstrated great accuracy in the external validation cohort (AUC=0.872), offering fresh information for the quick and early diagnosis of T2DM.

The most enriched GO categories, according to our analysis of KEGG pathway enrichment and GO enrichment, were responses to hypoxia, decreased oxygen levels, the RNA polymerase II transcription regulator complex, RNA polymerase II-specific DNA-binding transcription factor binding, and other processes. In addition, KEGG was also significantly enriched in the HIF-1 signaling pathway and ferroptosis. Erythropoietin is synthesized by hypoxia inducible factors (HIFs) (Jelkmann 2011). In diabetic nephropathy, HIF-2 expression is decreased while HIF-1 expression is elevated, as has been demonstrated to be directly associated with the dysregulation of HIF signaling (Olmos *et al.* 2018). In contrast, HIF-1 α inhibition and hypoxia-mimicking HIF-2 α activation slow the progression of diabetic nephropathy (Ohtomo *et al.* 2008; Bessho *et al.* 2019). In addition, hypoxia inhibits ferritin phagocytosis, increases mitochondrial ferritin, and protects from ferroptosis (Fuhrmann *et al.* 2020). Interestingly, the eight marker genes we screened contained HIF1A and were highly expressed in T2DM. Among the thirty HIF1A target drugs retrieved, PX-478, its small molecule inhibitor, preserved pancreatic β cell function and increased insulin levels in diabetic mice in the presence of high glucose metabolism overload (Ilegems *et al.* 2022). In addition, targeted knockdown of HIF1A in mice resulted in a reduction in the size of atherosclerotic lesions and a decrease in macrophage accumulation (Akhtar *et al.* 2015). Similar results were shown in PX-478-treated mice, who showed a considerable decrease in the amount of atherosclerotic plaque in their aorta (Villa-Roe *et al.* 2022). Therefore, the HIF1A inhibitor PX-478 has the potential to be used as a therapeutic agent against diabetes and its complications, but its exact mechanism remains to be investigated. HILPDA (hypoxia-inducible lipid droplet-associated protein) is another marker gene that we screened for associated with hypoxia. HILPDA is a novel peroxisome proliferator-activated receptor (PPAR) target that can be expressed in multiple tissues as a small lipid droplet-associated protein (Mattijssen *et al.* 2014). Different events, including as hypoxia, beta-adrenergic stimulation, and PPAR transcription factors increase the production of HILPDA (Gimm *et al.* 2010; Mattijssen *et al.* 2014; Dijk *et al.* 2017; de la Rosa Rodriguez *et al.* 2021). It was shown that HILPDA is regulated by PPAR α through upstream PPARE (PPAR response element), and targeted overexpression increases hepatic triglyceride (TG) storage by reducing TG secretion (Mattijssen *et al.* 2014). Rodriguez *et al.* also showed a moderate reduction in triglycerides in the liver of mice with nonalcoholic steatohepatitis when they were specifically deficient in HILPDA (de la Rosa Rodriguez *et al.* 2021). Norepinephrine (NE), a sympathetic neurotransmitter, increases extracellular fatty acid absorption and triglyceride storage in macrophages by

acting through HILPDA-activated beta2-adrenergic receptors (β 2ARs) and decreasing free fatty acid release from triglyceride-loaded macrophages(Petkevicius *et al.*2021). Thus, HILPDA could also provide new therapeutic directions for metabolism-related fatty liver disease as well as T2DM in the future. From the results of our study above, we can show that hypoxia has a very important place in the development of diabetes.

SCD is an enzyme that regulates lipids and helps to desaturate saturated fatty acids(Wang *et al.* 2015). The GSAV analysis of SCD also confirmed that it is involved in the biosynthesis of unsaturated fatty acids and the metabolism of fatty acids. Previous animal studies have shown that defects in SCD-1 isoforms expressed in human tissues result in reduced lipid synthesis, increased lipid oxidation, enhanced insulin sensitivity, reduced hepatic glucose output and increased systemic glucose uptake(Ntambi *et al.* 2002; Flowers *et al.* 2007; Igarashi *et al.*2021). Among the targeted agents in SCD, Rosiglitazone is a thiazolidinedione that enhances insulin sensitivity and was often used in the past for blood glucose control, but is now used sparingly due to its significant cardiovascular side effects(Leal *et al.*2013; Raveendran *et al.* 2021). Animal studies have shown that SCD is elevated in obese rats but returns to normal after rosiglitazone treatment(Song *et al.* 2008). MK-8245 is a potent liver-targeted SCD inhibitor that lowers blood lipids and blood glucose and has been used in therapeutic trials to study type 2 diabetes(Xu *et al.* 2007; Oballa *et al.*2011). CLOFIBRATE and ARAMCHOL can lower blood lipids and have been used to treat NAFLD, as well as to indirectly delay the progression of T2DM(Gustafson *et al.* 2002; Ratziu *et al.*2021). However, the mechanisms associated with SCD in T2DM need further investigation.

In addition, we verified the expression of two genes, CD44 and MYCN, in T2DM and non-diabetic tissues by Western Blotting, Immunohistochemistry staining and Immunofluorescence Staining, and the results were statistically significant. CD44 is a cell surface glycoprotein, and now more and more studies have indicated that CD44 is involved in the regulation of glucose metabolism(Bogdani *et al.* 2014). It has been shown that CD44 is elevated in diabetic tissues, which is also consistent with our analysis, and correlated with insulin resistance and glycemic control levels. (Kodama *et al.*2012; Kang *et al.* 2013; Liu *et al.*2015). Related studies have shown that hyaluronan (HA) activation of CD44 leads to increased vulnerability of β cells to damage and increased insulin resistance, resulting in elevated blood glucose. (Assayag-Asherie *et al.* 2015; Hasib *et al.* 2019). Conversely, disruption of HA-CD44 interaction reduces the inflammatory cascade involved in islet destruction and exerts an antidiabetic effect (Weiss *et al.* 2000). In addition, treatment of obese mice with anti-CD44 monoclonal antibody reduced fasting glucose levels, hepatic steatosis and insulin resistance to the level of treatment with metformin and pioglitazone(Kodama *et al.* 2015). MYCN is a member of the MYC family of proto-oncogenes and is associated with the development of many tumors, especially neuroblastoma (NB) (Alborzinia *et al.* 2022). In NB, MYCN maintains tumor growth by promoting fatty acid uptake(Tao *et al.* 2022). It has been shown that MYCN can increase glycolysis and is associated with non-obese diabetes mellitus(Wu *et al.* 2012; Zirath *et al.*2013). However, studies related to MYCN and diabetes are still scarce.

The development of T2DM is closely related to the disorders of immune status and function, abnormal immune cell activation and the subsequent inflammatory environment makes glycemic control more difficult(Donath and Shoelson 2011). In the present study, immune infiltration analysis revealed that dendritic cells activated were highly expressed in T2DM. In contrast, dendritic cells resting were lowly expressed in T2DM. A study by SURENDAR J *et al.*(Surendar *et al.* 2012). Showed that the activation state of myeloid dendritic cells(mDCs)and plasmacytoid dendritic cells(pDCs)in diabetic patients may be caused by increased levels of granulocyte-macrophage colony-stimulating factor (GM-CSF) and other pro-inflammatory cytokines. Hyperinsulinemia also stimulates dendritic cell (DC) activation and overexpression of the scavenger receptors SR-A, CD36, and LOX-1, which can boost DC oxidized low-density lipoprotein (oxLDL) absorption capacity(Lu *et al.* 2013; Lu *et al.* 2015). Cardiovascular disease (CVD) remains the leading cause of death in T2DM(Yun and Ko 2021). DC has an important role in the development of CVD and atherosclerosis(Bacci *et al.* 2008; Bobryshev 2010). Recent research on relevant animals have also demonstrated that DCs concentrate primarily in perivascular adipose tissue (PVAT) and are linked to an excess of pro-inflammatory cytokines, which impairs PVAT's capacity to increase vasorelaxation and perform anti-contractile action in T2DM(Qiu *et al.* 2018). For T2DM and its complications, immune infiltration can be thought of as one of the future therapy targets.

The ferroptosis genes we looked for in T2DM were SCD, CD44, HIF1A, BCAT2, MTF1, HILPDA, NR1D2, and MYCN. We mainly discussed five genes (SCD, CD44, HIF1A, HILPDA and MYCN) and selected two genes, CD44 and MYCN, for clinical validation. In conclusion, more research is needed to determine if our projected non-coding RNAs and gene-targeted medicines are involved in T2DM. Naturally, there are certain limitations to our study. First, we only have verified experimentally the protein expression of CD44 and MYCN in T2DM and normal tissues, and more experiments are needed in the future to further explore the mechanisms of these genes in T2DM. Second, a higher sample size of T2DM might increase its accuracy because the short sample size resulted in differential expression of several marker genes showing mistakes in external validation. Thirdly, new ferroptosis-related genes are still to be found, and the FerrDb database is always being updated.

Conclusion

In this study, we identified eight hub genes (SCD, CD44, HIF1A, BCAT2, MTF1, HILPDA, NR1D2 and MYCN) that are closely associated with ferroptosis in T2DM. The three iron death genes, HIF1A, HILPDA and SCD, are strongly related to T2DM, hypoxia and lipid metabolism, providing new research directions for the development and treatment of T2DM and its complications. Based on these eight genes, we constructed a model with a high ability to diagnose T2DM. We also predicted the drugs corresponding to these eight genes as well as constructed a ceRNA network map. In addition, we verified the protein expression of CD44 and MYCN in T2DM and non-diabetic tissues by Western Blotting, Immunohistochemistry staining and Immunofluorescence Staining, and the results were statistically significant. The above findings suggest that further studies of ferroptosis may offer new therapeutic goals and biomarkers for patients with T2DM.

Conflict of interest

The authors claim to have no conflicts of interest.

Ethics approval

All the experiments were approved by the ethics committee of Affiliated Hospital of Shandong University of Traditional Chinese Medicine & Shandong Provincial Hospital of Traditional Chinese Medicine (Approval Number: AF/SC-08/02.0). All procedures were carried out in strict accordance with the 1964 Declaration of Helsinki. All patients involved in this study had given their informed consent before study.

Author contributions

This work was designed by LF and SW. SW and YPL authored the text after analyzing the data. YXZ assisted in data collection and text revision. YLZ and HMG contributed to the data analysis. The final manuscript was reviewed and approved by all writers.

Consent for publication

The publishing of this paper has the approval of all authors.

Funding

This work was supported by the National Natural Science Foundation of China (No. 31900889).

Acknowledgments

We are grateful to everyone who took part in the study. We would like to extend a special thank you to the GEO database for the analytical data and the administrators of all open public websites.

References

Akhtar, S., P. Hartmann, E. Karshovska, F. A. Rinderknecht, P. Subramanian *et al.* , 2015 Endothelial Hypoxia-Inducible Factor-1alpha Promotes Atherosclerosis and Monocyte Recruitment by Upregulating MicroRNA-19a. *Hypertension* 66: 1220-1226.

- Alborzinia, H., A. F. Florez, S. Kreth, L. M. Bruckner, U. Yildiz *et al.* , 2022 MYCN mediates cysteine addiction and sensitizes neuroblastoma to ferroptosis. *Nat Cancer* **3**: 471-485.
- Ali, M. K., J. Pearson-Stuttard, E. Selvin and E. W. Gregg, 2022 Interpreting global trends in type 2 diabetes complications and mortality. *Diabetologia* **65**: 3-13.
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler *et al.* , 2000 Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29.
- Assayag-Asherie, N., D. Sever, M. Bogdani, P. Johnson, T. Weiss *et al.* , 2015 Can CD44 Be a Mediator of Cell Destruction? The Challenge of Type 1 Diabetes. *PLoS One* **10**: e0143589.
- Bacci, S., L. Pieri, A. M. Buccoliero, A. Bonelli, G. Taddei *et al.* , 2008 Smooth muscle cells, dendritic cells and mast cells are sources of TNFalpha and nitric oxide in human carotid artery atherosclerosis. *Thromb Res* **122**: 657-667.
- Bessho, R., Y. Takiyama, T. Takiyama, H. Kitsunai, Y. Takeda *et al.* , 2019 Hypoxia-inducible factor-1alpha is the therapeutic target of the SGLT2 inhibitor for diabetic nephropathy. *Sci Rep* **9**: 14754.
- Bobryshev, Y. V., 2010 Dendritic cells and their role in atherogenesis. *Lab Invest* **90**: 970-984.
- Bogdani, M., P. Y. Johnson, S. Potter-Perigo, N. Nagy, A. J. Day *et al.* , 2014 Hyaluronan and hyaluronan-binding proteins accumulate in both human type 1 diabetic islets and lymphoid tissues and associate with inflammatory cells in insulinitis. *Diabetes* **63**:2727-2743.
- Bruni, A., A. R. Pepper, R. L. Pawlick, B. Gala-Lopez, A. F. Gamble *et al.* , 2018 Ferroptosis-inducing agents compromise in vitro human islet viability and function. *Cell Death Dis* **9**: 595.
- Carbone, S., M. G. Del Buono, C. Ozemek and C. J. Lavie, 2019 Obesity, risk of diabetes and role of physical activity, exercise training and cardiorespiratory fitness. *Prog Cardiovasc Dis* **62**: 327-333.
- Cong, L., X. Dong, Y. Wang, Y. Deng, B. Li *et al.* , 2019 On the role of synthesized hydroxylated chalcones as dual functional amyloid-beta aggregation and ferroptosis inhibitors for potential treatment of Alzheimer's disease. *Eur J Med Chem* **166**: 11-21.
- de la Rosa Rodriguez, M. A., L. Deng, A. Gemmink, M. van Weeghel, M. L. Aoun *et al.* , 2021 Hypoxia-inducible lipid droplet-associated induces DGAT1 and promotes lipid storage in hepatocytes. *Mol Metab* **47**: 101168.
- Dijk, W., F. Mattijssen, M. de la Rosa Rodriguez, A. Loza Valdes, A. Loft *et al.* , 2017 Hypoxia-Inducible Lipid Droplet-Associated Is Not a Direct Physiological Regulator of Lipolysis in Adipose Tissue. *Endocrinology* **158**: 1231-1251.
- Dixon, S. J., K. M. Lemberg, M. R. Lamprecht, R. Skouta, E. M. Zaitsev *et al.* , 2012 Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* **149**: 1060-1072.
- Do Van, B., F. Gouel, A. Jonneaux, K. Timmerman, P. Gele *et al.* , 2016 Ferroptosis, a newly characterized form of cell death in Parkinson's disease that is regulated by PKC. *Neurobiol Dis* **94**:169-178.
- Donath, M. Y., and S. E. Shoelson, 2011 Type 2 diabetes as an inflammatory disease. *Nat Rev Immunol* **11**: 98-107.
- El Hout, M., L. Dos Santos, A. Hamai and M. Mehrpour, 2018 A promising new approach to cancer therapy: Targeting iron metabolism in cancer stem cells. *Semin Cancer Biol* **53**: 125-138.
- Flowers, J. B., M. E. Rabaglia, K. L. Schueler, M. T. Flowers, H. Lan *et al.* , 2007 Loss of stearyl-CoA desaturase-1 improves insulin sensitivity in lean mice but worsens diabetes in leptin-deficient obese mice. *Diabetes* **56**: 1228-1239.

- Friedman, J., T. Hastie and R. Tibshirani, 2010 Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* 33: 1-22.
- Fuhrmann, D. C., A. Mondorf, J. Beifuss, M. Jung and B. Brune, 2020 Hypoxia inhibits ferritinophagy, increases mitochondrial ferritin, and protects from ferroptosis. *Redox Biol* 36: 101670.
- Gautam, S., F. Alam, S. Moin, N. Noor and S. H. Arif, 2021 Role of ferritin and oxidative stress index in gestational diabetes mellitus. *J Diabetes Metab Disord* 20: 1615-1619.
- Gheibi, S., T. Singh, J. da Cunha, M. Fex and H. Mulder, 2020 Insulin/Glucose-Responsive Cells Derived from Induced Pluripotent Stem Cells: Disease Modeling and Treatment of Diabetes. *Cells* 9.
- Gimm, T., M. Wiese, B. Teschemacher, A. Deggerich, J. Schodel *et al.* , 2010 Hypoxia-inducible protein 2 is a novel lipid droplet protein and a specific target gene of hypoxia-inducible factor-1. *FASEB J* 24: 4443-4458.
- Gustafson, L. A., F. Kuipers, C. Wiegman, H. P. Sauerwein, J. A. Romijn *et al.* , 2002 Clofibrate improves glucose tolerance in fat-fed rats but decreases hepatic glucose consumption capacity. *J Hepatol* 37: 425-431.
- Hadian, K., and B. R. Stockwell, 2020 SnapShot: Ferroptosis. *Cell* 181: 1188-1188 e1181.
- Hanzelmann, S., R. Castelo and J. Guinney, 2013 GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14: 7.
- Hasib, A., C. K. Hennayake, D. P. Bracy, A. R. Bugler-Lamb, L. Lantier *et al.* , 2019 CD44 contributes to hyaluronan-mediated insulin resistance in skeletal muscle of high-fat-fed C57BL/6 mice. *Am J Physiol Endocrinol Metab* 317: E973-E983.
- Igarashi, Y., S. Iida, J. Dai, J. Huo, X. Cui *et al.* , 2021 Glavonoid-rich oil supplementation reduces stearyl-coenzyme A desaturase 1 expression and improves systemic metabolism in diabetic, obese KK-A(y) mice. *Biomed Pharmacother* 140: 111714.
- Ilegems, E., G. Bryzgalova, J. Correia, B. Yesildag, E. Berra *et al.* , 2022 HIF-1alpha inhibitor PX-478 preserves pancreatic beta cell function in diabetes. *Sci Transl Med* 14: eaba9112.
- Jelkmann, W., 2011 Regulation of erythropoietin production. *J Physiol* 589: 1251-1258.
- Kanehisa, M., and S. Goto, 2000 KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30.
- Kang, H. S., G. Liao, L. M. DeGraff, K. Gerrish, C. D. Bortner *et al.* , 2013 CD44 plays a critical role in regulating diet-induced adipose inflammation, hepatic steatosis, and insulin resistance. *PLoS One* 8: e58417.
- Kim, J. H., T. M. Lewin and R. A. Coleman, 2001 Expression and characterization of recombinant rat Acyl-CoA synthetases 1, 4, and 5. Selective inhibition by triacsin C and thiazolidinediones. *J Biol Chem* 276: 24667-24673.
- Kodama, K., M. Horikoshi, K. Toda, S. Yamada, K. Hara *et al.* , 2012 Expression-based genome-wide association study links the receptor CD44 in adipose tissue with type 2 diabetes. *Proc Natl Acad Sci U S A* 109: 7049-7054.
- Kodama, K., K. Toda, S. Morinaga, S. Yamada and A. J. Butte, 2015 Anti-CD44 antibody treatment lowers hyperglycemia and improves insulin resistance, adipose inflammation, and hepatic steatosis in diet-induced obese mice. *Diabetes* 64: 867-875.
- Leal, I., S. A. Romio, M. Schuemie, A. Oteri, M. Sturkenboom *et al.* , 2013 Prescribing pattern of glucose lowering drugs in the United Kingdom in the last decade: a focus on the effects of safety warnings about rosiglitazone. *Br J Clin Pharmacol* 75: 861-868.

- Li, X. Y., and P. S. Leung, 2020 Erastin-induced ferroptosis is a regulator for the growth and function of human pancreatic islet-like cell clusters. *Cell Regen* 9: 16.
- Li, Z., C. Zhang, Y. Liu, F. Wang, B. Zhao *et al.* , 2022 Diagnostic and Predictive Values of Ferroptosis-Related Genes in Child Sepsis. *Front Immunol* 13: 881914.
- Linkermann, A., R. Skouta, N. Himmerkus, S. R. Mulay, C. Dewitz *et al.* , 2014 Synchronized renal tubular cell death involves ferroptosis. *Proc Natl Acad Sci U S A* 111: 16836-16841.
- Liu, J., Q. Li, Y. Yang and L. Ma, 2020 Iron metabolism and type 2 diabetes mellitus: A meta-analysis and systematic review. *J Diabetes Investig* 11: 946-955.
- Liu, L. F., K. Kodama, K. Wei, L. L. Tolentino, O. Choi *et al.* , 2015 The receptor CD44 is associated with systemic insulin resistance and proinflammatory macrophages in human adipose tissue. *Diabetologia* 58: 1579-1586.
- Lu, H., D. Huang, K. Yao, C. Li, S. Chang *et al.* , 2015 Insulin enhances dendritic cell maturation and scavenger receptor-mediated uptake of oxidised low-density lipoprotein. *J Diabetes Complications* 29: 465-471.
- Lu, H., K. Yao, D. Huang, A. Sun, Y. Zou *et al.* , 2013 High glucose induces upregulation of scavenger receptors and promotes maturation of dendritic cells. *Cardiovasc Diabetol* 12: 80.
- Magliano, D. J., L. Chen, R. M. Islam, B. Carstensen, E. W. Gregg *et al.* , 2021 Trends in the incidence of diagnosed diabetes: a multicountry analysis of aggregate data from 22 million diagnoses in high-income and middle-income settings. *Lancet Diabetes Endocrinol* 9: 203-211.
- Marin-Penalver, J. J., I. Martin-Timon, C. Sevillano-Collantes and F. J. Del Canizo-Gomez, 2016 Update on the treatment of type 2 diabetes mellitus. *World J Diabetes* 7: 354-395.
- Mattijssen, F., A. Georgiadi, T. Andasarie, E. Szalowska, A. Zota *et al.* , 2014 Hypoxia-inducible lipid droplet-associated (HILPDA) is a novel peroxisome proliferator-activated receptor (PPAR) target involved in hepatic triglyceride secretion. *J Biol Chem* 289:19279-19293.
- Ntambi, J. M., M. Miyazaki, J. P. Stoehr, H. Lan, C. M. Kendziorski *et al.* , 2002 Loss of stearyl-CoA desaturase-1 function protects mice against adiposity. *Proc Natl Acad Sci U S A* 99:11482-11486.
- Oballa, R. M., L. Belair, W. C. Black, K. Bleasby, C. C. Chan *et al.* , 2011 Development of a liver-targeted stearyl-CoA desaturase (SCD) inhibitor (MK-8245) to establish a therapeutic window for the treatment of diabetes and dyslipidemia. *J Med Chem* 54: 5082-5096.
- Ohtomo, S., M. Nangaku, Y. Izuhara, S. Takizawa, C. Strihou *et al.* , 2008 Cobalt ameliorates renal injury in an obese, hypertensive type 2 diabetes rat model. *Nephrol Dial Transplant* 23:1166-1172.
- Olmos, G., J. M. Munoz-Felix, I. Mora, A. G. Muller, M. P. Ruiz-Torres *et al.* , 2018 Impaired erythropoietin synthesis in chronic kidney disease is caused by alterations in extracellular matrix composition. *J Cell Mol Med* 22: 302-314.
- Petkevicius, K., G. Bidault, S. Virtue, B. Jenkins, X. van Dierendonck *et al.* , 2021 Norepinephrine promotes triglyceride storage in macrophages via beta2-adrenergic receptor activation. *FASEB J* 35: e21266.
- Qiu, T., M. Li, M. A. Tanner, Y. Yang, J. R. Sowers *et al.* , 2018 Depletion of dendritic cells in perivascular adipose tissue improves arterial relaxation responses in type 2 diabetic mice. *Metabolism* 85: 76-89.
- Ratziu, V., L. de Guevara, R. Safadi, F. Poordad, F. Fuster *et al.* , 2021 Aramchol in patients with nonalcoholic steatohepatitis: a randomized, double-blind, placebo-controlled phase 2b trial. *Nat Med* 27: 1825-1835.
- Raveendran, A. V., C. J. Fernandez and K. Jacob, 2021 Efficacy and Cardiovascular Safety of Thiazolidinediones. *Curr Drug Saf* 16:233-249.

- Saeedi, P., I. Petersohn, P. Salpea, B. Malanda, S. Karuranga *et al.* , 2019 Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. *Diabetes Res Clin Pract* 157: 107843.
- Sha, W., F. Hu, Y. Xi, Y. Chu and S. Bu, 2021 Mechanism of Ferroptosis and Its Role in Type 2 Diabetes Mellitus. *J Diabetes Res* 2021:9999612.
- Song, J., H. Liu, H. W. Resson, S. Tiwari and C. M. Ecelbarger, 2008 Chronic rosiglitazone therapy normalizes expression of ACE1, SCD1 and other genes in the kidney of obese Zucker rats as determined by microarray analysis. *Exp Clin Endocrinol Diabetes* 116: 315-325.
- Surendar, J., V. Mohan, N. Pavankumar, S. Babu and V. Aravindhana, 2012 Increased levels of serum granulocyte-macrophage colony-stimulating factor is associated with activated peripheral dendritic cells in type 2 diabetes subjects (CURES-99). *Diabetes Technol Ther* 14:344-349.
- Tao, L., M. A. Mohammad, G. Milazzo, M. Moreno-Smith, T. D. Patel *et al.* , 2022 MYCN-driven fatty acid uptake is a metabolic vulnerability in neuroblastoma. *Nat Commun* 13: 3728.
- Tomic, D., J. E. Shaw and D. J. Magliano, 2022 The burden and risks of emerging complications of diabetes mellitus. *Nat Rev Endocrinol* 18: 525-539.
- Tsubouchi, K., J. Araya, M. Yoshida, T. Sakamoto, T. Koumura *et al.* , 2019 Involvement of GPx4-Regulated Lipid Peroxidation in Idiopathic Pulmonary Fibrosis Pathogenesis. *J Immunol* 203:2076-2087.
- Viigimaa, M., A. Sachinidis, M. Toumpourleka, K. Koutsampasopoulos, S. Alliksoo *et al.* , 2020 Macrovascular Complications of Type 2 Diabetes Mellitus. *Curr Vasc Pharmacol* 18: 110-116.
- Villa-Roel, N., K. Ryu, L. Gu, J. Fernandez Esmerats, D. W. Kang *et al.* , 2022 Hypoxia inducible factor 1alpha inhibitor PX-478 reduces atherosclerosis in mice. *Atherosclerosis* 344: 20-30.
- Wang, H., M. G. Klein, H. Zou, W. Lane, G. Snell *et al.* , 2015 Crystal structure of human stearyl-coenzyme A desaturase in complex with substrate. *Nat Struct Mol Biol* 22: 581-585.
- Weiss, L., S. Slavin, S. Reich, P. Cohen, S. Shuster *et al.* , 2000 Induction of resistance to diabetes in non-obese diabetic mice by targeting CD44 with a specific monoclonal antibody. *Proc Natl Acad Sci U S A* 97: 285-290.
- Weyer, C., C. Bogardus, D. M. Mott and R. E. Pratley, 1999 The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus. *J Clin Invest* 104:787-794.
- Wu, J., D. N. Kakoola, N. I. Lenchik, D. M. Desiderio, D. R. Marshall *et al.* , 2012 Molecular phenotyping of immune cells from young NOD mice reveals abnormal metabolic pathways in the early induction phase of autoimmune diabetes. *PLoS One* 7: e46941.
- Wu, T., E. Hu, S. Xu, M. Chen, P. Guo *et al.* , 2021 clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2: 100141.
- Wu, X., K. Qin, C. D. Iroegbu, K. Xiang, J. Peng *et al.* , 2022 Genetic analysis of potential biomarkers and therapeutic targets in ferroptosis from coronary artery disease. *J Cell Mol Med* 26:2177-2190.
- Xu, H., D. Wilcox, P. Nguyen, M. Voorbach, H. Smith *et al.* , 2007 Hepatic knockdown of stearyl-CoA desaturase 1 via RNA interference in obese mice decreases lipid content and changes fatty acid composition. *Front Biosci* 12: 3781-3794.
- Yang, C., J. Ren, B. Li, C. Jin, C. Ma *et al.* , 2019 Identification of gene biomarkers in patients with postmenopausal osteoporosis. *Mol Med Rep* 19: 1065-1073.
- Younossi, Z. M., P. Golabi, L. de Avila, J. M. Paik, M. Srishord *et al.* , 2019 The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis. *J Hepatol* 71: 793-801.

Yuan, H., X. Li, X. Zhang, R. Kang and D. Tang, 2016 Identification of ACSL4 as a biomarker and contributor of ferroptosis. *Biochem Biophys Res Commun* 478: 1338-1343.

Yun, J. S., and S. H. Ko, 2021 Current trends in epidemiology of cardiovascular disease and cardiovascular risk management in type 2 diabetes. *Metabolism* 123: 154838.

Zirath, H., A. Frenzel, G. Oliynyk, L. Segerstrom, U. K. Westermarck *et al.*, 2013 MYC inhibition induces metabolic changes leading to accumulation of lipid droplets in tumor cells. *Proc Natl Acad Sci U S A* 110: 10258-10263.

TABLE 1. 58 of 358 FRGs were differentially expressed between T2DM and normal samples, including 38 up-regulated and 20 down-regulated genes.

FIGURE 1. Overview of the differentially expressed ferroptosis genes with T2DM and controls. (A) Heat map of 58 ferroptosis-related DEGs. (B) The correlation of these genes. Most of these genes were strongly correlated with each other.

FIGURE 2. Enrichment analysis of ferroptosis-related DEGs. (A) Gene Ontology (GO) functional analysis showing enrichment of ferroptosis-related DEGs. (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of ferroptosis-related DEGs.

FIGURE 3. 8 ferroptosis-related DEGs were identified as diagnostic genes for T2DM. (A and B) 15 ferroptosis-related DEGs obtained using the LASSO algorithm based on the minimum lambda. (C and D) 18 ferroptosis-related DEGs obtained using the SVM-RFE algorithm (maximum precision = 0.731, minimum RMSE = 0.269). E Eight marker genes were obtained from the LASSO and SVM-RFE algorithm. F ROC curves for the 8 marker genes. G Logistic regression model to identify the AUC of disease samples.

FIGURE 4. The interaction relationships and expression of the eight marker genes in the GSE78721. (A) Interaction of 8 marker genes. (B-I) Comparison of the expression of 8 marker genes in T2DM and healthy samples.

FIGURE 5. Single gene KEGG-GSEA analysis of these 8 genes

FIGURE 6. High-and low-expression groups based on the expression levels of each marker gene combined with GSVA inHIF1A (A) and SCD (B).

FIGURE 7. Immune feature analysis. (A) Bar charts of 22 immune cell proportions in T2DM and controls. (B) Correlation heatmap depicting correlations between infiltrated immune cells in sepsis. The darker the colour is, the stronger the correlation. (C) Differential expression of different types of immune cell marker expression between T2DM and controls.

FIGURE 8. Correlation between the expression of these 8 marker genes and immune cells.

FIGURE 9. Construction of the nomogram based on the logistic regression model. (A) Nomogram specifically quantified the odds ratio of T2DM based on 8 ferroptosis characteristics. (B) The calibration curves of nomogram. (C) The decision curve of nomogram. (D) The clinical impact curves of nomogram.

FIGURE 10. Prediction of marker gene- targeted drugs. The drugs may target marker genes through the DGIdb database and the interaction relationship between the two.

FIGURE 11. A ceRNA network based on these 8 genes. The network includes 671 nodes (7 marker genes, 349 miRNAs and 315 lncRNAs).

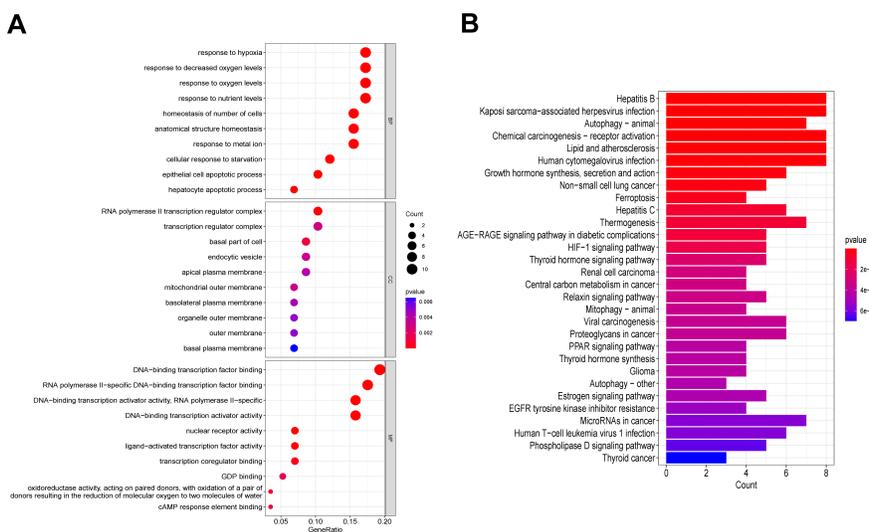
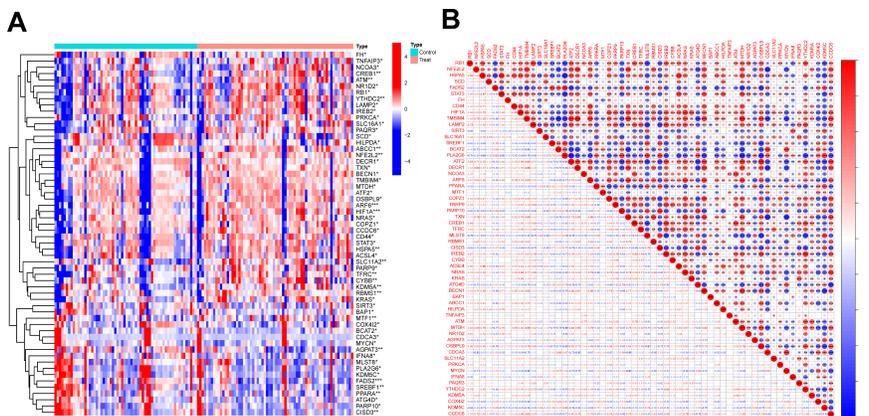
FIGURE 12. Expression status and diagnostic ability of marker genes in the validation set. (A) The expression of marker genes in the GSE76895 dataset. (B) ROC curves for the 8 marker genes in the GSE76895 dataset. (C) Logistic regression model to determine the AUC of the validation set T2DM samples.

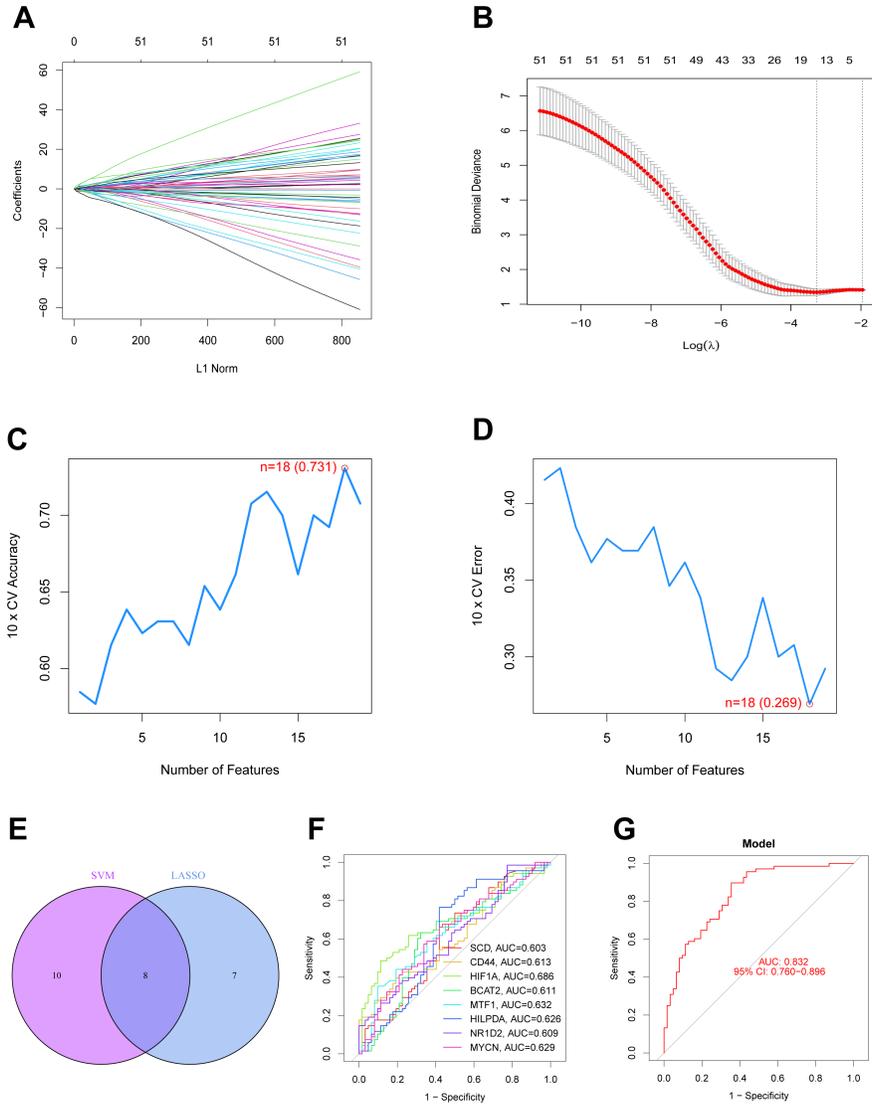
FIGURE 13. After ethical approval we took tissues from diabetic ulcers and non-diabetic patients and performed Western blot, immunofluorescence, and immunohistochemistry, and CD44 and MYCN differed in

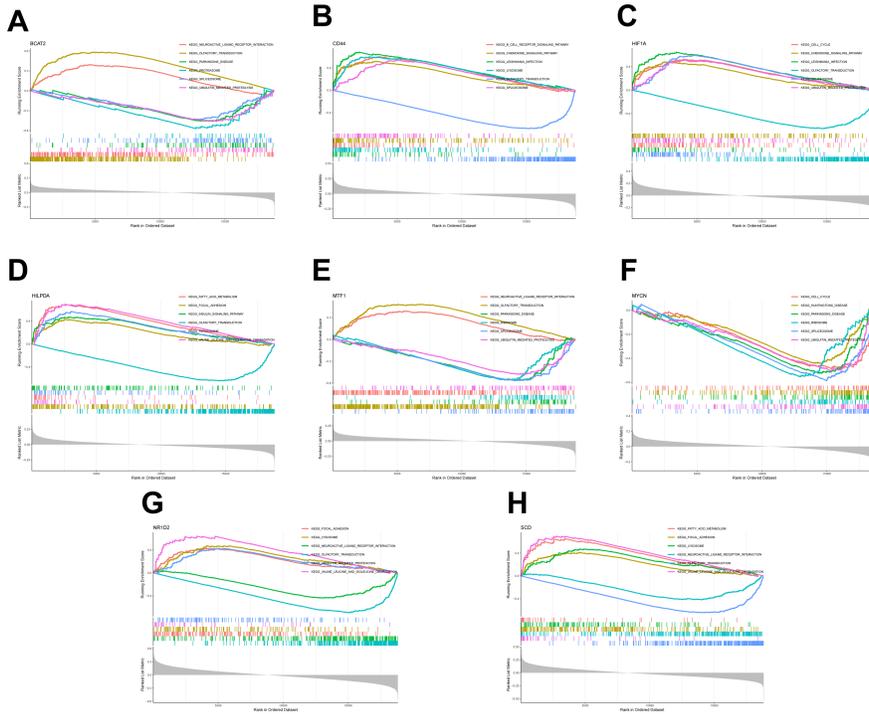
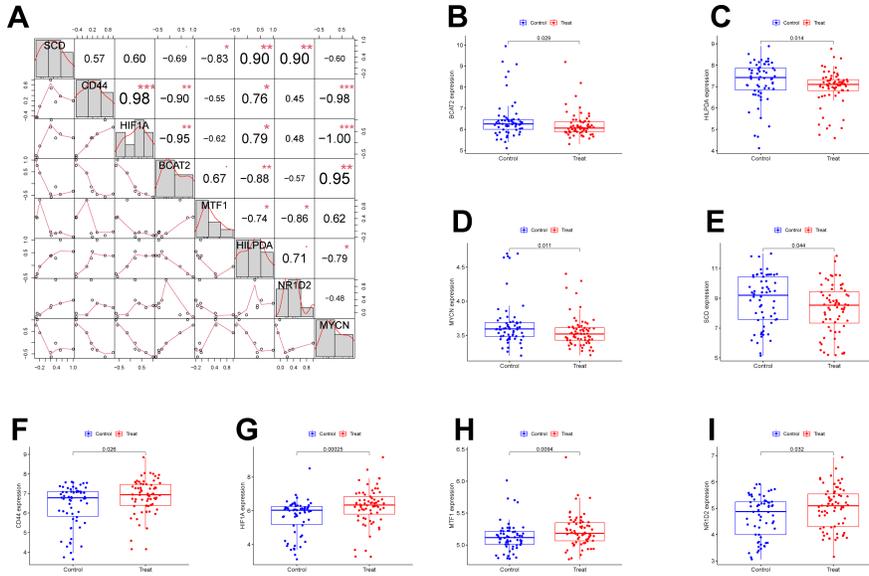
expression. (A, B) Protein blotting assays were conducted to detect changes in CD44 and MYCN expression in diabetic ulcers and normal skin tissue. (C, D, E, F) IHC analysis and quantification of percentage of CD44 with MYCN in diabetic ulcers and normal skin tissue positive cells. Scale bar = 50 μ m. (G, H, I, L) Immunofluorescence assay and quantification of the percentage of CD44 versus MYCN in diabetic ulcers and normal skin tissue positive cells. Scale bar = 50 μ m. *P < 0.05, **P < 0.01, ***P < 0.001.

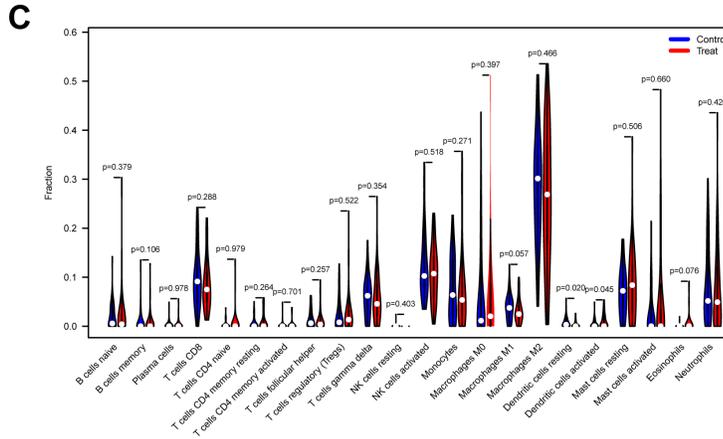
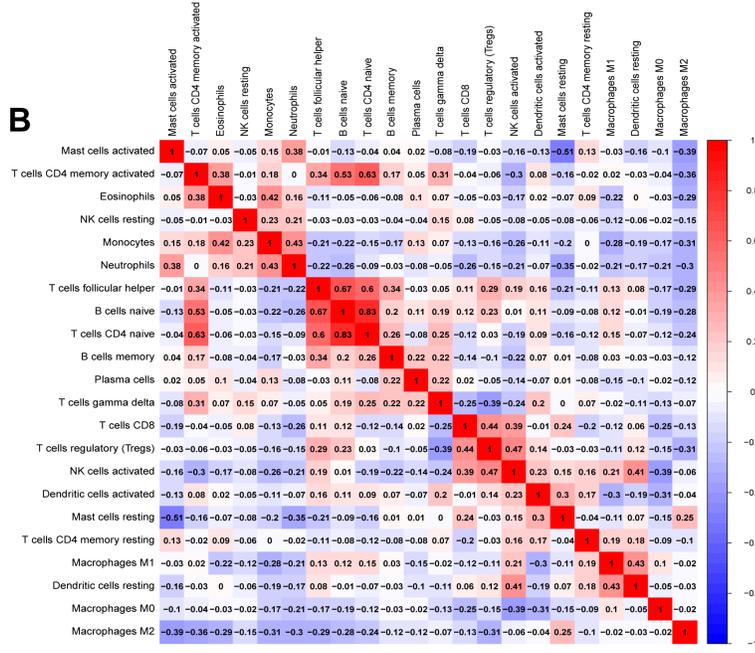
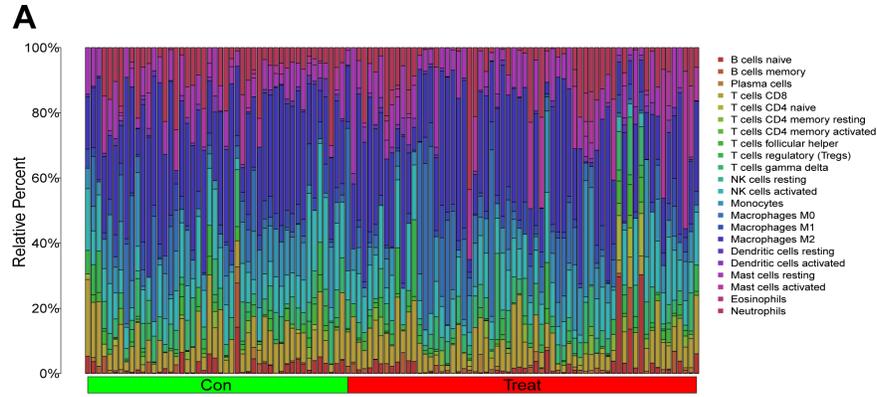
Hosted file

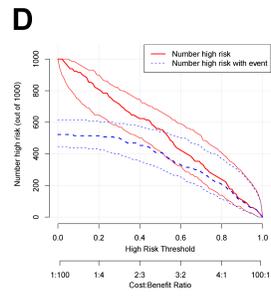
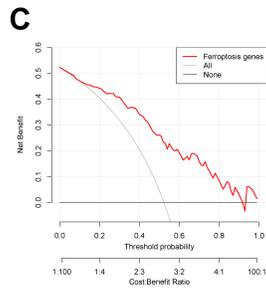
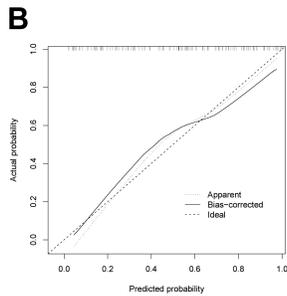
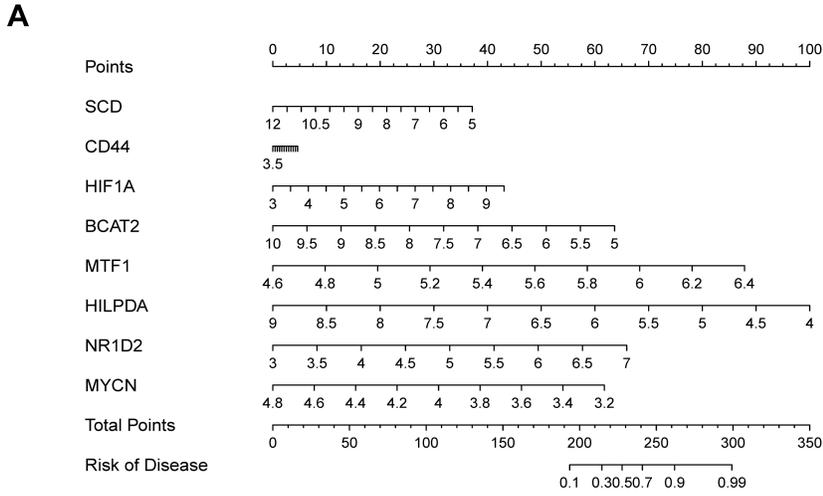
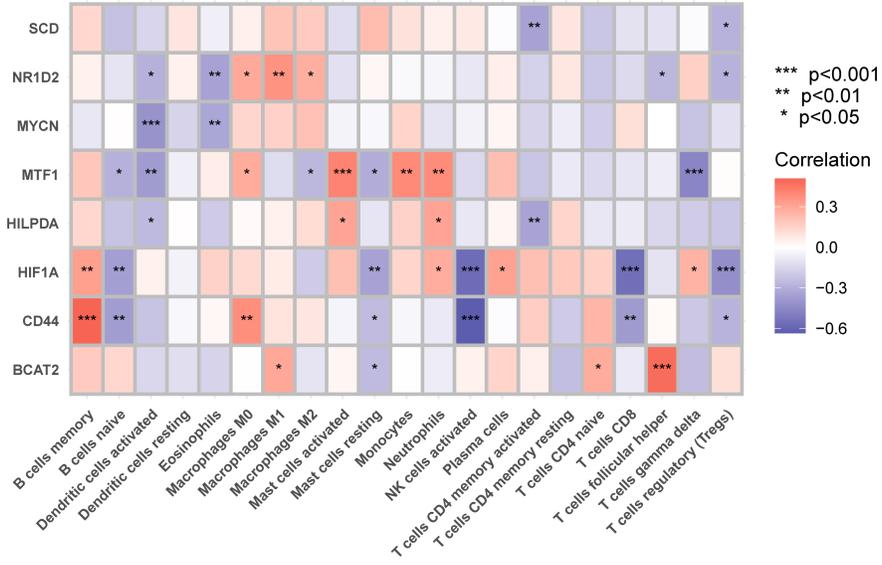
abbreviations list.docx available at <https://authorea.com/users/723366/articles/708327-identification-of-ferroptosis-related-genes-in-type-2-diabetes-mellitus-based-on-the-machine-learning>

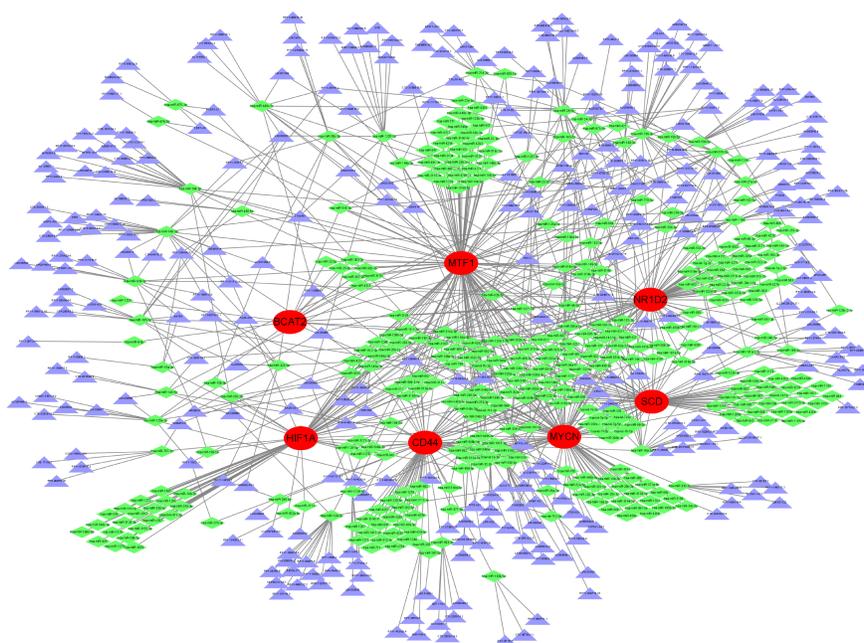
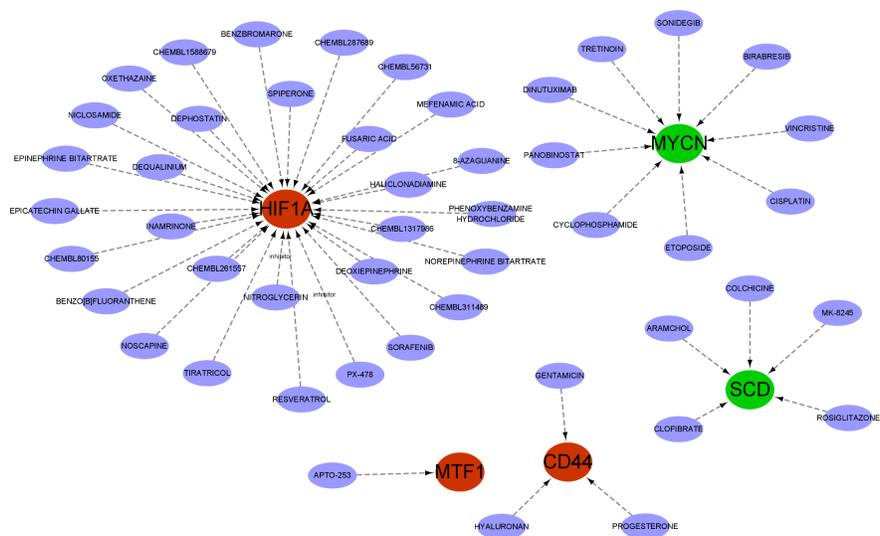


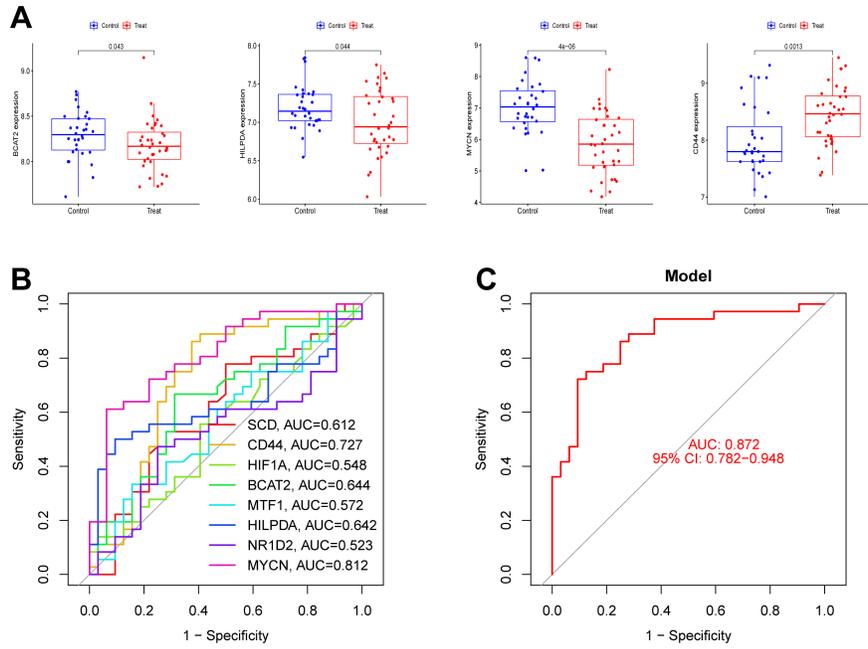




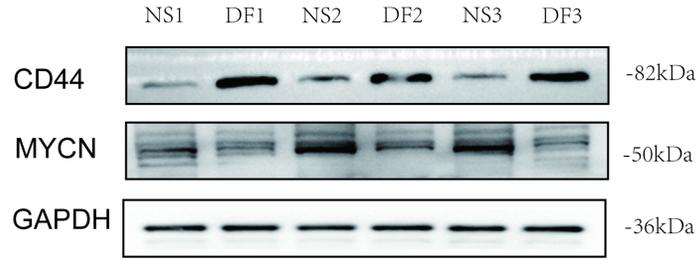




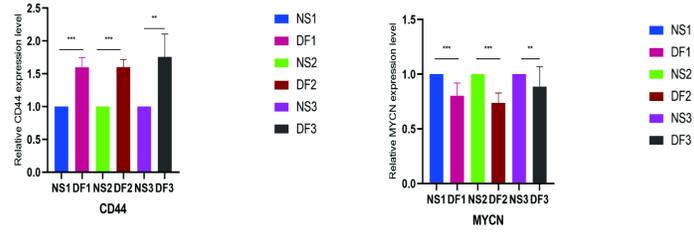




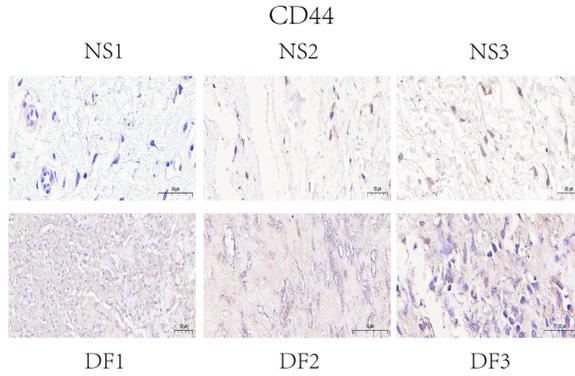
A



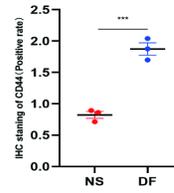
B



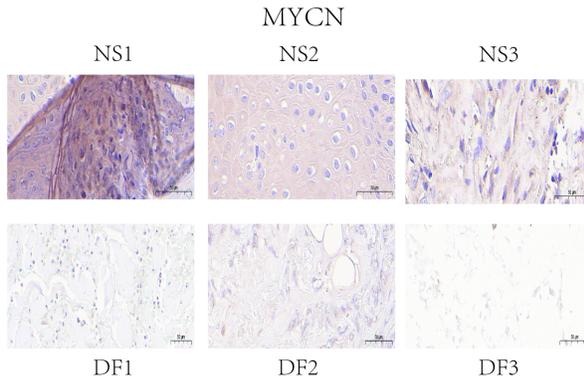
C



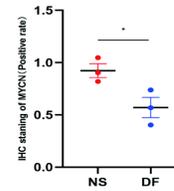
D



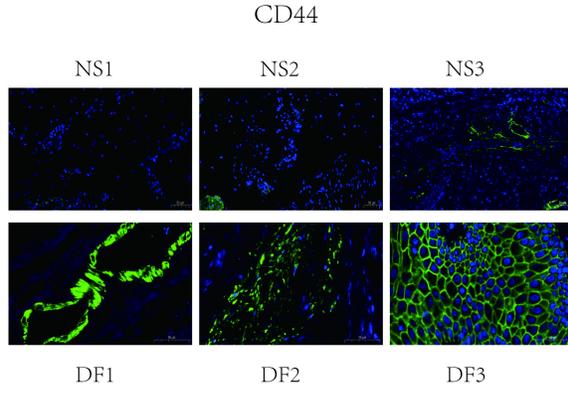
E



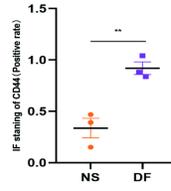
F



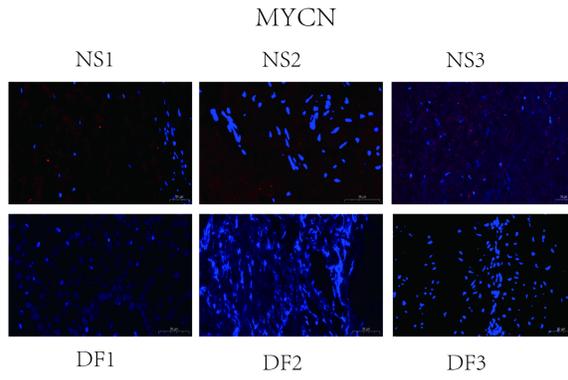
G



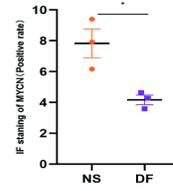
H

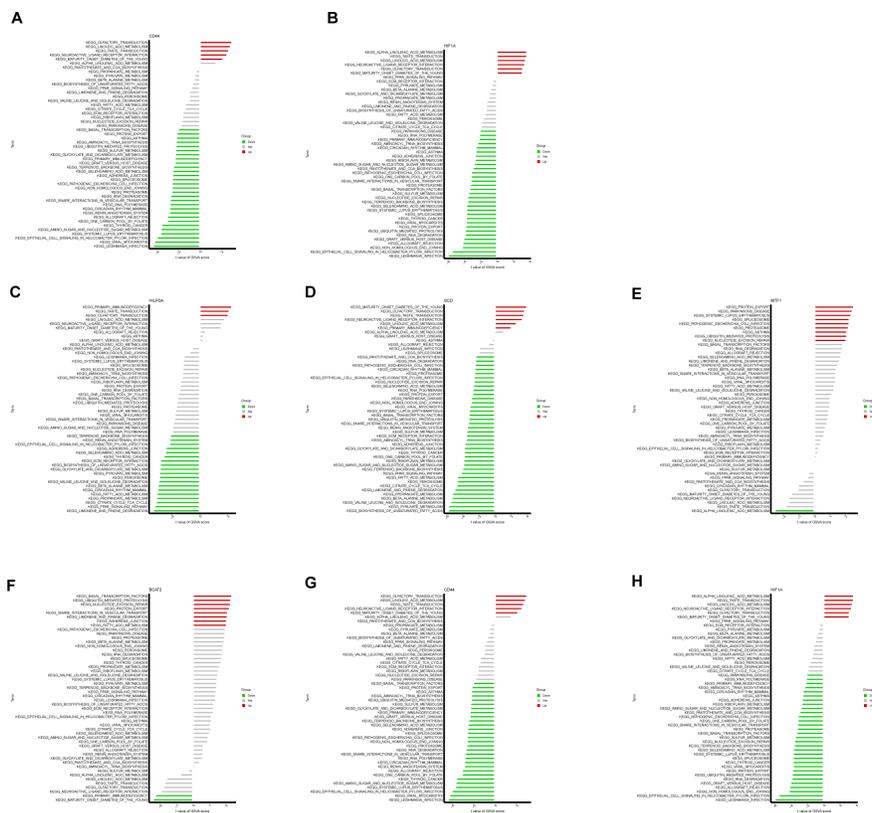


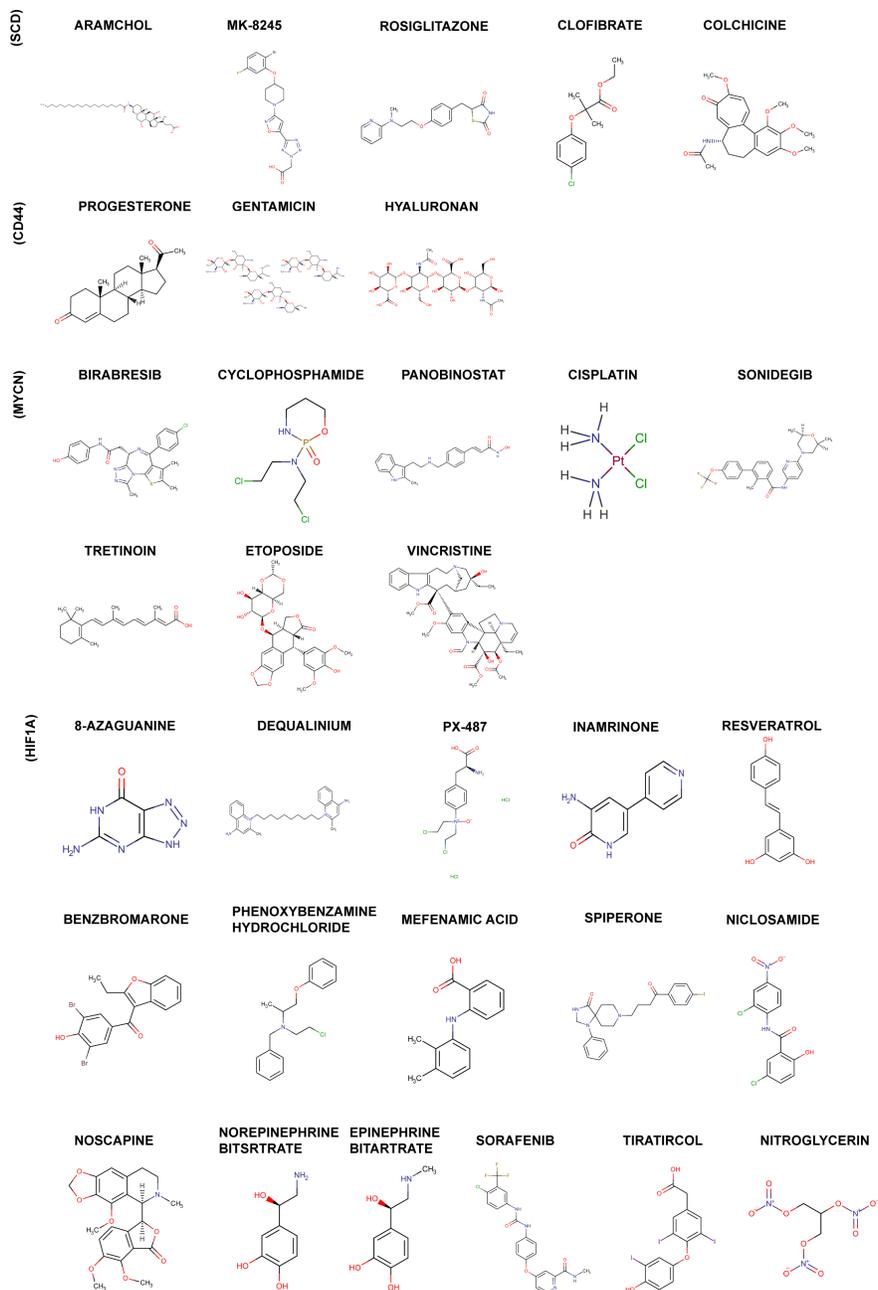
I



L







Hosted file

Table1.docx available at <https://authorea.com/users/723366/articles/708327-identification-of-ferroptosis-related-genes-in-type-2-diabetes-mellitus-based-on-the-machine-learning>