CHSY3 As A Potential Therapeutic Target For Gastric Cancer

Jiaxin Fan

The First Hospital of Shanxi Medical University

He Huang

The First Hospital of Shanxi Medical University

Chaowei Liang

The First Hospital of Shanxi Medical University

Jiajia Wang

The First Hospital of Shanxi Medical University

Chaojie Liang (✉ liangchaojie8@126.com)

The First Hospital of Shanxi Medical University

Research Article

Keywords: gastric cancer(GC), CHSY3, prognosis, biomarker, immune infiltration, nomogram, targeted therapy.

Posted Date: January 20th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1263603/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Chondroitin synthase 3(CHSY3) has been reported to be closely related to the occurrence, development, recurrence, and metastasis of a variety of cancers. But its role in gastric cancer (GC) has not been studied.

Methods: Tumor Immune Estimate Resource (TIMER), The Cancer Genome Atlas (TCGA), and Gene Expression Omnibus (GEO) were used to detect the expression of CHSY3 in gastric cancer. Analyze the correlation between CHSY3 expression and clinicopathological parameters. Draw a Kaplan-Meier survival curve to detect the clinical prognostic significance of CHSY3 in gastric cancer. Meta-analysis of the above results to determine the role of CHSY3 in GC. COX regression analysis was used to study whether CHSY3 is a predictor of GC patients. Analyze the methylation data from the TCGA database to determine the sites with a significant correlation. Through GO analysis and KEGG analysis to study the regulation mechanism of CHSY3 related genes. Screen 3 genes (APOD, GJA1, SERPINE1) from the related genes of CHSY3 as a predictive signature. Constructed a nomogram to predict the 1-year and 3-year survival probability of GC patients. The TIMER database was used to study the correlation between CHSY3 and tumor infiltrating immune cells. Study the drug sensitivity of CHSY3 related genes by using Gene Set Cancer Analysis (GSCA).

Results: CHSY3 was highly expressed in GC and significantly correlated with the T stage. The higher the expression level of CHSY3 in GC, the higher the methylation level of cg06610705 and cg11572844, which suggests that patients with gastric cancer have a poorer prognosis. Enrichment analysis identified multiple cancer-related pathways, for example, PI3K-Akt signaling pathway, MAPK signaling pathway, and TGF-beta signaling pathway. The risk score has a very high predictive value. CHSY3 is closely associated with CD4⁺ T cells, macrophages, neutrophils, and dendritic. Besides, different mutational forms of CHSY3 were associated with the immune infiltration of 6 leukocytes. In addition, CHSY3 may contribute to drug resistance in GC.

Conclusion: CHSY3 may be a potential prognostic biomarker for gastric cancer, and it can be used as a new therapeutic target for the treatment of gastric cancer.

Introduction

According to research, cancer plays an important role in affecting human morbidity and mortality[1]. At present, the awareness of cancer needs to be improved, and the mortality rate of cancer will increase in the future[2]. Gastric cancer is the fifth most common cancer, a heterogeneous malignant tumor[3], accounting for 8.2% of cancer deaths and the third leading cause of cancer deaths worldwide[4]. Almost half of GC cases are diagnosed in East Asia[5]. Although various treatment methods such as surgery, chemotherapy, radiotherapy, and immunotherapy have been developed, the 5-year survival rates of stage II and stage IIIA gastric cancer patients are approximately 34% and 20%, respectively[6]. In recent years, studying the occurrence and development of gastric cancer from the perspective of bioinformatics has

gradually become mainstream. Including microsatellite instability and HER2 mutation[7]. Many studies on molecular targeted therapy and related molecular pathways have clarified the pathogenesis of GC and have further improved the prognosis of gastric cancer patients. For example, Trastuzumab has a better effect on HER2-positive GC patients[8]. Therefore, it is very necessary to identify key new biomarkers and potential mechanisms related to GC.

Chondroitin synthase (CHSY) is one of the six enzymes responsible for CS biosynthesis in mammalian cells[9]. CHSY3 (Chondroitin Sulfate Synthase 3) is a Protein Coding gene, An important paralog of this gene is CHSY1. The methylation level of CHSY1 is related to the differentiation of T cells[10] and is also essential for bone development[11], and the lack of CHSY1 can cause temtamy preaxial brachydactyly syndrome[12]. Studies have shown that CHSY1 plays an important role in the occurrence and development of a variety of cancers. For example, CHSY1 is involved in the interaction between myeloma cells and osteoclasts[13]. The differential expression of CHSY1 was found in malignant soft tissue sarcomas[14]. Knockout of CHSY1 increased the expression of JAG2, a key molecule in glioblastoma cells[15]. In addition, the overexpression of CHSY1 enhanced the migration, invasion, and EMT of hepatocellular carcinoma[16]. Following CHSY1 knockdown, the proliferation of colorectal cancer cells will be significantly reduced[17]. However, there is no research on the role of CHSY3 in GC.

In this study, the expression of CHSY3, the correlation between CHSY3 expression and clinicopathological characteristics, degree of methylation, and prognosis were analyzed using the sequencing data set. The GO analysis and KEGG analysis reveal the potential regulatory mechanism of CHSY3. A prognostic signature was established through genes related to CHSY3, and then the risk score of each patient was calculated, and the relationship between the risk score and OS was explored. In addition, the relationship between CHSY3 and tumor-infiltrating immune cells was also studied. This study reveals the potential role of CHSY3 in GC, which will help us further understand the possible mechanism of gastric cancer.

Materials And Methods

Data Source and Preprocessing

Download gene expression data and clinical information of gastric cancer patients from the The Cancer Genome Atlas(TCGA)(<https://portal.gdc.cancer.gov/repository>) and Gene Expression Omnibus(GEO)(<https://www.ncbi.nlm.nih.gov/geo/>). We use Perl programming language to integrate data, match gene expression information with clinical information, and delete unknown or incomplete clinical information.

Since the sequencing data was obtained using public databases, there are no ethical issues.

CHSY3 gene expression analysis

The level of CHSY3 gene expression in a variety of cancers including GC was analyzed using Tumor Immune Estimation Resource (TIMER) database with The Cancer Genome Atlas (TCGA) data. For the data obtained from TCGA, survival package, Beeswarm package, and Limma package were used to

analyze the difference in CHSY3 expression between GC and normal tissues, and P value was measured by Wilcox test. The results were visualized using scatter difference and paired difference diagrams. In addition, we drew a box scatter plot to perform CHSY3 difference analysis on five datasets (GSE13911, GSE19826, GSE54129, GSE66229, GSE79973) from the GEO database. Finally, the above results were meta-analyzed. We then analyzed the data from TCGA and GEO using the survival package, survminer package, and timeROC package. Draw receiver operating characteristic (ROC) curve and SROC curve to assess the accuracy of CHSY3 in predicting GC patient survival. R software (V.4.0.4) Analyze the above.

Clinical correlation analysis

We selected GC tissue samples with clinicopathological data, including age, grade, stage, depth of invasion, lymph node metastasis, and distant metastasis, for correlation analysis with CHSY3 expression. Age and distant metastasis were analyzed using Wilcox-Test, the rest were analyzed using Kruskal-test.

Survival Analysis

According to the median expression of CHSY3, patients were divided into the high expression group and low expression group. Survival software was used for survival analysis, and Kaplan-Meier survival curve visualization results were drawn. Finally, meta-analysis was performed on the analysis results of TCGA and GEO data to determine the relationship between the expression of CHSY3 and the survival prognosis of GC patients.

Univariate and Multivariate Cox Regression Analyses.

We used the Cox proportional hazard regression model to perform Univariate and multivariate analyses on the data from TCGA and GEO datasets (GSE62254, GSE84437). Univariate COX analysis was used to evaluate whether CHSY3 is a predictor of survival and prognosis in GC patients. Multivariate Cox analysis was used to evaluate whether CHSY3 can be used as an independent prognostic factor for GC. The Survival and Survminer software packages were used to analyze the data, and the forest plots were used to visualize the results.

Methylation analysis of CHSY3

Methylation data obtained from the TCGA database were analyzed to obtain the methylation data of CHSY3, then the methylation degree of CHSY3 at each site was obtained. In addition, ggpubr package and ggplot2 package were used to analyze the correlation between the expression level of CHSY3 and the methylation level of each site. The relationship between methylation at two sites and overall survival in GC patients was analyzed.

The enrichment analysis of CHSY3

First, we identified the co-expression genes of CHSY3 using the Linkedomics platform, and then performed GO enrichment analysis and KEGG enrichment analysis using the R package "clusterProfiler". $P < 0.05$ was considered statistically significant.

Construction and Evaluation of the Prognostic Signatures of CHSY3

First, we obtained the 209 differential genes by constructing a co-expression network of CHSY3, ($\text{fdrFilter}=0.00000001$, $\text{logFCfilter}=10$), and then performed Univariate COX analysis on these DEGs, and determined DEGs that are significantly different from OS, ($\text{pFilter}=0.05$), and subsequently, using LASSO regression analysis to screen the DEGs that are significantly related to OS. By running 1000 times of cross-validation likelihood method, The optimal value for penalization coefficient lambda (λ) is screened out. This method can avoid overfitting the signature. In the case where λ is the smallest, select the most suitable gene to construct the signature.

Finally, we performed multivariate Cox regression analysis on DEGs screened by LASSO. 3 genes (APOD, GJA1, SERPINE1) were selected as a predictive signature. We used the risk score calculation formula to calculate the risk score of each patient: $\text{risk score} = \text{coef gene 1} \times \text{gene 1 expression} + \text{coef gene 2} \times \text{gene 2 expression} + \dots + \text{coef gene } N \times \text{gene } N \text{ expression}$. The risk score is obtained by weighting the expression level of the gene and the regression coefficient (coef). The risk ratio (HR) of the multivariate Cox regression analysis was log-transformed to calculate the coef value, and the expression of each gene involved in the prognostic characteristics was defined as the expression of N gene. GC patients in TCGA are divided into high-risk groups and low-risk groups according to the median risk score.

First, the difference of OS between the high-risk group and the low-risk group was analyzed by the K-M method, and the survival curve was obtained. We then plotted Receiver Operating Characteristic (ROC) curves using the "survivalROC" package to assess the ability of risk score and other clinical features to predict GC and to assess sensitivity and specificity by the area below the curve (AUC values). To determine whether the risk score can be used as an independent predictor of the prognosis of GC patients, we incorporated age, gender, grade, stage, T, M, N, and risk score into univariate and multivariate Cox regression analysis.

Construction and Validation of Nomogram Based on Risk Score

The nomogram can personally calculate the survival rate of GC patients, which is of great value in clinical practice. We screened the factors that affect the prognosis of GC patients and constructed a nomogram using the "survival" and "rms" packages to predict the 1-year and 3-year survival probability of GC patients. Then we drew calibration curves to evaluate the accuracy of the nomogram prediction. Finally, the time-varying ROC curves of the nomogram were drawn, and the AUC values were calculated.

Analysis of association between CHSY3 expression level and Immune Infiltrates

Tumor Immune Estimation Resource (TIMER) is a public site covering 32 cancer types, including 10,897 samples from the TCGA database, designed to assess immune invasion abundance

(<http://cistrome.org/TIMER/>). The association between CHSY3 expression and six types of infiltrating immune cells (CD8+ T cells, CD4+ T cells, B cells, dendritic cells macrophages, and neutrophils) was evaluated through the Correlation Module of the TIMER database. Meanwhile, the relationship between CHSY3 gene expression and tumor purity was analyzed. In addition, the SCNA module was used to explore the correlation between somatic copy number alteration (SCNA) of CHSY3 and the immune abundance of 6 leukocytes.

Drug sensitivity analysis of CHSY3-related drugs

We investigated drug sensitivity of CHSY3-related genes by using Gene Set Cancer Analysis (GSCA) (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>), a web-based platform for Gene Set Cancer Analysis GSCALite.

Statistical Analysis.

R software (version 4.0.4) was used for statistical analysis. Wilcox-Test and kruskal-Test were used to analyze clinicopathological parameters and the expression level of CHSY3. Kaplan-Meier analysis was used to investigate the relationship between survival rate and CHSY3 expression level. Univariate and multivariate survival analyses were performed using Cox proportional hazard regression model. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

The differential expression of CHSY3 between GC tumor tissue and normal tissue.

First, we used the TIMER database to analyze the TCGARNA-seq data to evaluate the transcription level of CHSY3 in different human tumors(Figure.1A). Compared with normal gastric tissue, the expression of CHSY3 in GC tissue is significantly higher. In addition, CHSY3 was also found to be highly expressed in the tissues of Breast invasive carcinoma(BRCA), Cholangiocarcinoma(CHOL), Colon adenocarcinoma(COAD), Esophageal carcinoma(ESCA), Head and Neck squamous cell carcinoma(HNSC), Kidney Chromophobe(KICH), Kidney renal clear cell carcinoma(KIRC), Prostate adenocarcinoma(PRAD) and Rectum adenocarcinoma(READ). CHSY3 is significantly lower in Bladder Urothelial Carcinoma(BLCA), Liver hepatocellular carcinoma(LIHC), Lung adenocarcinoma(LUAD), Lung squamous cell carcinoma(LUSC), Thyroid carcinoma(THCA), and Uterine Corpus Endometrial Carcinoma(UCEC) than in the respective normal tissues. These results indicate that CHSY3 is abnormally expressed in a variety of cancers. We used TCGA data to draw a Scatter difference plot($p < 0.001$, Figure.1B) and paired difference plot($p < 0.001$, Figure.1C) to further analyze the expression of CHSY3 in GC.and found that the expression of CHSY3 in GC was significantly higher than that in normal gastric tissues. We used TCGA data to draw ROC curves to evaluate the diagnostic value of CHSY3 in GC.Area under ROC curve (AUC): 0.618,95%-CI:0.867,0.644). Therefore, CHSY3 has considerable diagnostic value for GC. We obtained the immunohistochemical smear of CHSY3 in gastric cancer tissue through The

Human Protein Atlas (HPA)(<https://www.proteinatlas.org/>) to verify The differential expression of CHSY3 in GC(Figure.1E).

At the same time, we use five datasets from GEO to draw box diagrams(Figure.2A-E) (GSE13911,P=0.013;GSE19826,P=0.017;GSE54129,P<0.001;GSE66229,P<0.001;GSE79973,P=0.029) and ROC curves(Figure.2F-J)(GSE13911,AUC=7.200,95%-CI:0.750,0.703;GSE19826,AUC=6.957,95%-CI:0.933,0.917;GSE54129,AUC=2.278,95%-CI:1.000,0.739;GSE66229,AUC=0.567,95%-CI:0.690,0.797;GSE79973,AUC=2.836,95%-CI:1.000,0.800) to study the expression of CHSY3 in GC and the diagnostic value of CHSY3.Finally, we conducted a Meta analysis of the above results.Figure.3F shows a Meta-analysis containing 5 datasets from GEO and TCGA datas. The results show that CHSY is highly expressed in GC tissues(SMD=1.09,95% CI: 0.83-1.35; $I^2=44\%$, P=0.134).Studies that integrate diagnostic odds ratio (DOR)(0.83,95% CI: 1.84-3.18)(Figure.3A),negative likelihood ratio (NLR)(-1.12,95% CI: -1.34-0.89)(Figure.3B),positive likelihood ratio (PLR)(1.29,95% CI: 0.81-1.77) (Figure.3C),sensitivity(0.726,95% CI: 0.657-0.784)(Figure.3D),specificity(0.796,95% CI: 0.698-0.868) (Figure.3E),and the AUC of SROC of CHSY3(Figure.3G) have shown that CSHY3 can be used as a comparable GC diagnostic biomarker.

Relationship between expression of CHSY3 and clinicopathological features.

We downloaded the clinical data of gastric cancer patients from TCGA, and after deleting unknown and incomplete information, we performed the correlation analysis between CHSY3 expression and clinicopathological characteristics.The results showed that the expression of CHSY3 was significantly correlated with T stage(Figure.4F),and moderately correlated with grade(Figure.4B), but not with age (Figure.4A),pathologic stage(Figure.4C),M stage(Figure.4D), and N stage(Figure.4E).

High expression of CHSY3 in GC tumor tissue suggests poor overall survival

We used the Kaplan-Meier risk assessment method to evaluate the relationship between CHSY3 expression and prognosis in GC patients based on data from the TCGA and GEO databases(Figure.5). Then we conducted a Meta-analysis of the above results(HR =1.40,95%-CI:1.19-1.61) and drew a forest map to visualize the results of the Meta-analysis(Figure.6A). Because $I^2<50\%$ and $P>0.05(I^2=32.8\%,P=0.190)$,we chose the fixed effect model. The above analysis results show that the high expression of CHSY3 is related to the poor overall survival rate of GC patients, and it is indeed a high-risk gene for gastric cancer.

CHSY3 Is an Independent Predictor of Poor Survival in GC

After processing the clinical data obtained from GSE62254, GSE84437, and TCGA database, we performed univariate and multivariate Cox proportional hazards regression analysis. Univariate regression analysis of TCGA data shows that age(P=0.008), stage(P<0.001), and high CHSY3(P=0.049) expression are important predictors of the prognosis of GC patients(Figure.6C). Multivariate regression analysis of TCGA data shows that age(P<0.001) and stage(P<0.001) are independent predictors of

survival in GC patients, while CHSY3($P=0.149$) is not consistent with(Figure.6D). Univariate regression analysis of the GSE84437 dataset shows that CHSY3($P=0.024$),age($P=0.002$), T($P<0.001$), and N($P<0.001$) are important predictors of the prognosis of GC patients(Figure.6E).Multivariate regression analysis of the GSE84437 dataset shows that age($P<0.001$), T($P<0.001$) and N($P<0.001$) are independent predictors of survival in GC patients, while CHSY3($P=0.092$) is not consistent with(Figure.6F).Univariate regression analysis of the GSE62254 dataset shows that CHSY3($P<0.001$), T($P<0.001$), N($P<0.001$), M($P<0.001$), and stage($P<0.001$) are important predictors of the prognosis of GC patients(Figure.6G).Multivariate regression analysis of the GSE62254 dataset shows that CHSY3($P<0.001$), N($P=0.010$), and M($P=0.019$) are independent predictors of survival in GC patients, (Figure.6H). Finally, we conducted a Meta-analysis of the above studies, and the results showed that CHSY3 is an independent predictor of the prognosis of GC patients (ES=1.23,95% CI: 0.99-1.47; $I^2=0.00\%$, $P=0.527$)(Figure.6B).

Methylation analysis of CHSY3

Based on the methylation data downloaded from TCGA, the study found that the expression level of CHSY3 was positively correlated with the methylation degree on cg06610705(Figure.7B, $R=0.15$, $P=0.0064$), cg10678749(Figure.7C, $R=0.28$, $P<0.001$), and cg11572844(Figure.7D, $R=0.29$, $P<0.001$), while negatively correlated with the methylation degree on cg04729562(Figure.7A, $R=-0.23$, $P<0.001$) and cg18829263(Figure.7E, $R=-0.24$, $P<0.001$). Then we explored the relationship between the overall survival rate of GC patients and the degree of methylation on cg06610705(Figure.7F, $p=0.027$) and cg11572844(Figure.7G, $p=0.058$). The above studies showed that the higher the expression level of CHSY3 in GC, the higher the methylation level of cg06610705 (Figure.7F) and cg11572844 (Figure.7G), which would indicate the poor prognosis of GC patients.

Enrichment analysis of CHSY3 co-expressed genes

The co-expressed genes have similar functions and mechanisms. To further study the underlying mechanism of CHSY3 regulation, we used the Linkedomics platform to identify CHSY3 co-expressed genes. The volcano map(Figure.8A)shows a correlation between global genes and CHSY3 through Pearson's test. The heat map(Figure.8B) shows the top 50 genes in GC that are negatively correlated and positively correlated with CHSY3.Then we performed enrichment analysis first 50 positively and negatively correlated co-expressed genes. For KEGG pathway analysis(Figure.8C),We explored multiple cancer-related signaling pathways,such as PI3K-Akt signaling pathway,Human papillomavirus infection,ECM-receptor interaction,Proteoglycans in cancer,MAPK signaling pathway,TGF-beta signaling pathway,cGMP-PKG signaling pathway, and Small cell lung cancer.In addition,we explored three main types of GO enrichment(Figure.8D): biological process (BP), cellular component (CC),and molecular function (MF).In the BP category,we explored regulation of cell-substrate adhesion,regulation of cellular response to growth factor stimulus,transmembrane receptor protein serine/threonine and kinase signaling pathway.In the CC category,we explored contractile fiber,cell-substrate junction,focal adhesion,banded collagen fibril,fibrillar collagen trimer,complex of collagen trimers,collagen

trimer, endoplasmic reticulum lumen, basement membrane, collagen-containing and extracellular matrix. In the MF category, we explored platelet-derived growth factor binding, extracellular matrix binding, sulfur compound binding, integrin binding, heparin binding, extracellular matrix structural constituent conferring tensile strength, growth factor binding, glycosaminoglycan binding, collagen binding and extracellular matrix structural constituent.

Construction of the CHSY3 Prognostic Signature for GC Patients

Using the survival information of GC patients to perform univariate Cox regression analysis on the 209 DEGs, it was found that there are 95 DEGs with significant prognostic differences. Seven optimal candidate genes were obtained by LASSO analysis, including APOD, GJA1, RGS5, CTHRC1, SERPINE1, DUSP1, and CXCR4, which could reduce signature overfit (Figure.9A,B). Perform multivariate Cox analysis on 7 genes with prognostic significance, and construct a prognostic signature composed of 3 genes, including APOD, GJA1 and SERPINE1. Based on the prognosis signature, the risk score calculation formula was obtained: $\text{risk score} = (0.132417559 \times \text{GJA1 expression}) + (0.177130658 \times \text{SERPINE1 expression}) + (0.092985412 \times \text{APOD expression})$. The risk score was calculated for each GC patient based on the expression levels of 3 genes, and patients were divided into a high-risk group ($n = 184$) and a low-risk group ($n = 184$) based on the median. We constructed a heat map to show the expression of 3 genes in the high-risk group and the low-risk group, and the expression of 3 genes in the high-risk patients was higher than that in the low-risk patients (Figure.9C). Figure.9D shows the distribution of risk scores for GC patients. Patients are divided into two groups, with risk scores increasing from left to right. Figure.9E shows the distribution of survival status and survival time for patients with different risk scores. K-M curve was used to compare the difference in OS time between the high-risk group and the low-risk group (Figure.10A). Results showed that GC patients with a high-risk score had significantly lower OS than GC patients with a low-risk score ($P < 0.001$). We plotted a time-dependent ROC curve to predict survival in GC patients, showing that the risk score had high sensitivity and specificity. AUC of risk score (AUC=0.657) was higher than that of age (AUC=0.589), gender (AUC=0.461), grade (AUC=0.569), stage (AUC=0.605), T stage (AUC=0.561), N stage (AUC=0.579) and M stage (AUC=0.532) (Figure.10B). The results show that the prediction result of the risk score is reliable and accurate. Figure.10C reflects the univariate Cox analysis of the relationship between the clinical features, risk score and OS of GC patients. Age ($P = 0.006$), stage ($P < 0.001$), T ($P = 0.032$), N ($P = 0.006$), M ($P = 0.025$) and risk score ($P < 0.001$) significantly affect the prognosis of GC patients. Figure.10D reflects multivariate Cox analyzed the relationship between the clinical features, risk score, and OS of GC patients. Age ($P < 0.001$) and risk score ($P < 0.001$) are independent prognostic risk factors for GC.

Construction and Validation of the Nomogram

We used factors such as age, grade, stage, T, M, N, and risk score to construct a nomogram to predict the survival rate of GC patients more conveniently. According to the nomogram, the scores of GC patients are calculated and then added to obtain the total score, thereby predicting the survival probability of 1 year

and 3 years, which is beneficial to guide clinical decision-making. So the closer the calibration curve is to the diagonal, the more accurate the prediction result will be. The calibration curves of the nomogram show that the nomogram has good accuracy in predicting survival rates at 1 and 3 years (Figure.11B,C). The 1-year (AUC = 0.657) and 3-year (AUC = 0.608) ROC curves also show that the forecasting ability of the nomogram is very accurate. (Figure.11D)

CHSY3 Expression Is Correlated With Immune Infiltration Level in GC

Tumor infiltrating lymphocyte (TIL), a component of the tumor microenvironment, has been found to be associated with cancer prognosis and treatment response [18]. Therefore, we used the TIMER database to study the relationship between the expression of CHSY3 and the level of immune infiltration in gastric cancer. The results show that the high CHSY3 expression level had positive correlations with infiltrating levels of CD4⁺ T cells ($r=0.302, P=3.82e-09$), macrophages ($r=0.533, P=1.40e-28$), neutrophils ($r=0.224, P=1.28e-05$) and dendritic ($r=0.347, P=6.12e-12$). While CHSY3 expression has no significant correlations with tumor purity ($r=-0.123, P=1.69e-02$), B cell ($r=-0.071, P=1.72e-01$) and CD8⁺ T cells ($r=0.097, P=6.20e-02$) (Figure.12A). Besides, different mutational forms of CHSY3 were associated with immune infiltration of 6 leukocytes (B cell, CD4⁺ T cells, CD8⁺ T cells, macrophage, neutrophil, dendritic) (Figure.12B). These studies indicate that CHSY3 plays an important role in the immune infiltration of gastric cancer.

Analysis of drug sensitivity associated with CHSY3

We divided 209 DEGs into 4 groups and performed Spearman correlation analysis with small molecule/drug sensitivity (IC₅₀) to explore the correlation between DEGs and drug sensitivity. The results (Figure.13) showed that CHSY3 was significantly related to the drug resistance of many chemotherapeutic drugs and tumor-targeted drugs, including 5-Fluorouracil, Methotrexate, TAK715, Bleomycin, TG-101348, TPCA-1, PHA-793887, AT-7519, Belinostat, CUDC-101, vorinostat and so on.

Discussion

Research has shown that cancer plays an important role in affecting human morbidity and mortality [19]. Gastric cancer is the most common malignant tumor of the digestive system. At present, the diagnosis of gastric cancer mainly relies on endoscopy [20]. Many patients are diagnosed with advanced tumors as soon as they are checked, and have lost the opportunity for surgery. Obviously, this diagnostic method has certain limitations [21]. Early diagnosis of cancer is the key to improving the overall survival rate of patients, reducing disease-free progression, and reducing the risk of recurrence. Therefore, there is an urgent need to determine reliable predictive biomarkers and new therapeutic targets to truly improve the prognosis of patients. Studies have shown that abnormal gene expression is closely related to the occurrence and development of cancer [22, 23]. Nan Zhou's research shows that HTRA3 is abnormally highly expressed in gastric cancer tissues, and the high expression of HTRA3 indicates a poor

prognosis[24].Liya Hu's research found that PDPN is a prognostic biomarker of gastric cancer and is related to immune infiltration[25].Li Rong's research confirmed that COL1A2 is a prognostic biomarker of gastric cancer through bioinformatics and Meta-analysis, which can improve the clinical predictive ability of gastric cancer[26].Therefore, it is very necessary to carry out bioinformatics research on gastric cancer.

CHSY3 (Chondroitin Sulfate Synthase 3) is a Protein Coding gene,An important paralog of this gene is CHSY1.At present, the research on CHSY3 bioinformatics is relatively scarce.In our research, we mainly focus on the role of CHSY3 in gastric cancer. First, we studied the differential expression of CHSY3 in a variety of cancers in the TIMER database, and determined that CHSY3 is highly expressed in the tissues of Breast invasive carcinoma(BRCA),Cholangiocarcinoma(CHOL),Colon adenocarcinoma(COAD),Esophageal carcinoma(ESCA),Head and Neck squamous cell carcinoma(HNSC),Kidney Chromophobe(KICH),Kidney renal clear cell carcinoma(KIRC),Prostate adenocarcinoma(PRAD) and Rectum adenocarcinoma(READ).While it is significantly lower in Bladder Urothelial Carcinoma(BLCA),Liver hepatocellular carcinoma(LIHC),Lung adenocarcinoma(LUAD),Lung squamous cell carcinoma(LUSC),Thyroid carcinoma(THCA) and Uterine Corpus Endometrial Carcinoma(UCEC) than in the respective normal tissues.These results indicate that CHSY3 is abnormally expressed in a variety of cancers.Subsequently, we downloaded the transcriptome data and clinical information of gastric cancer patients from the TCGA database and GEO database. We drew a scatter plot, a paired plot, and studied five datasets(GSE13911,GSE19826,GSE54129,GSE66229,GSE79973) in the GEO database, and drew box scatter plots to explore the differential expression of CHSY3 in gastric cancer tissues. Meta-analysis of the data of TCGA and GEO has finally confirmed that CHSY3 is significantly higher in gastric cancer tissues than normal tissues(SWD=1.09,95% CI: 0.83-1.35; $I^2=44\%$, $P=0.134$).In order to judge the accuracy of CHSY3 in predicting the GC value, we performed ROC curve analysis and SROC curve analysis on the data of TCGA and GEO. The results show that CHSY3 does have a high diagnostic value.At the same time, our analysis also found that in gastric cancer, the expression of CHSY3 has a significant correlation with the T stage.Subsequently, the results of survival analysis showed that a high CHSY3 expression level in gastric cancer was related to poor OS.

We performed univariate and multivariate Cox proportional hazards regression analysis on the data of TCGA and GEO, and performed a meta-analysis on the results of the analysis, thereby confirming that the high expression of CHSY3 is an independent predictor of the prognosis of gastric cancer.Moreover,methylation studies showed that the higher the expression level of CHSY3 in GC, the higher the methylation level of CG06610705 and CG11572844, suggesting a poor prognosis of gastric cancer patients.In order to further investigate the potential regulatory mechanism of CHSY3 in gastric cancer.We identified the co-expressed genes of CHSY3 through the Linkedomics platform, and performed KEGG analysis and GO analysis on the top 50 positively related genes and negatively related genes. The results identified many pathways related to cancer. Including PI3K-Akt signaling pathway[27],Human papillomavirus infection[28],ECM-receptor interaction[29],Proteoglycans in cancer[30],MAPK signaling pathway[31],TGF-beta signaling pathway[32],cGMP-PKG signaling pathway[33],Small cell lung cancer,transmembrane receptor protein serine/threonine[34],kinase signaling pathway[35],focal

adhesion[36],endoplasmic reticulum lumen[37],basement membrane[38],extracellular matrix[39],integrin-binding[40],heparin-binding[41],growth factor binding[42] and glycosaminoglycan binding[43].A large number of studies have confirmed that it is related to the occurrence and development of tumors and tumor immune infiltration.

We can calculate the risk score of each patient based on the prognostic characteristics. The study found that risk score is an independent risk factor affecting prognosis and can be used to predict OS in patients with gastric cancer. High-risk patients have a poor prognosis for lower-risk patients. Therefore, we speculate that the three DEGs that constitute prognostic signals are involved in the progression of GC. Through the ROC curve analysis of OS of GC patients, we found that this prognostic signature has a good predictive value for gastric cancer (AUC = 0.657) and can be used to predict the prognosis of GC patients. In order to facilitate clinical application and better predict the prognosis of patients with gastric cancer, we constructed a nomogram to predict the survival rate of patients at 1 and 3 years. Judging from the calibration curve and ROC curve, the nomogram has a better predictive effect.

It is well known that tumor infiltration lymphol (TIL) plays an important role in the occurrence and development of tumors [44, 45]. The study by Yi-Ru Yu et al. showed that the metabolic challenges in tumors weaken the metabolic fitness and anti-tumor activity of tumor-infiltrating T lymphocytes (TILs), thereby promoting tumor progression[46].Our results show that the high CHSY3 expression level had positive correlations with infiltrating levels of CD4⁺ T cells($r=0.302, P=3.82e-09$), macrophages($r=0.533, P=1.40e-28$), neutrophils($r=0.224, P=1.28e-05$) and dendritic ($r=0.347, P=6.12e-12$).In addition,different mutational forms of CHSY3 were associated with the immune infiltration of 6 leukocytes.These results indicate that the role of CHSY3 in gastric cancer may be related to immune infiltration.

At present, studies have confirmed that abnormal gene expression is related to chemotherapy resistance[47, 48]. We used 209 DEGs to study in the GSCA database, and the results showed that the high expression of CHSY3 may be related to the resistance of gastric cancer chemotherapy and targeted therapy drugs, such as 5-Fluorouracil[49],Methotrexate[50],TAK715,Bleomycin[51],TG-101348[52],TPCA-1[53],PHA-793887[54],AT-7519[55],Belinostat,vorinostat[56],CUDC-101[57] and so on.Therefore, CHSY3 may become a target for the treatment of gastric cancer.

Our research has certain limitations. These studies are based on bioinformatics research, lack of experimental data and experiments are still needed to verify the role of CHSY3 in gastric cancer. In addition,It is necessary to conduct more in-depth prospective clinical research on the signature and nomogram.

In summary, CHSY3 is highly expressed in gastric cancer tissues, and the up-regulation of CHSY3 expression is closely related to T staging, indicating a poor prognosis, and the high expression of CHSY3 can be used as an independent predictor of the prognosis of gastric cancer.At the same time, we constructed a risk signature, including APOD, GJA1 and SERPINE1, which confirmed that risk score was an independent predictor of gastric cancer prognosis.For better clinical application, we have constructed

a nomogram with a higher predictive value. In addition, the high expression of CHSY3 is related to the increase of CD4 + T cells, macrophages, neutrophils, and dendritic, and it is also related to chemotherapy resistance. Therefore, CHSY3 may be a reliable prognostic biomarker in GC and may play an important role in tumor regression and immune cell infiltration. May serve as a new target for gastric cancer treatment.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article.

Competing interests

The authors declare that they have no competing interests.

Funding

Youth Science and Technology Research Fund of Shanxi Province [201901D211480]

Authors' contributions

Jiaxin Fan and Chaojie Liang: conception and design. Jiaxin Fan and He Huang : acquisition, analysis, and interpretation of data. Jiajia Wang and Chaowei Liang: figures drawing. Jiaxin Fan : writing and revision of manuscript. Chaojie Liang: study supervision. All authors read and approved the final manuscript.

Acknowledgements

My deepest gratitude is my teacher, Professor Liang Chaojie, for his continuous encouragement. He guided me through all stages of this thesis. Without his consistent and enlightening guidance, this paper would not have reached its current form. I am also very grateful to Teacher Huang He and Liang Chaowei for their guidance and assistance in the past two years. Of course, thanks to the support of the Youth Science and Technology Research Fund. Finally, I would like to thank my sweetheart, Ms. Wang Jiajia, for her care and great trust in me over the years, for giving me help and time during the difficult process of my thesis, listening to my opinions, and helping me to solve the problem.

Authors' information

First author: Fan Jiabin, male, doctor, master student, engaged in clinical and basic research of gastrointestinal tumor E-mail:fjx952859576@163.com.

Huang He, male, chief physician, master tutor, mainly engaged in clinical and basic research of gastrointestinal tumors. E-mail:hh93003@163.com.

Corresponding author: Chaojie Liang, male, chief physician, master supervisor, engaged in clinical and basic research of general surgery. E-mail:Email:liangchaojie8@126.com.

References

1. Wu S, Powers S, Zhu W, Hannun YA. Substantial contribution of extrinsic risk factors to cancer development. *Nature*. 2016;529:43-7.
2. Russnes HG, Lønning PE, Børresen-Dale AL, Lingjærde OC. The multitude of molecular analyses in cancer: the opening of Pandora's box. *Genome Biol*. 2014;15:447.
3. Yan H, Siu HC, Law S, Ho SL, Yue S, Tsui WY, et al. A Comprehensive Human Gastric Cancer Organoid Biobank Captures Tumor Subtype Heterogeneity and Enables Therapeutic Screening. *Cell Stem Cell*. 2018;23:882-97.e11.
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.
5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70:7-30.
6. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin*. 2019;69:363-85.
7. Pietrantonio F, Miceli R, Raimondi A, Kim YW, Kang WK, Langley RE, et al. Individual Patient Data Meta-Analysis of the Value of Microsatellite Instability As a Biomarker in Gastric Cancer. *J Clin Oncol*. 2019;37:3392-400.
8. Shitara K, Iwata H, Takahashi S, Tamura K, Park H, Modi S, et al. Trastuzumab deruxtecan (DS-8201a) in patients with advanced HER2-positive gastric cancer: a dose-expansion, phase 1 study. *Lancet Oncol*. 2019;20:827-36.
9. Ogawa H, Hatano S, Sugiura N, Nagai N, Sato T, Shimizu K, et al. Chondroitin sulfate synthase-2 is necessary for chain extension of chondroitin sulfate but not critical for skeletal development. *PLoS One*. 2012;7:e43806.

10. Hashimoto S, Ogoshi K, Sasaki A, Abe J, Qu W, Nakatani Y, et al. Coordinated changes in DNA methylation in antigen-specific memory CD4 T cells. *J Immunol.* 2013;190:4076-91.
11. Wilson DG, Phamluong K, Lin WY, Barck K, Carano RA, Diehl L, et al. Chondroitin sulfate synthase 1 (Chsy1) is required for bone development and digit patterning. *Dev Biol.* 2012;363:413-25.
12. Li Y, Laue K, Temtamy S, Aglan M, Kotan LD, Yigit G, et al. Temtamy preaxial brachydactyly syndrome is caused by loss-of-function mutations in chondroitin synthase 1, a potential target of BMP signaling. *Am J Hum Genet.* 2010;87:757-67.
13. Yin L. Chondroitin synthase 1 is a key molecule in myeloma cell-osteoclast interactions. *J Biol Chem.* 2005;280:15666-72.
14. Momose T, Yoshimura Y, Harumiya S, Isobe K, Kito M, Fukushima M, et al. Chondroitin sulfate synthase 1 expression is associated with malignant potential of soft tissue sarcomas with myxoid substance. *Hum Pathol.* 2016;50:15-23.
15. Tian J, Ling L, Shboul M, Lee H, O'Connor B, et al. Loss of CHSY1, a secreted FRINGE enzyme, causes syndromic brachydactyly in humans via increased NOTCH signaling. *Am J Hum Genet.* 2010;87:768-78.
16. Liu CH, Lan CT, Chou JF, Tseng TJ, Liao WC. CHSY1 promotes aggressive phenotypes of hepatocellular carcinoma cells via activation of the hedgehog signaling pathway. *Cancer Lett.* 2017;403:280-8.
17. Zeng L, Qian J, Luo X, Zhou A, Zhang Z, Fang Q. CHSY1 promoted proliferation and suppressed apoptosis in colorectal cancer through regulation of the NFκB and/or caspase-3/7 signaling pathway. *Oncol Lett.* 2018;16:6140-6.
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. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019;69:7-34.
20. Pasechnikov V, Chukov S, Fedorov E, Kikuste I, Leja M. Gastric cancer: prevention, screening and early diagnosis. *World J Gastroenterol.* 2014;20:13842-62.
21. Matsumoto K, Ueyama H, Yao T, Abe D, Oki S, Suzuki N, et al. Diagnostic limitations of magnifying endoscopy with narrow-band imaging in early gastric cancer. *Endosc Int Open.* 2020;8:E1233-1233E1242.
22. Jamali L, Tofigh R, Tutunchi S, Panahi G, Borhani F, Akhavan S, et al. Circulating microRNAs as diagnostic and therapeutic biomarkers in gastric and esophageal cancers. *J Cell Physiol.* 2018;233:8538-50.

23. Naeli P, Pourhanifeh MH, Karimzadeh MR, Shabaninejad Z, Movahedpour A, Tarrahimofrad H, et al. Circular RNAs and gastrointestinal cancers: Epigenetic regulators with a prognostic and therapeutic role. *Crit Rev Oncol Hematol*. 2020;145:102854.
24. Ji C, Sun LS, Xing F, Niu N, Gao HL, Dai JW, et al. HTRA3 Is a Prognostic Biomarker and Associated With Immune Infiltrates in Gastric Cancer. *Front Oncol*. 2020;10:603480.
25. Hu L, Zhang P, Sun W, Zhou L, Chu Q, Chen Y. PDPN is a prognostic biomarker and correlated with immune infiltrating in gastric cancer. *Medicine (Baltimore)*. 2020;99:e19957.
26. Rong L, Huang W, Tian S, Chi X, Zhao P, Liu F. COL1A2 is a Novel Biomarker to Improve Clinical Prediction in Human Gastric Cancer: Integrating Bioinformatics and Meta-Analysis. *Pathol Oncol Res*. 2018;24:129-34.
27. Singh SS, Yap WN, Arfuso F, Kar S, Wang C, Cai W, et al. Targeting the PI3K/Akt signaling pathway in gastric carcinoma: A reality for personalized medicine. *World J Gastroenterol*. 2015;21:12261-73.
28. de Sanjosé S, Brotons M, Pavón MA. The natural history of human papillomavirus infection. *Best Pract Res Clin Obstet Gynaecol*. 2018;47:2-13.
29. Wang J, Yu S, Chen G, Kang M, Jin X, Huang Y, et al. A novel prognostic signature of immune-related genes for patients with colorectal cancer. *J Cell Mol Med*. 2020;24:8491-504.
30. Vicente CM, da Silva DA, Sartorio PV, Silva TD, Saad SS, Nader HB, et al. Heparan Sulfate Proteoglycans in Human Colorectal Cancer. *Anal Cell Pathol (Amst)*. 2018;2018:8389595.
31. Lei L, An G, Zhu Z, Liu S, Fu Y, Zeng X, et al. C8orf48 inhibits the tumorigenesis of colorectal cancer by regulating the MAPK signaling pathway. *Life Sci*. 2021;266:118872.
32. Xiong R, Gao JL, Yin T. G3BP1 activates the TGF- β /Smad signaling pathway to promote gastric cancer. *Onco Targets Ther*. 2019;12:7149-56.
33. Xiang T, Yuan C, Guo X, Wang H, Cai Q, Xiang Y, et al. The novel ZEB1-upregulated protein PRTG induced by *Helicobacter pylori* infection promotes gastric carcinogenesis through the cGMP/PKG signaling pathway. *Cell Death Dis*. 2021;12:150.
34. McCarthy AJ, Chetty R. Smad4/DPC4. *J Clin Pathol*. 2018;71:661-4.
35. Rane CK, Minden A. P21 activated kinase signaling in cancer. *Semin Cancer Biol*. 2019;54:40-9.
36. Eke I, Cordes N. Focal adhesion signaling and therapy resistance in cancer. *Semin Cancer Biol*. 2015;31:65-75.
37. Wang Y, Wang K, Jin Y, Sheng X. Endoplasmic reticulum proteostasis control and gastric cancer. *Cancer Lett*. 2019;449:263-71.

38. Chang J, Chaudhuri O. Beyond proteases: Basement membrane mechanics and cancer invasion. *J Cell Biol.* 2019;218:2456-69.
39. Lu P, Weaver VM, Werb Z. The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol.* 2012;196:395-406.
40. Mitra SK, Schlaepfer DD. Integrin-regulated FAK-Src signaling in normal and cancer cells. *Curr Opin Cell Biol.* 2006;18:516-23.
41. Atallah J, Khachfe HH, Berro J, Assi HI. The use of heparin and heparin-like molecules in cancer treatment: a review. *Cancer Treat Res Commun.* 2020;24:100192.
42. Erdogan B, Webb DJ. Cancer-associated fibroblasts modulate growth factor signaling and extracellular matrix remodeling to regulate tumor metastasis. *Biochem Soc Trans.* 2017;45:229-36.
43. Morla S. Glycosaminoglycans and Glycosaminoglycan Mimetics in Cancer and Inflammation. *Int J Mol Sci.* 2019;20.
44. Callahan MK, Wolchok JD. Recruit or Reboot? How Does Anti-PD-1 Therapy Change Tumor-Infiltrating Lymphocytes. *Cancer Cell.* 2019;36:215-7.
45. Demaria O, Vivier E. Immuno-Oncology beyond TILs: Unleashing TILCs. *Cancer Cell.* 2020;37:428-30.
46. Yu YR, Imrichova H, Wang H, Chao T, Xiao Z, Gao M, et al. Disturbed mitochondrial dynamics in CD8(+) TILs reinforce T cell exhaustion. *Nat Immunol.* 2020;21:1540-51.
47. Klintman M, Buus R, Cheang MC, Sheri A, Smith IE, Dowsett M. Changes in Expression of Genes Representing Key Biologic Processes after Neoadjuvant Chemotherapy in Breast Cancer, and Prognostic Implications in Residual Disease. *Clin Cancer Res.* 2016;22:2405-16.
48. Litviakov NV, Cherdyntseva NV, Tsyganov MM, Denisov EV, Garbukov EY, Merzliakova MK, et al. Changing the expression vector of multidrug resistance genes is related to neoadjuvant chemotherapy response. *Cancer Chemother Pharmacol.* 2013;71:153-63.
49. Hashimoto Y, Yoshida Y, Yamada T, Aisu N, Yoshimatsu G, Yoshimura F, et al. Current Status of Therapeutic Drug Monitoring of 5-Fluorouracil Prodrugs. *Anticancer Res.* 2020;40:4655-61.
50. Widemann BC, Adamson PC. Understanding and managing methotrexate nephrotoxicity. *Oncologist.* 2006;11:694-703.
51. Constantin D, Widmann C. ASH2L drives proliferation and sensitivity to bleomycin and other genotoxins in Hodgkin's lymphoma and testicular cancer cells. *Cell Death Dis.* 2020;11:1019.
52. Liu Q, Li A, Wang L, He W, Zhao L, Wu C, et al. Stomatin-like Protein 2 Promotes Tumor Cell Survival by Activating the JAK2-STAT3-PIM1 Pathway, Suggesting a Novel Therapy in CRC. *Mol Ther Oncolytics.*

53. Cataldi M, Shah NR, Felt SA, Grdzlishvili VZ. Breaking resistance of pancreatic cancer cells to an attenuated vesicular stomatitis virus through a novel activity of IKK inhibitor TPCA-1. *Virology*. 2015;485:340-54.

54. Alzani R, Pedrini O, Albanese C, Ceruti R, Casolaro A, Patton V, et al. Therapeutic efficacy of the pan-cdk inhibitor PHA-793887 in vitro and in vivo in engraftment and high-burden leukemia models. *Exp Hematol*. 2010;38:259-69.e2.

55. Xi C, Wang L, Yu J, Ye H, Cao L, Gong Z. Inhibition of cyclin-dependent kinases by AT7519 is effective to overcome chemoresistance in colon and cervical cancer. *Biochem Biophys Res Commun*. 2019;513:589-93.

56. Kenny RG, Ude Z, Docherty JR, Marmion CJ. Vorinostat and Belinostat, hydroxamate-based anti-cancer agents, are nitric oxide donors. *J Inorg Biochem*. 2020;206:110981.

57. Ji M, Li Z, Lin Z, Chen L. Antitumor activity of the novel HDAC inhibitor CUDC-101 combined with gemcitabine in pancreatic cancer. *Am J Cancer Res*. 2018;8:2402-18.

Table

TABLE:1 The multivariate Cox analysis of the 7 DEGs previously screened resulted in a prognostic signature composed of 3 DEGs. The coef value of each DEGs expression in the risk score calculation formula was the regression coefficient of the prognostic signature.

id	coef	HR	HR.95L	HR.95H	pvalue
APOD	0.09298541	1.09744572	1.00913806	1.19348101	0.02981838
GJA1	0.13241755	1.14158489	0.95537573	1.36408748	0.14497942
SERPINE1	0.17713065	1.19378706	1.05032874	1.35683951	0.00669428

Figures

Figure 1

Differential expression levels of CHSY in different malignancies and CHSY-related differentially expressed genes (DEGs). (A) Increased or decreased CHSY of different cancers compared with normal tissues in the TCGA. (B, C) Differential expression levels of CHSY in GC. (D) A ROC curve to test the value of CHSY to identify GC tissues was created. (E) The immunohistochemical smear of CHSY in gastric cancer tissue.

Figure 2

(A-E) CHSY levels in GC tissues and normal gastric tissues in the GSE13911(A),GSE19826(B),GSE54129(C),GSE66229(D)andGSE79973(E)datasets.(F-J) ROC curves were drawn using GSE13911(F),GSE19826(G),GSE54129(H),GSE66229(I)andGSE79973(J) datasets.

Figure 3

Forest plots exhibit diagnostic performance of CHSY3 in GC.

Figure 4

Association with CHSY3 expression and clinicopathological characteristics, including (A) age, (B) grade, (C) pathologic stage, (D) M stage (E) N stage, and (F) T stage in GC patients in TCGA cohort.

Figure 5

Survival analysis based on TCGA and GEO database.(A)TCGA, $P=0.015$,(B)GSE15459, $P=0.031$,
(C)GSE22377, $P=0.95$ (D)GSE51105, $P=0.1$ (E)GSE62254, $P<0.001$ (F)GSE84437, $P<0.05$.

Figure 6

(A)Meta-analysis of survival analysis results;(B)Meta-analysis of Cox regression analysis results; (C)Univariate regression analysis and (D)multivariate regression analysis based on TCGA database; (E)Univariate regression analysis and (F)multivariate regression analysis based on GSE84437 dataset; (G)Univariate regression analysis and (H)multivariate regression analysis based on GSE62254 dataset.

Figure 7

The correlation between CHSY expression level and methylation degree on cg04729562(A),cg06610705(B),cg10678749(C),cg11572844(D) and cg18829263(E) .Relationship between the overall survival rate of GC patients and the degree of methylation on cg06610705(F) and cg11572844(G).

Figure 8

CHSY co-expression genes in GC.(A) The global CHSY significantly correlated genes in the GC cohort were identified by LinkedOmics.(B) Heatmaps showing top 50 genes positively and negatively correlated with CHSY in GC. Red dot,positively correlated gene; blue dot, negatively correlated genes.(D)KEGG analysis of CHSY co-expressed genes,(E)GO analysis of CHSY co-expressed genes.

Figure 9

Construction of the DEGs prognostic signature for GC patients.(A)(B)Perform LASSO analysis on 95 DEGs determined by univariate Cox analysis, and 7 best candidate DEGs were obtained for constructing a prognostic signature.(C) The heatmap of the 3 related gene expression profiles in high- and low-risk GC patients.(D)Distribution of risk scores of high- and low-risk GC patients.(E)Scatter plot shows the correlation between survival time and risk score.

Figure 10

(A) The K-M curve reflects that the OS of high-risk GC patients is significantly lower than that of low-risk patients ($P < 0.001$). (B) ROC curve reflects that the prediction accuracy of the risk score is higher than other clinical features ($AUC = 0.657$). (C) The forest plot reflects the univariate Cox analysis of the relationship between the clinical features, risk score and OS of GC patients. Age ($P = 0.006$), stage ($P < 0.001$), T ($P = 0.032$), N ($P = 0.006$), M ($P = 0.025$) and risk score ($P < 0.001$) significantly affect the prognosis of GC patients. (D) Forest plot reflects multivariate Cox analyzed the relationship between the clinical features, risk score, and OS of GC patients. Age ($P < 0.001$) and risk score ($P < 0.001$) are independent prognostic risk factors for GC.

Figure 11

Construction and validation of the nomogram.(A)Calculate the scores of each item of GC patients according to the nomogram, and the total scores obtained after addition can predict the 1- and 3-year survival probability.(B)(C) The 1- and 3-year calibration curves of the nomogram.(D)The ROC curves of 1- and 3-year nomogram (AUC = 0.657 of 1 year, AUC = 0.608 of 3 years).

Figure 12

(A) Correlation of CHSY expression with immune infiltration level in STAD (stomach adenocarcinoma). CHSY expression is significant positive correlations with infiltrating levels of CD4 + T cells, macrophages, neutrophils, and dendritic cells in STAD (stomach adenocarcinoma). While CHSY expression has no significant correlations with tumor purity, B cell, and CD8 + T cells. (B) The relationships between infiltration levels of 6 immune cells and copy number of CHSY3.

Figure 13

Analysis of drug sensitivity associated with CHSY3. Red shows a positive correlation, with higher gene expression associated with greater drug sensitivity, while blue shows the opposite.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Abbreviations.xlsx](#)