



OPEN

A bioinformatics approach to identify a disulfidptosis-related gene signature for prognostic implication in colon adenocarcinoma

Gunchu Hu¹, Hongliang Yao¹, Zuxing Wei¹, Linye Li¹, Zhuowen Yu^{3,4}, Jian Li¹, Xiong Luo¹ & Zhushu Guo^{1,2}✉

Colon adenocarcinoma (COAD) is a type of cancer that arises from the glandular epithelial cells that produce mucus in the colon. COAD is influenced by various factors, including genetics, environment and lifestyle. The outcome of COAD is determined by the tumor stage, location, molecular characteristics and treatment. Disulfidptosis is a new mode of cell death that may affect cancer development. We discovered genes associated with disulfidptosis in colon adenocarcinoma and proposed them as novel biomarkers and therapeutic targets for COAD. We analyzed the mRNA expression data and clinical information of COAD patients from The Cancer Genome Atlas (TCGA) database and Xena databases, extracted disulfidptosis-related genes from the latest reports on disulfidptosis. We used machine learning to select key features and build a signature and validated the risk model using data from the Gene Expression Omnibus (GEO) database and Human Protein Atlas (HPA). We also explored the potential biological functions and therapeutic implications of the disulfidptosis-related genes using CIBERSORTx and GDSC2 databases. We identified four disulfidptosis-related genes: TRIP6, OXSM, MYH3 and MYH4. These genes predicted COAD patient survival and modulated the tumor microenvironment, drug sensitivity and immune microenvironment. Our study reveals the importance of disulfidptosis-related genes for COAD prognosis and therapy. Immune infiltration and drug susceptibility results provide important clues for finding new personalized treatment options for COAD. These findings may facilitate personalized cancer treatment.

Colon adenocarcinoma (COAD) is a type of cancer that affects the cells that line the colon. It accounts for most cases of colon cancer¹. The best treatment for COAD depends on how advanced the cancer is, how healthy the patient is and what they prefer. COAD can be cured with surgery and other treatments if it is found early². But it can come back or resist drugs and make treatment harder³. COAD has become more common and deadly cancer in China in recent years⁴⁻⁶. Studying the molecular features of COAD can help doctors predict how it will behave and choose the best treatment for each patient⁷. It is important to educate people about COAD and how to prevent it.

Disulfidptosis is a new way of cell death that happens when cells have too many disulfide bonds in their proteins⁸. Disulfide bonds are links that hold proteins together, but too many of them can make proteins work poorly or break down. Disulfidptosis is caused by low glucose in cells that make a lot of SLC7A11, a protein that helps cells get cystine and make glutathione⁹. Cystine and glutathione are important for keeping cells healthy and balanced^{10,11}. Low glucose makes SLC7A11 stop working and mess up the cell balance, causing oxidative stress and more disulfide bonds^{12,13}. Disulfidptosis mainly affects the actin cytoskeleton, which is a structure

¹Department of General Surgery, The Second Xiangya Hospital of Central South University, Changsha 410011, China. ²Clinical Nursing Teaching and Research Section, The Second Xiangya Hospital Central South University, Changsha 410011, China. ³Department of Orthopedics, The Second Xiangya Hospital of Central South University, Changsha 410011, China. ⁴Hunan Key Laboratory of Tumor Models and Individualized Medicine of Hunan Province, Changsha 410011, China. ✉email: guozhushu@csu.edu.cn

made of protein threads that gives cells shape and strength^{14,15}. Disulfide bonds make the actin threads fall apart and clump together, killing the cell. Disulfidptosis is different from other types of cell death because it involves disulfide proteins and actin collapse. Disulfidptosis can influence cancer prognosis and treatment efficacy by modulating tumor microenvironment, drug sensitivity and immune microenvironment.

Prognostic biomarkers are indicators of the future outcome of a disease, such as survival, recurrence, or treatment response. They include genetic, epigenetic, transcriptomic, proteomic, and imaging features. However, many existing biomarkers are inaccurate, inconsistent, non-standardized, and poorly validated in clinics. Therefore, more research is needed to find and verify new and reliable biomarkers for different diseases and populations. Our study focuses on biomarkers that reflect the overall cancer outcome regardless of therapy. They can help diagnose and monitor tumors, and guide personalized treatment strategies.

We got the COAD sample data from the TCGA database. We then used previous studies to find genes related to disulfidptosis. We built and tested a model to predict these genes in COAD using the GEO database. We also looked at immune cells and drug responses in COAD. The goal of this study is to make prognostic markers for COAD and explore how they work and what they mean for COAD patients. We also want to suggest new ways to treat COAD.

Materials and methods

Data collection and preparation. We collected and prepared the data as follows. We downloaded the mRNA-sequencing data and clinical information of 426 COAD patients from The Cancer Genome Atlas Program (TCGA) database (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>, accessed on 28 February 2023). We extracted mRNA expression data for COAD patients from TCGA using R. The data contained 60,660 genes in total. After averaging and removing duplicates based on gene names, we curated 59,427 genes with expression values in Transcripts Per Kilobase per Million (TPM). We also curated the survival data of the patients, including progression-free survival (PFS), overall survival (OS), T stage, N stage and M stage, from the UCSC Xena (downloaded from: <https://xenabrowser.net/datapages/>). We only included COAD patients with survival information in this study. Moreover, we obtained the microarray data GSE39582 (GSE39582 is a dataset of gene expression and clinical data for 566 colon cancer and 19 normal samples from France that was used to identify six molecular subtypes of colon cancer) based on the same platform GPL570 and the corresponding clinical information from Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>, accessed on 1 March 2023). Based on previous literature⁸, we identified potential disulfidptosis-related genes (DRGs). See Supplementary Table S1 for a detailed list of genes.

Functional enrichment analysis. We used the “org.Hs.eg.db” R package to convert the symbol ID of each DRG to Entrez Gene ID. Then, we used the “cluster-Profiler” R package¹⁶ to perform Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis^{17,18} on the DEGs. We used the “ggplot2” and “ComplexHeatmap” packages¹⁹ to display the results.

Filter of the four OS-related genes. We used the R package “forestplot” to select all DRGs based on univariate cox analysis in the total dataset. Then, we used the “riskRegression” package to screen 20 potential prognostic genes. Finally, we used multivariate Cox regression to include only 4 genes in the risk signature.

Different tissue expression and single gene K–M analysis. We used the R package “ggplot2” to plot box plots to compare the expression of DRGs in different datasets. We used the R package “ggsurvplot” to plot the Kaplan–Meier survival curve to compare the OS between high-expression and low-expression groups for each of the four genes. We obtained the immunohistochemistry (IHC) data from two patients in the Human Protein Atlas (HPA) database. HPA analysis is a method of studying how proteins work and where they are in human cells and tissues using data from the Human Protein Atlas database. Supplementary Table S2 shows their basic information.

Construction and validation of a prognostic disulfidptosis-related gene signature. We randomly selected 300 samples from the 426 COAD samples in the TCGA-COAD dataset as the train cohort. The other 126 samples were the test cohort. Table 1 shows the baseline clinical characteristics of both cohorts. We used the “glmnet”²⁰ function in the R package to use the LASSO algorithm to shrink variables and make specific regression coefficients zero. This gave us a clear model. We used tenfold cross-validation to choose the penalty parameter of the prognosis model based on minimum criteria. Then, we calculated the risk score for each patient as follows: risk score = coefficient Σ (Genei) \times expression(Genei), where Genei is the expression of genes and coefficient is the coefficient of multivariate Cox regression. We split all samples into low-risk and high-risk groups based on the median risk score of samples from the train set. We used Kaplan–Meier analysis to compare the OS differences between low- and high-risk groups. We used the “timeROC” package to check how accurate our model predictions were. Then, we used the test set to verify how reliable our model was.

Construction and validation of a nomogram. We used the “rms” R package and the “survival” package to build a nomogram to predict OS. We used the “pROC”²¹ package to do time-dependent ROC analysis to check how well nomogram predicted. We used calibration curves to check how close the actual survival rates and predicted survival rates were.

	Train set	Test set
Cases	300	126
Age (median, range)	68 (31–90)	70.5 (37–90)
Gender		
Female	138 (46.00%)	59 (46.83%)
Male	162 (54.00%)	67 (53.17%)
T stage		
T1	6 (2.00%)	5 (3.97%)
T2	51 (17.00%)	23 (18.25%)
T3	207 (69.00%)	85 (67.46%)
T4	36 (12.00%)	13 (10.32%)
N stage		
N0	178 (59.33%)	73 (57.94%)
N1	68 (22.67%)	33 (26.19%)
N2	54 (18.00%)	20 (15.87%)
M stage		
M0	256 (85.33%)	109 (86.51%)
M1	44 (14.67%)	17 (13.49%)
Survival status		
Alive	241 (80.33%)	93 (73.81%)
Dead	59 (19.67%)	33 (26.19%)

Table 1. Clinical characteristics of the TCGA-COAD (train set and test set).

Independent datasets validation. We used survival curves from the validation set (GSE39582) to check how reliable our group risk assessment model was. We only included those with OS.time \leq 80 in the validation set. Table 2 shows the main clinical features of GSE39582. We also predicted the 1-, 2-, 3-, and 5-year survival rates of COAD patients from the validation sets.

Immune analysis. We split the whole cohort into two groups of high_risk and low_risk based on the average risk signature point. We used Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORTx) (<https://cibersortx.stanford.edu>) to analyze how 22 types of immune cells infiltrated the high_risk and low_risk groups. We used the “estimate” R package to do Estima of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) analysis to measure the stroma and immune scores. We also used the TIMER database^{22,23} for model gene analysis in COAD.

Drug sensitivity analysis. We used the “OncoPredict”²⁴ R package to find susceptibility data in the GDSC2 database and to predict how COAD would respond to drugs in the GDSC2 database. We used Spearman correlation analysis to find drugs related to risk outcomes and used the “ggstatsplot” package to plot the correlation scatterplot.

Statistical analysis. We used R 4.2.2 for all statistical analyses. We used the Wilcoxon test to compare non-parametric data from two independent samples. We used t-test and one-way ANOVA to analyze parametric data. We considered a p value $<$; 0.05 as statistically significant (p-value $<$ 0.05; *p-value $<$ 0.01; ***p-value $<$ 0.001). We used related R packages such as ggplot2, ggpubr, survival, sur-vminer, and others from the Bioconductor package or the R package. We considered a p value $<$ 0.05 as statistically significant unless we said otherwise.

Results

GO and KEGG analysis. We made a list of 105 genes (Supplementary Table S1) that work closely with Disulfidptosis. Figure 1 shows our study. Figure 2a shows the heatmap of disulfidptosis-related genes in tumor and normal tissues in the TCGA cohort. GO analysis showed that these 107 DRGs were mostly involved in biological processes like muscle filament sliding, actin-myosin filament sliding and actin filament-based movement, cellular components like actin cytoskeleton myosin II complex and myosin filament, and molecular functions like actin filament binding, actin binding and microfilament motor activity (Fig. 2b–d) (Supplementary Table S2). The KEGG pathway enrichment analysis showed that these DRGs were involved in pathways like amyotrophic lateral sclerosis, tight junction, diabetic cardiomyopathy, chemical carcinogenesis-reactive oxygen species and oxidative phosphorylation (Fig. 2e,f) (Supplementary Table S3).

Filter disulfidptosis-related genes. We compared the gene expression levels between normal and tumor tissues of COAD patients from TCGA and found that OXSM was significantly downregulated ($p = 5.20 \times 10^{-9}$), while TRIP6 ($p = 5.00 \times 10^{-12}$), MYH3 ($p = 6.70 \times 10^{-3}$), and MYH4 ($p = 6.41 \times 10^{-7}$) were upregulated in COAD tissues (Fig. 3e,f). These genes were selected for the forest plot of univariate Cox regression analysis (Fig. 3a,b).

GSE39582	
Cases	394
Age (median, range)	69 (22–97)
Gender	
Female	170 (43.15%)
Male	224 (56.85%)
T stage	
T1	7 (1.78%)
T2	35 (8.88%)
T3	255 (64.72%)
T4	97 (24.62%)
N stage	
N0	207 (52.54%)
N1	99 (25.13%)
N2	88 (22.34%)
M stage	
M0	337 (85.53%)
M1	57 (14.47%)
MMR status	
pMMR	310 (78.68%)
dMMR	60 (15.23%)
NA	24 (6.09%)
Survival status	
Alive	235 (59.64%)
Dead	159 (40.36%)

Table 2. Clinical characteristics of the GSE39582 (OS.time ≤ 80).

Then, we performed stepwise backward regression on the results separately, using two different sets of variables: one screened by univariate Cox ($p < 0.05$) and another by all-subsets regression (BSR). Four genes were finally chosen for the forest plot of multicox regression analysis after comparing the two methods (Fig. 3c,d).

We also assessed the prognostic value of these genes in COAD. All of them were strongly associated with overall survival (OS). The hazards ratios of TRIP6 (HR 1.36 (1.10–1.68), $p = 0.006$), OXSM (HR 0.57 (0.38–0.87), $p = 0.009$), MYH3 (HR 1.50 (1.01–2.23), $p = 0.029$), and MYH4 (HR 0.70 (0.51–0.96), $p = 0.048$) were statistically significant (Fig. 4a–d). Moreover, we searched the HPA database for immunohistochemical staining data of TRIP6 and MYH3 in tumor tissues and observed higher protein expression of these genes in analyzed tumor tissues (Fig. 4e,f). In addition, we investigated the correlation between the expression of different genes and found some associations (Fig. 4g).

Establishment of the four-gene risk signature. We divided 426 cases randomly into a train set and a test set (Supplementary Table S6). Table 1 shows their clinical features. We used Lasso algorithm to select four DRGs as candidate genes (Fig. 5a,b). We generated risk scores based on these genes to predict survival and prognosis in COAD patients:

$$\text{Score} = 0.30291 \times \text{TRIP6} - 0.50344 \times \text{OXSM} + 0.39184 \times \text{MYH3} + 0.40015 \times \text{MYH4}.$$

KM analysis revealed that higher risk scores were associated with poorer overall survival in both the train set (HR 2.75 (1.59–4.74), log-rank $p = 2.86 \times 10^{-4}$) and the test set (HR 2.91 (1.35–6.26), log-rank $p = 6.38 \times 10^{-3}$) (Fig. 5c,d, Supplementary Table S7). We assessed the accuracy of the model predictions using the “timeROC” package (Fig. 5e,f). The results confirmed that higher risk scores indicated worse prognosis and survival time (Fig. 5g,h). We obtained clinical information tables and risk scores for the train set (Supplementary Table S8) and the test set (Supplementary Table S9).

Building and validating a predictive nomogram. We plotted the expression of model-genes in the train set and test set (Fig. 6a,b). We used risk score, age, T_stage, N_stage and M_stage to build a nomogram for predicting OS (Fig. 6c). The nomogram had a C-index of 0.739 and a good calibration curve (Fig. 6d). The AUCs of the nomogram for 1-year, 2-year, 3-year, and 5-year OS were 0.779, 0.696, 0.743 and 0.700 in train set and 0.859, 0.892, 0.892 and 0.895 in test set (Fig. 6e,f). The decision curves for 1-year OS, 2-year OS, 3-year OS and 5-year OS showed that the nomogram model had a high net benefit for predicting OS of CO-AD patients (Fig. 6g,h). This section should have subheadings to describe the results, their interpretation, and the conclusions.

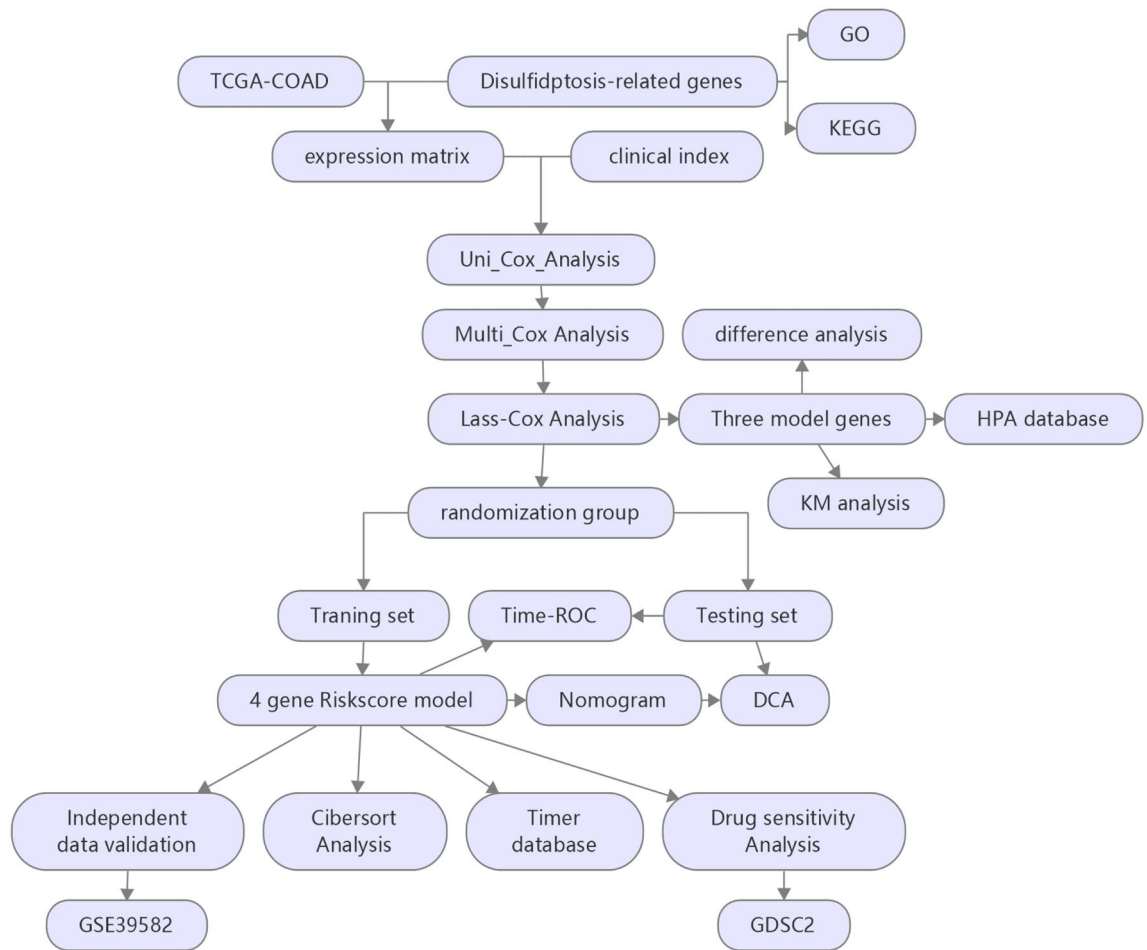


Figure 1. The study exploration workflow.

Validation of the DRGs signature model based on GEO dataset. We used GSE39582 as the validation data set and selected cases with OS.time ≤ 80 (Table 2, Supplementary Table S10). Figure 7a shows a heatmap of the gene expression distribution of 394 COAD samples. KM analyses indicated that higher risk scores were associated with poorer overall survival in the validation set (Fig. 7b). We compared the expression of model genes in tumor and normal tissues (Fig. 7c). We also displayed the survival status and risk score distributions of these COAD samples using scatterplots and histograms (Fig. 7d). The AUCs for the 1-year, 2-year, 3-year, and 5-year ROC curves were 0.567, 0.576, 0.574, and 0.556, respectively (Fig. 7e).

Immune analysis. To quantify the immune infiltration of tumor samples, we used CIBERSORT, a computational method that estimates the relative abundance of immune cell types from gene expression data. The tumor microenvironment affects COAD occurrence and development, and T cells. Memory CD4 cells, macrophages M0, and plasma cells are the most common tumor-infiltrating immune cells (Fig. 8a,b). We found significant differences in the frequencies of 3 types of immune cells between the favorable and poor prognosis groups (Fig. 8c,d). The ratio of mesenchymal cells to immune cells might differ between the two groups, possibly affecting tumor purity. These results suggest that immune infiltration and immune microenvironment are important for OS in COAD patients.

We investigated whether DRGs would influence immune cell recruitment in the tumor microenvironment and affect COAD prognosis. We examined the relationships between TRIP6, OXSM, MYH3 and MYH4 and immune infiltration in COAD. TRIP6 expression was positively associated with CD4+ T cells ($p = 1.48 \times 10^{-2}$) and negatively correlated with Neutrophil ($p = 9.32 \times 10^{-3}$) (Fig. 8e). MYH4 expression was negatively correlated with Dendritic cells ($p = 7.46 \times 10^{-4}$) (Fig. 8f). MYH3 expression was positively associated with CD4+ T cells ($p = 4.51 \times 10^{-16}$), Macrophage (1.69×10^{-4}) and Neutrophil ($p = 4.86 \times 10^{-5}$) (Fig. 8g). OXSM expression was positively associated with Putity ($p = 1.81 \times 10^{-2}$), B cells ($p = 6.71 \times 10^{-3}$) and CD8+ T Cells ($p = 5.89 \times 10^{-3}$) (Fig. 8h). These results suggest that immune parameters can be used as potential biomarkers or therapeutic targets for cancer patients.

Drug sensitivity. We used the drug sensitivity data from GDSC2 to validate our analytical model and identified some drugs that were associated with drug sensitivity. We assessed the risk scores' value in predicting drug

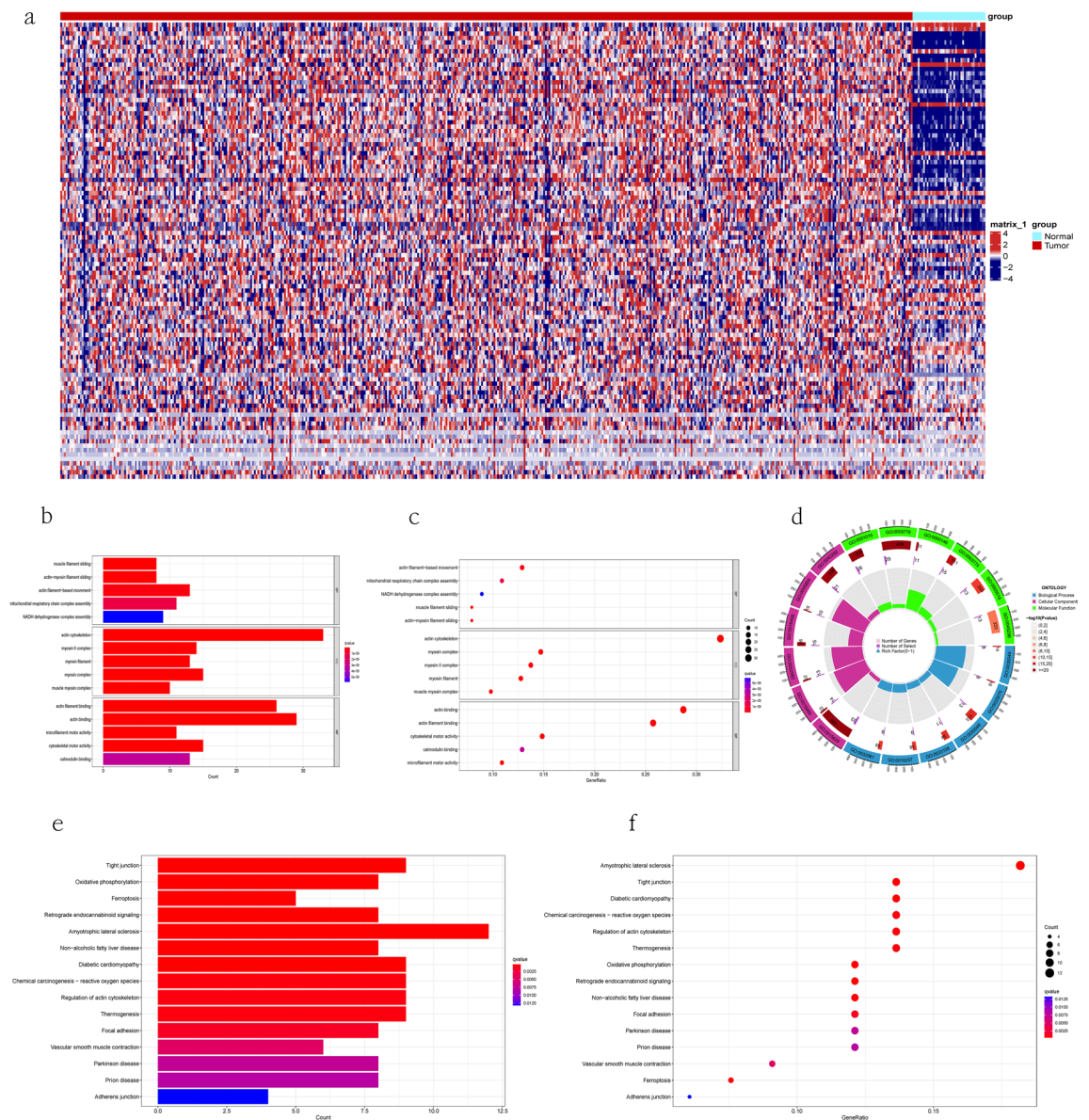


Figure 2. Analysis of disulfidptosis-related genes with TCGA (a) Heatmap of DRGs identified in integrated microarray (tumor in red and normal in blue), (b–d) Results of GO analysis with DRGs, (e, f) Results of KEGG analysis with DRGs.

sensitivity in different cancer types. We selected 30 drugs from the GDSC2 database that had significant Spearman correlation between risk scores and drug sensitivity. The risk score was positively correlated with sensitivity to 30 drugs, such as Borte-zomib, 5-Fluorouracil, Cytarabine, Fludarabine, and others (Fig. 9a,b,f–h). The risk score can help us choose appropriate and effective treatment strategies. The IC50 of three drugs was lower in the high-risk group than low-risk group (Fig. 9c–e). These three drugs were Borte-zomib, 5-Fluorouracil, Cytarabine, which could be potential drugs for COAD treatment.

Discussion

Disulfidptosis happens when cells face disulfide stress, which can be caused by lack of glucose or oxidative stress. Disulfide stress makes disulfide bonds accumulate in actin cytoskeleton proteins, which are important for cell shape and survival. The excess disulfide bonds make actin filaments collapse and aggregate, leading to cell death. Disulfidptosis is different from other cell death types, such as apoptosis, necrosis, autophagy or ferroptosis, because it has a specific target and a specific trigger. Disulfidptosis can be detected by using a fluorescent probe called DCP-Bio1 that binds to disulfide bonds. DCP-Bio1 can label both inside and outside disulfides and can be used to watch disulfidptosis in live cells or tissues. Disulfidptosis can also be stopped by using antioxidants or reducing agents that stop or undo disulfide bond formation. The finding of disulfidptosis has implications for understanding how cell death and survival work in different normal and abnormal conditions. For example, disulfidptosis may have a role in cancer growth and therapy resistance, as some cancer cells have high levels of

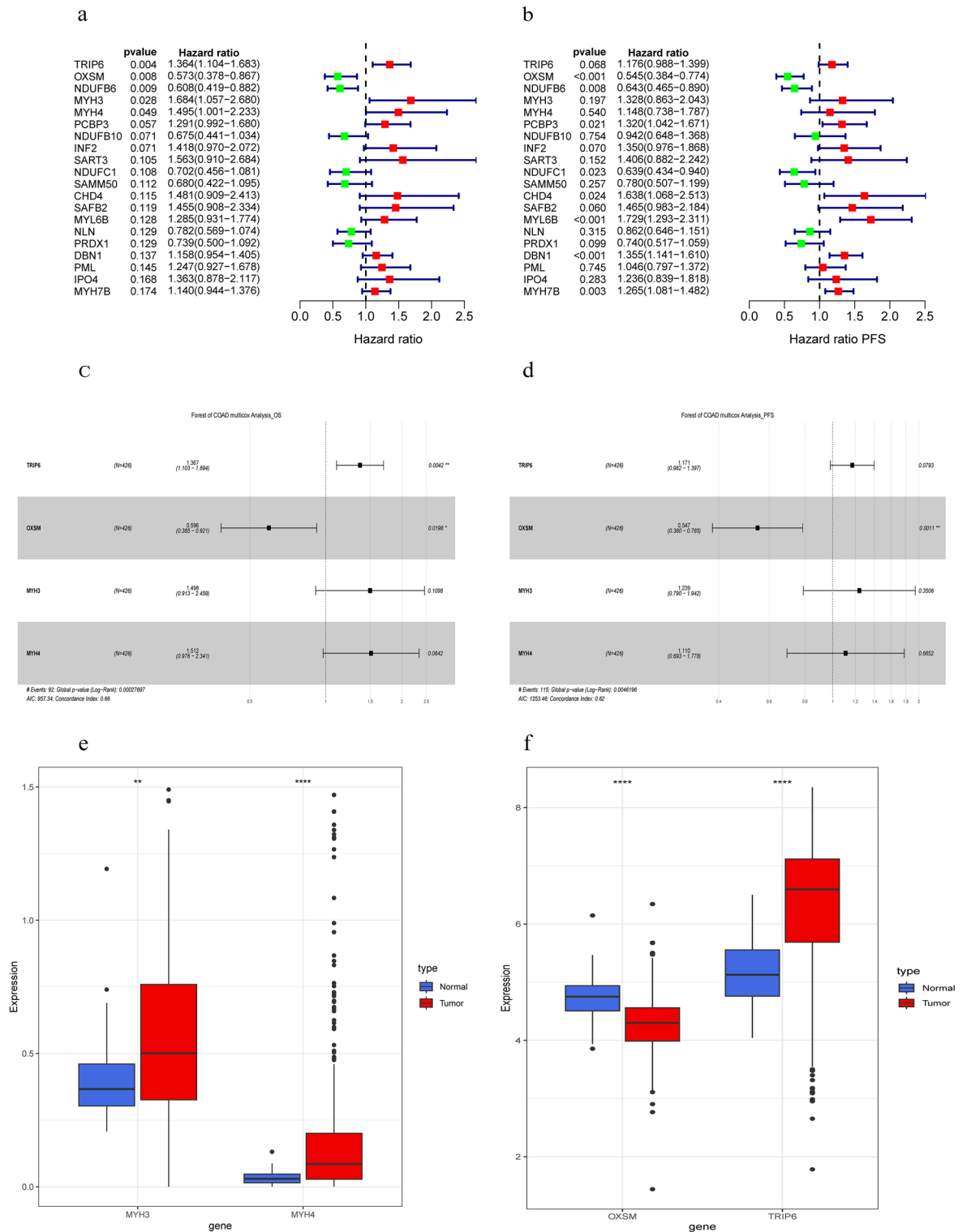


Figure 3. Forest plots DRGs from univariate and multivariate Cox proportional hazards mode. (a) The effect of 20 DRGs on the overall survival of TCGA datasets, (b) The effect of 20 DRGs on the progress free survival of TCGA datasets, For (c) OS and (d) PFS, hazard ratios and p-value of the constituents involved in multivariate Cox regression in COAD. (e) The expression of MYH3 and MYH4 in COAD and normal tissues (tumor in red and normal in blue). (f) The expression of TRIP6 and OXSM in COAD and normal tissues.

disulfide stress and may use disulfidptosis to survive. In fact, Liu et al. did a genetic analysis of 33 kinds of cancers and found that genes related to disulphidesoptosis were often mutated or changed in different cancer types. These genetic changes may affect how cancer cells respond to agents that induce disulphidesoptosis. Therefore, targeting

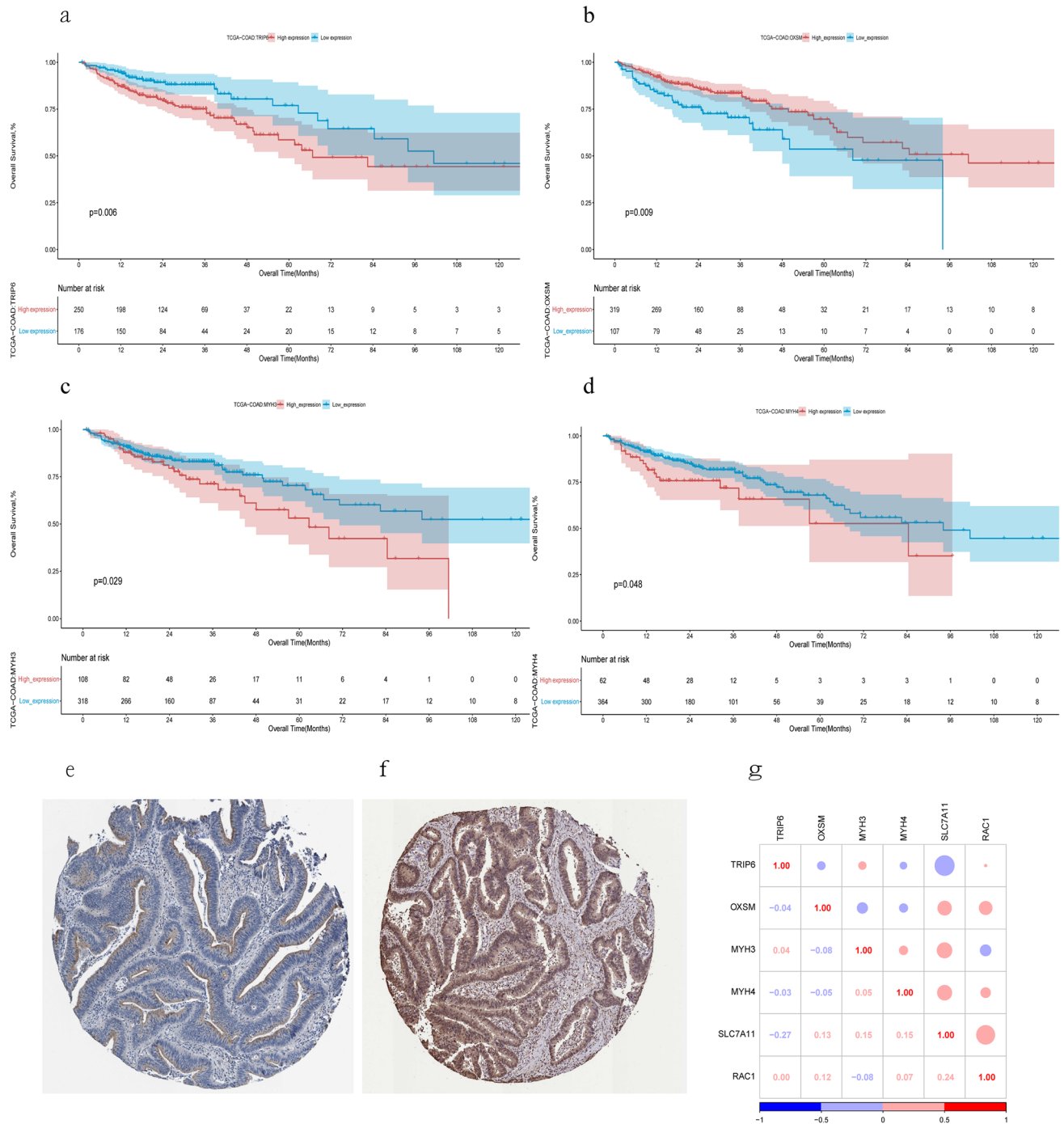


Figure 4. Clinical relevance of DRGs in the COAD patients of TCGA. For OS outcome, Kaplan Meier plot for the expression of (a) TRIP6 (b) OXSM (c) MYH3 and (d) MYH4 and overall survival. Representative IHC images of MYH3 (e) and TRIP6 (f) in COAD tissues. The data was retrieved from HPA database. (g) Correlations between the expression of DRGs.

di-sulphidesoptosis may be a potential way to treat cancer. In summary, disulphides are important for controlling cell fate under stress conditions. Disulphides can cause a new form of cell death called disulphidesoptosis by affecting the actin cytoskeleton structure. Disulphidesoptosis has potential implications for various diseases that involve cell death or survival mechanisms. More research on this new pathway may give new insights into the biology of cell death and offer new therapeutic options.

TRIP6 is frequently overexpressed in COAD tissues and cell lines compared to normal tissues that is associated with poor prognosis, advanced tumor stage, lymph node metastasis and distant metastasis in COAD patients²⁵. The molecular mechanisms by which TRIP6 promotes COAD tumorigenesis have been investigated by several studies. One study found that TRIP6 can be stabilized by another gene called TTPAL, which is preferentially

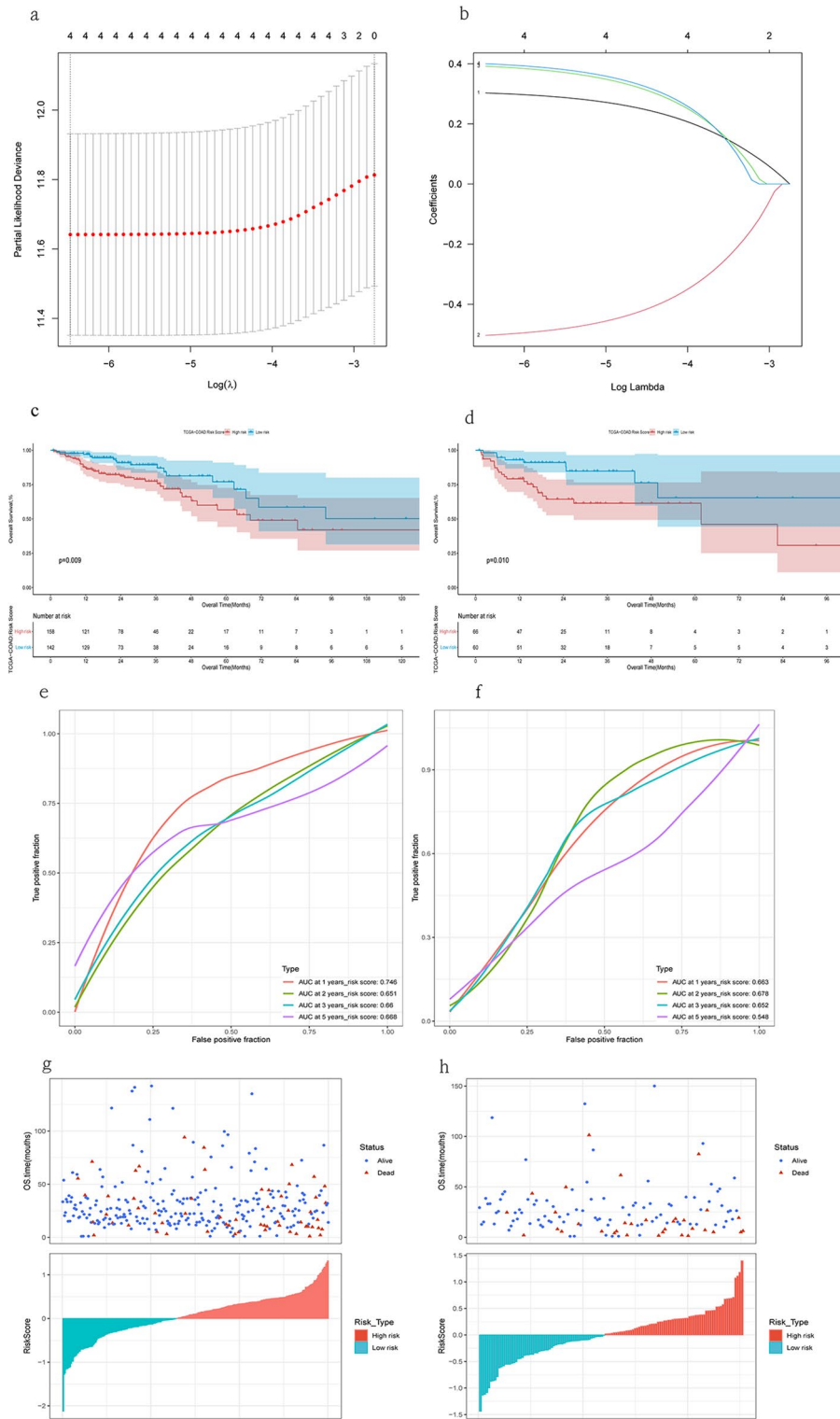


Figure 5. Construction and evaluation of a risk score model. (a,b) DRGs associated with prognosis were selected using the least absolute shrinkage and selection operator (LASSO) method. Kaplan–Meier plots of the risk score and overall survival were generated for the train set (c) and test set (d). ROC curves of the risk score were generated for the train set (e) and test set (f). The distribution of risk scores and overall survival status was shown for the train set (g) and test set (h).

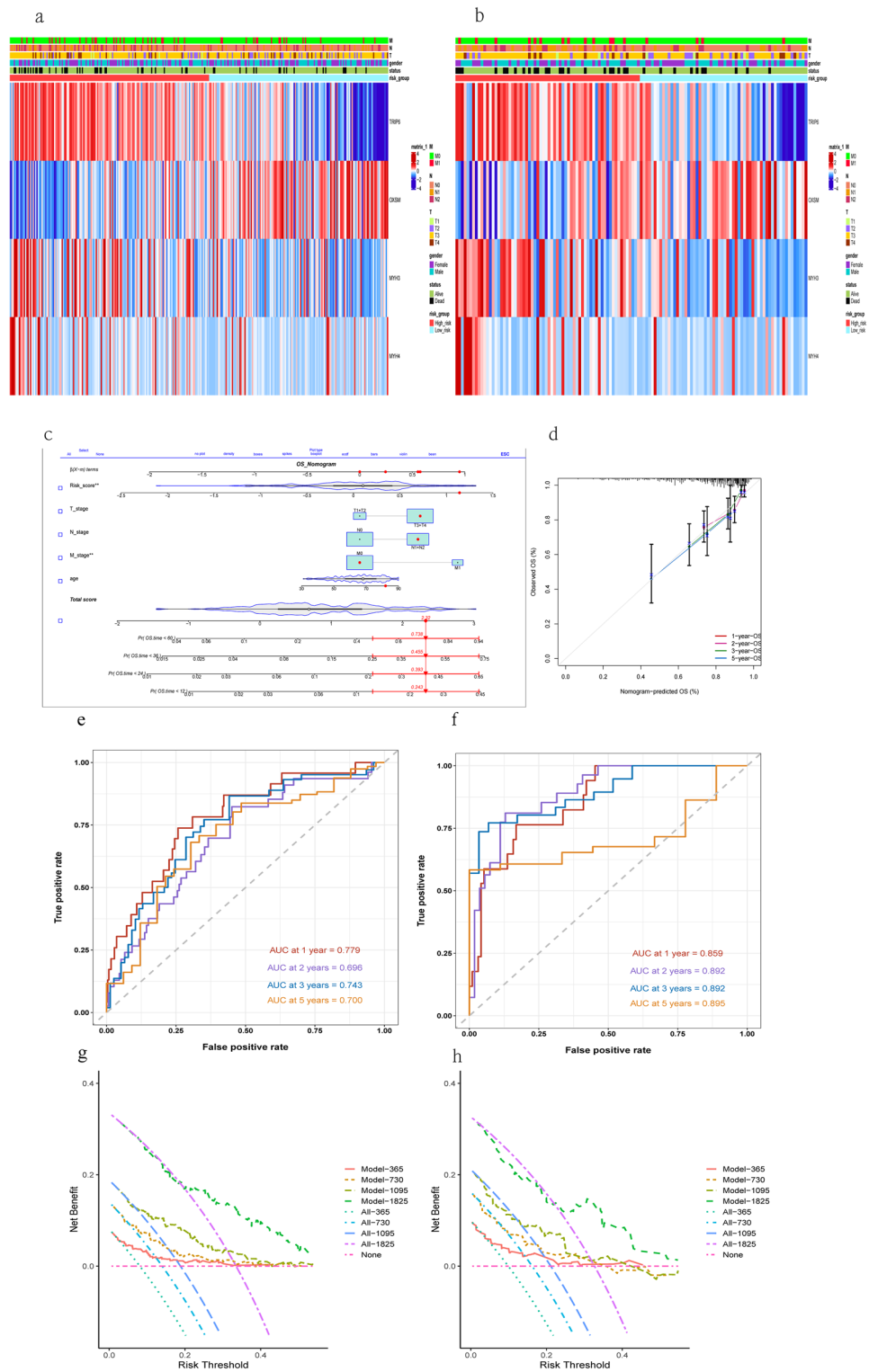


Figure 6. Nomogram development, evaluation and validation. (a,b) Heatmap of model-genes identified in the train set and test set. (c) Nomogram was developed based on the risk scores and different clinical features. (d) Calibration plots were constructed to evaluate the predictive performance of overall survival. The ROC curve of the nomogram in the train cohort (e) and test cohort (f). The decision curves of the nomogram in the train set (g) and test set (h).

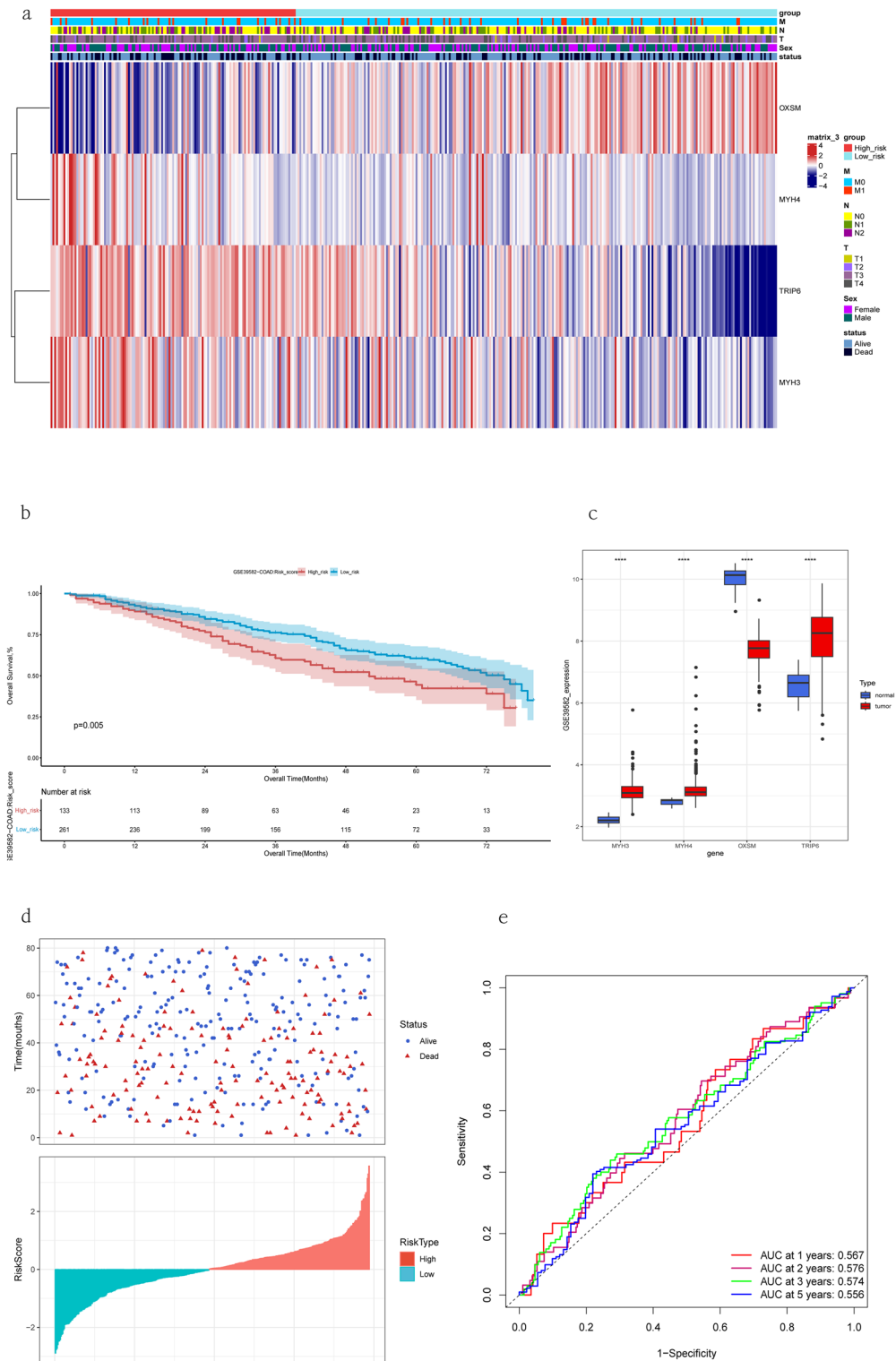


Figure 7. Validation of the seven-gene signature. **(a)** Heatmap of model-genes identified in the validation set. **(b)** Kaplan–Meier plot of risk score and overall survival in the validation set. **(c)** Expression of four genes in COAD and normal tissues in the validation set. **(d)** Distribution of risk score and overall survival status in the validation set. **(e)** ROC curve of risk score in the validation cohort.

amplified in COAD²⁶. TTPAL prevents TRIP6 from being degraded by the proteasome and enhances its interaction with β -catenin, a key mediator of the Wnt/ β -catenin pathway. The Wnt/ β -catenin pathway is known to play a crucial role in colorectal cancer initiation and progression by regulating genes involved in cell growth, invasion

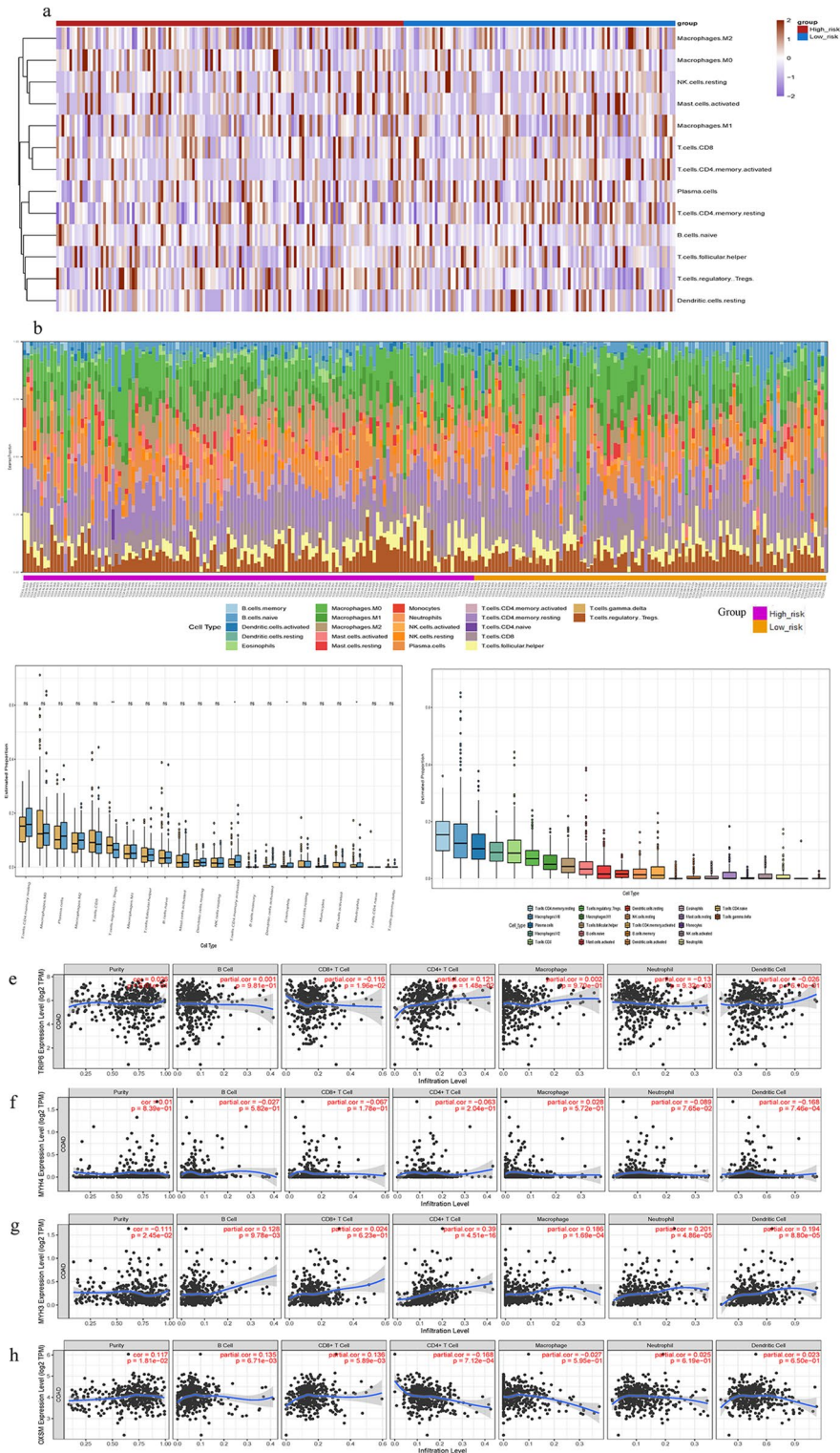


Figure 8. Immune infiltration analysis. **(a)** Heatmap of various immune cell types. **(b)** Infiltration of 22 different immune cell types. **(c,d)** Immune cell component between high-risk and low-risk groups. Correlation between **(e)** TRIP6, **(f)** MYH4, **(g)** MYH3, and **(h)** OXSM expression and immune infiltration in COAD in the TIMER database.

and stemness. Therefore, TTPAL-TRIP6- β -catenin axis can activate the Wnt/ β -catenin pathway and promote COAD. 3-oxoacyl-ACP synthase, mitochondrial (OXSM), also known as FASN2D, is a protein that is involved in the elongation of fatty acid chains in mitochondria²⁷. The cis-regulation of enhancers on target genes depends

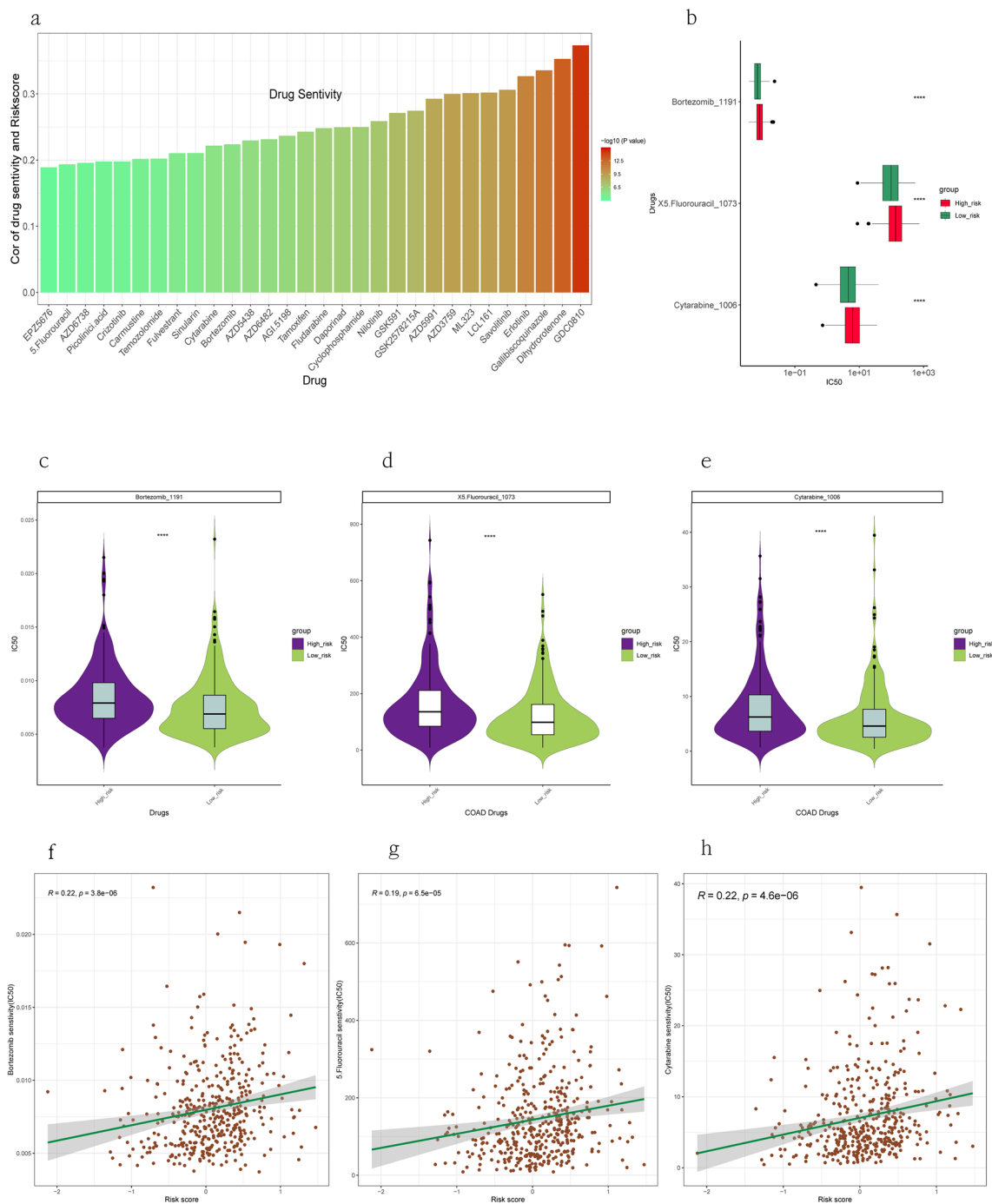


Figure 9. Drug response prediction. **(a)** Comparison of 30 chemotherapy drugs for COAD on the Genomics of Drug Sensitivity in Cancer (GDSC2) database. **(b)** Assessment of drug sensitivity of COAD tumor based on the risk score. **(c–e)** Boxplot visualizing different drugs. **(f–h)** Correlation analysis between risk score and the IC50 of 3 chemotherapy drugs.

on the action of transcription factors²⁸, which are proteins that bind to specific DNA sequences and regulate gene expression. Previous studies²⁹ have identified a key transcription factor, core-binding factor subunit beta (CBFB), which is strongly correlated with OXSM expression. In colorectal cancer, CBFB deficiency has been shown to enhance cell resistance to MEK inhibitors³⁰, which are drugs that target the mitogen-activated protein kinase (MAPK) pathway. Additionally, hsa-miR-338-3p has been found to participate in the regulation of fatty acid biosynthesis by regulating OXSM levels³¹, and to affect the biological occurrence and rapid proliferation of glioma cells, which are a type of brain tumor.

The instructions for producing myosin-3, a protein that belongs to the myosin family of proteins, which are involved in cellular movement and the transportation of materials within and between cells, are provided by the MYH3 gene. THZ1, a novel covalent CDK7 inhibitor, has been utilized as an anti-tumor drug. It suppresses the

expression of differentiation-related transcription factors and muscle structural proteins, including myogenic protein myh3³². Encoded by the MYH4 gene, myosin-4 is a protein that belongs to a group of proteins involved in movement and transport within and between cells. It is also expressed in skeletal muscle fibers. A previous study³³ found that long noncoding RNA Neat1 promoted myoblast proliferation mainly by decreasing the expression of the cyclin-dependent kinase inhibitor P21 gene, but inhibited myoblast differentiation by suppressing the transcription of myogenic marker genes, such as Myh4.

We investigated how disulfidptosis, a type of cell death induced by oxidative stress, affects the genetic and transcriptional landscape of colorectal adenocarcinoma (COAD) and how these changes are linked to each other. We also developed a risk assessment model based on the expression of disulfidptosis-related genes to estimate the prognosis and treatment response of COAD patients. This model can help us reveal the molecular mechanisms and clinical implications of disulfidptosis in COAD and devise more precise and effective immunotherapy strategies. In our experiment, we applied the risk assessment model to assess the immunotherapy response of COAD patients and compared the immune cell composition in tumors with high or low risk scores. CIBERSORT is a computational method that can estimate the relative proportions of immune cell types from bulk RNA sequencing data. It has been widely used to characterize the tumor microenvironment and its impact on cancer biology and treatment. The relationship between disulfidptosis genes and immune cell infiltration is complex and context-dependent. On one hand, disulfidptosis can affect the recruitment and activation of immune cells by modulating the release of damage-associated molecular patterns (DAMPs), cytokines, chemokines, and reactive oxygen species (ROS). On the other hand, immune cells can influence the expression and function of disulfidptosis genes by altering the tumor microenvironment and providing pro- or anti-inflammatory signals. In this study, we found that different immune cell subsets had different correlations with disulfidptosis gene expression and survival outcomes. Our results show that this model can be useful for identifying immunotherapy candidates for COAD patients. GDSC2 is a database that contains drug sensitivity data and genomic features for over 1000 human cancer cell lines. It is a valuable resource for identifying molecular determinants of drug response and developing predictive biomarkers for personalized medicine. The relationship between disulfidptosis genes and drug sensitivity is complex and context-dependent. On one hand, disulfidptosis genes can affect the response and resistance to drugs by modulating the cellular redox state, protein stability, apoptosis induction, and DNA damage repair. On the other hand, drugs can influence the expression and function of disulfidptosis genes by altering the intracellular or extracellular redox environment, inducing oxidative stress or endoplasmic reticulum stress, or activating or inhibiting signaling pathways. In this study, we found that different drugs had different correlations with disulfide death gene expression and survival outcomes. We also examined the relationship between risk score and drug sensitivity in COAD patients and found that apoptosis, signal transduction and metabolism pathways are involved in COAD treatment.

We devised a risk score model with 4 disulfide death genes for survival prediction in COAD. We also uncovered the complex interplay between disulfide death genes and immune cells in the tumor microenvironment and their impact on tumor immunity. Our model has some caveats. It relied on bulk gene expression data, which may not reflect tumor heterogeneity and dynamics. It did not adjust for other confounding factors, such as treatment, comorbidities, or lifestyle. It was based on retrospective data, which may have biases and limitations. Prospective studies are warranted to validate our findings and assess our model's clinical utility. We intend to investigate the mechanisms and functions of disulfide death genes in cancer and therapy. We also aim to target or modulate disulfide death genes as a novel therapeutic strategy for cancer. We aspire to discover better biomarkers based on disulfide death gene expression and other omics data.

Conclusions

We developed and tested a COAD risk score based on the expression of four genes related to disulfidptosis. This risk score can help us predict the prognosis and immunotherapy response of COAD patients. By analyzing the risk score, we can gain more in-sight into the molecular mechanisms and clinical implications of disulfidptosis in COAD and design more personalized and precise treatment options.

Data availability

The data presented in this study are openly available in the TCGA database [<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>, (accessed on 28 February 2023)] and GEO database [<https://www.ncbi.nlm.nih.gov/geo/>, (accessed on 1 March 2023)].

Received: 1 May 2023; Accepted: 27 July 2023

Published online: 31 July 2023

References

1. Ai, D. *et al.* Regularized survival learning and cross-database analysis enabled identification of colorectal cancer prognosis-related immune genes. *Front. Genet.* **14**, 1148470 (2023).
2. Zhou, X. Y., Wang, J. Q., Chen, J. X. & Chen, J. S. The expression of PPAR pathway-related genes can better predict the prognosis of patients with colon adenocarcinoma. *PPAR Res.* **2022**, 1285083 (2022).
3. Hossain, M. S. *et al.* Colorectal cancer: A review of carcinogenesis, global epidemiology, current challenges, risk factors, preventive and treatment strategies. *Cancers (Basel)*. **14**(7), 1732 (2022).
4. Feng, R. M., Zong, Y. N., Cao, S. M. & Xu, R. H. Current cancer situation in China: Good or bad news from the 2018 Global Cancer Statistics?. *Cancer Commun. (Lond)*. **39**(1), 22 (2019).
5. Chen, W. *et al.* Cancer statistics in China, 2015. *CA Cancer J. Clin.* **66**(2), 115–132 (2016).
6. Xia, C. *et al.* Cancer statistics in China and United States, 2022: Profiles, trends, and determinants. *Chin. Med. J. (Engl)*. **135**(5), 584–590 (2022).

7. Van Cutsem, E. *et al.* ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann. Oncol.* **27**(8), 1386–1422 (2016).
8. Liu, X. *et al.* Actin cytoskeleton vulnerability to disulfide stress mediates disulfidptosis. *Nat. Cell Biol.* **25**(3), 404–414 (2023).
9. Jyotsana, N., Ta, K. T. & DelGiorno, K. E. The role of cystine/glutamate antiporter SLC7A11/xCT in the pathophysiology of cancer. *Front. Oncol.* **12**, 858462 (2022).
10. Yu, X. & Long, Y. C. Crosstalk between cystine and glutathione is critical for the regulation of amino acid signaling pathways and ferroptosis. *Sci. Rep.* **6**, 30033 (2016).
11. Wu, G., Fang, Y. Z., Yang, S., Lupton, J. R. & Turner, N. D. Glutathione metabolism and its implications for health. *J. Nutr.* **134**(3), 489–492 (2004).
12. Koppula, P., Zhuang, L. & Gan, B. Cystine transporter SLC7A11/xCT in cancer: Ferroptosis, nutrient dependency, and cancer therapy. *Protein Cell.* **12**(8), 599–620 (2021).
13. Koppula, P., Zhang, Y., Zhuang, L. & Gan, B. Amino acid transporter SLC7A11/xCT at the crossroads of regulating redox homeostasis and nutrient dependency of cancer. *Cancer Commun. (Lond.)* **38**(1), 12 (2018).
14. Xu, K., Zhong, G. & Zhuang, X. Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons. *Science* **339**(6118), 452–456 (2013).
15. Radja, A., Horsley, E. M., Lavrentovich, M. O. & Sweeney, A. M. Pollen cell wall patterns form from modulated phases. *Cell* **176**(4), 856–68 e10 (2019).
16. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb.)* **2**(3), 100141 (2021).
17. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
18. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**(D1), D457–D462 (2016).
19. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**(18), 2847–2849 (2016).
20. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**(1), 1–22 (2010).
21. Robin, X. *et al.* pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
22. Aran, D. & Butte, A. J. Digitally deconvolving the tumor microenvironment. *Genome Biol.* **17**(1), 175 (2016).
23. Li, T. *et al.* TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* **77**(21), e108–e110 (2017).
24. Maeser, D., Gruener, R. F., Huang, R. S. oncoPredict: An R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief Bioinform.* **22**(6) (2021).
25. Wang, H. *et al.* Identification of gene modules and hub genes in colon adenocarcinoma associated with pathological stage based on WGCNA analysis. *Cancer Genet.* **242**, 1–7 (2020).
26. Gou, H. *et al.* TTPAL promotes colorectal tumorigenesis by stabilizing TRIP6 to activate Wnt/beta-catenin signaling. *Cancer Res.* **79**(13), 3332–3346 (2019).
27. Zhang, L., Joshi, A. K., Hofmann, J., Schweizer, E. & Smith, S. Cloning, expression, and characterization of the human mitochondrial beta-ketoacyl synthase. Complementation of the yeast CEM1 knock-out strain. *J. Biol. Chem.* **280**(13), 12422–12429 (2005).
28. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: Transcriptional enhancers in development and evolution. *Cell* **167**(5), 1170–1187 (2016).
29. Zhang, L., Zhao, S., Liu, Y., Lv, F. & Geng, X. Identification and validation of transcription factor-driven enhancers of genes related to lipid metabolism in metastatic oral squamous cell carcinomas. *BMC Oral Health* **22**(1), 126 (2022).
30. Sustic, T., Bosdriesz, E., van Wageningen, S., Wessels, L. F. A. & Bernards, R. RUNX2/CBFB modulates the response to MEK inhibitors through activation of receptor tyrosine kinases in KRAS-mutant colorectal cancer. *Transl. Oncol.* **13**(2), 201–211 (2020).
31. Wang, W. Y. & Lu, W. C. Reduced expression of hsa-miR-338-3p contributes to the development of glioma cells by targeting mitochondrial 3-oxoacyl-ACP synthase (OXSM) in glioblastoma (GBM). *Onco Targets Ther.* **13**, 9513–9523 (2020).
32. Ma, X. *et al.* Covalent CDK7 inhibitor THZ1 inhibits myogenic differentiation. *J. Cancer* **9**(17), 3149–3155 (2018).
33. Wang, S. *et al.* Long noncoding RNA Neat1 modulates myogenesis by recruiting Ezh2. *Cell Death Dis.* **10**(7), 505 (2019).

Author contributions

Conceptualization, G.H. and Z.G.; software, G.H.; resources, G.H. and Z.W.; data curation, L.L.; writing—original draft preparation, Z.Y.; writing—review and editing, J.L.-X.L., and L.L.; visualization, G.H. and H.Y.; supervision, H.Y.; all authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-39563-y>.

Correspondence and requests for materials should be addressed to Z.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023