


RESEARCH PAPER



Identification of endoplasmic reticulum stress-related genes as prognostic markers in colon cancer

Wenjing Xu, Wei Li, Dayu Kuai, Yaqiang Li, Wei Sun, Xian Liu, and Baohong Xu 

Department of Gastroenterology, Beijing Luhe Hospital Affiliated to Capital Medical University, Beijing, China

ABSTRACT

Endoplasmic reticulum stress (ERS) has been implicated in the pathogenesis of various cancers, including colon cancer, by regulating tumor cell survival, growth, and immune response. However, the specific genes involved in ERS that could serve as prognostic markers in colon cancer remain underexplored. This study aims to identify and validate endoplasmic reticulum stress related genes (ERSRGs) in colon cancer that correlate with patient prognosis, thereby enhancing the understanding of ERS in oncological outcomes and potential therapeutic targeting. We utilized bioinformatics analyses to identify ERSRGs from publicly available colon cancer datasets. Differential expression analysis and survival analysis were performed to assess the prognostic significance of these genes. Validation was conducted through quantitative real-time PCR (RT-qPCR) on selected colon cancer cell lines. Our study identified nine ERS related genes (ASNS, ATF4, ATF6B, BOK, CLU, DDIT3, MANF, SLC39A14, TRAF2) involved in critical pathways including IL-12, PI3K-AKT, IL-7, and IL-23 signaling, and linked to 1-, 3-, and 5-year survival of patients with colon cancer. A multivariate Cox model based on these ERS related genes demonstrated significant prognostic power. Further, TRAF2 strongly correlated with immune cells infiltration, suggesting its potential roles in modulating immune responses in the tumor microenvironment. The RT-qPCR validation confirmed the differential expression of these genes in human colon cancer cell lines versus human normal colonic epithelial cell line. The identified ERSRGs could serve as valuable prognostic markers and may offer new insights into the therapeutic targeting of ERS in colon cancer.

ARTICLE HISTORY

Received 26 March 2024
Revised 20 December 2024
Accepted 22 January 2025

KEYWORDS

Colon adenocarcinoma;
endoplasmic reticulum
stress; bioinformatics
analysis; prognostic model



Introduction


Colon adenocarcinoma (COAD) is a prevalent malignancy in the digestive realm.¹ According to data from the International Cancer Research Center in 2020, COAD is the third most common type of cancer worldwide in terms of incidence and ranks second in mortality rates.² Due to population aging and shifts in dietary habits, COAD is poised to surpass gastric and liver cancers, potentially becoming the most prevalent digestive tract malignancy in some countries.³ The prognosis of COAD is influenced by factors such as lymph node involvement, tumor invasion depth, histological type, and vascular infiltration.⁴ Despite surgical resection being standard, recurrence rates remain high, and chemotherapy often results in adverse effects like granulocytopenia.⁵ The overall 5-year survival rate for COAD ranges from 12.5% to 50%, with common metastasis to the liver, peritoneum, and lungs.⁶ Early stages of COAD frequently present no clear symptoms, leading to late-stage diagnosis in about half of the cases.⁷

Recent advances in the diagnosis, therapy, and prognosis of colon adenocarcinoma have significantly improved patient outcomes, yet challenges persist. Diagnostic techniques, including colonoscopy and liquid biopsies, have enhanced early detection, while molecular characterization has facilitated personalized treatment approaches, particularly through the

identification of biomarkers like KRAS and HER2, which inform the efficacy of targeted therapies.^{8–10} Immunotherapy, particularly immune checkpoint inhibitors, has shown promise, especially in tumors with microsatellite instability.⁸ However, the majority of patients do not benefit from these advancements, as only a small percentage exhibit these favorable biomarkers.^{8,10} Furthermore, disparities in outcomes, particularly among minority populations, and the complexities of tumor heterogeneity and metastasis remain significant hurdles.¹¹ Addressing these challenges requires ongoing research to enhance biomarker utility and develop novel therapeutic strategies.^{9,12}

The development and progression of COAD are driven by a complex interplay of molecular pathways characterized by genetic mutations and dysregulated signaling mechanisms.¹³ Key pathways implicated include Wnt/ β -catenin, Notch, TGF- β , EGFR/MAPK, and PI3K/AKT, which are often altered through mutations in tumor suppressor genes and genes involved in DNA repair.^{14–16} These pathways contribute to malignant phenotypes by affecting processes such as cell cycle regulation, apoptosis, and angiogenesis.^{13,17} Recent studies have identified hub genes associated with COAD progression, revealing potential biomarkers for early diagnosis and therapeutic targets.¹⁷ Furthermore, understanding the genomic and epigenomic instability in these pathways is crucial for

CONTACT Baohong Xu  bhxu22@ccmu.edu.cn  Department of Gastroenterology, Beijing Luhe Hospital Affiliated to Capital Medical University, 82 Xinhua Rd, Beijing, 101149, China

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/15384047.2025.2458820>

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

developing effective treatment strategies and improving patient outcomes.^{13,16}

The endoplasmic reticulum (ER) is crucial for protein folding, modification, and assembly. ER stress (ERS) occurs when its homeostasis is disturbed by factors like ischemia, hypoxia, or imbalances in calcium ions, leading to the unfolded protein response (UPR).^{18,19} Research indicates that ERS-related genes (ERSRGs) can serve as effective prognostic markers, with studies identifying specific gene signatures that correlate with patient outcomes and immune response.²⁰ Downregulation of ERO1 α promote cell apoptosis induced by ER stress by regulating the miR-101/ESZ2 axis.²¹ Exacerbating ERS promoted cell apoptosis and increased the sensitivity of mucinous colon cancer to chemotherapy.²² A novel ERS gene signature showed its predictive capability for overall survival and response to immunotherapy in COAD patients.²³ Additionally, the establishment of an ERS-related prognostic risk model has shown promise in assessing survival rates and immune cell infiltration in colorectal cancer.²⁴ The classification of colorectal cancer subtypes based on ERS has revealed distinct clinical characteristics and prognostic implications.^{25–27} Overall, these findings underscore the importance of ERS in cancer prognosis and treatment strategies. Targeted therapies toward ERS have shown potential in various therapeutic contexts. For example, innovative nano-engineering approaches have been explored for delivering therapeutic agents directly to sites of ER stress in cancer cells, demonstrating a novel method to potentially enhance treatment efficacy.²⁸

The prognostic value of ERSRGs in other cancer types has been intensively evaluated.^{29–32} Given the limited studies evaluating the prognostic value of ERS gene signature and its association with immune response in colon cancer,^{33,34} it would be necessary to conduct a thorough analysis to provide deeper understanding of ERSRGs in colon cancer with well-established approaches and novel insights. In this study, we utilized public databases to explore the abnormally expressed ERSRGs and screen for key biomarkers. We identified nine new ERSRGs such as ASNS, ATF4, and TRAF2, which are involved in crucial signaling pathways like IL-12 and PI3K-AKT, linking them to the survival of colon cancer patients over 1-, 3-, and 5-year timelines. At the same time, a prognostic model was constructed to validate its predictive value. We additionally conducted RT-qPCR detection on these key genes in the laboratory to verify their differential expression in human colon cancer cell lines (SW48, HCT-15) and human normal colonic epithelial cellline(CCD-841CoN).

Materials and methods

Transcriptome data acquisition and pre-processing

We used the “TCGA biolinks” package of R³⁵ to download the Colon Adenocarcinoma (COAD) dataset from the Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>). The dataset includes 462 colon cancer samples (group: COAD) with prognostic OS clinical information and 41 control samples (group: Control). The corresponding clinical data was obtained from the UCSC Xena database³⁶ (<http://genome.ucsc.edu>).

Table 1. Endoplasmic reticulum stress-related genes list.

GENE	GENE	GENE
HSPA5	DERL1	ATP2A2
DDIT3	DERL2	OS9
HERPUD1	FICD	UMOD
EIF2AK3	WFS1	BOK
ERN1	MBTPS1	PDIA4
ATF6	HMG81	CANX
ATF4	SIRT1	ERP29
XPB1	CREB3L1	BCL2L11
EIF2S1	PTPN1	CCDC88B
CASP4	TMEM117	CASP7
PPP1R15A	RNF183	NCK1
ATF3	CASP9	ASNS
BCL2	RTN3	TXNDC12
MAP3K5	TMBIM6	CYCS
HSP90B1	CREB3L3	TGFB1
CASP3	EIF4G1	SGPP2
CEBPB	DNAJB9	PPP1CB
MAPK8	EIF2AK2	SIGMAR1
SCAMP5	TMEM214	GSK3B
TMED4	BAX	TMEM33
CREB3	HYOU1	UBA5
DNAJC3	CCL2	SERPINA1
TP53	PRNP	CDK5RAP3
NFE2L2	DDX3X	AIFM1
DDRGRK1	TRIB3	UFM1
SESN2	CASP8	HERPUD2
ATF6B	SLC39A14	SREBF1
P4HB	PDIA3	MTOR
CALR	PPP1CA	TLR4
TNFRSF10B	NUPR1	CLU
MBTPS2	JUN	CREB3L2
DNAJC10	TRAF2	SEL1L
QRICH1	DNAJB11	AKT1
VCP	PPP1CC	ATP2A3
RNF186	NLRP3	PARP1
MANF	DERL3	

In addition, the expression profile datasets GSE39582³⁷ and GSE44076³⁸ for COAD patients from the Gene Expression Omnibus (GEO) database³⁹ were also downloaded by using the R package “GEOquery”.⁴⁰ The GSE39582 dataset consists of 566 colon samples (group: COAD) and 19 non-tumor colon samples (group: Control). The GSE44076 dataset consists of 98 colon samples (group: COAD) and 148 normal colon samples (group: Control).

We collected ERSRGs from the “GeneCards”⁴¹ database (<https://www.genecards.org/>), with the search keyword “endoplasmic reticulum stress”. After retaining only protein coding genes, we screened ERSRGs by using “Relevance score > 5” to obtain 107 ERSRGs (Table 1).

Sample-grouping based on the ERSRGs risk model

To obtain the characteristic genes of ERSRGs, we used “glmnet” package of R⁴² to perform the Least absolute shrinkage and selection operator (LASSO) regression⁴³ using parameter set. seed (2022) and ten-fold cross validation, with each iteration leaving one subset as the validation set and the remaining nine as the training set, based on the expression matrix of ERSRGs in the COAD samples in the test dataset and ran 1000 cycles to prevent overfitting. To ensure robustness and validity of the prognostic model, we balanced model complexity and fit by prioritizing minimizing the mean squared error (MSE) across the validation sets during the cross-validation process, recorded the MSE for each lambda

and selected the lambda that resulted in the smallest average validation MSE. The formula of calculating risk score is:

$$\text{riskScore} = \sum_i \text{Coefficient}(\text{hubgene}_i) * \text{mRNA Expression}(\text{hubgene}_i)$$

Subsequently, we extracted the coefficient of each prognostic ERSRG and completed the risk-score formula, which was a tool to calculate the risk scores of all samples in the dataset. Next, with the median of the scores as the critical value, we divided the COAD samples into two risk groups, which were high/low ERS-risk group (HERS/LERS group). We used the “survival” package for Kaplan Meier (KM curve) analysis to identify their impact over the survival rate, respectively.

GSEA and GSVA

The Gene Set Enrichment Analysis (GSEA)⁴⁴ determines the contribution of a pre-defined gene set to a specific phenotype. We ranked all the DEGs between HERS and LERS group based on their logFC values. Next, the “clusterProfiler” package was used to perform GSEA on all the DEGs, with the parameter “seed” set to 2022, “calculation frequency” set to 1000, “gene number in each gene set” set to 10, and “max gene number” set to 500. The *p*-value correction method was Benjamin Hochberg (BH). We obtained the reference gene set from the MSigDB (Molecular signatures database 3.0)⁴⁵ database and set the significantly enriched screening criteria as *P*.adj < 0.05 and FDR value (*q*.value) < 0.05.

The Gene Set Variation Analysis (GSVA)⁴⁶ evaluates whether a pathway is enriched between different strains. We also obtained reference gene set from MSigDB and performed GSVA on the gene expression matrix of the dataset to obtain all the enrichment pathway. We retained only the 10 pathways with highest and lowest log₂FC value for subsequent analysis.

Differential expression of ERSRGs

We compared the expression differences of prognostic ERSRGs in the TCGA-COAD, GSE44076, and GSE39582 dataset between group COAD and Control. The results were presented in a group comparison graph using R package “ggplot2”. The genes showed expression difference in all three datasets were chosen as ERS-related DEGs (ERSRDEGs).

GO and KEGG

The Gene Ontology (GO) analysis⁴⁷ and the Kyoto Encyclopedia of Genes and Genomes (KEGG)⁴⁸ were the tools to identify the pathways that the genes were enriched. We used R package “clusterProfiler”⁴⁹ for both GO and KEGG annotation on the ERSRDEGs, with the item screening criteria: *P*.adj < 0.05 and *q* < 0.05, and BH to correct the *p* value.

Protein-protein interaction network (PPI network)

In this study, we used the STRING database to construct the PPI network of ERSRDEGs with low confidence (0.150).⁵⁰ The CHIPBase database (version 3.0) (<https://rna.sysu.edu.cn/chipbase>) predicts all the interactions between transcription factors (TF) and genes. The HTFtarget⁵¹ database (<http://bioinfo.life.hust.edu.cn/hTFtarget>) consists data on TFs and their corresponding regulatory targets. We searched for all the TFs that bind to ERSRDEGs by using both the two databases and filter them with “Number of samples found (upstream)” > 0 and “Number of samples found (downstream)” > 0. The ENCORI database⁵² (<https://starbase.sysu.edu.cn/>) contains information about the small molecules including miRNA, ncRNA, mRNA. We searched in the database for all the miRNAs that interact with the ERSRDEGs. All the networks were visualized by using Cytoscape.

ROC curve

We used the “pROC” package of R to perform the ROC analysis and calculated the area under the curve (AUC) to evaluate the predictive effect of each ERSRDEG.

The differential expression analysis of genes among clinical features

We compared the expression differences of ERSRDEGs among different clinical features, including TNM stage, age, gender, and survival outcome, to determine if ERSRDEGs are related to the clinical features.

Constructing a prognostic model

We used the TCGA-COAD dataset to conduct a univariate Cox regression analysis on ERSRDEGs to filter all significant factors (genes) with *p* < .1 as the threshold. Before conducting the multivariate Cox regression analysis, we verified the proportional hazards assumption using graphical diagnostics with log-minus-log survival plots, supplemented by Schoenfeld residuals tests. We also examined the linearity assumption between the log hazard and each continuous covariate using Martingale residuals. To address potential multicollinearity among the ERSRGs, we calculated the Variance Inflation Factor (VIF) for each gene included in the model. Next, to assess each variable's impact on survival outcomes while adjusting for potential confounders, we performed the multivariate Cox regression analysis on variables that met the threshold and constructed the prognosis model. To assess the accuracy of the model at different time points, we constructed calibration curves to compare the predicted probabilities of survival outcomes with the observed outcomes over a given timeframe. Furthermore, we drew the forest plot and nomogram⁵³ based on the results and introduced the Decision curve analysis (DCA)⁵⁴ to evaluate the effect of the model which was done by the R package “ggDCA”.⁵⁵

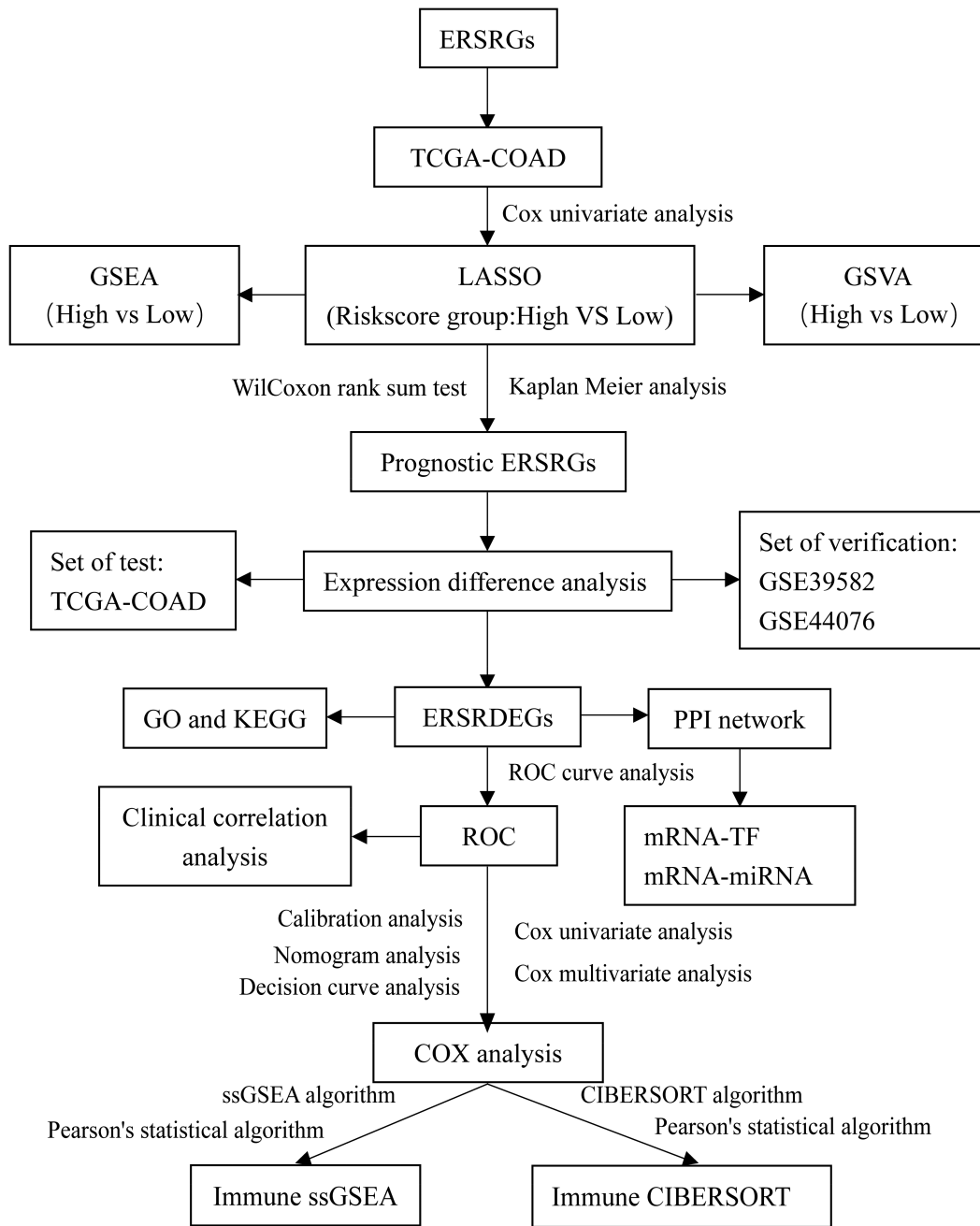


Figure 1. The workflow of the study. ERSRGs, endoplasmic reticulum stress related genes; TCGA, the cancer genome atlas; COAD, colon adenocarcinoma; LASSO, least absolute shrinkage and selection operator; GSEA, gene set enrichment analysis; GSVA, gene set Variation analysis; ERSRDEGs, endoplasmic reticulum stress-related differentially expressed genes; ssGSEA, single-sample gene-set enrichment analysis.

The analysis of immune infiltration

We analyzed and calculated the enrichment scores of each immune cell by applying the ssGSEA algorithm in the R package “GSVA” (version 1.46.0).^{56,57} The differences in the infiltration abundance of immune cells (IAIC) between the two risk groups of TCGA-COAD were displayed through boxplot plots. We also calculated the abundance correlation between the immune cells and visualized it with a heat map by using “corplot” package. The correlation between each IAIC and gene was visualized through the bubble plot drawn by R package “ggplot2”.

CIBERSORT⁵⁸ is an algorithm to estimate the composition of IAIC. We uploaded the expression matrix data of samples between different groups to CIBERSORT (<https://cibersortx.stanford.edu/>) to screen immune cells with enrichment scores greater than zero. The differences in the IAIC between different groups were displayed through group comparison graphs. The correlation heat map and the bubble plot were also drawn.

RNA extraction and quantitative real-time PCR (RT-qPCR)

Cell lines SW48 (human colon cancer cells), HCT-15 (human colorectal adenocarcinoma cells) were purchased from Wuhan

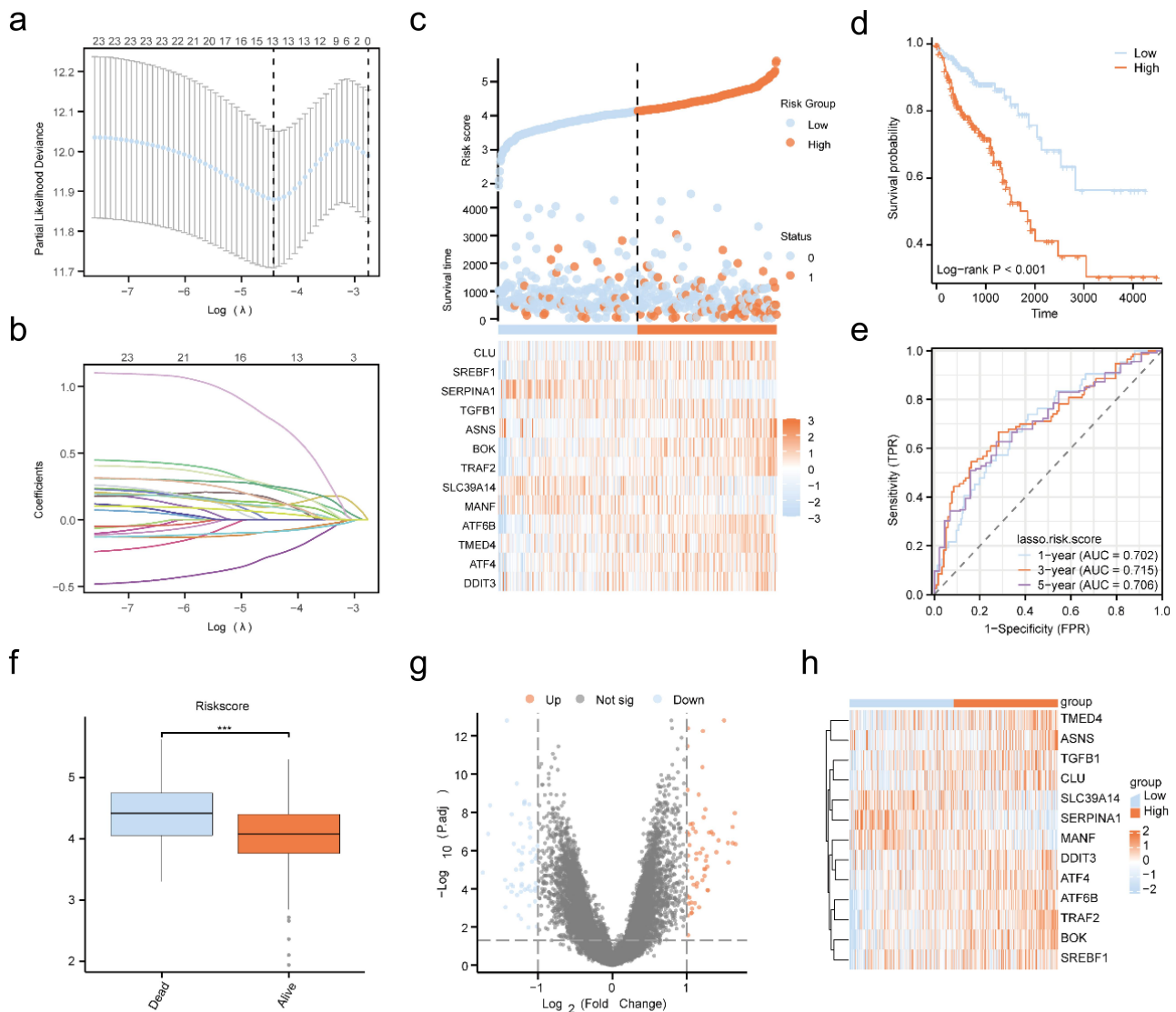


Figure 2. Construction of the ESRGs-risk model and grouping of the samples. a) the risk model diagram. b) the variable trajectory diagram. c) the risk factor diagram. d) the prognostic KM curve. e) the ROC curve. f) the grouping comparison chart. g) the volcano map of the difference analysis results between the two risk groups. h) a simplified numerical heatmap of ESRGs between the two risk groups.

Shangen Biotechnology Co., Ltd., China. The CCD-841CoN (human colonic epithelial cells) were obtained from Xiamen Immocell Biotechnology Co., Ltd., China. Total RNA was extracted from three cells using Trizol reagent (Cat. #15596026, Invitrogen) following the manufacturer's instructions. RNA quality was evaluated by 1% agarose gel electrophoresis, RNA concentration and purity were determined by Nanodrop 2000 with approximately concentration of 100–250 ng/ μ L and an A260/A280 ratio of 1.8–2.0. Reverse transcription was conducted using iScript™ cDNA Synthesis Kit (Cat. #1708890, Bio-Rad). QPCR reaction system was configured for qPCR experiment using iQ™ SYBR® Green Supermix (Cat. #1708880, Bio-Rad). The primers of RT qPCR were designed using primer-blast function in National Center for Biotechnology Information and were synthesized by Sangon Biotech (Shanghai) as follows:

GAPDH-F: AGAAGGCTGGGGCTCATTTG
GAPDH-R: GCAGGAGGCATTGCTGATGAT

ASNS-F2: TTCCAGTCCTTGCCCTCATC
ASNS-R2: ATGTTGAGCCAAAGCCACT
ATF4-F2: CGCAACATGACCGAAATGA
ATF4-R2: TCTCCAGCGACAAGGCTAAG
ATF6B-F3: TGTCTCCTGGGAGATGACCC
ATF6B-R3: TCTGTCTTCACTTCCAGGACCT
BOK-F4: TTCGTGCTGCTGCCAGAGA
BOK-R4: CTGGGTCTGCGGAGGAAC
CLU-F1: TTTTATGTTGAGTTGCTGCTTC
CLU-R1: AGGGACTGTCATACCAGTGAA.
DDIT3-F2: CCACTCTTGACCCGTCTTC
DDIT3-R2: GGAAACGGAAACAGAGTGG
MANF-F3: ACCAGGACCTCAAAGACAGAGA
MANF-R3: GTCCTTCTTCTTAAGCTTCTCACAG
SLC39A14-F2: GGGTGCCATTTATGCTCTA
SLC39A14-R2: TGCGATGCCCTAAGTGAAG
TRAF2-F1: GCTGCCCTTGCTGTCTCTGT
TRAF2-R1: GACAGGCAGAAACGAGGGC

Table 2. Differentially expressed genes in TCGA-COAD.

GENE	logFC	P.Value	adj.P.Val	B	group
REG4	-1.74398	6.88E-07	1.41E-05	5.587351	down
TCN1	-1.67141	1.64E-09	1.34E-07	11.32208	down
FER1L6	-1.645	1.66E-11	4.33E-09	15.71145	down
L1TD1	-1.44292	2.2E-06	3.46E-05	4.493539	down
REG1A	-1.43916	.000148	.000921	0.568054	down
SERPINA1	-1.4191	9.87E-18	1.55E-13	29.49511	down
CLDN18	-1.41152	7.29E-06	8.57E-05	3.36834	down
B3GNT6	-1.36122	4.58E-06	5.99E-05	3.804424	down
SPINK4	-1.35662	1.4E-05	.000144	2.760718	down
XKR9	-1.354	7.51E-11	1.26E-08	14.26738	down
MUC2	-1.33781	6.01E-06	7.4E-05	3.548258	down
SLC18A1	-1.32652	7.41E-09	4.39E-07	9.883265	down
HEPACAM2	-1.2956	4.36E-07	9.89E-06	6.017571	down
RAB27B	-1.28612	4.76E-13	3.1E-10	19.11884	down
CTSE	-1.2821	2.25E-05	.000211	2.31332	down
REG3A	-1.26402	.000496	.00246	-0.54352	down
CA8	-1.26009	2.12E-08	9.99E-07	8.885605	down
KLK12	-1.23478	2.04E-07	5.63E-06	6.734587	down
CXCL8	-1.23451	3.03E-10	3.88E-08	12.93158	down
PIGR	-1.23048	1.4E-05	.000145	2.755298	down
REG1B	-1.22842	.00084	.003742	-1.02359	down
ATOH1	-1.22456	4.15E-07	9.56E-06	6.063899	down
PCSK1	-1.22223	5.83E-06	7.21E-05	3.577493	down
CCDC60	-1.22026	1.67E-07	4.8E-06	6.923554	down
CXCL5	-1.21911	4.38E-06	5.77E-05	3.845476	down
SDR16C5	-1.21635	3.09E-08	1.31E-06	8.52594	down
TFF1	-1.21018	1.93E-07	5.38E-06	6.789662	down
SULT1C3	-1.2001	1.26E-08	6.63E-07	9.379792	down
REP15	-1.18551	1.22E-11	3.54E-09	16.00449	down
TRPA1	-1.18478	9.85E-12	3.08E-09	16.21167	down
FAM169A	-1.18031	7.59E-12	2.6E-09	16.46158	down
SLITRK6	-1.1778	1E-05	.00011	3.069072	down
TFF2	-1.14032	8.84E-06	9.94E-05	3.187774	down
CLCA1	-1.12909	.002991	.010474	-2.16943	down
IL1A	-1.12542	1.34E-10	2.05E-08	13.71161	down
PADI3	-1.11839	5.91E-06	7.29E-05	3.564685	down
IL1B	-1.11767	1.35E-11	3.82E-09	15.91137	down
DMBT1	-1.11746	.000116	.000758	0.792245	down
FAM177B	-1.10921	2.92E-09	2.03E-07	10.77167	down
AGBL4	-1.10594	2.79E-10	3.7E-08	13.01328	down
GALNTL6	-1.10292	3.04E-08	1.3E-06	8.542144	down
ZBTB7C	-1.09899	5.57E-10	6.11E-08	12.35137	down
ANXA10	-1.09695	1.2E-05	.000127	2.90005	down
LYZ	-1.08097	1.18E-07	3.74E-06	7.255722	down
AC005833.1	-1.06037	1.76E-08	8.67E-07	9.06209	down
ABCA12	-1.04199	8.48E-06	9.64E-05	3.226363	down
FCGBP	-1.04014	6.11E-05	.000457	1.385483	down
GABRP	-1.03816	1.17E-05	.000125	2.928627	down
CLCA2	-1.03311	7.36E-08	2.59E-06	7.702652	down
CBLIF	-1.03168	3.62E-06	5.04E-05	4.024506	down
SLC16A7	-1.01844	4.5E-08	1.76E-06	8.16812	down
OLFM4	-1.01542	.002578	.009326	-2.03634	down
HSPA4L	-1.00719	1.03E-08	5.63E-07	9.569198	down
SLC4A4	-1.00023	6.66E-06	8E-05	3.452598	down
PRDM13	1.008285	5.28E-07	1.15E-05	5.8371	up
NRXN2	1.008343	1.28E-09	1.11E-07	11.5561	up
SLC13A3	1.011005	2.58E-05	.000234	2.187991	up
RNF208	1.011904	1.42E-15	3.47E-12	24.70729	up
PABPC1L	1.013379	1.48E-12	6.7E-10	18.03265	up
NOTUM	1.015555	.0005	.002474	-0.55137	up
KIFC2	1.017578	1.03E-16	4.02E-13	27.23622	up
PRAC1	1.017868	.009486	.02653	-3.19081	up
SMTNL2	1.018777	4.78E-08	1.83E-06	8.111644	up
C6orf15	1.02825	8.46E-05	.000592	1.083942	up
RASL10B	1.028416	1.85E-09	1.45E-07	11.20767	up
B4GALNT4	1.037876	3.36E-05	.000287	1.942217	up
KCNT1	1.038877	9.1E-09	5.14E-07	9.687793	up
NAT8L	1.049882	3.41E-10	4.16E-08	12.82018	up
SHISA9	1.062912	7.27E-05	.000526	1.224798	up
PLA2G12B	1.063161	.000357	.001882	-0.24402	up
IGF2	1.067601	.00034	.001811	-0.19918	up
GDF10	1.06962	1.56E-06	2.64E-05	4.815638	up
MAP7D2	1.071725	.000182	.00109	0.376583	up
CHST13	1.079086	3.1E-10	3.93E-08	12.9113	up

(Continued)

Table 2. (Continued).

GENE	logFC	P.Value	adj.P.Val	B	group
NKD2	1.084929	6.8E-10	6.98E-08	12.16116	up
LY6G6F-LY6G6D	1.093004	4.59E-06	6E-05	3.80169	up
FAM155B	1.110526	7.95E-05	.000563	1.141355	up
ASCL2	1.121157	1.13E-09	1.05E-07	11.67733	up
SLC22A31	1.128049	8.2E-08	2.81E-06	7.599712	up
BRSK2	1.132111	3.4E-08	1.42E-06	8.435932	up
DLX3	1.13333	6.34E-06	7.68E-05	3.498023	up
SLC6A4	1.141529	3.68E-06	5.09E-05	4.009661	up
VGF	1.161524	2.93E-12	1.16E-09	17.37398	up
ELF5	1.167268	.000149	.000927	0.560894	up
CAMKV	1.18319	3.12E-08	1.32E-06	8.516039	up
CPNE7	1.21168	3.91E-14	4.32E-11	21.52075	up
COMP	1.212366	2.26E-06	3.53E-05	4.46708	up
SLC38A3	1.213889	1.55E-08	7.97E-07	9.178635	up
CAPS	1.217736	1.7E-16	5.54E-13	26.75249	up
TH	1.229189	1.13E-10	1.82E-08	13.87834	up
LY6G6D	1.237273	3.89E-07	9.19E-06	6.126155	up
TRIM54	1.239163	1.18E-06	2.13E-05	5.0761	up
TNNC2	1.247038	4.05E-07	9.41E-06	6.0878	up
HSPB6	1.251382	3.24E-11	6.95E-09	15.07187	up
GAL	1.256787	1.02E-07	3.34E-06	7.395465	up
ISM2	1.269116	5.15E-08	1.95E-06	8.041455	up
SLC35D3	1.27667	2.73E-07	6.91E-06	6.461026	up
URAD	1.280378	1.08E-05	.000118	3.001203	up
WIF1	1.286552	1.1E-05	.00012	2.979357	up
WNT11	1.301448	1.32E-09	1.14E-07	11.52562	up
KLRG2	1.33361	1.17E-09	1.06E-07	11.64508	up
F7	1.355957	1.7E-09	1.36E-07	11.28924	up
DRD2	1.400218	2.74E-08	1.22E-06	8.641686	up
SLC30A2	1.427893	4.51E-09	2.92E-07	10.35648	up
CLDN9	1.502353	2.38E-17	1.55E-13	28.64598	up
CACNG4	1.555269	6.54E-09	3.96E-07	10.00229	up
CEL	1.561806	1.36E-07	4.15E-06	7.12362	up
MAGEB17	1.617644	5.06E-10	5.82E-08	12.44246	up
GNG4	1.62035	6.79E-09	4.05E-07	9.966799	up
PCSK1N	1.647402	5.96E-11	1.09E-08	14.48875	up
CYP2W1	1.656789	7.91E-09	4.57E-07	9.821275	up

TCGA: The cancer genome atlas. COAD: Colon adenocarcinoma.

Statistical analysis

All data processing and analysis in the study were based on R software (Version 4.1.2). The differences between non-normal distribution variables were analyzed through Mann Whitney U test, while distributed variables were analyzed by Student T test. Pearson correlation analysis was introduced to calculate the coefficients. All statistical *p*-values were bilateral with *p* < .05 significant.

Results

Construction of a risk model for ERSRGs and grouping of samples

The workflow chart of this study was depicted in Figure 1. We conducted a univariate COX analysis on 107 ERSRGs. A total of 23 ERSRGs were obtained: *DDIT3*, *EIF2AK3*, *ATF4*, *XBPI*, *PPP1R15A*, *CASP3*, *MAPK8*, *TMED4*, *DNAJC3*, *ATF6B*, *DNAJC10*, *MANF*, *CASP8*, *SLC39A14*, *TRAF2*, *BOK*, *CANX*, *ASNS*, *TGFB1*, *TMEM33*, *SERPINA1*, *SREBF1*, *CLU*. These genes were input into the LASSO to further screening and ERSRGs-risk model construction (Figure 2a). The result and the variable trajectory map were visualized (Figure 2b). We retained 13 ERSRGs (*DDIT3*, *ATF4*, *TMED4*, *ATF6B*, *MANF*,

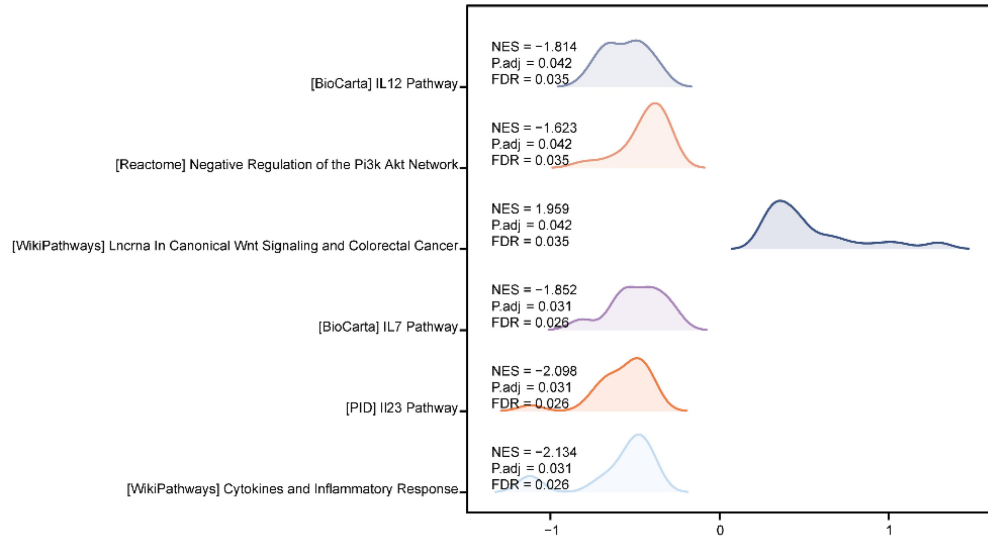
SLC39A14, *TRAF2*, *BOK*, *ASNS*, *TGFB1*, *SERPINA1*, *SREBF1*, *CLU*) as characteristic genes. We grouped the COAD samples based on their risk scores calculated by the formula and exhibited them with a risk factor map (Figure 2c).

$$\begin{aligned} \text{Risk score} = & \text{MANF} * -0.269340666 + \text{SERPINA1} \\ & * -0.091310941 + \text{SLC39A14} * -0.087571533 \\ & + \text{CLU} * 0.050409538 + \text{TGFB1} * 0.095258705 \\ & + \text{TRAF2} * 0.099151578 + \text{DDIT3} * 0.122236994 \\ & + \text{SREBF1} * 0.14089734 + \text{ATF4} * 0.152580471 \\ & + \text{TMED4} * 0.23394756 + \text{BOK} * 0.249088116 \\ & + \text{ATF6B} * 0.251062516 + \text{ASNS} * 0.743247137 \end{aligned}$$

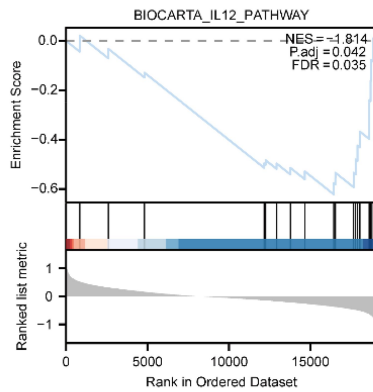
KM curve analysis showed significant differences in predicting the survival rate of COAD patients between the two risk groups (*p* < .001, Figure 2d). The ERSRGs risk model demonstrated promising predictive outcomes for 1-year, 3-year, and 5-year survival, with AUC values of 0.702, 0.715, and 0.706, respectively (Figure 2e). Survivors had significant lower risk scores compared with the non-survivors (Figure 2f).

We used the “limma” package for differential analysis to obtain the DEGs between the HERS group and LERS group and obtained a total of 19,538 genes with 111 genes met the threshold (Table 2). A total of 57 genes were higher, while 54 genes were lower in the HERS group (Figure 2g,h).

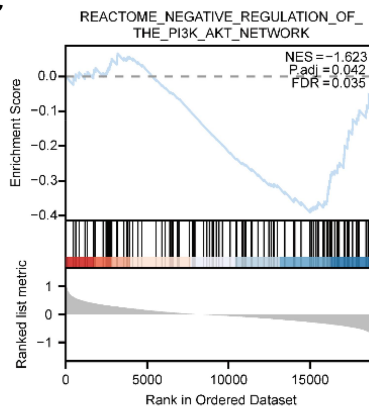
a



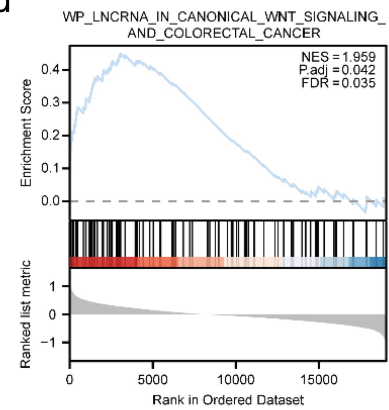
b



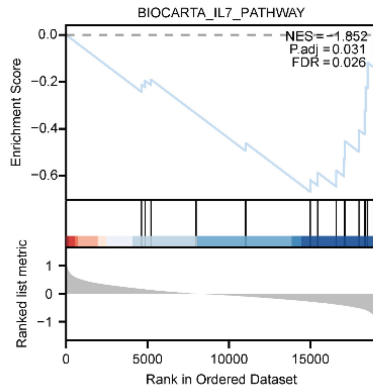
c



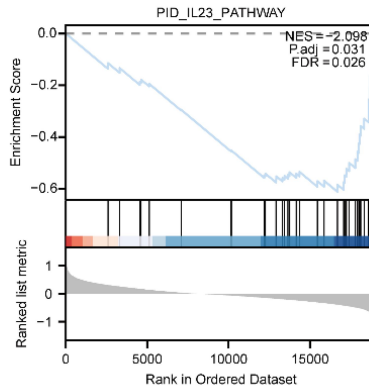
d



e



f



g

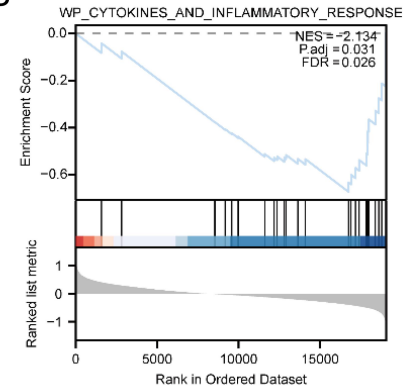


Figure 3. GSEA between the two groups. a) Mountain map of GSEA results. b-g) Genes are significantly enriched in biocarta_IL12 (b), PI3K/AKT (c), Wnt_signaling (d), biocarta_IL7 (e), pid_IL23 (f), cytokines_and_inflammatory_response (g). The screening criteria for GSEA are $p_{adj} < .05$ and FDR value (q.value) < 0.05 .

Table 3. GSEA results of TCGA-COAD dataset.

Description	setSize	enrichmentScore	NES	p.adjust	qvalue
WP_CYTOKINES_AND_INFLAMMATORY_RESPONSE	26	-0.6728	-2.1336	0.0312	0.0260
PID_IL23_PATHWAY	37	-0.6120	-2.0981	0.0312	0.0260
BIOCARTA_IL7_PATHWAY	16	-0.6688	-1.8524	0.0312	0.0260
WP_LNCRNA_IN_CANONICAL_WNT_SIGNALING_AND_COLORECTAL_CANCER	93	0.4493	1.9591	0.0424	0.0353
REACTOME_NEGATIVE_REGULATION_OF_THE_PI3K_AKT_NETWORK	113	-0.3906	-1.6228	0.0424	0.0353
BIOCARTA_IL12_PATHWAY	19	-0.6224	-1.8135	0.0424	0.0354

GSEA: Gene Set Enrichment Analysis; TCGA: The cancer genome atlas; ESRGs: Endoplasmic reticulum stress related genes.

Table 4. GSVA results of TCGA-COAD dataset.

	logFC	AveExpr	adj.P.Val
REACTOME_NEGATIVE_REGULATION_OF_TCF_DEPENDENT_SIGNALING_BY_DVL_INTERACTING_PROTEINS	-0.30554	-0.01027	2.55E-10
REACTOME_ARACHIDONATE_PRODUCTION_FROM_DAG	-0.30141	-0.00116	5.99E-09
BARRIER_CANCER_RELAPSE_NORMAL_SAMPLE_DN	-0.2829	-0.17489	1.09E-16
REACTOME_VEGF_LIGAND_RECEPTOR_INTERACTIONS	-0.28204	-0.02039	3.2E-09
MIKI_COEXPRESSED_WITH_CYP19A1	-0.2814	0.007306	8.44E-09
WP_MED_AND_PSEUDOACHONDROPLASIA_GENES	-0.25965	-0.017	4.73E-08
BLANCO_MELO_INFLUENZA_A_INFECTION_A594_CELLS_DN	-0.258	-0.204	1.39E-21
REACTOME_DISEASES_OF_BASE_EXCISION_REPAIR	-0.25637	-0.02188	9.57E-07
SHEDDEN_LUNG_CANCER_GOOD_SURVIVAL_A5	-0.25241	-0.06838	5.24E-18
WP_HEDGEHOG_SIGNALING_PATHWAY_WP47	-0.25058	-0.02067	1.76E-10
SPIRA_SMOKERS_LUNG_CANCER_UP	0.34095	-0.08982	2.02E-28
JONES_TCOF1_TARGETS	0.346685	-0.03404	2.03E-11
MYLLYKANGAS_AMPLIFICATION_HOT_SPOT_18	0.348213	-0.02944	4.1E-09
BARRIER_CANCER_RELAPSE_NORMAL_SAMPLE_UP	0.350544	-0.03728	1.37E-26
ZHAN_EARLY_DIFFERENTIATION_GENES_UP	0.356973	-0.00263	4.63E-15
HERNANDEZ_ABERRANT_MITOSIS_BY_DOCETACEL_4NM_DN	0.362646	-0.03538	1.36E-14
REACTOME_FASL_CD95L_SIGNALING	0.363093	-0.03228	6.61E-12
REACTOME_WAX_AND_PLASMALOGEN_BIOSYNTHESIS	0.379241	0.030213	4.87E-18
REACTOME_PROPIONYL_COA_CATABOLISM	0.381947	-0.04128	3.92E-14
REACTOME_ACTIVATION_OF_CASPASES_THROUGH_APOPTOSOME_MEDIATED_CLEAVAGE	0.432089	-0.01617	1.1E-19

GSVA: Gene Set Variation Analysis; TCGA: The cancer genome atlas; COAD: Colon adenocarcinoma; ERSRGs: Endoplasmic reticulum stress related genes.

GSEA analysis

GSEA analysis (Figure 3a) showed that the genes were significantly enriched in biocarta IL-12 (Figure 3b), PI3K/AKT (Figure 3c), Wnt signaling (Figure 3d), biocarta IL-7 (Figure 3e), PID IL-23 (Figure 3f), cytokines and inflammatory response (Figure 3g). ((Table 3) for specific pathways).

GSVA analysis

We conducted GSVA and selected the 10 pathways with the highest and the lowest log₂FC value (a total of 20). The results of the genes were enriched in TCF dependent signaling, arachidonate production from DAG, VEGF ligand receptor interactions (Table 4, Figure 4a,b).

Differential expression analysis of ERSRGs

Next, we compared the expression of 13 ERSRGs (*DDIT3*, *ATF4*, *TMED4*, *ATF6B*, *MANF*, *SLC39A14*, *TRAF2*, *BOK*, *ASNS*, *TGFB1*, *SERPINA1*, *SREBF1*, *CLU*) between COAD and control group in the TCGA-COAD, GSE39582, GSE44076 datasets. To avoid any potential inconsistency and batch effect between different datasets, the comparison of ERSRGs were conducted independently. In the TCGA-COAD dataset (Figure 5a), the expression levels of 11 ERSRGs (*ASNS*, *ATF4*, *ATF6B*, *BOK*, *CLU*, *DDIT3*, *MANF*, *SLC39A14*, *TGFB1*, *TMED4*, *TRAF2*) exhibited significant change, with 9 ERSRGs upregulated and 2 downregulated ($p < .05$). In the dataset GSE39582 (Figure 5b), there was a statistically significant difference ($p < .05$) in the expression levels of 11 ERSRGs (*ASNS*, *ATF4*, *ATF6B*, *BOK*, *CLU*, *DDIT3*, *MANF*, *SLC39A14*, *TGFB1*, *TMED4*, *TRAF2*), with 8 ERSRGs upregulated and 3 downregulated. In the dataset GSE44076 (Figure 5c), *ASNS*, *ATF4*, *ATF6B*, *BOK*, *CLU*, *DDIT3*, *MANF*, *SERPINA1*, *SLC39A14*, *TGFB1*, and *TRAF2* showed statistical differences ($p < .05$) with 9 upregulated and 2 downregulated. We intersected the three sets of significantly expressed ERSRGs and obtained 9 ERSRDEGs,

ASNS, *ATF4*, *ATF6B*, *BOK*, *CLU*, *DDIT3*, *MANF*, *SLC39A14*, *TRAF2* for further analysis.

GO and KEGG analysis

The results of GO and KEGG analysis showed that 9 ERSRDEGs were mainly enriched in GO items as “response to ERS”, “response to topologically incorrect protein”, “regulation of response to ERS”, “response to unfolded protein”, “intrinsic apoptotic signaling pathway”, “ubiquitin protein ligase binding”, “DNA-binding transcription activator activity” and KEGG pathways such as “Protein processing in endoplasmic reticulum”, “TNF signaling pathway”, “Apoptosis”, “MAPK signaling”, etc. (Table 5, Figure 6(a-e)).

The construction of PPI network

The PPI network of 9 ERSRDEGs was constructed based on STRING database (Figure 7a). A total of 38 pairs of interaction between 6 ERSRDEGs (*ATF4*, *BOK*, *DDIT3*, *MANF*, *SLC39A14*, *TRAF2*) and 28 TFs were obtained (Figure 7b and Table 6). A total of 45 pairs of mRNA-miRNA interactions consists of 5 mRNA (*CLU*, *DDIT3*, *MANF*, *SLC39A14*, *TRAF2*) and 28 miRNA molecules were obtained and visualized through the network (Figure 7c and Table 7). In addition, we also used the “RCircos” package to annotate the chromosome localization of 9 ERSRDEGs (Figure 7d).

Clinical analysis of ERSRDEGs

Then we plotted a diagnostic ROC curve based on the expression of 9 ERSRDEGs (*ASNS*, *ATF4*, *ATF6B*, *BOK*, *CLU*, *DDIT3*, *MANF*, *SLC39A14*, *TRAF2*) in the TCGA-COAD dataset. From the graph, *CLU* (AUC = 0.936,) showed highest accuracy in predicting survivorship in COAD patients, followed by *ASNS* (AUC = 0.873), *ATF4* (AUC = 0.732), *BOK* (AUC = 0.725), *DDIT3* (AUC = 0.845), *MANF* (AUC = 0.850), and *TRAF2* (AUC = 0.878), *ATF6B* (AUC = 0.604) and *SLC39A14* (AUC = 0.604) (Figure 8(a-h)). The 7

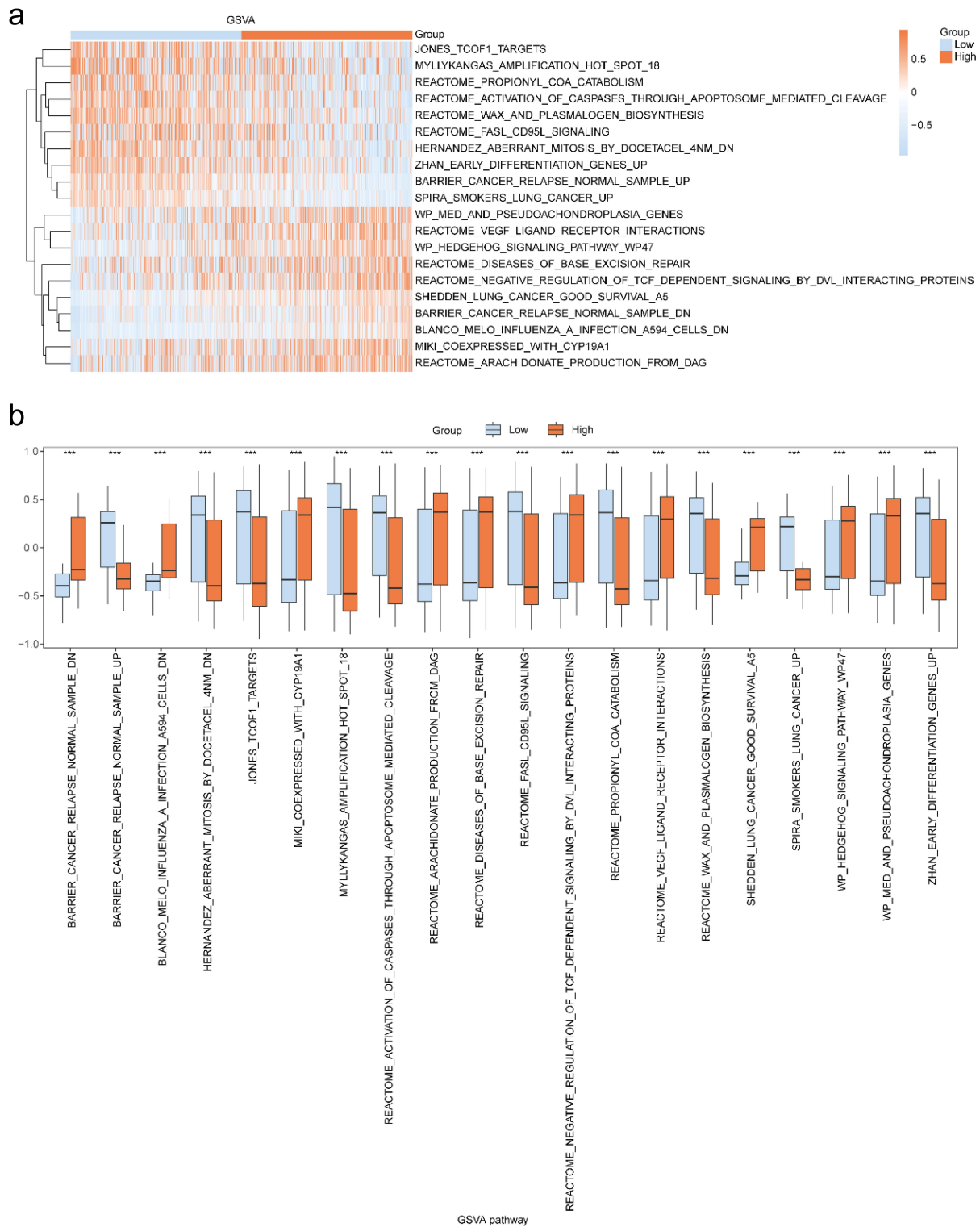


Figure 4. GSVA between the two groups. a-b) the heatmap (a) and group comparison (b). * $p < .05$, ** $p < .01$, *** $p < .001$.

ERSRDEGs (*ASNS*, *ATF4*, *BOK*, *CLU*, *DDIT3*, *MANF*, *TRAF2*) with high AUC were used in the subsequent analysis.

We compared the expression differences of 7 ERSRDEGs among different clinical features (Figure 9(a-f)). The expression of *DDIT3* was found to be significantly different between T1+T2 and T3 samples, also T1+T2 and T4 samples (Figure 9a). The expression of *BOK*, *CLU*, and *DDIT3*

showed significant differences between N0 and N2 samples; additionally, the expression of *DDIT3* was significantly different between N0 and N1 samples (Figure 9b). The expression of *ASNS*, *CLU*, *DDIT3*, and *MANF* between M0 and M1 samples were significantly different (Figure 9c). The expression levels of *CLU* in patients with age > 65 were significantly lower than its expression in patients with age ≤ 65

Table 5. GO and KEGG enrichment analysis results of endoplasmic reticulum stress.

ONTOLOGY	ID	Description	p.adjust	qvalue
BP	GO:0034976	response to endoplasmic reticulum stress	1.69E-09	7.82E-10
BP	GO:0035966	response to topologically incorrect protein	8.29E-09	3.85E-09
BP	GO:1905897	regulation of response to endoplasmic reticulum stress	3.37E-08	1.56E-08
BP	GO:0006986	response to unfolded protein	3.69E-07	1.71E-07
BP	GO:0097193	intrinsic apoptotic signaling pathway	5.65E-06	2.62E-06
CC	GO:0090575	RNA polymerase II transcription regulator complex	0.008401	0.005225
CC	GO:0005667	transcription regulator complex	0.024624	0.015316
CC	GO:0016234	inclusion body	0.016432	0.010221
CC	GO:0031901	early endosome membrane	0.043701	0.027182
KEGG	hsa04141	Protein processing in endoplasmic reticulum	0.000804	0.000385
KEGG	hsa04668	TNF signaling pathway	0.004418	0.002114
KEGG	hsa04210	Apoptosis	0.004627	0.002214
KEGG	hsa04010	MAPK signaling pathway	0.009955	0.004763
KEGG	hsa04911	Insulin secretion	0.012157	0.005817
MF	GO:0031625	ubiquitin protein ligase binding	0.005314	0.002276
MF	GO:0044389	ubiquitin-like protein ligase binding	0.005314	0.002276
MF	GO:0046982	protein heterodimerization activity	0.005314	0.002276
MF	GO:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific	0.010182	0.00436
MF	GO:0001216	DNA-binding transcription activator activity	0.010182	0.00436

GO: Gene Ontology; BP: biological process; CC: cellular component; MF: molecular function; KEGG: Kyoto Encyclopedia of Genes and Genomes.

(Figure 9d). No difference was found between male and female (Figure 9e). We divided the TCGA-COAD samples into survivor group and non-survivor group based on survival outcome, the expression of *BOK*, *DDIT3*, and *TRAF2* were lower in the survivor group compared with the non-survivor group (Figure 9f).

Prognostic model building

We have collected the clinical information of samples with COAD that obtained from the TCGA-COAD dataset after removing duplicate samples (Table 8). Then we conducted Cox regression analysis on the expression of ERSRDEGs (*ASNS*, *ATF4*, *BOK*, *CLU*, *DDIT3*, *MANF*, *TRAF2*) based on the dataset and constructed a prognosis model (Table 9). To address potential multicollinearity among the ERSRDEGs, we calculated the Variance Inflation Factor (VIF) for each gene included in the model (Supplementary Table S1). In our analysis, all VIF values were below the commonly accepted threshold (typically a VIF of 5 or 10), indicating a low risk of multicollinearity among the variables. To evaluate the predictive power of the model, we calculated Harrell's concordance index (C-index). Our model achieved a C-index of 0.66 (95% CI: 0.629–0.692), indicating a reasonable predictive accuracy. A forest map was to display the results (Figure 10a) and a formula was formed.

$$\begin{aligned} \text{Riskscore} = & -2.972992754 + \text{ASNS} * 0.727036866 + \text{ATF4} \\ & * 0.24622324 + \text{BOK} * 0.394914969 + \text{CLU} \\ & * 0.081710475 + \text{DDIT3} * 0.182170641 + \text{MANF} \\ & * -0.454273837 + \text{TRAF2} * 0.359211043 \end{aligned}$$

Then we constructed a nomogram with the coefficients obtained by the formula (Figure 10b). According to the nomogram, *BOK* has the highest contribution to the model compared to other variables.

In addition, the calibration curves of the nomogram were plotted (Figure 10(c-e)). The results of DCA showed that the

clinical predictive effect of our constructed model was 5- year > 3- year > 1- year (Figure 10(f-h)).

Immune differences comparison based on ssGSEA algorithm

After applying the ssGSEA algorithm, we compared the differences of IAIC between the two risk groups and showed the results in comparison chart (28 immune cell types) (Figure 11a), of which 11 immune cell types of IAIC showed significance ($p < .05$). Pearson statistical algorithm showed that in both risk groups, positive correlations were found between the 11 types of IAICs, with the most significant difference being B cell and the Activated dendritic cell (Figure 11(b-c)). The study found significant correlations between the expression levels of 7 ERSRDEGs and 11 types of immune cells across both risk groups (Figure 11(d-e)). In the LERS group (Figure 11d), the strongest correlation existed between CD56dim natural killer cells and *TRAF2*; CD56dim natural killer cells showed a significant positive correlation with all ERSRDEGs except for *ASNS*; The Effector memory CD4⁺ T cell, Eosinophil, and Memory B cell were significantly negatively correlated with all ERSRDEGs except for *ASNS* and *CLU*. In the HERS group (Figure 11e), the strongest correlation existed between Immature B cells and *CLU*.

Immune difference comparison based on CIBERSORT

The results of CIBERSORT showed that 22 types of IAICs in the TCGA-COAD sample (Figure 12a), in which 9 types of IAICs exhibited significance ($p < .05$). Pearson statistical algorithm showed that in the LERS score group, the correlation between Neutrophils and Mast cells activated exhibited the most significant difference (Figure 12b); In the HERS group, the correlation between CD8⁺ T cell and CD4⁺ memory T cell resting exhibited the strongest significance (Figure 12). The analysis revealed a consistent pattern of

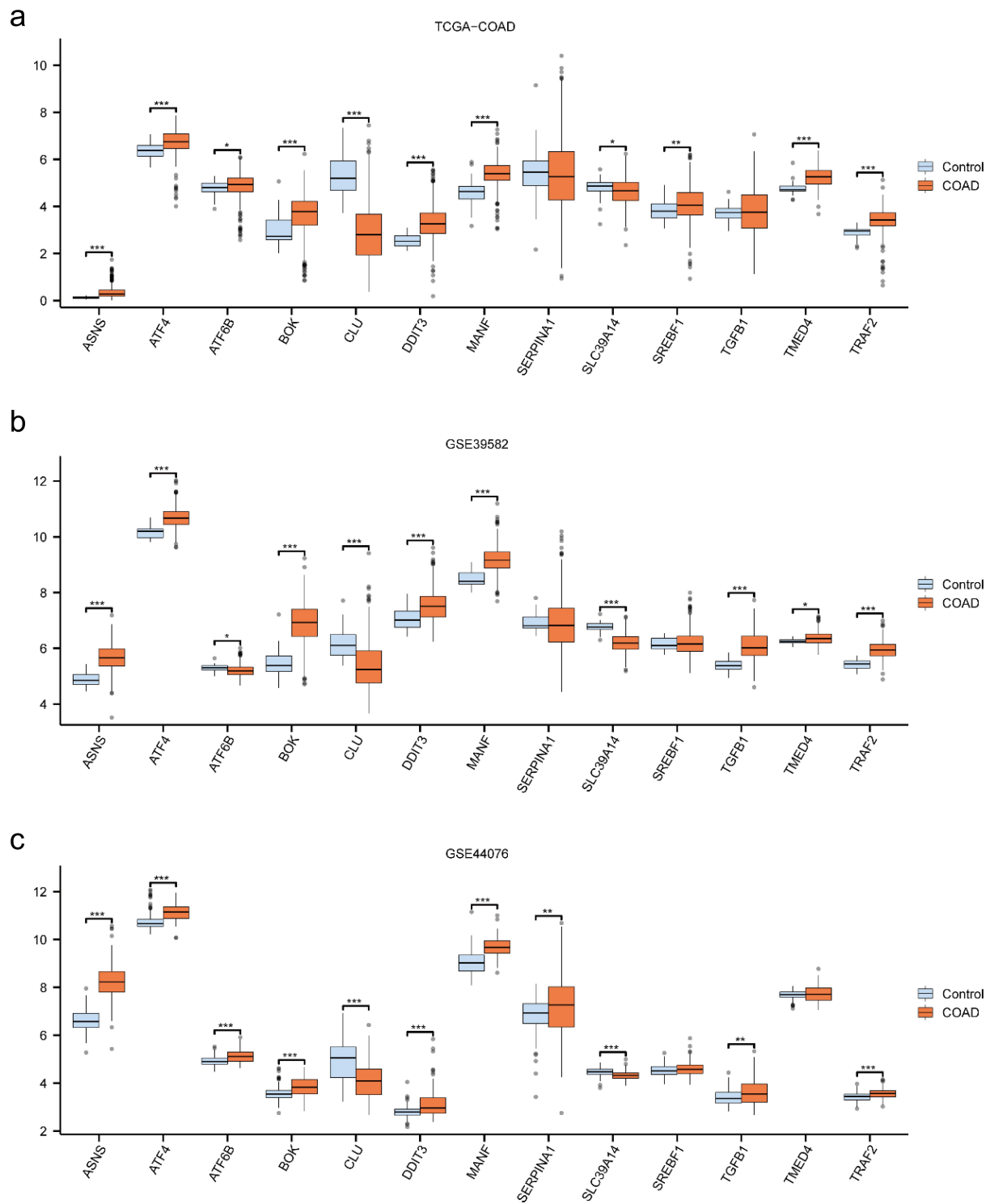


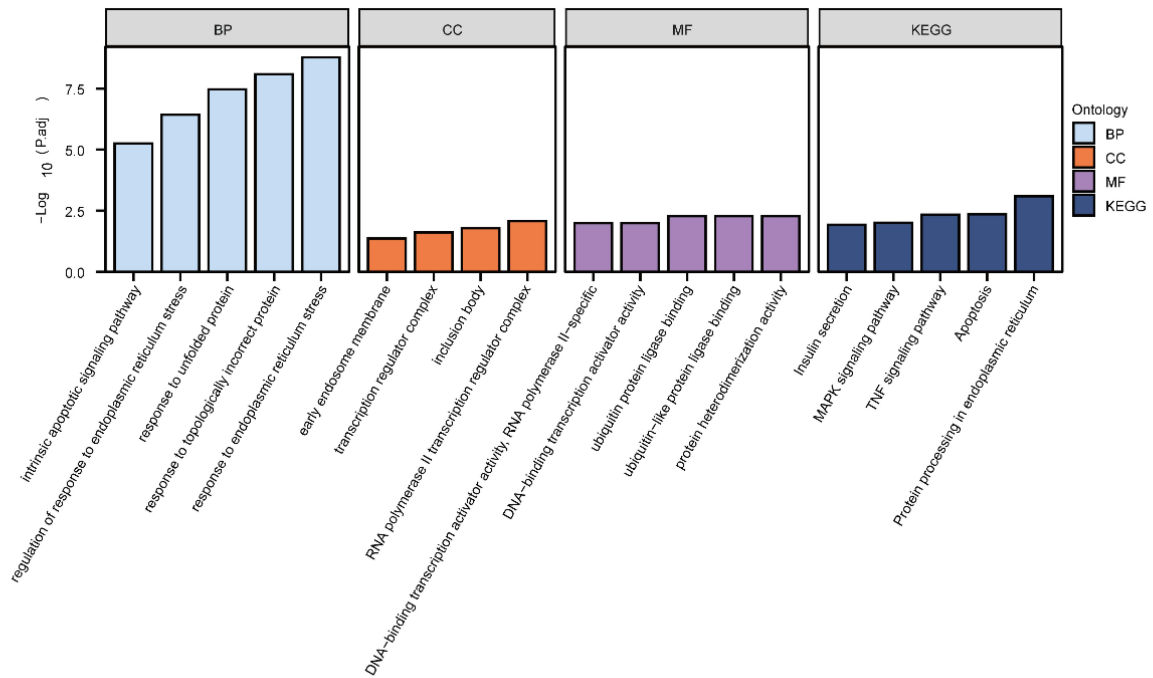
Figure 5. Differential expression analysis of ERSRGs between COAD and Control group from three datasets a-c) group comparison of expression levels of ERSRGs in the TCGA-COAD (a), GSE39582 (b), and GSE44076 (c) dataset. * $p < .05$, ** $p < .01$, *** $p < .001$.

positive and negative correlations between 9 types of immune cells and 7 ERSRDEGs within both the LERS and HERS groups (Figure 12(d-e)), with the strongest correlation between CD4⁺ resting memory T cell and ASNS in the LERS group, T cell regulation (Tregs) and TRAF2 in the HERS group.

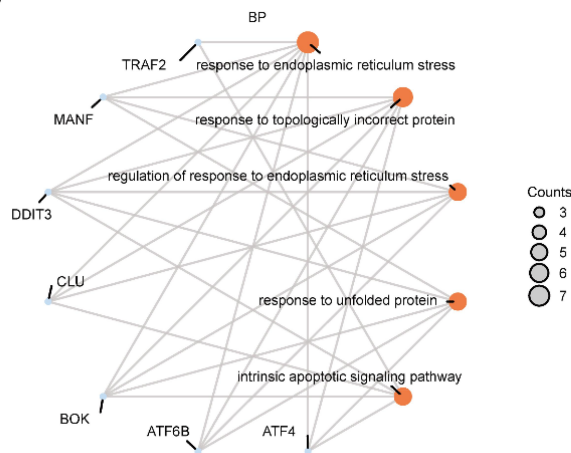
Laboratory verification of the expression of ERSRDEGs

Next, we examined the expression of ERSRDEGs in human normal colonic epithelial and CRC cancer cells. The gene expression of ASNS, ATF4, BOK, DDIT3, MANF and TRAF2 in SW-48 and HCT-15 cells was significantly higher than that in CCD-841 CoN cells, while ATF6B, CLU and SLC39A14

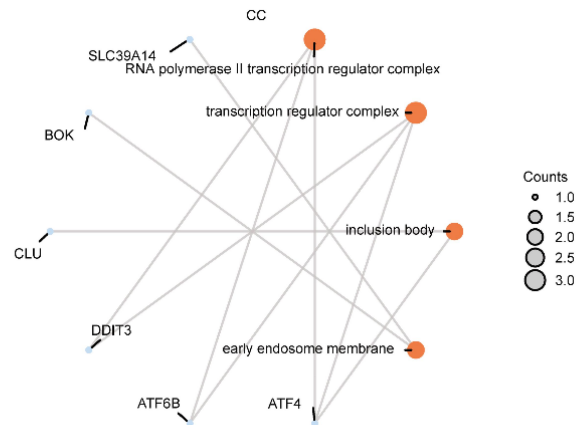
a



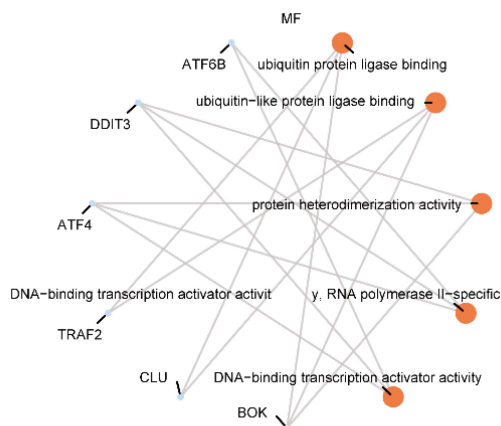
b



c



d



e

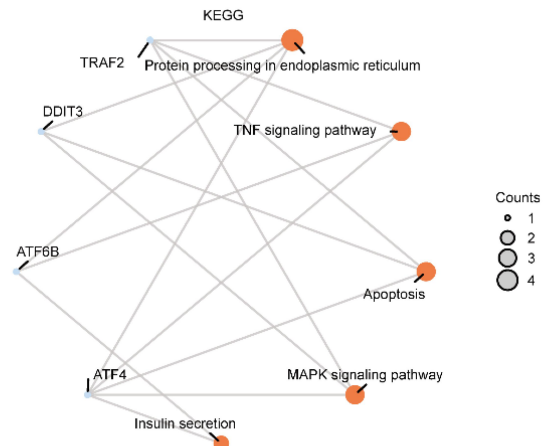


Figure 6. GO and KEGG analysis of ERSRDEGs. a) Bar graph of the results. b-e) the circular network diagrams of results, BP (b), CC (c), MF (d), and KEGG (e). In the bar chart (a), the horizontal axis represents GO terms or KEGG terms, and the bar height represents the p value of each term. The light blue dots in the network diagram (b, c, d) represent specific genes, while the orange dots represent specific pathways. BP, biological process; CC, cellular component; MF, molecular function; the screening criteria for the enriched items are $p < .05$.

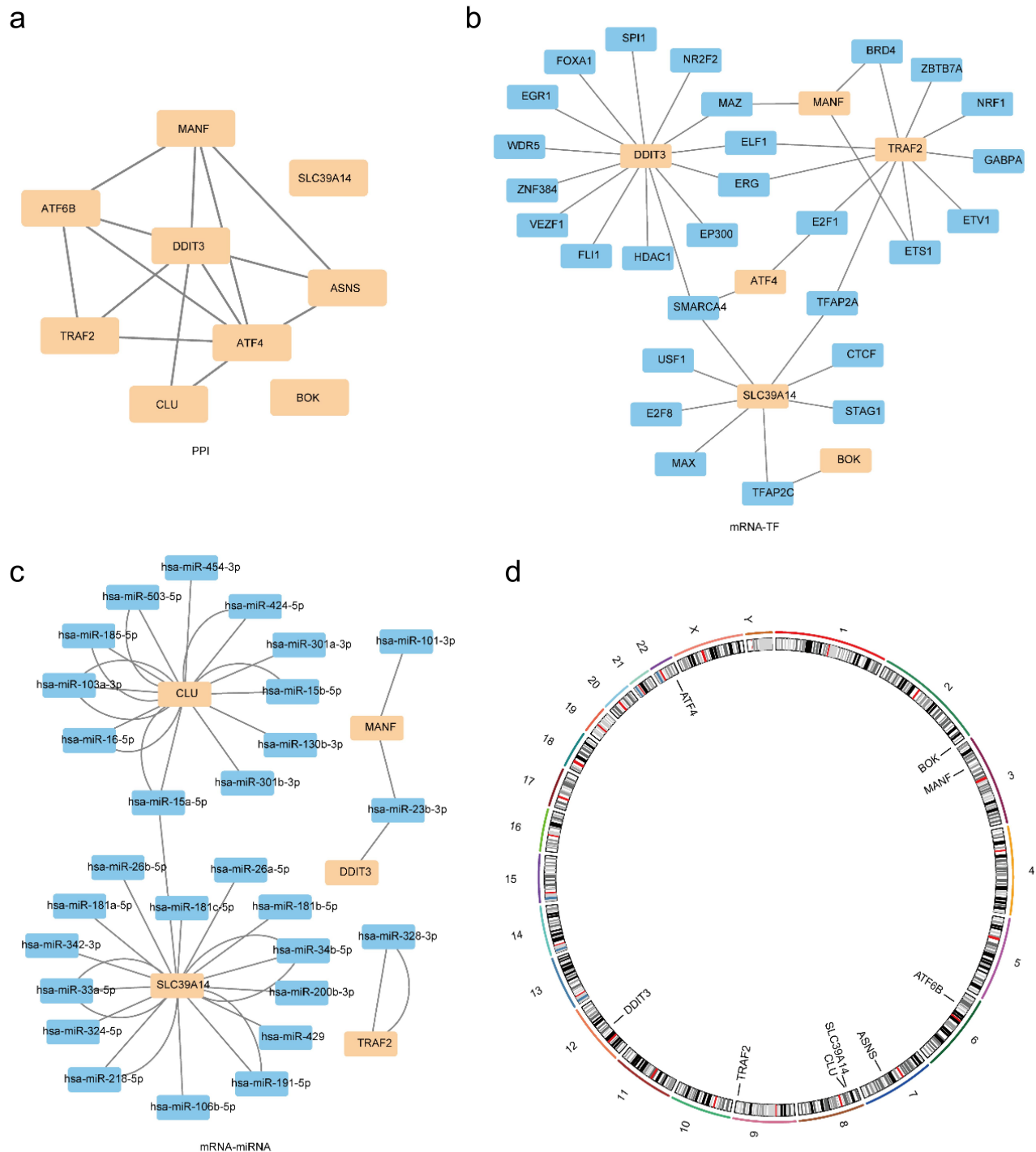


Figure 7. The PPI networks. a) the PPI network of 9 ERSRDEGs. b) the mRNA-TF interaction network, with a light orange rectangle as the mRNA and the blue rectangle is TF. c) mRNA-miRNA networks, with light orange rectangular shaped mRNA and the blue rectangle represents miRNA. d) chromosomal mapping of 9 ERSRDEGs.

genes was lower than that of CCD-841 CoN cells (Figure 13), which is consistent with the former data analysis.

Discussion

In recent years, although the treatment of COAD has made significant progress, its incidence rate has increased year by year, and the five-year survival rate is still low. ERS is widely present in various tumors and is closely related to the survival and drug resistance of tumor cells. Therefore, designing and developing anti-tumor targeted drugs targeting the ERS

signaling pathway has become a current research hotspot. In our study, we identified several ERSRGs in colon adenocarcinoma that significantly correlate with patient prognosis. These genes provide potential targets for improving therapeutic strategies and enhancing prognostic accuracy in this challenging disease.

Recent studies have explored the role of ER stress in tumor progression and as a potential therapeutic target.^{59–62} The ER regulates cellular homeostasis, and disruptions in its function activate the unfolded protein response (UPR), particularly under conditions like hypoxia and chemotherapy.⁶³ ER stress

Table 6. mRNA-TF interaction network.

mRNA	Transcription factor	mRNA	Transcription factor
ATF4	E2F1	MANF	BRD4
ATF4	SMARCA4	SLC39A14	CTCF
BOK	TFAP2C	SLC39A14	E2F8
DDIT3	EGR1	SLC39A14	MAX
DDIT3	ELF1	SLC39A14	SMARCA4
DDIT3	EP300	SLC39A14	STAG1
DDIT3	ERG	SLC39A14	TFAP2A
DDIT3	FLI1	SLC39A14	TFAP2C
DDIT3	FOXA1	SLC39A14	USF1
DDIT3	HDAC1	TRAF2	E2F1
DDIT3	MAZ	TRAF2	ELF1
DDIT3	NR2F2	TRAF2	ERG
DDIT3	SMARCA4	TRAF2	ETS1
DDIT3	SP1	TRAF2	ETV1
DDIT3	VEZF1	TRAF2	GABPA
DDIT3	WDR5	TRAF2	NRF1
DDIT3	ZNF384	TRAF2	BRD4
MANF	ETS1	TRAF2	TFAP2A
MANF	MAZ	TRAF2	ZBTB7A

TF: Transcription factors.

Table 7. mRNA-miRNA interaction network.

mRNA	miRNA	mRNA	miRNA
CLU	hsa-miR-15a-5p	SLC39A14	hsa-miR-26a-5p
CLU	hsa-miR-15a-5p	SLC39A14	hsa-miR-26b-5p
CLU	hsa-miR-16-5p	SLC39A14	hsa-miR-33a-5p
CLU	hsa-miR-16-5p	SLC39A14	hsa-miR-33a-5p
CLU	hsa-miR-103a-3p	SLC39A14	hsa-miR-33a-5p
CLU	hsa-miR-103a-3p	SLC39A14	hsa-miR-181a-5p
CLU	hsa-miR-103a-3p	SLC39A14	hsa-miR-181b-5p
CLU	hsa-miR-15b-5p	SLC39A14	hsa-miR-181c-5p
CLU	hsa-miR-15b-5p	SLC39A14	hsa-miR-218-5p
CLU	hsa-miR-185-5p	SLC39A14	hsa-miR-218-5p
CLU	hsa-miR-185-5p	SLC39A14	hsa-miR-200b-3p
CLU	hsa-miR-301a-3p	SLC39A14	hsa-miR-191-5p
CLU	hsa-miR-130b-3p	SLC39A14	hsa-miR-191-5p
CLU	hsa-miR-424-5p	SLC39A14	hsa-miR-106b-5p
CLU	hsa-miR-424-5p	SLC39A14	hsa-miR-34b-5p
CLU	hsa-miR-503-5p	SLC39A14	hsa-miR-34b-5p
CLU	hsa-miR-503-5p	SLC39A14	hsa-miR-34b-5p
CLU	hsa-miR-454-3p	SLC39A14	hsa-miR-342-3p
CLU	hsa-miR-301b-3p	SLC39A14	hsa-miR-324-5p
DDIT3	hsa-miR-23b-3p	SLC39A14	hsa-miR-429
MANF	hsa-miR-101-3p	TRAF2	hsa-miR-328-3p
MANF	hsa-miR-23b-3p	TRAF2	hsa-miR-328-3p
SLC39A14	hsa-miR-15a-5p		

contributes to tumor growth, inflammation, invasion, and treatment resistance.⁶⁴ In endometrial cancer, specific ER stress-related genes correlate with immune response and chemotherapy efficacy, providing a prognostic risk signature.⁶⁵ Additionally, ER stress-induced protein misfolding supports cancer cell proliferation and survival.⁶⁶ Modulating ER stress may also enhance tumor immunity, as it has both pro- and anti-tumorigenic effects, depending on the cellular context.⁶⁷ In gynecologic tumors, ER stress impacts apoptosis pathways, influencing cancer cell survival and highlighting the need for understanding ER-associated molecular mechanisms.⁶⁸

For lung adenocarcinoma, a prognostic model based on ER stress-related genes links ERS to patient survival, guiding therapy decisions and underscoring its prognostic impact.⁶⁹ A recent study reported an endoplasmic reticulum stress-responsive genes-based prognostic model for colon cancer.³³ In endometrial cancer, an ER stress gene-based risk

signature has been associated with patient outcomes and immune response, offering insights into prognosis.⁷⁰ Furthermore, ER stress modulation is shown to improve cancer immunotherapy by enhancing anti-tumor immunity, reflecting its dual role in both pro- and anti-tumorigenic processes.⁶⁷ Additionally, markers like GRP78 in endometrial cancer may predict occult carcinoma and treatment response, underscoring ERS's relevance in prognosis.⁷¹ These findings collectively highlight the potential of ERS in developing therapeutic and prognostic strategies across various cancers.

In this study, we collected 107 ESRGs and conducted COX univariate analysis by using the expression matrix of these genes in the TCGA-COAD dataset. We further used LASSO to screen for genes with prognostic significance and validated them in the GEO dataset, ultimately identifying 9 genes closely related to the prognosis of colon cancer: *ASNS*, *ATF4*, *ATF6B*, *BOK*, *CLU*, *DDIT3*, *MANF*, *SLC39A14*, *TRAF2*.

These genes are all associated with ERS. Among them, *ATF6B/DDIT3/MANF* are genes directly related to ERS. The protein encoded by *ATF6B* is a TF in the UPR pathway during ERS and is an important ERS sensing protein. Under resting state, *ATF6B* protein binds to GRP78/BIP protein, presenting an inactive state. During ERS, *ATF6B* protein dissociates from GRP78 and transfers to Golgi apparatus, where it is enzymatically hydrolyzed and activated, releasing functional fragments containing alkaline leucine zipper domains that enter the nucleus and activate transcription of downstream genes.^{72,73} *DDIT3* encodes a TF called CCAAT/enhancer binding protein (C/EBP), associated with adipogenesis and erythropoiesis, activated by ERS and promoting cell apoptosis.⁷⁴ The protein encoded by the *MANF* gene is located in the endoplasmic reticulum (ER) and Golgi apparatus, reducing its expression increases susceptibility to ERS-induced death and leads to cell proliferation.⁷⁵

The other five genes are indirectly related to ERS and belong to downstream activating proteins of ERS. *ASNS* encodes asparagine synthase, *ATF4* encodes proteins that belong to DNA binding proteins, and *BOK* encodes proteins that belong to the BCL2 family;⁷⁶ The protein encoded by the *CLU* gene is a secret chaperone, which can also be found in the cytoplasm of cells under certain stress conditions; *SLC39A14* encodes a zinc transporter protein; *TRAF2* encodes a TNF receptor related factor.⁷⁷ The ERS response caused by glucose deprivation will lead to a significant increase in the transcription of *ASNS* gene.⁷⁸ Inhibition of disulfide bonds caused by cell hypoxia can damage protein folding in the endoplasmic reticulum, thereby activating *ATF4* and generating adaptive signaling pathways.⁷⁹ B-cell lymphoma type 2 (BCL-2) killer (*BOK*) can selectively regulate the apoptotic response to ER stress. The specific expression of these proteins reflects that the ERS process is a self-protective function of cells. During ERS, cells increase the expression of stress protein genes, upregulate endoplasmic reticulum chaperone proteins, inhibits protein translation, and initiates endoplasmic reticulum related protein degradation, improving cellular physiological status, and strengthening the self-repairing function of the endoplasmic reticulum.⁸⁰ These genes are all ERS-related and help us dive further into the deviation of ERS to tumors.

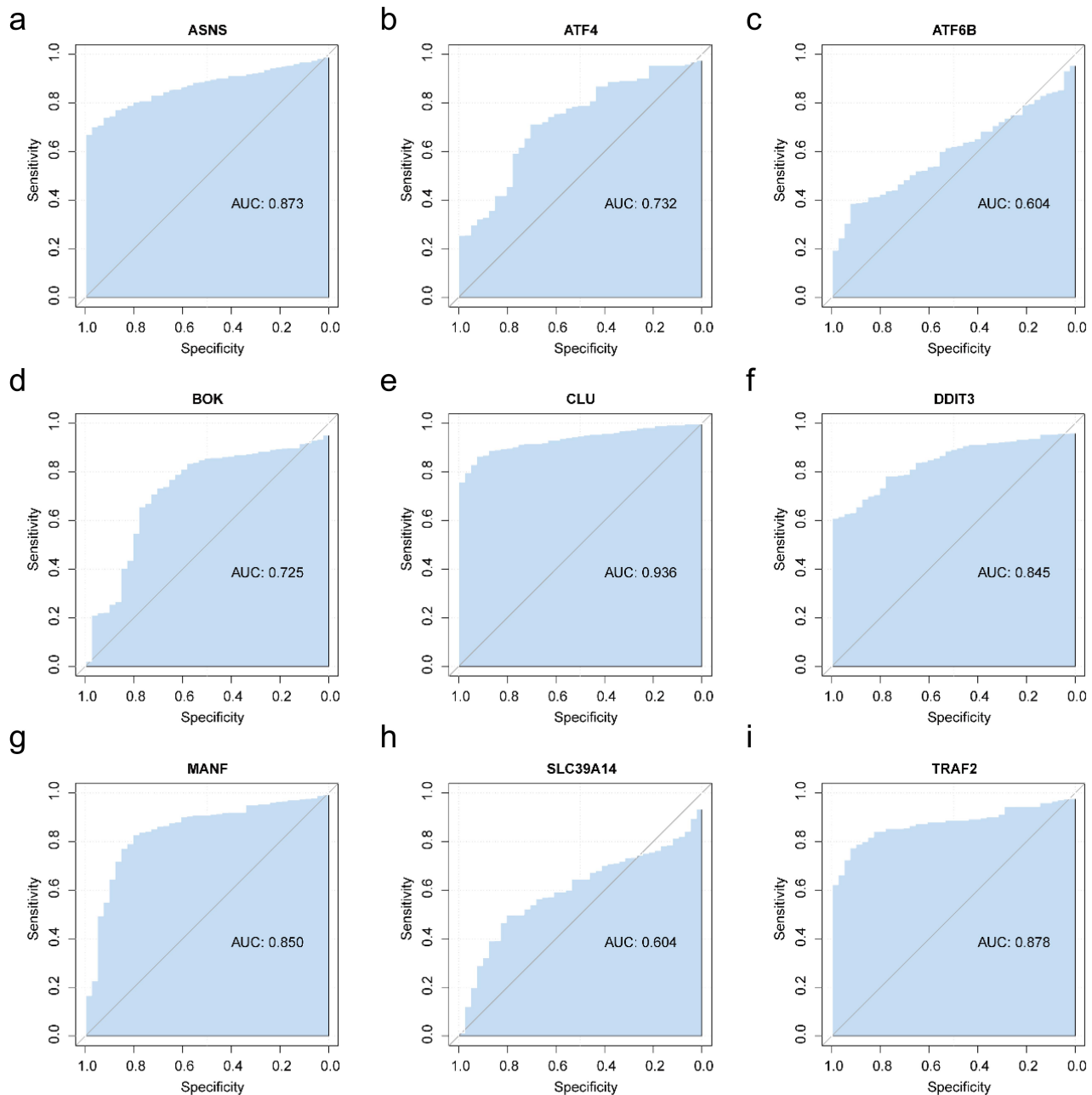


Figure 8. ROC curve of ERSRDEGs. a-l) the ROC curves of ASNS (a), ATF4 (b), ATF6B(c), BOK (d), CLU (e), DDIT3(f), MANF(g),SLC39A14 (h), TRAF2(i)in the TCGA-COAD dataset with sample information (OS survival outcome).

In the GO/KEGG functional analysis of 9 ERSRDEGs, we found that these genes are mainly enriched in the response to ERS, response to topological error proteins, regulation of ERS response, response to unfolded proteins, protein processing, TNF signaling pathway, and intrinsic apoptosis signaling pathway, which is consistent with the function analysis of each gene. In tumor cells, ERS regulate autophagy and apoptosis to induce cell death. When ERS is over severe or persistent, the UPR initiates endoplasmic reticulum related cell apoptosis mediated by TF C/EBP homologous proteins and caspase 12, clearing damaged cells and preventing further damage.⁸¹

We plotted the survival ROC curve and found that all nine genes were able to accurately predict patient death. Among

them, *ASNS*, *ATF4*, *BOK*, *CLU*, *DDIT3*, *MANF*, and *TRAF2* had a prediction rate of AUC > 0.7 for patient death. Through differential analysis, it was found that there were differences in the expression of certain specific genes among different clinical features (case stage T, N, M, age), which revealed a correlation between these seven genes and the clinical characteristics of patients. Therefore, we constructed a prognostic model for colon cancer using these seven genes and plotted a column chart. This prognostic model has accurate predictive effects on the 1-year, 3-year, and 5-year outcomes of patients, especially for the best 5-year survival rate. The prognosis prediction model has good application value.

To explore the biological function and involvement pathways of ERSRDEGs in colon cancer, we constructed a risk

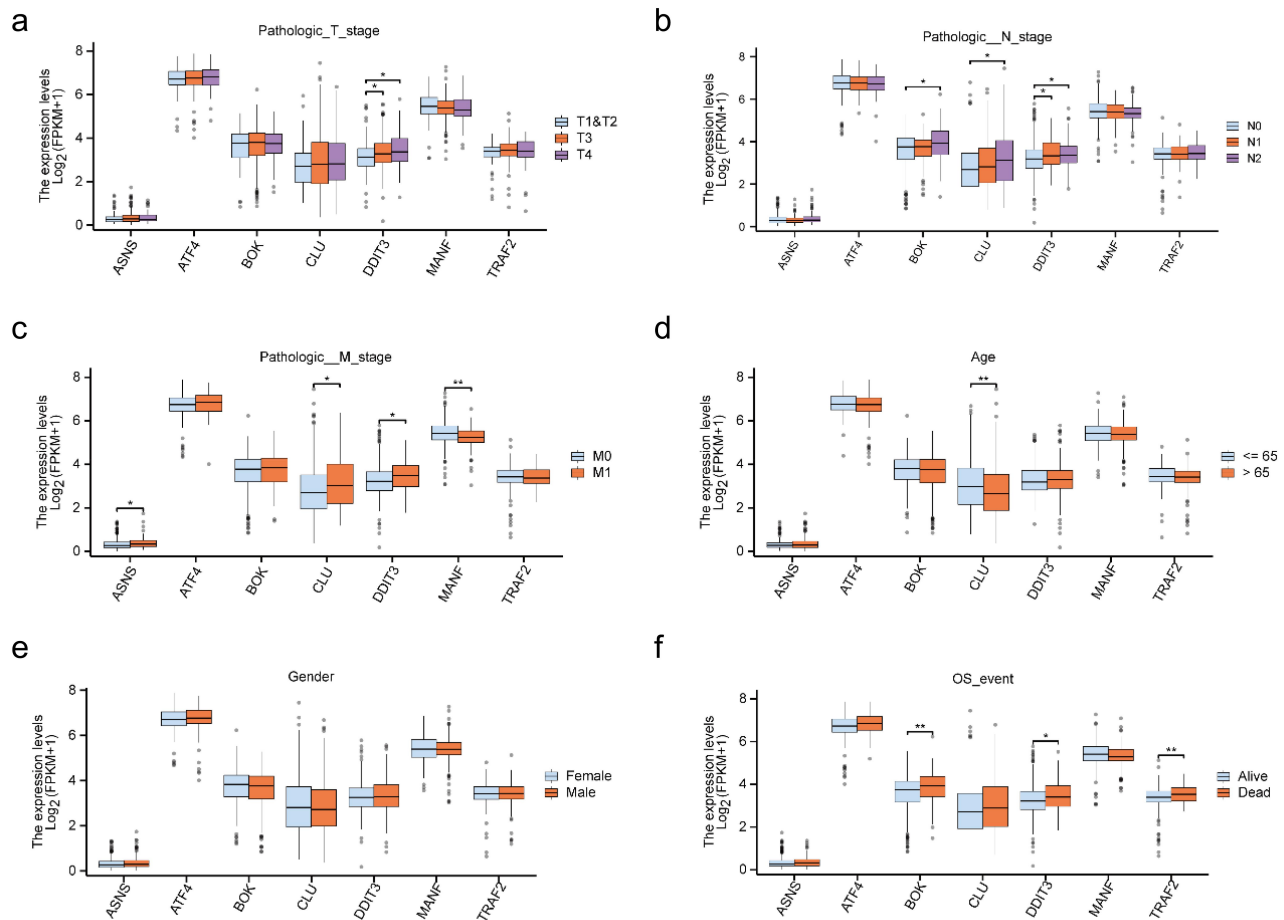


Figure 9. Clinical correlation analysis of ERSRDEGs. a-f) the expression of ERSRDEGs were compared among different groups in pathological stage T (a), pathological stage N (b), pathological stage M (c), age (d), gender (e), and survival outcome (f). * $p < .05$, ** $p < .01$

Table 8. Patient characteristics of COAD patients in the TCGA datasets.

characteristics	overall
Pathologic T stage, n (%)	
T1&T2	94 (19.7%)
T3	323 (67.7%)
T4	60 (12.6%)
Pathologic N stage, n (%)	
N0	284 (59.4%)
N1	108 (22.6%)
N2	86 (18%)
Pathologic M stage, n (%)	
M0	349 (84.1%)
M1	66 (15.9%)
Age, n (%)	
≤65	194 (40.6%)
>65	284 (59.4%)
Gender, n (%)	
Female	226 (47.3%)
Male	252 (52.7%)
ASNS, median (IQR)	0.28167 (0.17932, 0.44055)
ATF4, median (IQR)	6.7543 (6.453, 7.0782)
BOK, median (IQR)	3.7826 (3.2001, 4.2203)
CLU, median (IQR)	2.7959 (1.9633, 3.66)
DDIT3, median (IQR)	3.2636 (2.8485, 3.7115)
MANF, median (IQR)	5.3896 (5.1095, 5.729)
TRAF2, median (IQR)	3.424 (3.1726, 3.7304)

TCGA: The cancer genome atlas; COAD: Colon adenocarcinoma.

model based on LASSO analysis and divided the sample of colon cancer patients into LERS/HERS groups for comparison. GSEA enrichment analysis found that the pathway with

significantly enriched differential genes is biocarta_IL12, PI3K_AKT, biocarta_IL7, pid_IL23, these pathways are all related to cytokines and inflammatory responses. Chronic inflammation caused by the activation of inflammasome plays a central role in all stages of tumor development, including immune suppression, proliferation, angiogenesis, and metastasis. Inflammatory factors can cause ERS and activate UPR. Metabolic factors can all induce ERS and inflammatory responses in various cells. The possible mechanism is the imbalance of calcium ion balance in the endoplasmic reticulum caused by cytokines and metabolic factors, as well as the interference of free radical accumulation with protein folding and mitochondrial metabolism.⁸²

The PI3K/AKT/mTOR pathway is important for tumor cell proliferation and metabolism, standing for a target of tumor. A study by Hsu HS et al.⁸³ on pulmonary fibrosis suggests that PI3K/AKT plays an upstream role in ER stress, affecting the proliferation of lung fibroblasts and leading to bleomycin induced pulmonary fibrosis. It is speculated that PI3K/AKT can phosphorylate I κ B kinase, leading to the degradation of I κ B and the release of NF- κ B, which transfers into the nucleus, activating the transcription of inflammatory genes, and causing ERSs. PI3K/AKT has become a target pathway for inhibiting ERS. Zhang Y et al.⁸⁴ found in their study that leptin inhibit the activation of ERS through the PI3K/Akt pathway,

Table 9. COX regression analysis to ESRGs associated with OS in TCGA-COAD.

Characteristics	Total(N)	HR (95% CI) Univariate analysis	p value Univariate analysis	HR (95% CI) Multivariate analysis	p value Multivariate analysis
ASNS	462	2.036 (1.027–4.038)	0.042	2.069 (0.883–4.848)	0.094
ATF4	462	1.413 (0.952–2.098)	0.086	1.279 (0.801–2.042)	0.302
BOK	462	1.429 (1.106–1.846)	0.006	1.484 (1.123–1.962)	0.006
CLU	462	1.151 (0.992–1.335)	0.064	1.085 (0.924–1.274)	0.319
DDIT3	462	1.278 (1.002–1.630)	0.048	1.200 (0.897–1.605)	0.220
MANF	462	0.749 (0.545–1.028)	0.074	0.635 (0.419–0.962)	0.032
TRAF2	462	2.015 (1.286–3.157)	0.002	1.432 (0.877–2.338)	0.151

COAD: Colon adenocarcinoma; ESRGs: Endoplasmic reticulum stress related genes.

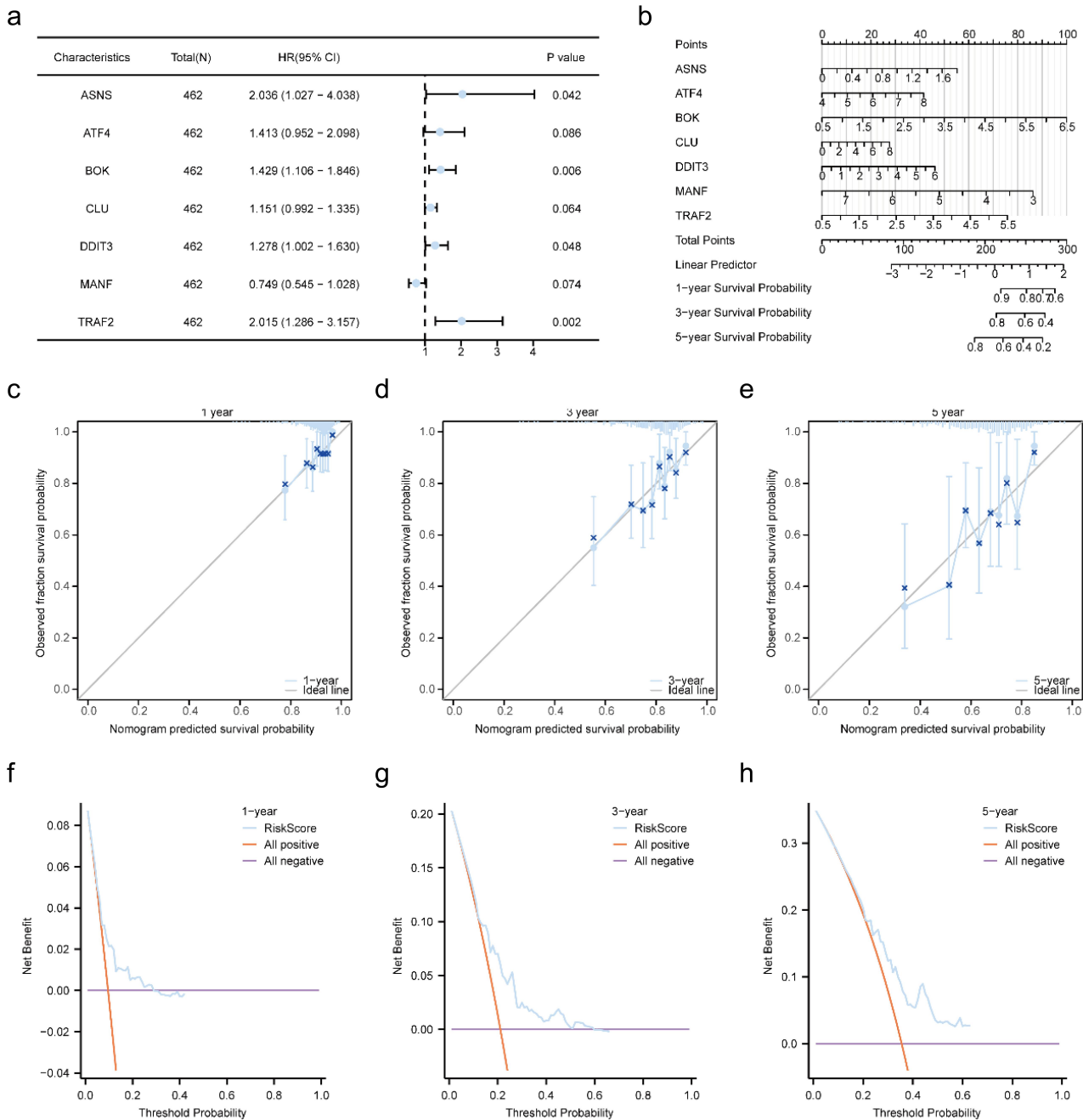


Figure 10. Construction of the prognostic model. a) the forest map. b) the column chart of the model. c-e) the calibration curves of nomogram for 1-year (c), 3-year (d), and 5-year (e). f-h) DCA plots of the model for 1-year (f), 3-year (g), and 5-year (h). DCA, decision curve analysis.

thereby alleviating symptoms of cerebral ischemia. However, no application has been seen in colorectal cancer. The GSVA results found that the differential genes between LERS/HERS risk groups were enriched in TCF dependent

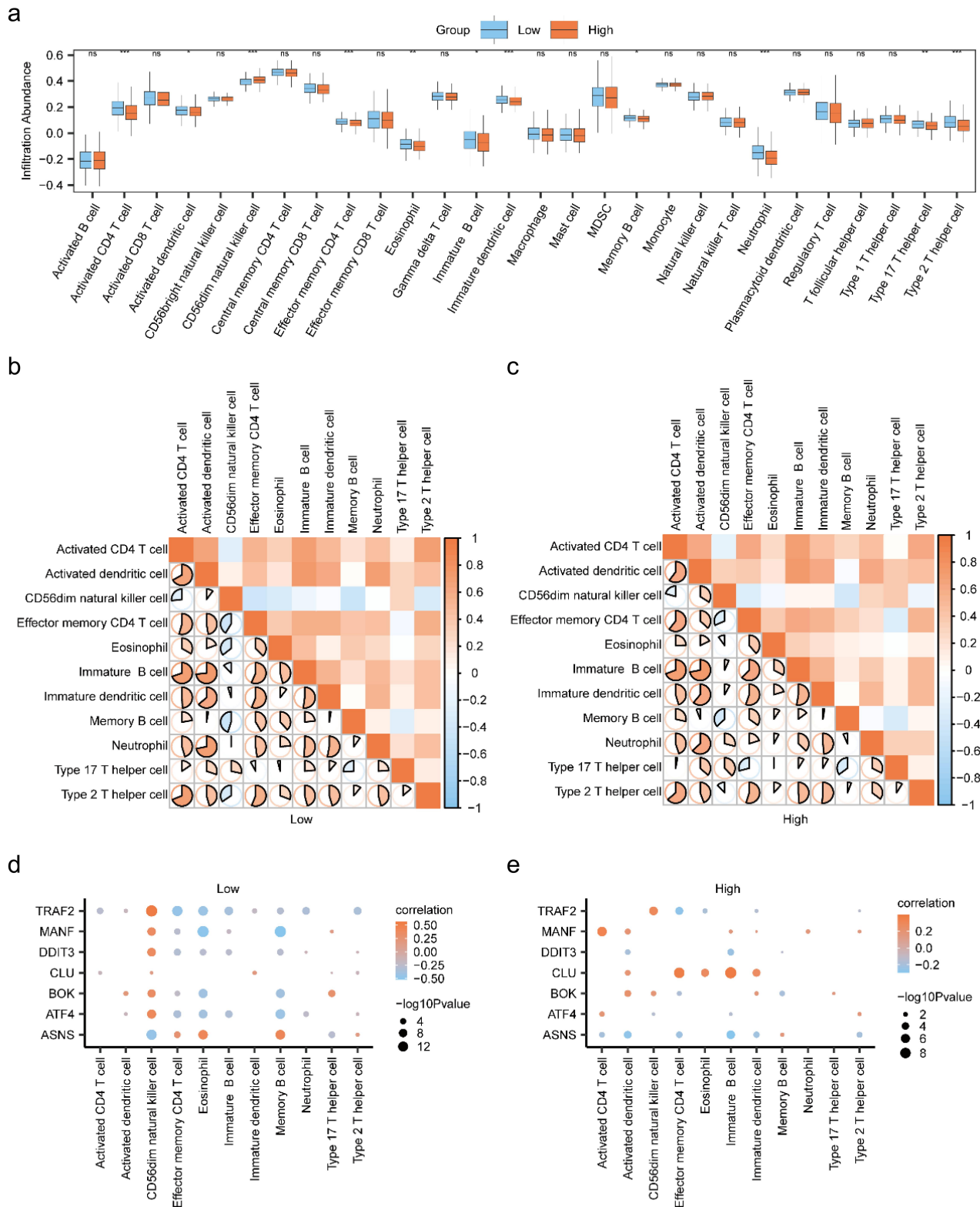


Figure 11. Immune characteristics between the two groups by ssGSEA. a) the comparison chart of immune infiltration. b-c) the correlation analysis results between immune cells in the low ers-risk group (b) and high ers-risk group (c). d-e) correlation plot between immune cells and ERSRDEGs in the low ers-risk group (d) and high ers-risk group (e). ns, no significance; * $p < .05$, ** $p < .01$, *** $p < .001$.

signaling by dvl, arachidonate production from DAG, VEGF ligand receiver interactions, etc. It is speculated that these pathways are mostly related to the cellular repair function after ERS response. However, the interaction between ERS and these pathways still requires in-depth analysis.

We used the ssGSEA algorithm to calculate and compare the infiltration abundance between LERS/HERS groups of 28 immune cells and found 11 differentially expressed. Correlation analysis showed a strong correlation between *TRAF2* and CD56dim natural killer cells, while CD56dim

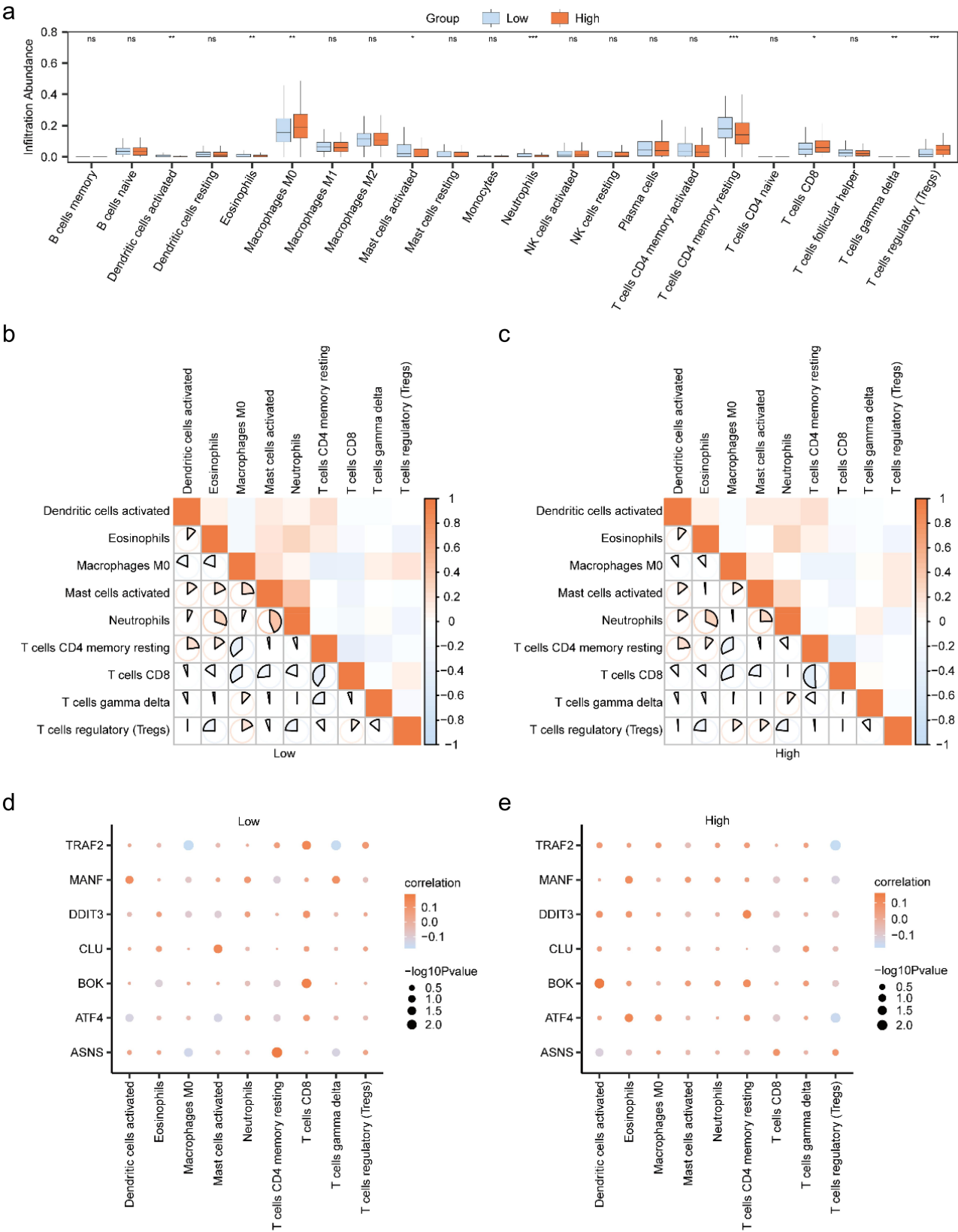


Figure 12. Immune characteristics between the two groups by CIBERSORT. a) the comparison chart of immune infiltration with abundance > 0. b-c) the correlation analysis results between immune cells in the low ers-risk group (b) and high ers-risk group (c). d-e) correlation plot between immune cells and ERSRDEGs in the low ers-risk group (d) and high ers-risk group (e). ns, no significance; * $p < .05$, ** $p < .01$, *** $p < .001$.

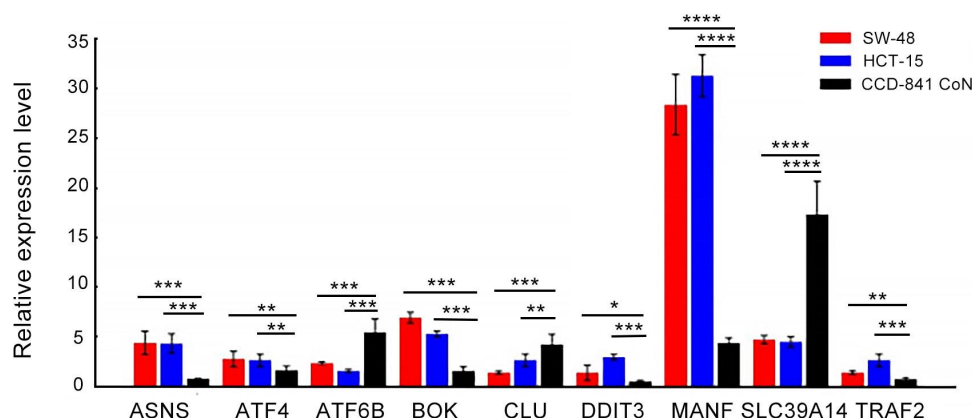


Figure 13. Laboratory verification of the expression of ERSRDEGs. RT-qPCR was performed to measure the relative expression of 9 genes in SW-48, HCT-15, and CCD-841 CoN cell lines. * $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$.

natural killer cells were correlated with almost all ERSRDEGs. This suggests that CD56dim natural killer cells and *TRAF2* gene may become targets for immunotherapy. We also used the CIBERSORT algorithm and found that the *TRAF2* gene has the strongest correlation with T cell regulation, indicating that *TRAF2* is a promising target for immunotherapy.

The ERSRGs identified as prognostic markers in our study could serve as direct targets for novel therapeutic agents. Drugs designed to modulate the activity of these genes might help in alleviating endoplasmic reticulum stress, potentially reducing tumor viability and enhancing the effectiveness of existing treatments. Understanding the role of ERSRGs in the pathogenesis of colon adenocarcinoma could lead to more effective combination therapy strategies. For example, combining ER stress-modulating drugs with conventional therapies such as chemotherapy or radiation might enhance therapeutic outcomes by attacking multiple pathways critical for cancer cell survival and proliferation.

There are still some limitations in this study. Firstly, we identified specific ERS gene biomarkers with great prognostic value for COAD, but without confirmation in large-scale clinical studies. We intend to collaborate with clinical partners to obtain a broader range of colon cancer samples. This will allow us to validate our findings across a more diverse patient cohort, thereby enhancing the generalizability of our results. Secondly, some deep-seated regulatory pathways and axes still need to be further explored, such as the impact of P3IK/AKT on ERS. We need to conduct *in vitro* and *in vivo* assays to examine the effect of these genes on cancer cell proliferation, migration, apoptosis, and explore the therapeutic implications of ERSRDEGs in targeted treatments or immunotherapy. Thirdly, we have developed a promising prognostic prediction model based on the ERSRGs, but this model still needs to be proven in practical use.

Conclusions

Our study identifies several ERS-related genes that demonstrate significant associations with prognosis in colon adenocarcinoma. In addition, these findings also offer promising avenues for further research, potentially guiding the

development of novel therapeutic strategies, which will require extensive further research to determine their viability and efficacy in clinical settings.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Baohong Xu  <http://orcid.org/0009-0009-5837-659X>

Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

References

1. Mahmoud NN. Colorectal cancer: preoperative evaluation and staging. *Surg Oncol Clin N Am*. 2022;31(2):127–141. doi:10.1016/j.soc.2021.12.001.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660.
3. Baidoun F, Elshiwly K, Elkeraie Y, Merjaneh Z, Khoudari G, Sarmini MT, Gad M, Al-Husseini M, Saad A. Colorectal cancer epidemiology: recent trends and impact on outcomes. *Curr Drug Targets*. 2021;22(9):998–1009. doi:10.2174/1389450121999201117115717.
4. Chen K, Collins G, Wang H, Toh JWT. Pathological features and prognostication in colorectal cancer. *Curr Oncol*. 2021;28(6):5356–5383. doi:10.3390/curroncol28060447.
5. Luo XJ, Zhao Q, Liu J, Zheng J-B, Qiu M-Z, Ju H-Q, Xu R-H. Novel genetic and epigenetic biomarkers of prognostic and predictive significance in stage II/III colorectal cancer. *Mol Ther*. 2021;29(2):587–596. doi:10.1016/j.ymthe.2020.12.017.

6. Mahar AL, Compton C, Halabi S, Hess KR, Weiser MR, Groome PA. Personalizing prognosis in colorectal cancer: a systematic review of the quality and nature of clinical prognostic tools for survival outcomes. *J Surg Oncol.* **2017**;116(8):969–982. doi:10.1002/jso.24774.
7. Jin M, Frankel WL. Lymph node metastasis in colorectal cancer. *Surg Oncol Clin N Am.* **2018**;27(2):401–412. doi:10.1016/j.soc.2017.11.011.
8. Ashouri K, Wong A, Mittal P, Torres-Gonzalez L, Lo JH, Soni S, Algaze S, Khoulakz T, Zhang W, Yang Y, et al. Exploring predictive and prognostic biomarkers in colorectal cancer: a comprehensive review. *Cancers (Basel).* **2024**;16(16):2796. doi:10.3390/cancers16162796.
9. Chauhan S, Sharma S. Recent approaches on molecular markers, treatment and novel drug delivery System used for the management of colorectal cancer: a comprehensive review. *Curr Pharm Biotechnol.* **2024**;25(15):1969–1985. doi:10.2174/0113892010270975231208113157.
10. Pathak PS, Chan G, Deming DA, Chee CE. State-of-the-art management of colorectal cancer: treatment advances and innovation. *Am Soc Clin Oncol Educ Book.* **2024**;44(3):e438466. doi:10.1200/EDBK_438466.
11. Yang G, Yu XR, Weisenberger DJ, Lu T, Liang G. A multi-Omics overview of colorectal cancer to address mechanisms of disease, metastasis, patient disparities and outcomes. *Cancers (Basel).* **2023**;15(11):2934. doi:10.3390/cancers15112934.
12. Alrushaid N, Khan FA, Al-Suhaimi E, Elaissari A. Progress and perspectives in colon cancer pathology, diagnosis, and treatments. *Diseases.* **2023**;11(4):148. doi:10.3390/diseases11040148.
13. Malki A, ElRuz RA, Gupta I, Allouch A, Vranic S, Al Moustafa AE. Molecular mechanisms of colon cancer progression and metastasis: recent insights and advancements. *Int J Mol Sci.* **2020**;22(1):130. doi:10.3390/ijms22010130.
14. Maurya NS, Mani A. Navigating molecular pathways: an update on drugs in colorectal cancer treatment. *Curr Top Med Chem.* **2023**;23(30):2821–2843. doi:10.2174/1568026623666230614165548.
15. Ebru Nur A. Therapeutic targeting of molecular pathways in colorectal cancer. *Exp Oncol.* **2022**;44(1):2–6. doi:10.32471/exp-oncology.2312-8852.vol-44-no-1.17455.
16. Koveitpour Z, Panahi F, Vakilian M, Peymani M, Seyed Forootan F, Nasr Esfahani MH, Ghaedi K. Signaling pathways involved in colorectal cancer progression. *Cell Biosci.* **2019**;9(1):97. doi:10.1186/s13578-019-0361-4.
17. Xu S, Liu D, Cui M, Zhang Y, Zhang Y, Guo S, Zhang H. Identification of hub genes for early diagnosis and predicting prognosis in colon adenocarcinoma. *Biomed Res Int.* **2022**;2022(1):1893351. doi:10.1155/2022/1893351.
18. Chen X, Cubillos-Ruiz JR. Endoplasmic reticulum stress signals in the tumour and its microenvironment. *Nat Rev Cancer.* **2021**;21(2):71–88. doi:10.1038/s41568-020-00312-2.
19. Oakes SA, Papa FR. The role of endoplasmic reticulum stress in human pathology. *Annu Rev Pathol Mech Dis.* **2015**;10(1):173–194. doi:10.1146/annurev-pathol-012513-104649.
20. Cheng M, Fan X, He M, Dai X, Liu X, Hong J, Zhang L, Liao L. Identification of an endoplasmic reticulum stress-related prognostic risk model with excellent prognostic and clinical value in oral squamous cell carcinoma. *Aging (Albany NY).* **2023**;15(19):10010–10030. doi:10.18632/aging.204983.
21. Wang G, Han J, Wang G, Wu X, Huang Y, Wu M, Chen Y. ERO1 α mediates endoplasmic reticulum stress-induced apoptosis via microRNA-101/EZH2 axis in colon cancer RKO and HT-29 cells. *Hum Cell.* **2021**;34(3):932–944. doi:10.1007/s13577-021-00494-3.
22. Dilly AK, Honick BD, Lee YJ, Bartlett DL, Choudry HA. Synergistic apoptosis following endoplasmic reticulum stress aggravation in mucinous colon cancer. *Orphanet J Rare Dis.* **2020**;15(1):211. doi:10.1186/s13023-020-01499-1.
23. Xu Y, Xie YM, Sun WS, Zi R, Lu H-Q, Xiao L, Gong K-M, Guo S-K. Exploration of an prognostic signature related to endoplasmic reticulum stress in colorectal adenocarcinoma and their response targeting immunotherapy. *Technol Cancer Res Treat.* **2023**;22:15330338231212073. doi:10.1177/15330338231212073.
24. Yang X, Zhang C, Yan C, Ma L, Ma J, Meng X. System analysis based on the ER stress-related genes identifies WFS1 as a novel therapy target for colon cancer. *Aging (Albany NY).* **2022**;14(22):9243–9263. doi:10.18632/aging.204404.
25. Qu Z, Chu J, Wu Y, Zhuang J, Liu J, Han S, Wu W, Han S. Classification of colorectal cancer subtypes based on endoplasmic reticulum stress. *Digestion.* **2024**;105(2):107–130. doi:10.1159/000535230.
26. Wang H, Li Z, Tao Y, Ou S, Ye J, Ran S, Luo K, Guan Z, Xiang J, Yan G, et al. Characterization of endoplasmic reticulum stress unveils ZNF703 as a promising target for colorectal cancer immunotherapy. *J Transl Med.* **2023**;21(1):713. doi:10.1186/s12967-023-04547-z.
27. Geng J, Guo Y, Xie M, Li Z, Wang P, Zhu D, Li J, Cui X. Characteristics of endoplasmic reticulum stress in colorectal cancer for predicting prognosis and developing treatment options. *Cancer Med.* **2023**;12(10):12000–12017. doi:10.1002/cam4.5874.
28. Huang H, Feng X, Feng Y, Peng Z, Jiao C, Chen H, Fu CR, Xu F, Wang Y, Su X, et al. Bone-targeting HUVEC-derived exosomes containing miR-503-5p for osteoporosis therapy. *ACS Appl Nano Mater.* **2023**;7(1):1156–1169. doi:10.1021/acsanm.3c05056.
29. Qiao E, Ye J, Huang K. An endoplasmic reticulum stress related signature for clinically predicting prognosis of breast cancer patients. *Hum Mol Genet.* **2024**. doi:10.1093/hmg/ddae170.
30. Wan L, Chen Z, Yang J, Wu G, Xu Y, Cui J, Zhao X. Identification of endoplasmic reticulum stress-related signature characterizes the tumor microenvironment and predicts prognosis in lung adenocarcinoma. *Sci Rep.* **2023**;13(1):19462. doi:10.1038/s41598-023-45690-3.
31. Zhang G, Sun J. Endoplasmic reticulum stress-related signature for predicting prognosis and immune features in hepatocellular carcinoma. *J Immunol Res.* **2022**;2022:1–18. doi:10.1155/2022/1366508.
32. Liu L, Yao D, Chen Z, Duan S. A comprehensive signature based on endoplasmic reticulum stress-related genes in predicting prognosis and immunotherapy response in melanoma. *Sci Rep.* **2023**;13(1):8232. doi:10.1038/s41598-023-35031-9.
33. Yuan Z, Wang Y, Xu S, Zhang M, Tang J. Construction of a prognostic model for colon cancer by combining endoplasmic reticulum stress responsive genes. *J Proteomics.* **2024**;309:105284. doi:10.1016/j.jprot.2024.105284.
34. Wang B, Yang J, Wu J, Hu X, Zhu J, Fang J, Han B, Zhou B. Identification and validation of endoplasmic reticulum stress-related genes that enhance immunotherapy in colon cancer. *Transl Cancer Res.* **2024**;13(7):3760–3770. doi:10.21037/tcr-23-2227.
35. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, et al. Tcgabiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **2016**;44(8):e71. doi:10.1093/nar/gkv1507.
36. Goldman MJ, Craft B, Hastie M, Repčeka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, et al. Visualizing and interpreting cancer genomics data via the xena platform. *Nat Biotechnol.* **2020**;38(6):675–678. doi:10.1038/s41587-020-0546-8.
37. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-Grimaldi M-C, Schiappa R, Guenot D, Ayadi M, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLOS Med.* **2013**;10(5):e1001453. doi:10.1371/journal.pmed.1001453.
38. Díez-Villanueva A, Sanz-Pamplona R, Solé X, Cordero D, Crous-Bou M, Guinó E, Lopez-Doriga A, Berenguer A, Aussó S, Paré-Brunet L, et al. COLONOMICs - integrative omics data of one hundred paired normal-tumoral samples from colon cancer patients [published correction appears in Sci data. **2022** Nov 18;9(1): 712]. *Sci Data.* **2022**;9(1):595. doi:10.1038/s41597-022-01697-5.

39. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R. NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res.* 2007;35(Database issue):D760–D765. doi:10.1093/nar/gkl887.
40. Davis S, Meltzer PS. Geoquery: a bridge between the gene expression omnibus (GEO) and BioConductor. *Bioinformatics.* 2007;23(14):1846–1847. doi:10.1093/bioinformatics/btm254.
41. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinf.* 2016;54(1):1.30.1–1.30.33. doi:10.1002/cpbi.5.
42. Engebretsen S, Bohlén J. Statistical predictions with glmnet. *Clin Epigenet.* 2019;11(1):123. doi:10.1186/s13148-019-0730-1.
43. Cai W, van der Laan M. Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator [published online ahead of print, 2020 Aug 10]. *Int J Biostat.* 2020;16(2). doi:10.1515/ijb-2017-0070.
44. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545–15550. doi:10.1073/pnas.0506580102.
45. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 2011;27(12):1739–1740. doi:10.1093/bioinformatics/btr260.
46. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* 2013;14(1):7. doi:10.1186/1471-2105-14-7.
47. Yu G. Gene Ontology semantic similarity analysis using GOSemSim. *Methods Mol Biol.* 2020;2117:207–215. doi:10.1007/978-1-0716-0301-7_11.
48. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000;28(1):27–30. doi:10.1093/nar/28.1.27.
49. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–287. doi:10.1089/omi.2011.0118.
50. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* 2003;31(1):258–261. doi:10.1093/nar/gkg034.
51. Zhang Q, Liu W, Zhang HM, Xie G-Y, Miao Y-R, Xia M, Guo A-Y. hTfTarget: a comprehensive database for regulations of human transcription factors and their targets. *Genomics Proteomics Bioinf.* 2020;18(2):120–128. doi:10.1016/j.gpb.2019.09.006.
52. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 2014;42(Database issue):D92–D97. doi:10.1093/nar/gkt1248.
53. Park SY. Nomogram: an analogue tool to deliver digital knowledge. *J Thorac Cardiovasc Surg.* 2018;155(4):1793. doi:10.1016/j.jtcvs.2017.12.107.
54. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, Roobol MJ, Steyerberg EW. Reporting and interpreting decision curve analysis: a Guide for investigators. *Eur Urol.* 2018;74(6):796–804. doi:10.1016/j.eururo.2018.08.038.
55. Tataranni T, Piccoli C. Dichloroacetate (DCA) and cancer: an overview towards clinical applications. *Oxid Med Cell Longev.* 2019;2019:1–14. doi:10.1155/2019/8201079.
56. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 2017;18(1):248–262. doi:10.1016/j.celrep.2016.12.019.
57. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, et al. Systematic RNA interference reveals that oncogenic kras-driven cancers require TBK1. *Nature.* 2009;462(7269):108–112. doi:10.1038/nature08460.
58. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243–259. doi:10.1007/978-1-4939-7493-1_12.
59. Zhang W, Shi Y, Oyang L, Cui S, Li S, Li J, Liu L, Li Y, Peng M, Tan S, et al. Endoplasmic reticulum stress—a key guardian in cancer. *Cell Death Discov.* 2024;10(1):343. doi:10.1038/s41420-024-02110-3.
60. Yadav RK, Chae SW, Kim HR, Chae HJ. Endoplasmic reticulum stress and cancer. *J Cancer Prev.* 2014;19(2):75–88. doi:10.15430/JCP.2014.19.2.75.
61. Xu D, Liu Z, Liang MX, Fei Y-J, Zhang W, Wu Y, Tang J-H. Endoplasmic reticulum stress targeted therapy for breast cancer. *Cell Commun Signal.* 2022;20(1):174. doi:10.1186/s12964-022-00964-7.
62. Oakes SA. Endoplasmic reticulum stress signaling in cancer cells. *Am J Pathol.* 2020;190(5):934–946. doi:10.1016/j.ajpath.2020.01.010.
63. Moenner M, Pluquet O, Bouchecareilh M, Chevet E. Integrated endoplasmic reticulum stress responses in cancer. *Cancer Res.* 2007;67(22):10631–10634. doi:10.1158/0008-5472.CAN-07-1705.
64. Urrea H, Dufey E, Avril T, Chevet E, Hetz C. Endoplasmic reticulum stress and the hallmarks of cancer. *Trends Cancer.* 2016;2(5):252–262. doi:10.1016/j.trecan.2016.03.007.
65. Zhang TA, Zhang Q, Zhang J, Zhao R, Shi R, Wei S, Liu S, Zhang Q, Wang H. Identification of the role of endoplasmic reticulum stress genes in endometrial cancer and their association with tumor immunity. *BMC Med Genomics.* 2023;16(1):261. doi:10.1186/s12920-023-01679-5.
66. Clarke HJ, Chambers JE, Liniker E, Marciniak SJ. Endoplasmic reticulum stress in malignancy. *Cancer Cell.* 2014;25(5):563–573. doi:10.1016/j.ccr.2014.03.015.
67. Mohamed E, Cao Y, Rodriguez PC. Endoplasmic reticulum stress regulates tumor growth and anti-tumor immunity: a promising opportunity for cancer immunotherapy. *Cancer Immunol Immunother.* 2017;66(8):1069–1078. doi:10.1007/s00262-017-2019-6.
68. Liu K, Fang W, Sun E, Chen Y. Roles of endoplasmic reticulum stress and apoptosis signaling pathways in gynecologic tumor cells: a systematic review. *Oncol Transl Med.* 2017;3(3):131–135. doi:10.1007/s10330-016-0201-1.
69. Song Y, Ma J, Fang L, Tang M, Gao X, Zhu D, Liu W. Endoplasmic reticulum stress-related gene model predicts prognosis and guides therapies in lung adenocarcinoma. *BMC Bioinf.* 2023;24(1):255. doi:10.1186/s12859-023-05384-z.
70. Zhang TA, Zhang Q, Zhang J, Zhao R, Shi R, Wei S, Liu S, Zhang Q, Wang H. Identification of the role of endoplasmic reticulum stress genes in endometrial cancer and their association with tumor immunity. *BMC Med Genomics.* 2023;16(1):261. doi:10.1186/s12920-023-01679-5.
71. Tierney KE, Ji L, Dralla SS, Yoo E, Yessaian A, Pham HQ, Roman L, Spoto R, Mhawech-Fauceglia P, Lin YG. Endoplasmic reticulum stress in complex atypical hyperplasia as a possible predictor of occult carcinoma and progesterone response. *Gynecol Oncol.* 2016;143(3):650–654. doi:10.1016/j.ygyno.2016.10.015.
72. Nguyen DT, Le TM, Hattori T, Takarada-Iemata M, Ishii H, Roboon J, Tamatani T, Kannon T, Hosomichi K, Tajima A, et al. The ATF6 β -calreticulin axis promotes neuronal survival under endoplasmic reticulum stress and excitotoxicity. *Sci Rep.* 2021;11(1):13086. doi:10.1038/s41598-021-92529-w.
73. Park TJ, Kim JH, Pasaje CF, Park B-L, Bae JS, Uh S-T, Kim Y-H, Kim M-K, Choi IS, Choi BW, et al. Polymorphisms of ATF6B are potentially associated with FEV1 decline by aspirin provocation in asthmatics. *Allergy Asthma Immunol Res.* 2014;6(2):142–148. doi:10.4168/aaair.2014.6.2.142.
74. Li M, Thorne RF, Shi R, Zhang XD, Li J, Li J, Zhang Q, Wu M, Liu L. DDIT3 directs a dual mechanism to balance glycolysis and oxidative phosphorylation during glutamine deprivation. *Adv Sci (Weinh).* 2021;8(11):e2003732. doi:10.1002/advs.202003732.
75. Yu Y, Liu DY, Chen XS, Zhu L, Wan LH. MANF: a novel endoplasmic reticulum stress response protein—the role in

- neurological and metabolic disorders. *Oxid Med Cell Longev*. 2021;2021(1):6467679. doi:[10.1155/2021/6467679](https://doi.org/10.1155/2021/6467679).
76. Gong T, Liu Y, Tian Z, Zhang M, Gao H, Peng Z, Yin S, Cheung CW, Liu Y. Identification of immune-related endoplasmic reticulum stress genes in sepsis using bioinformatics and machine learning. *Front Immunol*. 2022;13:995974. doi:[10.3389/fimmu.2022.995974](https://doi.org/10.3389/fimmu.2022.995974).
77. Vargas MF, Tapia-Pizarro AA, Henríquez SP, Quezada M, Salvatierra AM, Noe G, Munroe DJ, Velasquez LA, Croxatto HB. Effect of single post-ovulatory administration of levonorgestrel on gene expression profile during the receptive period of the human endometrium. *J Mol Endocrinol*. 2012;48(1):25–36. doi:[10.1530/JME-11-0094](https://doi.org/10.1530/JME-11-0094).
78. Pan Y, Chen H, Siu F, Kilberg MS. Amino acid deprivation and endoplasmic reticulum stress induce expression of multiple activating transcription factor-3 mRNA species that, when overexpressed in HepG2 cells, modulate transcription by the human asparagine synthetase promoter. *J Biol Chem*. 2003;278(40):38402–38412. doi:[10.1074/jbc.M304574200](https://doi.org/10.1074/jbc.M304574200).
79. Rozpedek W, Pytel D, Mucha B, Leszczynska H, Diehl JA, Majsterek I. The role of the PERK/eIF2 α /ATF4/CHOP signaling pathway in tumor progression during endoplasmic reticulum stress. *Curr Mol Med*. 2016;16(6):533–544. doi:[10.2174/1566524016666160523143937](https://doi.org/10.2174/1566524016666160523143937).
80. Rana SVS. Endoplasmic reticulum stress induced by toxic elements-a review of Recent developments. *Biol Trace Elem Res*. 2020;196(1):10–19. doi:[10.1007/s12011-019-01903-3](https://doi.org/10.1007/s12011-019-01903-3).
81. Akman M, Belisario DC, Salaroglio IC, Kopecka J, Donadelli M, De Smaele E, Riganti C. Hypoxia, endoplasmic reticulum stress and chemoresistance: dangerous liaisons. *J Exp Clin Cancer Res*. 2021;40(1):28. doi:[10.1186/s13046-020-01824-3](https://doi.org/10.1186/s13046-020-01824-3).
82. Krebs J, Agellon LB, Michalak M. Ca(2+) homeostasis and endoplasmic reticulum (ER) stress: an integrated view of calcium signaling. *Biochem Biophys Res Commun*. 2015;460(1):114–121. doi:[10.1016/j.bbrc.2015.02.004](https://doi.org/10.1016/j.bbrc.2015.02.004).
83. Hsu HS, Liu CC, Lin JH, Hsu T-W, Hsu J-W, Su K, Hung S-C. Involvement of ER stress, PI3K/AKT activation, and lung fibroblast proliferation in bleomycin-induced pulmonary fibrosis. *Sci Rep*. 2017;7(1):14272. doi:[10.1038/s41598-017-14612-5](https://doi.org/10.1038/s41598-017-14612-5).
84. Zhang Y, Cheng D, Jie C, Liu T, Huang S, Hu S. Leptin alleviates endoplasmic reticulum stress induced by cerebral ischemia/reperfusion injury via the PI3K/Akt signaling pathway. *Biosci Rep*. 2022;42(12):BSR20221443. doi:[10.1042/BSR20221443](https://doi.org/10.1042/BSR20221443).