RESEARCH

Evaluation of vital genes correlated with CD8 + T cell infiltration as prognostic biomarkers in stomach adenocarcinoma

Dun Pan[1†], Hui Chen[1†], Jiaxiang Xu[1], Xin Lin[1] and Liangqing Li[1*]

Abstract

Background Infiltration of CD8 + T cells in the tumor microenvironment is correlated with better prognosis in various malignancies. Our study aimed to investigate vital genes correlated with CD8 + T cell infiltration in stomach adenocarcinoma (STAD) and develop a new prognostic model.

Methods   Using the STAD dataset, differentially expressed genes (DEGs) were analyzed, and co-expression networks were constructed. Combined with the CIBERSORT algorithm, the most relevant module of WGCNA with CD8 + T cell infiltration was selected for subsequent analysis. The vital genes were screened out by univariate regression analysis to establish the risk score model. The expression of the viral genes was verified by lasso regression analysis and in vitro experiments.

Results Four CD8 + T cell infiltration-related genes (CIDEC, EPS8L3, MUC13, and PLEKHS1) were correlated with the prognosis of STAD. Based on these genes, a risk score model was established. We found that the risk score could well predict the prognosis of STAD, and the risk score was positively correlated with CD8 + T cell infiltration. The validation results of the gene expression were consistent with TCGA. Furthermore, the risk score was significantly higher in tumor tissues. The high-risk group had poorer overall survival (OS) in each subgroup.

Conclusions Our study constructed a new risk score model for STAD prognosis, which may provide a new perspective to explore the tumor immune microenvironment mechanism in STAD.

Keywords Stomach adenocarcinoma, CD8 + T cells, Immune infiltration, Tumor microenvironment, Prognostic biomarkers

Introduction

 Gastric cancer (GC) is the third leading cause of cancer death worldwide [1–3]. According to latest cancer statistics information, 1,958,310 new cancer cases and 609,820

†Dun Pan and Hui Chen contributed equally to this work.

*Correspondence:
Liangqing Li
liliangqing@fjmu.edu.cn
[1] Department of Gastrointestinal Surgery, The First Affiliated Hospital of Fujian Medical University, No.20, ChaZhong Road, TaiJiang District, Fuzhou 350000, Fujian Province, China

cancer deaths are projected to occur in the United States [4]. The vast majority (about 90%) of GCs are adenocarcinomas arising from the glands in the most superficial or mucosal layers of the stomach [5]. The early symptoms of gastric adenocarcinoma (STAD) are not obvious, but the late symptoms are poor digestion, anorexia nervosa and celialgia [6]. Despite major advances in the treatment of STAD, the prognosis of STAD patients remains unsatisfactory [7–9]. In recent years, the development of immunotherapy has a significant impact on the treatment of STAD [10]. It is worth noting that the immunotherapy effect mainly depends on the immune response

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 2 of 17

[11], which is markedly affected by the tumor microenvironment [12]. CD8 + T cells are core effector cells in the tumor microenvironment, and highly infiltrating CD8 + T cells have prognostic value in most tumors [13]. However, the mechanism of CD8 + T cell infiltration in the STAD tumor microenvironment remains unclear. Therefore, identification of novel biomarkers correlated with CD8 + T cell infiltration is helpful to explore the mechanisms of immune infiltration in STAD.

There are various deconvolution methods for quantifying the cellular composition of immune cells, such as xCell, TIMER, EPIC, QuanTIseq and CIBERSORT. xCell, a novel gene signature-based method, can be used to infer 64 immune and stromal cell types [14]. TIMER was able to estimate the abundance of 6 immune cell types in 32 cancer types [15]. EPIC is to estimate the proportion of immune and cancer cells from bulk tumor gene expression data, and can accurately detect all major cell types in tumors [16]. QuanTIseq is a method to quantify the fractions of ten immune cell types from bulk RNA-sequencing data [17]. CIBERSORT is a common method for calculating immune cell infiltration, providing expression data for 22 common immune infiltrating cells [18]. CIBERSORT consistently outperformed other methods in some cases, and unknown content and lab-specific factors had little effect on CIBERSORT performance [19]. It was successfully applied to estimate the level of immune cell infiltration in various cancers, including renal cell carcinoma [20], colon cancer [21]. Therefore, CIBERSORT was selected in our study to quantify the cellular composition of immune cells.

In our current research, we obtained differentially expressed genes (DEGs) in the Cancer Genome Atlas (TCGA) database. The important modules and genes associated with CD8 + T cell infiltration were identified with WGCNA and CIBERSORT methods. The risk score model was constructed by univariate regression analysis and LASSO regression analysis. And the expression levels of these genes in risk score model were verified in validation set and in vitro experiment. This may provide a theoretical basis for the exploration of prognostic biomarkers of STAD.

Materials and methods
Data set sources and preprocessing
All data in this study were obtained from the TCGA and GEO database. In TCGA database, STAD expression and clinical data from 375 STAD patients and 32 adjacent normal tissue samples were downloaded from UCSC Xena. After excluding samples with overall survival (OS) less than 30 days, 335 STAD samples were ultimately included in the TCGA cohort. TCGA dataset was used to screen vital genes and the construct the

prognostic model. The GSE26899 dataset and GSE29272 dataset were downloaded from the GEO database for the expression validation of vital genes. The GSE26899 dataset included 96 tumor tissues and 12 adjacent normal tissue samples, and the GSE29272 dataset included 134 tumor tissues and 134 adjacent normal tissue samples. GSE84437 dataset was downloaded from the GEO database to verify the prognostic ability of the constructed model, which included 431 tumor samples. Among them, the TCGA dataset was regarded as the training set, and the GEO dataset was regarded as the validation set.

Identification of DEGs
DEGs were identified using the limma package in R, and visualized with volcano plots and heatmaps. False discovery rate (FDR) < 0.05 and $|\log_2$ Fold Change (FC)| >2 were considered as the criteria for identifying DEGs. Heatmaps and volcano plots were drawn by the R packages "pheatmap" and "ggplot", respectively.

WGCNA
WGCNA is a typical phylogenetic algorithm to describe the correlation patterns between gene expression profiles and build gene co-expression networks. The DEGs co-expression network analysis was performed by the R package "WGCNA", and a scale-free gene co-expression network was built. First, outliers were detected by clustering sample data with the "hclust" function. Then, the "pickSoftThreshold" function was applied to select a suitable soft-threshold power regulator to construct a scale-free topology with a soft threshold of 5. The adjacency matrix was calculated from this value, which was transformed into a topological overlap matrix (TOM) and the corresponding dissimilarity matrix (1-TOM). Genes were clustered by the mean linkage hierarchical clustering method. The minimum number of genes per gene network module was set to 30 based on the criteria of the hybrid dynamic clipping tree method. The signature genes of each module were calculated in turn, and the modules were clustered.

Analysis of immune infiltration and functional enrichment
The proportion of immune cells in the sample was calculated by "CIBERSORT". Correlations between WGCNA module genes and T cell subtypes were calculated using Pearson's test. The modules most significantly associated with CD8 + T cells were used for subsequent analyses. Enrichment analysis of genes in the modules most significantly associated with CD8 + T cells was performed by the R package "clusterProfiler" with a threshold of $p <$ 0.05.

Pan *et al. BMC Gastroenterology*   *(2023) 23:399*

Page 3 of 17

Table 1 Clinical Information

Number	Age (years)	Gender	Height (cm)	Weight (Kg)	BMI	Clinical stages			Stage	Grade	Tumor metastasis	Alcohol history	Smoking history	Family history
						T	N	M						
1	54	Female	158	61	24.44	T4a	N2	M0	IIIB	G3	No	No	No	No
2	72	Male	164	50	18.59	pT4a	N0	M0	?B	G3	No	Yes	Yes	No
3	68	Male	165	56	20.57	T4a	N3	M0	IIIC	G3	No	No	No	No
4	68	Male	167	60	21.51	T2	N0	M0	IB	G2	No	No	No	No
5	72	Female	150	47.5	21.11	T4a	N0	M0	?B	G2	No	No	No	No
6	68	Male	162	58	22.10	T1a	N0	M0	IA	G1	No	No	No	No
7	68	Male	170	70	24.22	T3	N0	M0	?A	G3	No	No	No	No
8	61	Female	155	42	17.48	T1a	N0	M0	IA	G4	No	No	No	No
9	59	Male	168	60	21.25	T3	N2	M0	IIIA	G3	No	No	No	No

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 4 of 17

Construction and verification of prognostic model

The genes significantly associated with CD8 + T cells in the module were analyzed by univariate Cox regression, and the prognostic genes were screened ($p < 0.05$). The prognostic model was further constructed by LASSO Cox regression analysis. The risk score was calculated as follows: risk score = (ßA × gene A expression) + (ßB × gene B expression) + (ßN × gene N expression). The genes used to construct the risk score were defined as vital genes. The median risk score served as a cut-off point to divide patients into high- and low-risk groups. The risk score was also calculated in the GSE84437 validation set. OS curves of risk score were analyzed by Kaplan-Meier, and the accuracy of the model was verified using the time-dependent ROC curve. Additionally, univariate Cox regression analysis and multivariate Cox regression analysis were used to identify whether the risk score was an independent prognostic factor for OS in patients with STAD.

Expression and verification of vital genes

The expression of genes in prognostic model was analyzed in the TCGA validation set by rank sum test. It was displayed by boxplot. Furthermore, we also validated the expression of genes in the GSE26899 and GSE29272 validation sets. The Human Protein Atlas (HPA) (https://www.proteinatlas.org/) provided immunohistochemical (IHC) results of genes in adjacent normal tissues and tumor tissues.

In vitro expression validation of vital genes by real time polymerase chain reaction (RT-PCR)

Nine STAD patients were recruited from First Affiliated Hospital of Fujian Medical University. Tumor tissues and adjacent normal tissues were collected from 9 STAD patients. Clinical information of individuals was displayed in Table 1, mainly including the age, gender, stage, grade, drinking, smoking and family history.

Inclusion criteria of STAD patients: (1) The patients were initially diagnosed with STAD; (2) The patients did not undergo other therapy before diagnosis; (3) The patients had no other malignant tumor; (4) The patients had no other autoimmune diseases; (5) The patients were 18 to 70 years old. The exclusion criteria of STAD patients: (1) Patients had other malignancy; (2) Patients received other treatment before surgery; (3) Patients had incomplete clinical data; (4) Patients had a history of STAD; (5) Patients with recurrence. The Ethics Committee of First Affiliated Hospital of Fujian Medical University approved this study (2,020,219). Informed consent of patients and their families was obtained.

Total RNA was extracted from tissue samples using TRIzol® Reagent. FastKing cDNA first-strand synthesis kit (KR116) was used for reverse transcription of mRNA, and Gene-9660 fluorescence quantitative PCR instrument was used for relative quantitative analysis of data by $2^{-??ct}$ method. GAPDH and ACTB was used for internal reference genes. The primer information for RT-PCR is shown in Table 2.

Subgroup analyses to evaluate model performance

Risk score for different subgroups of tumor (T), node (N), metastasis (M), and grade (G) was assessed in the TCGA and GEO datasets to test the performance of the model. In addition, the Kaplan-Meier method and the log-rank test were applied to assess the ability of prognosis prediction in different subgroups. $p < 0.05$ was statistically significant.

Evaluation of drug therapy and risk score

To assess the response of risk score for drug therapy, the "pRRophetic" R package was applied to calculate the half-maximal inhibitory concentration (IC50) of samples. Based on clinical recommendations, molecular drugs such as Axitinib, Dasatinib, Etoposide, Midostaurin, Pyrimethamine and Sunitinib were selected as drug candidates. IC50 was compared between high- and low-risk groups by Wilcoxon signed-rank test.

Statistical analysis

All statistics were performed with R software. "Limma" and "WGCNA" were applied to screen genes correlated with CD8 + T cells. Univariate Cox regression analysis was performed by the "survival" and "survminer" package. Lasso analysis and model construction were performed by the "glmnet" software package. The Wilcox test was applied to determine statistical differences. Kaplan-Meier curves were plotted and log-rank was applied to test for significant differences in OS among groups. ROC analysis was applied to assess the prognostic performance of risk score. ROC-AUC was an indicator for judging the accuracy of prognosis. $P < 0.05$ was statistically significant in all analyses.

Results

Identification of DEGs and WGCNA

In total, 1013 DEGs were identified, including 700 up-regulated genes and 313 down-regulated genes (Fig. 1A). In addition, the heatmap of the top 100 DEGs is shown in the Fig. 1B. The 1013 DEGs were subjected to construct a weighted gene co-expression network. First, the samples were clustered by the mean linking method in WGCNA (Fig. 2A). 120 was chosen as the cutting tree height to remove outliers (red line). After re-clustering, the number of samples below the red line was 406. The dendrogram and rating feature heatmap for the 406 samples

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 5 of 17

Table 2 Primer sequence in the RT-PCR

Primer name	Primer sequence (5' to 3')
GAPDH-F (Internal reference)	GGAGCGAGATCCCTCCAAAAT
GAPDH-R (Internal reference)	GGCTGTTGTCATACTTCTCATGG
ACTB-F (Internal reference)	CATGTACGTTGCTATCCAGGC
ACTB-R (Internal reference)	CTCCTTAATGTCACGCACGAT
CIDEC-F	TTCCCCAGTGAAGGACTGACT
CIDEC-R	GACCAGTCTGGATGGGCTAAG
EPS8L3-F	CAGAAGCTGTTCGAGATGGATG
EPS8L3-R	GCTGTCTAGGCGGTAAGAGTC
MUC13-F	AGCGCTTGTCAGAGAGGTG
MUC13-R	ACCTCCACAGTTGATGCGT
PLEKHS1-F	GACGTGGTGGTTCATGCCT
PLEKHS1-R	CATGTGCCACAATGCCCAG

in the study are shown in Fig. 2B. ß = 5 was chosen as the soft threshold to construct the scale-free network (Fig. 2C), and 7 modules were confirmed (Fig. 2D).

Identification of key modules

The proportions of 7 T cell subtypes including CD8 + T cells in the samples were calculated by CIBERSORT and their association with the WGCNA module were analyzed. The highest correlation was identified between genes in the yellow module (65 genes) and CD8 + T cells. Therefore, genes in yellow module were selected for subsequent analysis (Fig. 3).

Functional enrichment analysis of genes

Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) function enrichment analysis were performed on the 65 genes in the yellow module using the R package "clusterProfiler" at a screening criteria of $p < 0.05$. Top 15 of biological process (BP), cellular component (CC) and molecular function (MF) analyzed by GO enrichment are shown in Fig. 4A-C, respectively. DEGs were mostly enriched in the BP of O-glycan processing and protein O-linked glycosylation, and CC of apical part of cell. KEGG analysis indicated that DEGs were mostly enriched in maturity onset diabetes of the young, retinol metabolism and GC. The top 5 enriched pathways were used to construct a network. From the network, regenerating family member 4 (REG4), caudal type homeobox 2 (CDX2), cadherin 17 (CDH17) were supposed to have a direct relationship with GC (Fig. 4D).

Construction and verification of a prognostic model

After univariate Cox regression analysis of 65 genes, 4 prognostic genes were obtained ($p < 0.05$) (Fig. 5A). Afterwards, a four-gene risk model consisting of CIDEC, EPS8L3, MUC13 and PLEKHS1 was constructed (Fig. 5B). Risk score was calculated for each tumor sample using the coefficients obtained by the LASSO algorithm. The formula was as follows: risk score = CIDEC * 0.104 + PLEKHS1* (-0.035) + MUC13 * (-0.059) + EPS8L3 * (-0.052). STAD patients were divided into high- and low-risk groups based on a median risk score. Patients in the low-risk group had a higher survival proportion
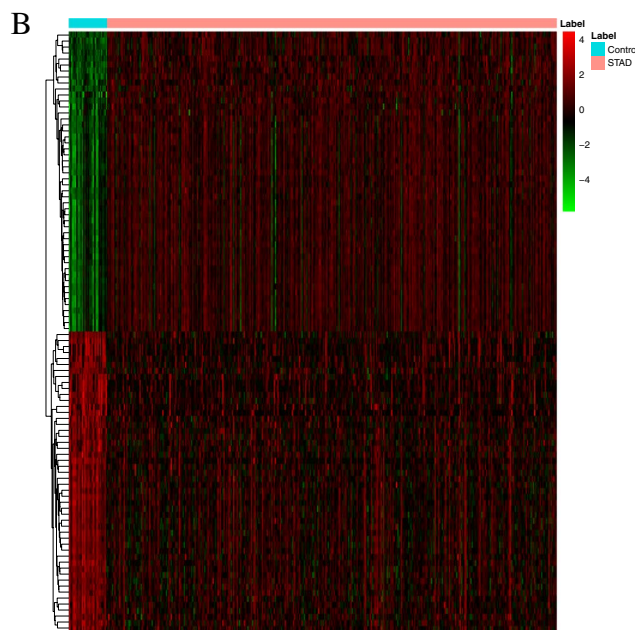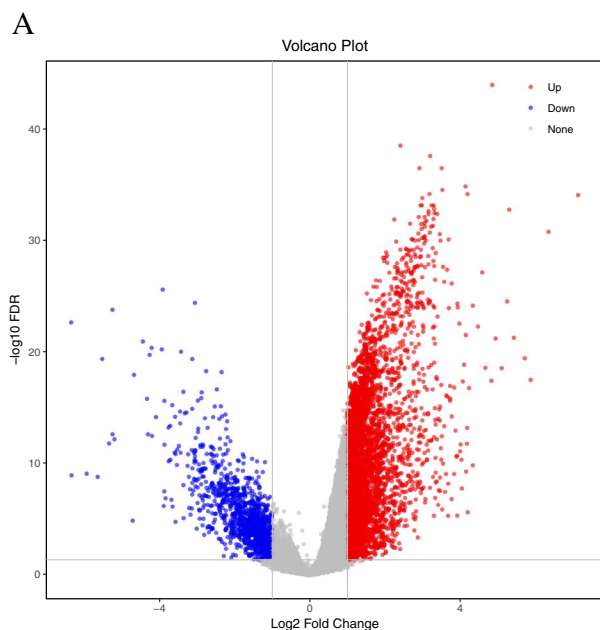
Fig. 1 Volcano plot (**A**) and heat map of top 100 DEGs (**B**)

Pan *et al. BMC Gastroenterology* (2023) 23:399

Page 6 of 17
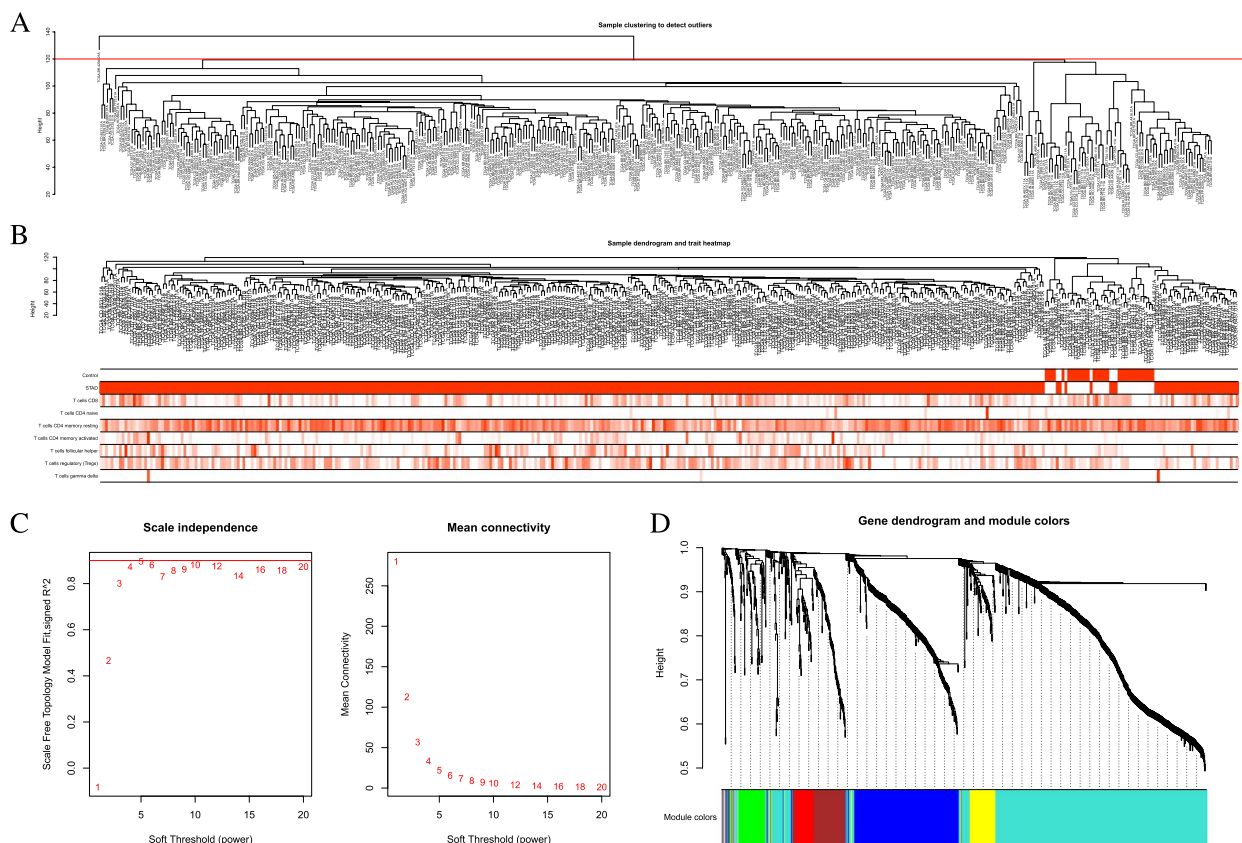
Fig. 2 Weighted Gene Co-expression Network Analysis (WGCNA) analysis. **A** Sample clustering to detect outliers; **B** Sample dendrogram and scoring feature heatmap; **C** Soft threshold screening: **D** Genes clustering

(Fig. 5C). Kaplan-Meier survival analysis showed significantly lower OS in the low-risk group patients (Fig. 5D). To evaluate the predictive efficiency of the model in 1-, 3-, and 5-year survival, ROC curves were performed on the training set, indicating that the prognostic model has good sensitivity and specificity (Fig. 5E).

To further assess the robustness of the prognostic model constructed by the four vital genes, the risk score was calculated in the GSE84437 validation set. Patients with low-risk score had longer OS, which was in line with the results in training set (Fig. 5F). Kaplan-Meier survival analysis showed that the OS of patients in the low-risk group was lower than that in the high-risk group (Fig. 5G). Survival prediction by risk score was assessed by ROC analysis at 1, 3 and 5 years (Fig. 5H).

Expression and verification of vital genes

In addition, the relative expression of four vital genes between tumor tissues and adjacent normal tissues was explored by the TCGA dataset and further validated by the GSE26899 and GSE29272 dataset. CIDEC was significantly down-regulated in tumor tissues, while EPS8L3, MUC13 and PLEKHS1 were significantly highly

expressed in tumor tissues (Fig. 6A-C). In addition, RT-PCR results displayed that CIDEC was significantly reduced in tumor tissues, MUC13 was significantly increased in tumor tissues, EPS8L3 showed an up-regulated tendency, while PLEKHS1 was down-regulated in tumor tissues (Fig. 6D). We speculated that the small sample size may cause this difference. IHC further confirmed the differential expression of CIDEC, EPS8L3 and MUC13 in tumor tissues and adjacent normal tissue. The expression of CIDEC in tumor tissue was significantly lower than that in adjacent normal tissues (Fig. 6E), while the expression of EPS8L3 and MUC13 was significantly higher in tumor tissues (Fig. 6F-G).

Risk score was an Independent prognostic factor for STAD

The correlation analysis shows that the risk score was positively connected with CD8 + T cell infiltration (Fig. 7A). Univariate analysis revealed that, except for gender and G stage, the other factors were correlated with OS (Fig. 7B). Incorporating these factors into multivariate analysis, risk score remained significantly related to OS (Fig. 7C). These results indicate that risk score was an independent prognostic factor for STAD.

Pan *et al. BMC Gastroenterology* (2023) 23:399

Page 7 of 17

Module−trait relationships
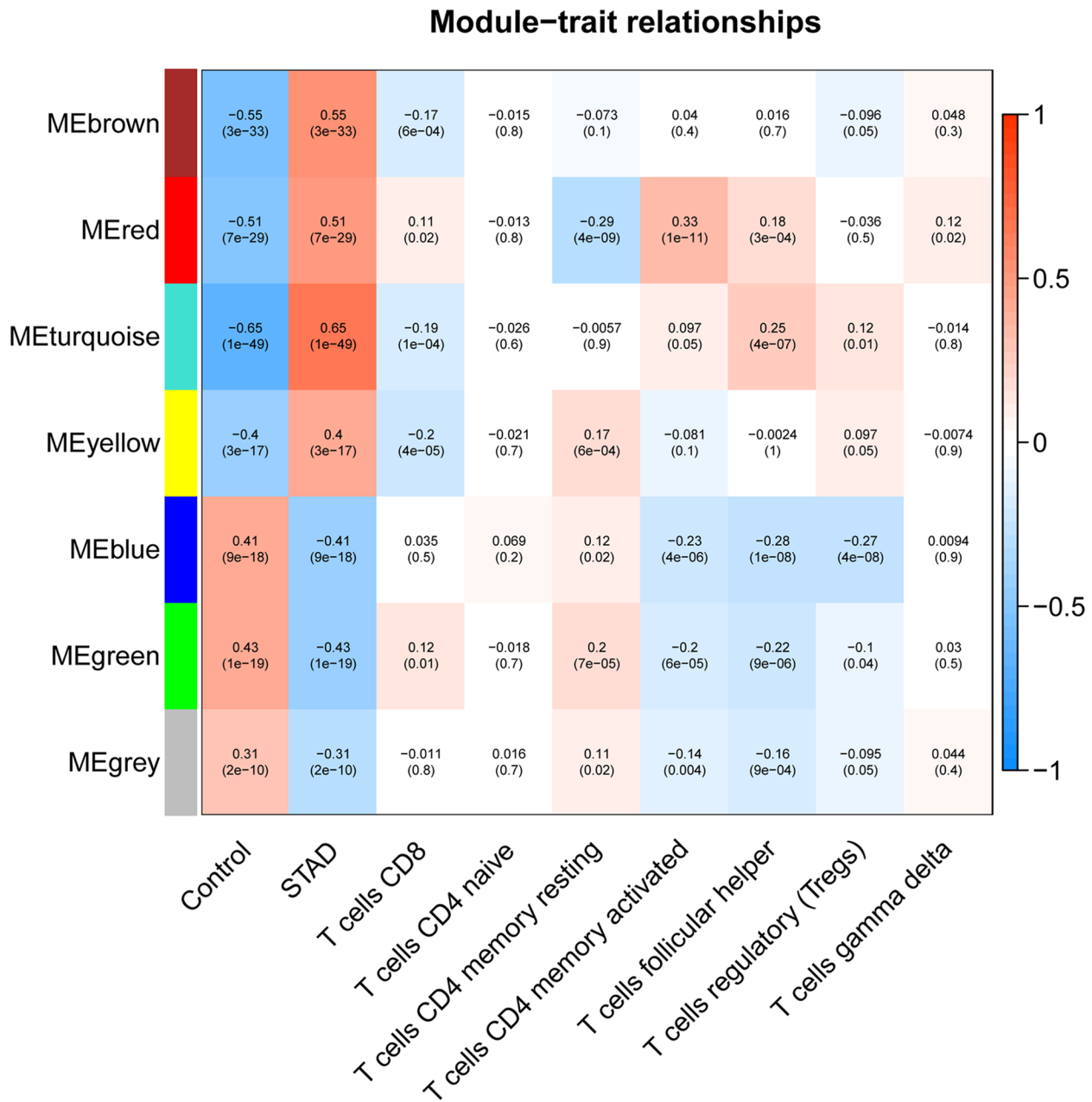
Fig. 3 Correlation of T cell subtypes infiltration with different gene modules

Correlation analysis of vital genes and immune cell subtypes

The correlation analysis of vital genes and CD8 + T cells was performed based on the TIMER2.0 database. With XCELL algorithm, TIMER algorithm, CIBER-SORT algorithm and EPIC algorithm, CIDEC, EPS8L3, and MLC13 were remarkable negatively correlation with CD8 + T cells, but PLEKHS1 were positively cor-relation with CD8 + T cells (Fig. 8).

Subgroup analysis

Kaplan-Meier curves were plotted for subgroups including age, gender, clinical stages in the training set. In the clinical indicators such as age = 60, T3 + T4, the high-risk groups have poor survival prognosis (Fig. 9). Survival was significantly different in age = 60 and T3 + T4 subgroups between high- and low-risk groups in GSE84437 validation set (Fig. 10). The result

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 8 of 17
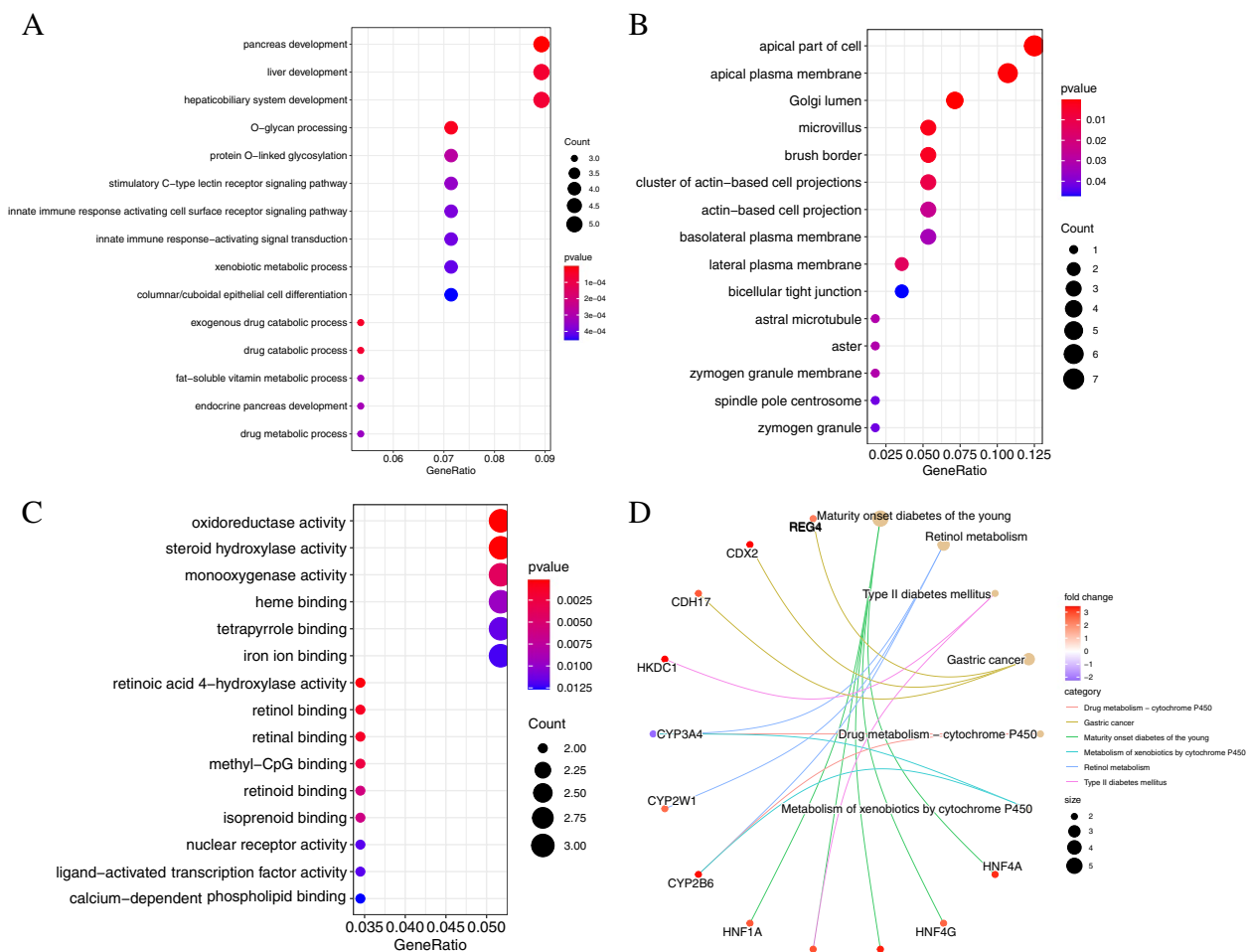
Fig. 4 GO and KEGG enrichment analysis of genes in yellow module. **A** Biological process (BP); **B** Cellular component (CC); **C** Molecular function (MF); **D** KEGG enrichment analysis (https://www.kegg.jp/kegg/kegg1.html). Appropriate copyright permission' to use the signalling pathways was obtained

suggests that risk score model could identify outcomes in different subgroups of patients.

Evaluation of drug therapy and risk score

The IC50 of Axitinib, Dasatinib, Midostaurin, and Sunitinib was higher in the low-risk group and the IC50 of Etoposide and Pyrimethamine was lower in low-rish group based on the evaluation of risk score model and drug treatment sensitivity (Fig. 11). In addition, the IC50 of Cisplatin and Docetaxel, which are commonly used in the treatment of gastric adenocarcinoma, was higher in the low-rish group in GEO dataset (Supplementary Fig. 1). These results suggest that risk score has potential predictive value for chemotherapy and targeted therapies.

Discussion

STAD is a heterogeneous malignancy with high risk of locoregional recurrence and distant metastasis after surgery [22, 23]. In recent years, increasing studies have explored the prognostic biomarkers of STAD. For instance, Jiang et al. mined the differentially expressed genes (DEGs) in the early stage of STAD and constructed a prognostic signature with 10 early-stage specific mRNAs [24]. Shen et al. constructed a STAD prognosis prediction model based on 8 chemotherapy-related characteristic genes to provide a reference for the treatment and prognosis improvement of STAD patients [25]. Immunotherapy is increasingly recognized for its potential therapeutic effects on various tumors [26, 27]. CD8 + T cells are the central effector cells of anti-tumor immunity [28]. Identification of vital

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 9 of 17
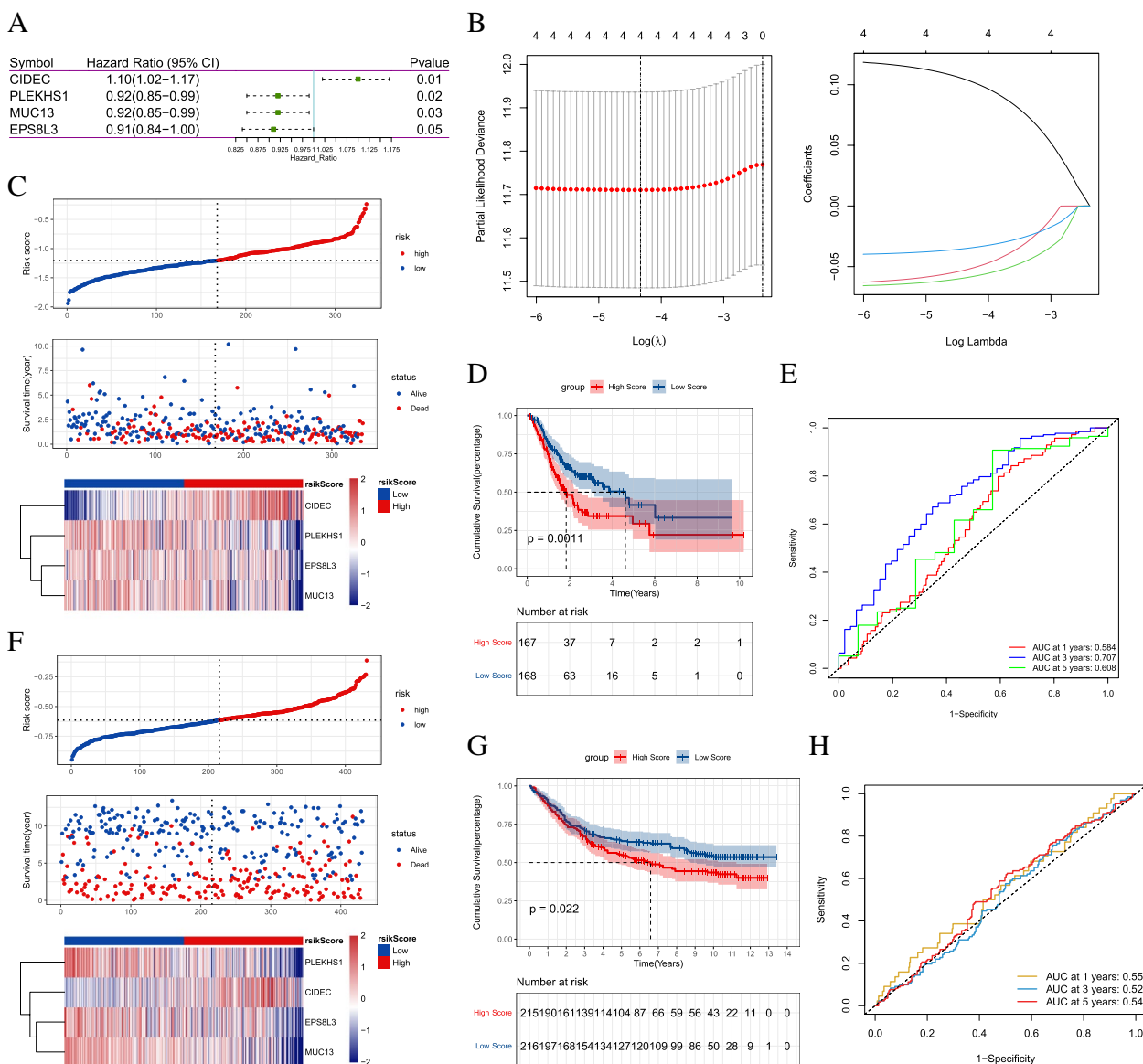
Fig. 5 Construction and validation of a prognostic model of CD8 + T cell infiltration-related genes. **A** Screening of vital genes; **B** LASSO Cox regression analysis; **C** Distribution map of risk score in TCGA training set; **D** Survival Curve in TCGA training set; **E** ROC curve in TCGA training set; **F** Distribution map of risk score in GSE84437 validation set; **G** Survival curve in GSE84437 validation set; **H** ROC curve in GSE84437 validation set

genes associated with CD8 + T cell infiltration may provide novel ideas for the prevention and treatment of STAD.

CD8 T cell infiltration is correlated with better clinical outcomes in many cancers. Triple Negative Breast Cancer (TNBC) patients with high infiltration of CD8 T cells had better survival and highly immunoreactivity [29]. High expression of CD8 + T lymphocytes has a strong association with prognostic in GC [30]. The study by Huang et al. in breast cancer further confirms

that CD8 T cells are regarded as the main effector cells of anti-tumor immunity [31]. Therefore, tumor-infiltrating CD8 T cells may be an important biomarker for predicting cancers.

In total, 700 up- and 313 down-regulated DEGs were identified in STAD. Subsequently, essential modules and genes most significantly associated with CD8 + T cells were confirmed. Finally, we identified four vital genes (CIDEC, EPS8L3, MUC13 and PLEKHS1) correlated with prognosis and CD8 + T cell infiltration by Lasso Cox regression analysis.
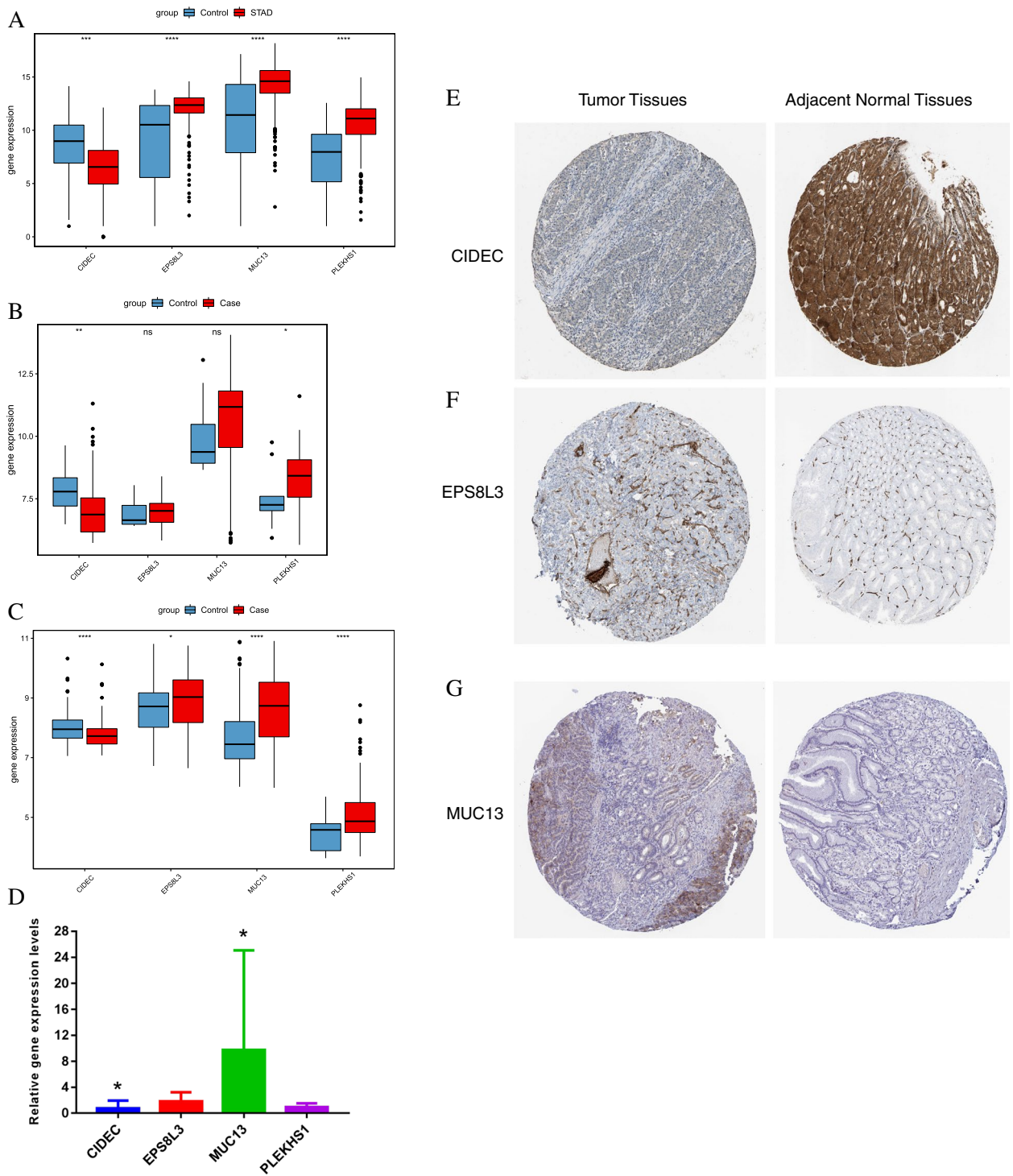
Fig. 6 Expression of CD8 + T cell infiltration-related genes in tumor tissues and tumor adjacent tissues. **A** Expression of genes in TCGA training set; **B** Expression of genes in GSE26899 validation set; **C** Expression of genes in GSE29272 validation set; **D** Expression of genes in RT-PCR; **E**-**G** Immunohistochemical analysis of genes. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$ and $****p < 0.0001$; ns, not significant

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 11 of 17

A

Correlation between riskScore and T cells CD8+

$t_{Student}(333) = 2.10$, $p = 0.037$, $\hat{r}_{Pearson} = 0.11$, $CI_{95\%}$ [7.05e−03, 0.22], $n_{pairs} = 335$

$\log_e(BF_{01}) = 0.31$, $\hat{r}_{Pearson}^{posterior} = 0.11$, $CI_{95\%}^{HDI}$ [0.02, 0.19], $r_{beta}^{JZS} = 1.41$

B

Factor	Hazard Ratio (95% CI)	P−value	
riskScore	3.25(1.91−5.55)	1.49e−05	
Age	1.02(1.01−1.04)	0.002	
Gender	0.77(0.54−1.10)	0.148	
T	1.28(1.04−1.58)	0.022	
N	1.31(1.13−1.53)	0.000465258	
M	1.96(1.08−3.56)	0.026	
G	1.29(0.94−1.78)	0.12	
stage	1.50(1.22−1.85)	0.000130569	

C

Factor	Hazard Ratio (95% CI)	P−value	
riskScore	3.99(2.20−7.24)	<0.001	
Age	1.04(1.02−1.06)	<0.001	
Gender	0.61(0.41−0.91)	0.015	
G	1.22(0.85−1.77)	0.283	
M	1.18(0.52−2.67)	0.699	
N	1.07(0.84−1.35)	0.588	
T	1.13(0.83−1.54)	0.439	
stage	1.36(0.90−2.06)	0.149	

Fig. 7 Correlation of risk score with CD8 + T cells and OS. **A** Correlation analysis; **B** Univariate regression analysis; **C** Multivariate regression analysis

Cell death-inducing DFFA-like effctor c (CIDEC) is predominantly expressed in white adipose tissue, which is essential in lipid metabolism and energy regulation [32]. Relevant literature studies have revealed that CIDEC is over-expressed in various cancers, such as hepatocellular carcinoma [33] and clear cell renal cell carcinoma [34]. Wang et al. reported that the expression of CIDEC increase may reduce fatty acid oxidation and promotes de novo lipogenesis in human adenovirus-infected primary cultured skeletal muscle [35]. Nevertheless, there is no study about the expression of CIDEC in STAD. The low expression of CIDEC in STAD was first found in this study. Epidermal growth factor receptor kinase substrate 8 like 3 (EPS8L3) has reported in various cancers, sunch as live cancer, pancreatic cancer and gastric cancer, which involved in regulating cell proliferation, differentiation and migration [36–38].The expression of EPS8L3 is markedly up-regulated in liver cancer tissues and cell lines, and patients with high expression of EPS8L3 have shorter survival [36, 39]. In the hepatocellular carcinoma (HCC) cell line, overexpression EPS8L3 can enhance cell proliferation [40]. This indicates that EPS8L3 is correlated with poor clinical prognosis. Mucins are recognized as potential oncogenes and possible therapeutic targets in various malignancies [41, 42]. MUC13, a high molecular weight transmembrane glycoprotein, is frequently overexpressed in various epithelial cancers [43]. Shimamura et al. detected up-regulation of MUC13 at GC's mRNA and protein levels [44]. Cai et al. showed that the expression of MUC13 increase can enhance GC cell proliferation and invasion [45]. In addition, MUC13 can be used as a marker for early cancer screening, providing a promising target for targeted therapy. PLEKHS1 is up-regulated in the majority of cancers. Highly recurrent mutations in PLEKHS1 may lead to tumorigenesis [46]. Chessa et al. reported that PLEKHS1 can escape homeostasis, actovate the PIP3 signaling, and support tumour progression in cells absenting PTEN [47]. Furthermore, Xing et al. showed that overexpression PLEKHS1 enhanced anaplastic thyroid carcinomas (ATC) invasion in cell experiment [48]. Besides, Over-expressed PLE-KHS1 increases the risk of disease progression in bladder
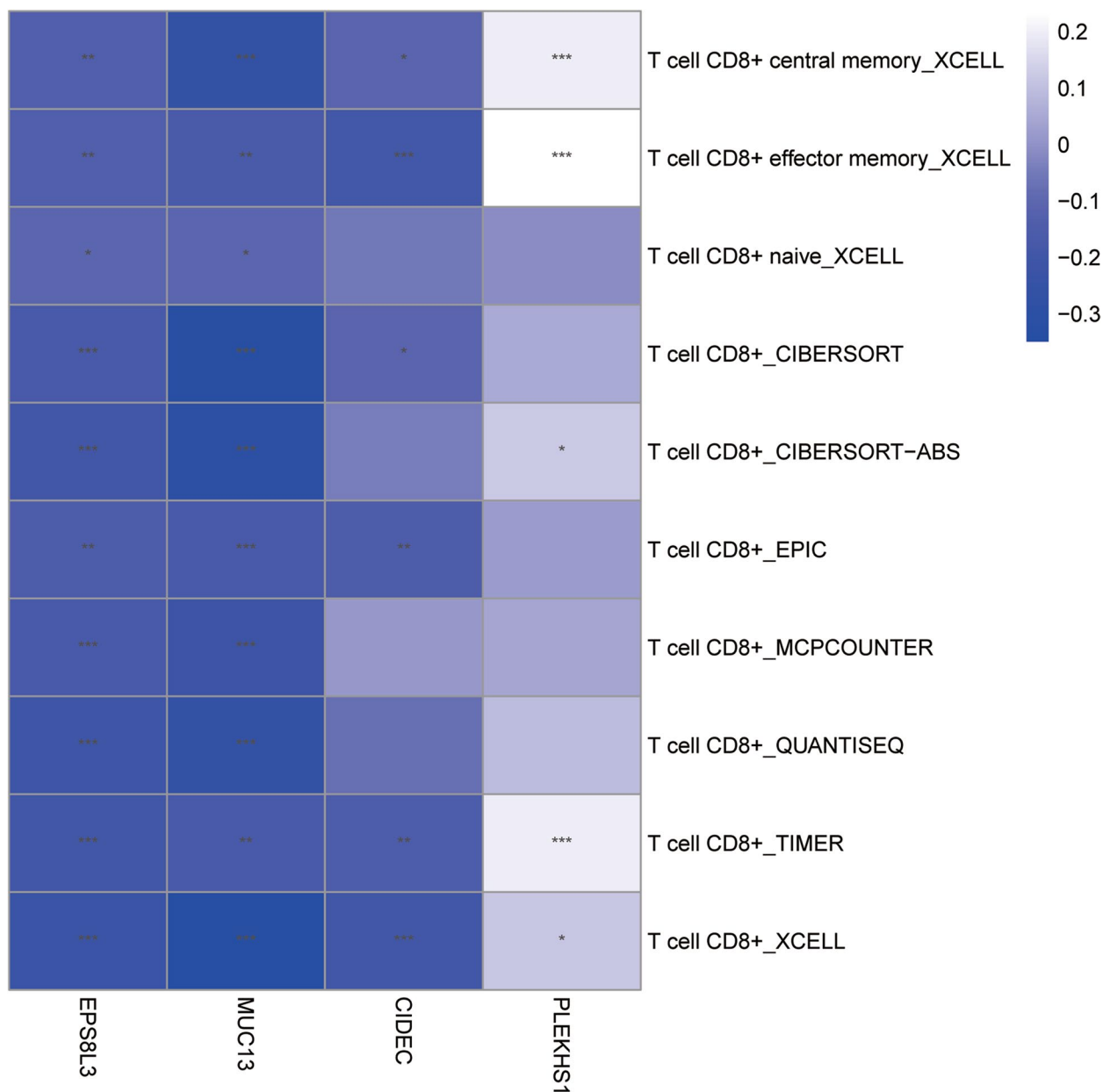
**Fig. 8** Correlation analysis of genes and immune cell subtypes in the TIMER database

cancer [49]. PLEKHS1 is one of the up-regulated DEGs in hepatocellular carcinoma and a poor prognostic indicator [50]. Therefore, the expression of PLEKHS1 may be correlated with cancer prognosis.

We built a risk score model based on four genes. Risk score was an independent prognostic factor for OS in patients with STAD and was positively connected with CD8 + T cell infiltration. These four genes may play a role in CD8 + T cell infiltration in the immune microenvironment. CD8 + tissue-resident memory T (Trm)

cells depended on fatty acid oxidation for cell survival in STAD patients. Absention of fatty acid leads to Trm cell death [51]. Furthermore, Joseph et al. evidenced that CD8 + T cells suppress tumour metastasis in mouse tumour models [52]. Correlation analysis indicated that CIDEC, EPS8L3, and MLC13 were remarkably negatively correlated with CD8 + T cells. Li et al.'s study showed a correlation between CD8 T cells and CIDEC [53]. The correlation between the other two genes and CD8 T cells in tumours has not been reported. Based on the
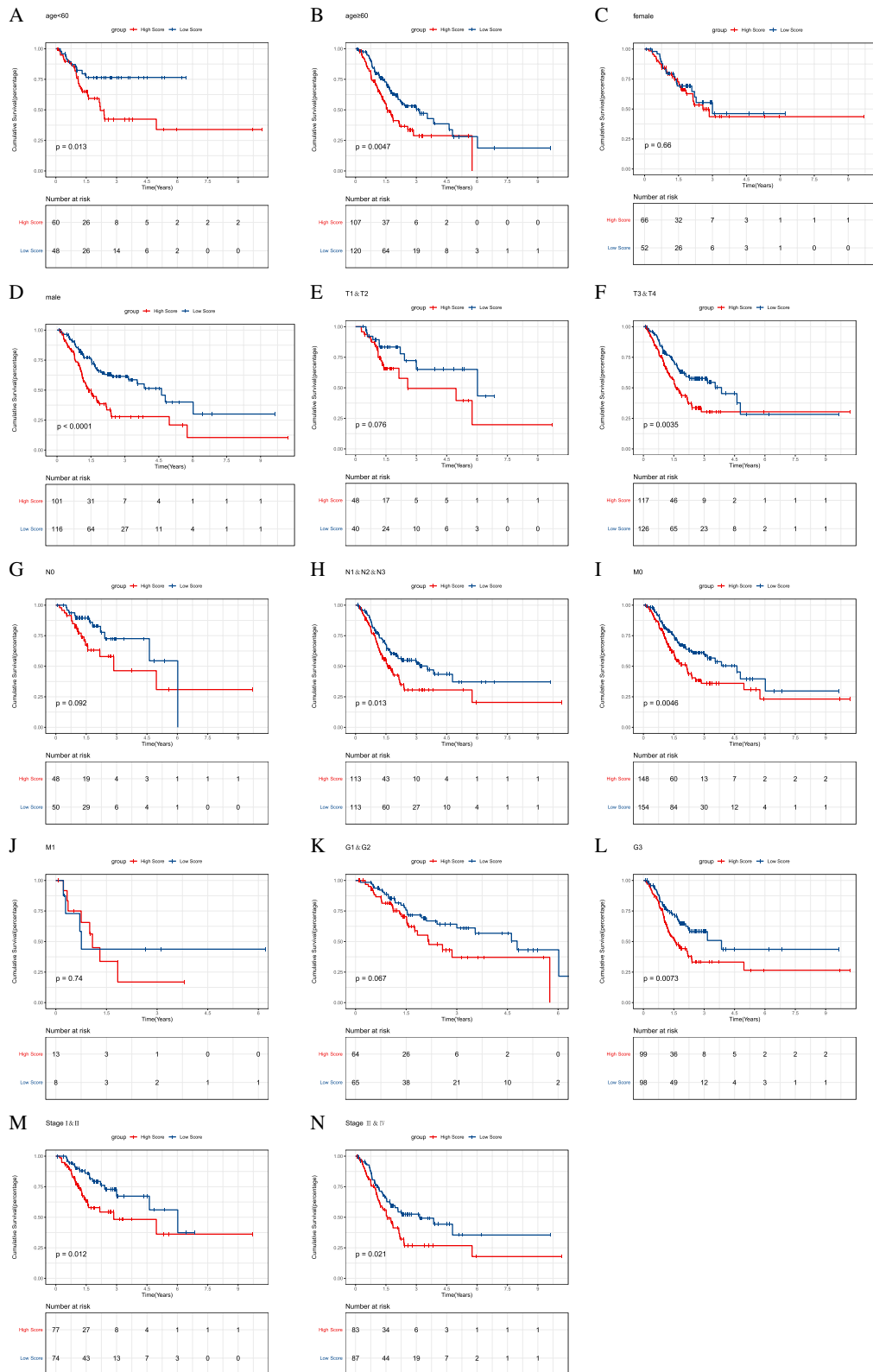
Fig. 9 Subgroup survival analysis in TCGA training set
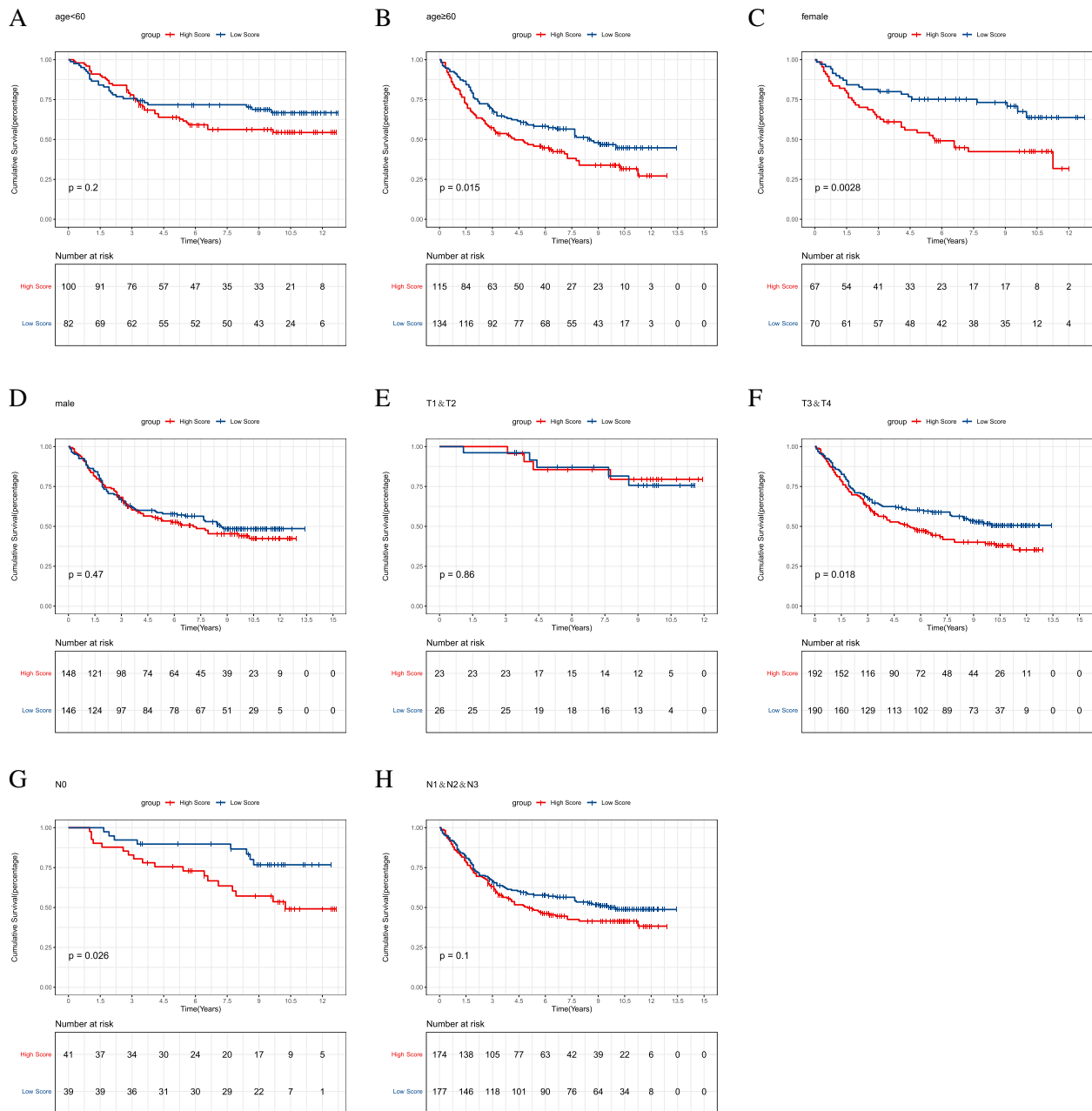
Fig. 10 Subgroup survival analysis in GSE84437 validation set

above results, we speculated that CIDEC expression may influence fatty acid metabolism and reduce the number of CD8 T cells. In addition, EPS8L3, MLC13, and PLEKHS1 can enhance tumour invasion. Therefore, CIDEC, EPS8L3, MLC13, and PLEKHS1 may reduce the number of CD8 T cells and increase the aggressiveness of STAD, leading to poor prognosis. CIDEC, EPS8L3, MLC13 and PLEKHS1 may be potential prognostic factors for STAD and can be used to assess the level of immune cell infiltration in tumour tissues. However, we look forward to collecting more samples for mechanistic analysis to demonstrate the correlation between these genes and CD8 + T cell infiltrating tumours in future studies.

Additionally, we assessed the sensitivity of the risk score model to drug therapy. IC50 is the concentration at which an anticancer drug kills half of the inhibitory concentration of cancer cells [54]. It helps to quantify the therapeutic ability of a drug to induce apoptosis
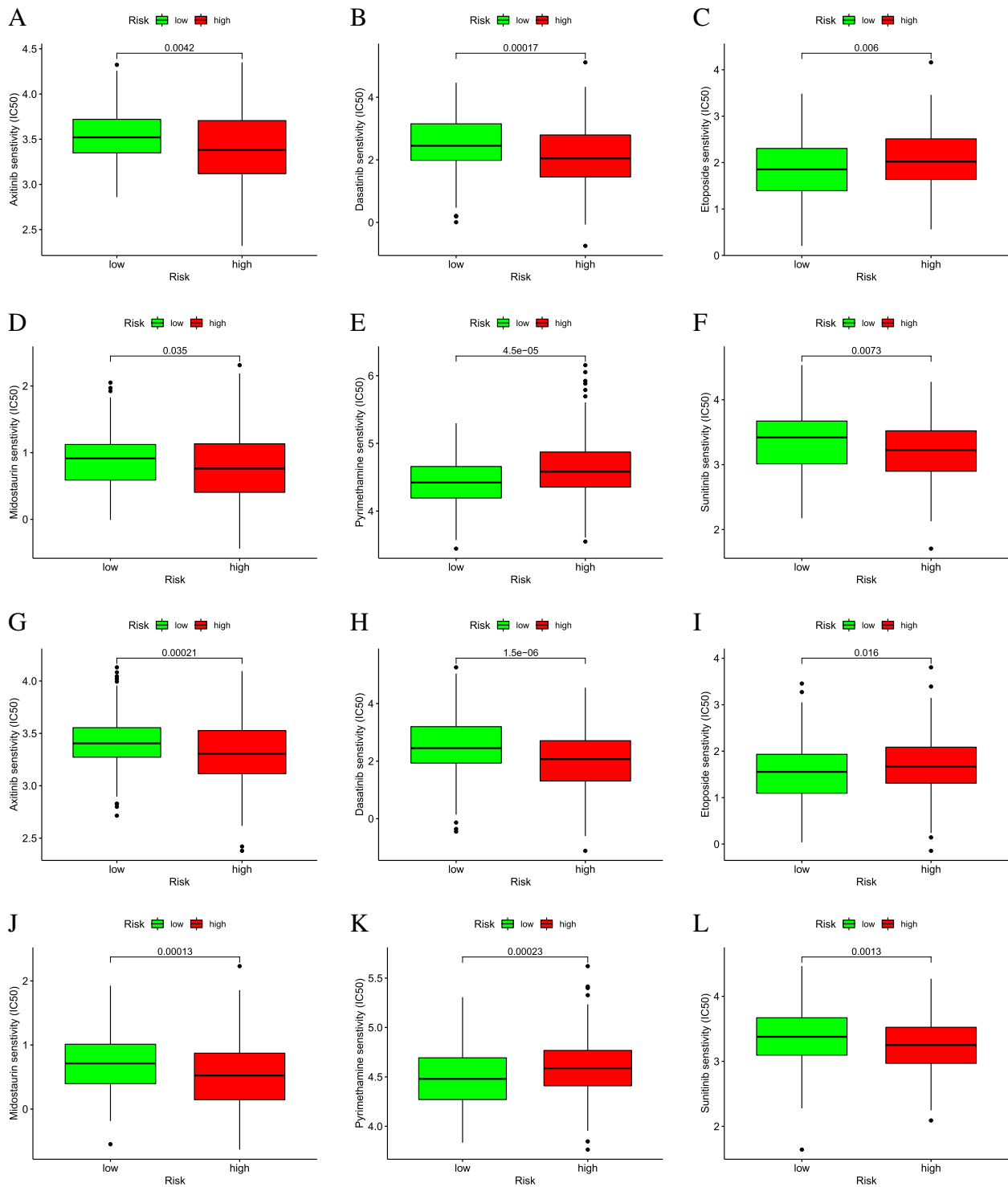
Fig. 11 Sensitivity to drugs in high- and low-risk groups of risk score. **A**-**F**: in TCGA training set; **G**-**I**: in GSE84437 validation set

in cancer cells, which is inversely proportional to the sensitivity of small molecule drugs [55]. Among them, the IC50 of Axitinib, Dasatinib, Midostaurin and Sunitinib in the low-risk group was significantly higher than that in the high-risk group. In contrast, the IC50 of Etoposide and Pyrimethamine was higher in low-risk group. Above results suggested that the drugs may have a good therapeutic effect in STAD patients. The

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 16 of 17

results of related studies also showed that these drugs had lower IC50 values in STAD high-risk group compared with low-risk group, suggesting that these drugs may be more sensitive to high-risk patients [55–58]. In addition, the IC50 of Cisplatin and Docetaxel in GEO dataset was lower in higer-rick group, suggesting that Cisplatin and Docetaxel are certain effective in high-risk subgroups of gastric adenocarcinoma patients. In summary, we speculate that the four-gene risk score model can not only divide patients into different risk groups, but also may help clinicians in clinical decision making for patients in different risk groups.

However, some limitations in our study should be pointed out. First, the identification of four risk genes and construction of the risk score model were based on the systematic bioinformatics analysis of gene expression profiles and public database. Further experiments, such as immunostaining, are needed to evaluate the association between risk score and the extent of CD8 + T cells. Second, the mechanisms underlying the impact of these genes on CD8 + T cell infiltration were not investigated in this study; which will be included in our further study. Third, validation were only conducted in a public database, and a prospective study with a substantial sample size is essential to confirm the clinical utility of the risk score.

Conclusions
In short, we built a four-gene risk score model correlated with CD8 + T cell infiltration, which may provide some guidance for future prognosis prediction and molecular targeted therapy of STAD.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12876-023-03003-y.

Additional file 1: Supplementary Figure 1. Sensitivity to drugs in high- and low-risk groups of risk score. A-B: in TCGA training set; C-D: in GSE84437 validation set.

References
1. Matsuoka T, Yashiro M. Biomarkers of gastric cancer: current topics and future perspective. World J Gastroenterol. 2018;24:2818–32.
2. Alessandrini L, Manchi M, De Re V, Dolcetti R, Canzonieri V. Proposed Molecular and miRNA Classification of Gastric Cancer. Int J Mol Sci. 2018;19:1683.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016;66:7–30.
4. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics. CA Cancer J Clin. 2023;73:233–54.
5. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev. 2014;23:700–13.
6. Wagner AD, Unverzagt S, Grothe W, et al. Chemotherapy for advanced gastric cancer. Cochrane Database Syst Rev. 2010;2019:CD004064.
7. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014;64:9–29.
8. Isobe Y, Nashimoto A, Akazawa K, et al. Gastric cancer treatment in Japan: 2008 annual report of the JGCA nationwide registry. Gastric Cancer. 2011;14:301–16.
9. Cunningham SC, Kamangar F, Kim MP, et al. Survival after gastric adenocarcinoma resection: eighteen-year experience at a single institution. J Gastrointest Surg. 2005;9:718–25.
10. Li K, Zhang A, Li X, Zhang H, Zhao L. Advances in clinical immunotherapy for gastric cancer. Biochim Biophys Acta Rev Cancer. 2021;1876:188615.
11. Zhao Q, Cao L, Guan L, et al. Immunotherapy for gastric cancer: dilemmas and prospect. Brief Funct Genomics. 2019;18:107–12.
12. Zeng D, Li M, Zhou R, et al. Tumor Microenvironment characterization in gastric Cancer identifies prognostic and immunotherapeutically relevant Gene signatures. Cancer Immunol Res. 2019;7:737–50.
13. Chen Y, Sun Z, Chen W, et al. The Immune subtypes and Landscape of Gastric Cancer and to Predict based on the whole-slide images using deep learning. Front Immunol. 2021;12:685992.
14. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18:220.
15. Li T, Fan J, Wang B, et al. TIMER: a web server for Comprehensive Analysis of Tumor-infiltrating Immune cells. Cancer Res. 2017;77:e108–e10.
16. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Elife. 2017;6:e26476.

Pan *et al. BMC Gastroenterology*   (2023) 23:399

Page 17 of 17

17. Finotello F, Mayer C, Plattner C, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med. 2019;11:34.
18. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor infiltrating Immune cells with CIBERSORT. Methods Mol Biol. 2018;1711:243–59.
19. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7.
20. Zhang S, Zhang E, Long J, et al. Immune infiltration in renal cell carcinoma. Cancer Sci. 2019;110:1564–72.
21. Zhou R, Zhang J, Zeng D, et al. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon Cancer. Cancer Immunol Immunother. 2019;68:433–42.
22. Chen J, Wang A, Ji J, et al. An innovative prognostic model based on four genes in Asian patient with gastric Cancer. Cancer Res Treat. 2021;53:148–61.
23. National Health Commission Of The People's Republic Of C. Chinese guidelines for diagnosis and treatment of gastric cancer 2018 (English version). Chin J Cancer Res. 2019;31:707–37.
24. Jiang F, Lin H, Yan H, Sun X, Yang J, Dong M. Construction of mRNA prognosis signature associated with differentially expressed genes in early stage of stomach adenocarcinomas based on TCGA and GEO datasets. Eur J Med Res. 2022;27:205.
25. Shen Y, Chen K, Gu C. Identification of a chemotherapy-associated gene signature for a risk model of prognosis in gastric adenocarcinoma through bioinformatics analysis. J Gastrointest Oncol. 2022;13:2219–33.
26. Song W, Shen L, Wang Y, et al. Synergistic and low adverse effect cancer immunotherapy by immunogenic chemotherapy and locally expressed PD-L1 trap. Nat Commun. 2018;9:2237.
27. Wang Z, Wang K, Yu X, Chen M, Du Y. Comprehensive analysis of expression signature and immune microenvironment signature of biomarker endothelin receptor type A in stomach adenocarcinoma. J Cancer. 2022;13:2086–104.
28. van der Leun AM, Thommen DS, Schumacher TN. CD8+ T cell states in human cancer: insights from single-cell analysis. Nat Rev Cancer. 2020;20:218–32.
29. Oshi M, Asaoka M, Tokumaru Y, et al. CD8 T Cell Score as a Prognostic Biomarker for Triple Negative Breast Cancer. Int J Mol Sci. 2020;21:6968.
30. Mansuri N, Birkman EM, Heuser VD, et al. Association of tumor-infiltrating T lymphocytes with intestinal-type gastric cancer molecular subtypes and outcome. Virchows Arch. 2021;478:707–17.
31. Huang Y, Ma C, Zhang Q, et al. CD4 + and CD8 + T cells have opposing roles in Breast cancer progression and outcome. Oncotarget. 2015;6:17462–78.
32. Li J, Liu G, Zhang F, Zhang Z, Xu Y, Li Q. Role of glycoprotein 78 and cidec in hepatic steatosis. Mol Med Rep. 2017;16:1871–77.
33. Xu MJ, Cai Y, Wang H, et al. Fat-Specific protein 27/CIDEC promotes development of alcoholic steatohepatitis in mice and humans. Gastroenterology. 2015;149:1030–41e6.
34. Yu M, Wang H, Zhao J, et al. Expression of CIDE proteins in clear cell renal cell carcinoma and their prognostic significance. Mol Cell Biochem. 2013;378:145–51.
35. Wang ZQ, Yu Y, Zhang XH, Floyd EZ, Cefalu WT. Human adenovirus 36 decreases fatty acid oxidation and increases de novo lipogenesis in primary cultured human skeletal muscle cells by promoting Cidec/FSP27 expression. Int J Obes (Lond). 2010;34:1355–64.
36. Li P, Hu T, Wang H, et al. Upregulation of EPS8L3 is associated with tumorigenesis and poor prognosis in patients with Liver cancer. Mol Med Rep. 2019;20:2493–99.
37. Fan Z, Li M, Xu Y, Ge C, Gu J. EPS8L3 promotes Pancreatic cancer proliferation and Metastasis by activating GSK3B. J Med Biochem. 2023;42:105–12.
38. You L, Xiao L, Jin S. EPS8L3 suppresses apoptosis and autophagy of gastric cancer through PI3K/AKT/mTOR signaling. Mol Cell Toxicol. 2023;19:373–81.
39. Xuan Z, Zhao L, Li Z, et al. EPS8L3 promotes hepatocellular carcinoma proliferation and Metastasis by modulating EGFR dimerization and internalization. Am J Cancer Res. 2020;10:60–77.
40. Zeng CX, Tang LY, Xie CY, et al. Overexpression of EPS8L3 promotes cell proliferation by inhibiting the transactivity of FOXO1 in HCC. Neoplasma. 2018;65:701–07.
41. He L, Qu L, Wei L, Chen Y, Suo J. Reduction of miR–132–3p contributes to gastric cancer proliferation by targeting MUC13. Mol Med Rep. 2017;15:3055–61.
42. Luo H, Guo W, Wang F, et al. miR-1291 targets mucin 1 inhibiting cell proliferation and invasion to promote cell apoptosis in esophageal squamous cell carcinoma. Oncol Rep. 2015;34:2665–73.
43. Maher DM, Gupta BK, Nagata S, Jaggi M, Chauhan SC. Mucin 13: structure, function, and potential roles in cancer pathogenesis. Mol Cancer Res. 2011;9:531–37.
44. Shimamura T, Ito H, Shibahara J, et al. Overexpression of MUC13 is associated with intestinal-type gastric cancer. Cancer Sci. 2005;96:265–73.
45. Cai T, Peng B, Hu J, He Y. Long noncoding RNA BBOX1-AS1 promotes the progression of gastric cancer by regulating the miR-361-3p/Mucin 13 signaling axis. Bioengineered. 2022;13:13407–21.
46. Weinhold N, Jacobsen A, Schultz N, Sander C, Lee W. Genome-wide analysis of noncoding regulatory mutations in cancer. Nat Genet. 2014;46:1160–5.
47. Chessa TAM, Jung P, Anwar A, et al. PLEKHS1 drives PI3Ks and remodels pathway homeostasis in PTEN-null prostate. Mol Cell. 2023;83:2991–3009e13.
48. Xing X, Mu N, Yuan X, et al. PLEKHS1 Over-Expression is Associated with Metastases and Poor Outcomes in Papillary Thyroid Carcinoma. Cancers (Basel). 2020;12:2133.
49. Pignot G, Le Goux C, Vacher S, et al. PLEKHS1: a new molecular marker predicting risk of progression of non-muscle-invasive Bladder cancer. Oncol Lett. 2019;18:3471–80.
50. Deng Z, Wang J, Xu B, et al. Mining TCGA Database for Tumor Microenvironment-Related Genes of Prognostic Value in Hepatocellular Carcinoma. Biomed Res Int. 2019;2019:2408348.
51. Lin R, Zhang H, Yuan Y, et al. Fatty acid oxidation controls CD8(+) tissue-Resident memory T-cell survival in gastric adenocarcinoma. Cancer Immunol Res. 2020;8:479–92.
52. Joseph R, Soundararajan R, Vasaikar S, et al. CD8(+) T cells inhibit Metastasis and CXCL4 regulates its function. Br J Cancer. 2021;125:176–89.
53. Li WH, Zhang L, Li YY, et al. Apolipoprotein A-IV Has Bi-Functional Actions in Alcoholic Hepatitis by Regulating Hepatocyte Injury and Immune Cell Infiltration. Int J Mol Sci. 2022;24:670.
54. Liu C, Wei D, Xiang J, et al. An improved Anticancer Drug-Response Prediction based on an Ensemble Method integrating Matrix Completion and Ridge Regression. Mol Ther Nucleic Acids. 2020;21:676–86.
55. Xiao C, Dong T, Yang L, et al. Identification of Novel Immune ferropotosis-related genes Associated with Clinical and Prognostic features in gastric Cancer. Front Oncol. 2022;12:904304.
56. Chang J, Wu H, Wu J, et al. Constructing a novel mitochondrial-related gene signature for evaluating the tumor immune microenvironment and predicting survival in stomach adenocarcinoma. J Transl Med. 2023;21:191.
57. Tong X, Yang X, Tong X, Zhai D, Liu Y. Complement system-related genes in stomach adenocarcinoma: prognostic signature, immune landscape, and drug resistance. Front Genet. 2022;13:903421.
58. Ke P, Bao X, Liu C, et al. LPCAT3 is a potential prognostic biomarker and may be correlated with immune infiltration and ferroptosis in acute Myeloid Leukemia: a pan-cancer analysis. Transl Cancer Res. 2022;11:3491–505.

Publisher's Note