MEDICAL SCIENCE MONITOR

LAB/IN VITRO RESEARCH

e-ISSN 1643-3750
© Med Sci Monit, 2019; 25: 6213-6220
DOI: 10.12659/MSM.914815

Received: 2018.12.26
Accepted: 2019.04.03
Published: 2019.08.18

A 5-Gene Prognostic Combination for Predicting Survival of Patients with Gastric Cancer

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

BCDE 1 Liang Song
DE 2 Xiao-Yan Wang
AEF 3 Xiao-Feng He

1 Endoscopy Room, Heping Hospital Affiliated to Changzhi Medical College, Changzhi, Shanxi, P.R. China
2 Department of Epidemiology and Health Statistics, Basic Medical College of Zhejiang University of Traditional Chinese Medicine, Hangzhou, Zhejiang, P.R. China
3 Department of Science and Education, Heping Hospital Affiliated to Changzhi Medical College, Changzhi, Shanxi, P.R. China

Corresponding Author: Xiao-Feng He, e-mail: 393120823@qq.com
Source of support: Departmental sources

Background: The aim of the study was to identify a multigene prognostic factor in patients with gastric cancer (GC).

Material/Methods: Random survival forest (RSF) was performed to screen survival-related genes and develop a multigene combination based on the cumulative hazard function of each GC patient in TCGA-STAD and GSE15459. Kaplan-Meier curve and univariate and multivariable Cox proportional hazards regression model were applied to evaluate the prognostic performance of the 5-gene combination. C-index was used to compare the prognostic performance of the 5-gene combination and another 9-gene signature in GC. Gene set enrichment analysis (GSEA) was conducted.

Results: We obtained 19 survival-related genes through univariate Cox proportional hazards analysis in the training set, 5 of which were identified and were used to develop a 5-gene combination through RSF. Patients in the 5-gene combination low-risk group had better overall survival (OS) than those in the 5-gene combination high-risk group, and the 5-gene combination was demonstrated to be an independent prognostic factor in patients with GC. The 5-gene combination outperformed the 9-gene signature in predicting the OS of GC patients, and it might affect the prognosis of GC patients through E2F signaling, MYC signaling, and G2M checkpoint.

Conclusions: We introduce a 5-gene combination that can predict the survival of GC patients and might be an independent prognostic factor in GC.

MeSH Keywords: Prognosis • Stomach Neoplasms • Survival Analysis

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/914815

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Background

Gastric cancer (GC) is the third leading cause of cancer-related deaths, with about one million newly diagnosed GC annually [1,2]. GC can be classified into diffuse and intestinal types on the basis of its histological appearance, and can also be categorized into cardia and non-cardia tumor based on its location [3,4]. For patients with early-stage GC, surgical treatment with adequate extended lymphadenectomy can bring about good prognosis [5–7]. Nevertheless, more than half of GC patients with advanced disease will eventually develop local relapse or distant metastases or have unresectable disease [6–8]. As a result, the clinical outcomes of GC patients remain extremely poor (the 5-year survival remains about 5–10%) [4,7,8]. These dismal clinical outcomes stress the need to develop biomarkers that help to identify at-risk patients who might benefit from early interventions.

Random survival forest (RSF), which has been widely used to identify risk features of different diseases in clinical settings, is an exploratory analysis algorithm for right-censored highly correlated complex survival data, and uses a collection of decision trees for prediction and to rank variables by their importance for time to event. It can handle very high-dimensional data without having to make feature selection, and can internally produce an unbiased estimate of the error after generalization when constructing a forest. It is a good method for estimating missing data and maintaining accuracy if a large portion of the data is lost, and calculates the intimacy in each case, which is very useful for data mining, detecting outliers, and visualizing data. Moreover, the learning process is very fast [9,10].

Thanks to advancements in genetic technology, various GC gene expression studies have been published and promoted for prognostic biomarker identification [11–14]. However, the prognostic performance of established biomarkers is controversial and limited. In the present study, we identified and developed a 5-gene combination to predict the survival of patients with GC by using RSF.

Material and Methods

GC gene expression study

The GC gene expression study TCGA-STAD (level 3 data) [13], measured by using the Illumina HiSeq 2000 RNA Sequencing platform by the University of North Carolina TCGA genome characterization center, was shown as in log2(x+1) transformed RSEM normalized count and downloaded from UCSC Xena (*https://xenabrowser.net/datapages/*). The related clinical information (including age, gender, pathological stage, survival information, and relapse information) of samples in TCGA-STAD was also downloaded. GSE15459 [11,12,15,16] was measured by Affymetrix Human Genome U133 Plus 2.0 Array and included 200 GC samples, and we downloaded the MAS.5.0 normalized gene expression profile and clinical information (age, gender, pathological stage, and histological type) of these samples from the Gene Expression Omnibus (GEO) database (*https://www.ncbi.nlm.nih.gov/geo/*). Expression profiles of these GC studies were scaled, and genes with near-zero variances were removed for subsequent analysis.

Statistical analysis

At first, we screened survival-related genes in the TCGA-STAD using univariate Cox proportional hazards regression model. Genes at P value <0.05 and familywise error rate less than 0.01 were included for subsequent analysis. Then, patients in the TCGA-STAD were randomly divided into a training set and a test set in a 1: 1 ratio. Thus, we employed the RSF method introduced by Ishwaran et al. [17,18] to perform survival analysis on the basis of the survival-related genes and the overall survival (OS) of GC patients in the training set. As introduced previously, the RSF [17,18], which extends Breiman's random forests (RF) method [19], was an ensemble tree algorithm for the analysis of right-censored survival data. We grew 1000 trees for each RSF, and each tree of the forests was grown by splitting patients by comparing survival differences via log-rank test based on a randomly selected subset of variables at each node. We selected the features using a variable hunting method, with the number of Monte Carlo iterations (nrep) and value to control step size used in the forward process (nstep) set as 100 and 5, respectively. The test sets were dropped down to the trees for prediction when trees were constructed. The cumulative hazard function, or Nelson–Aalen estimator [19], was derived from each tree, and an ensemble cumulative hazard function with an average over 1000 survival trees was determined in the training set and test set. The RSF and variable hunting algorithms were implemented in the R package "randomForestSRC" [20]. Mortality was obtained as a weighted sum over ensemble cumulative hazard functions for each individual for all unique death time-points. Higher mortality values correspond to higher risk. Therefore, we applied mortality as a risk score to divide patients into a low-risk group and a high-risk group according to the optimal cutoff derived from time-dependent survival receiver operating characteristic (ROC) analysis using the R package "survivalROC" [21] in the training set, test set, and validation set. We performed Kaplan-Meier survival curve analyses to evaluate the OS of GC in the 2 groups. Furthermore, we constructed univariate and multivariable Cox proportional hazards regression models to investigate whether the risk score was an independent prognostic factor in patients with GC.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Comparing prognostic performance with other study

Wang et al. introduced a 9-gene combination for predicting the survival of patients with GC [22]. Thus, we tried to compare the prognostic performance of the 9-gene combination and our 4-gene combination based on the concordance index (C-index) calculated based on the mRNA expression levels of the 9 genes (NR1I2, LGALSL, C1ORF198, CST2, LAMP5, FOXS1, CES1P1, MMP7 and COL8A1) Wang et al. introduced and the 5 genes in the present study. Higher C-index implied better prognostic performance. The R package "survcomp" was used to calculate and compare the C-indexes between the 2 prognostic combinations [23], and the *t* test for dependent samples was used to compare the C-indexes for different models.

Gene set enrichment analysis (GSEA)

To investigate the potentially relevant molecular mechanisms that the risk score affected, the survival of GC patients, the GSEA [24,25] introduced by Subramanian et al., was used based on the high-risk group and low-risk group in the training set. The Hallmark gene set (h.all.v6.2) was used as a reference. Gene sets at normal P value <0.05 and false discovery rate <0.05 were considered significantly enriched.

Results

Characteristics of GC patients in the training set, test set, and validation set

A total of 388 GC patients with survival information were included in the present study, with half in each set. As shown in Supplementary Table 1, the characteristics of patients in the training set and test set were balanced (median age in years was 66 [range 34–86], 69 females [35.57%], and 125 males [64.43%] in the training set; median age in years was 68 [range 30–90], 67 females [34.54%], and 127 males [65.46%] in the test set; median age in years was 66.55 [range 23.4–92.4], 67 females [34.9%], and 125 males [65.1%] in the training set).

Development of the prognostic multigene combination and validation of its prognostic performance

A total of 19 survival-related genes were identified in the TCGA-STAD at the inclusion criteria P value <0.05 and familywise error rate less than 0.01. Then, the 19 survival-related genes were included in the RSF (Supplementary Table 2). After variable hunting, 5 genes were finally identified: FRMD7 (FERM domain containing 7), FLJ16779 (LOC100192386), PRR20A (proline rich 20A), SLC7A2 (solute carrier family 7 member 2), and SLC22A16 (solute carrier family 22 member 16). Thus, we included these 5 genes in the final RSF, from which we calculated the risk score of the 5-gene combination based on the sum of the cumulative hazard functions for each individual for all unique death time-points. According to the cutoff of 12.199 (shown in Figure 1A), we classified patients into the 5-gene combination high-risk group and 5-gene combination low-risk group in the training set, test set, and validation set. The results of Kaplan-Meier curve and univariate Cox proportional hazards regression models suggested that patients in the 5-gene combination low-risk group had better OS compared with those in the 5-gene combination high-risk group in the training set (hazards ratio (HR)=1.1136, 95% CI: 1.0952–1.1324, P<0.0001, Figure 1B and Supplementary Table 3), test set (HR=1.034, 95% CI: 11.0174–1.0511, P=0.0001, Figure 1C and Supplementary Table 4), and validation set (HR=1.0618, 95% CI: 1.0059–1.1209, P=0.0299, Figure 1D and Supplementary Table 5). Meanwhile, the results of multivariable Cox proportional hazards regression model suggested that the 5-gene combination might be an independent prognostic factor after adjusting for the other clinical factors such as gender, pathological stage, grade, and age (Supplementary Tables 3–5). These results suggested that the 5-gene signature could predict the OS of gastric cancer patients.

The prognostic performance of the 5-gene combination was better than the established prognostic signature in GC

As shown in Figure 2, the C-index of the 5-gene combination predicting the OS of patients was significantly higher compared to the 9-gene signature in the training set (0.73±0.01 *vs.* 0.54±0.01, P<0.001), test set (0.69±0.02 *vs.* 0.52±0.01, P<0.001), and validation set (0.68±0.02 *vs.* 0.53±0.01, P<0.001). These results suggested that the prognostic performance of the 5 gene combination was better compared with that of the 9-gene signature.

The results of the gene set enrichment analysis

We tried to investigate the potential mechanisms involved in the influence of the 5-gene combination on the OS of GC patients, and we classified the GC patients in the training set into the 5-gene combination low-risk group and the 5-gene combination high-risk group based on the risk score of each patient. The results suggested that GC samples in the 5-gene combination low-risk group were significantly enriched in E2F signaling (P=0.0038, FDR=0.0408), MYC signaling (P<0.0001, FDR=0.0285) and G2M checkpoint (P=0.0058, FDR=0.0282) (Table 1). These results suggested that the 5-gene combination might affect the OS of GC patients through E2F signaling, MYC signaling, and G2M checkpoint.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

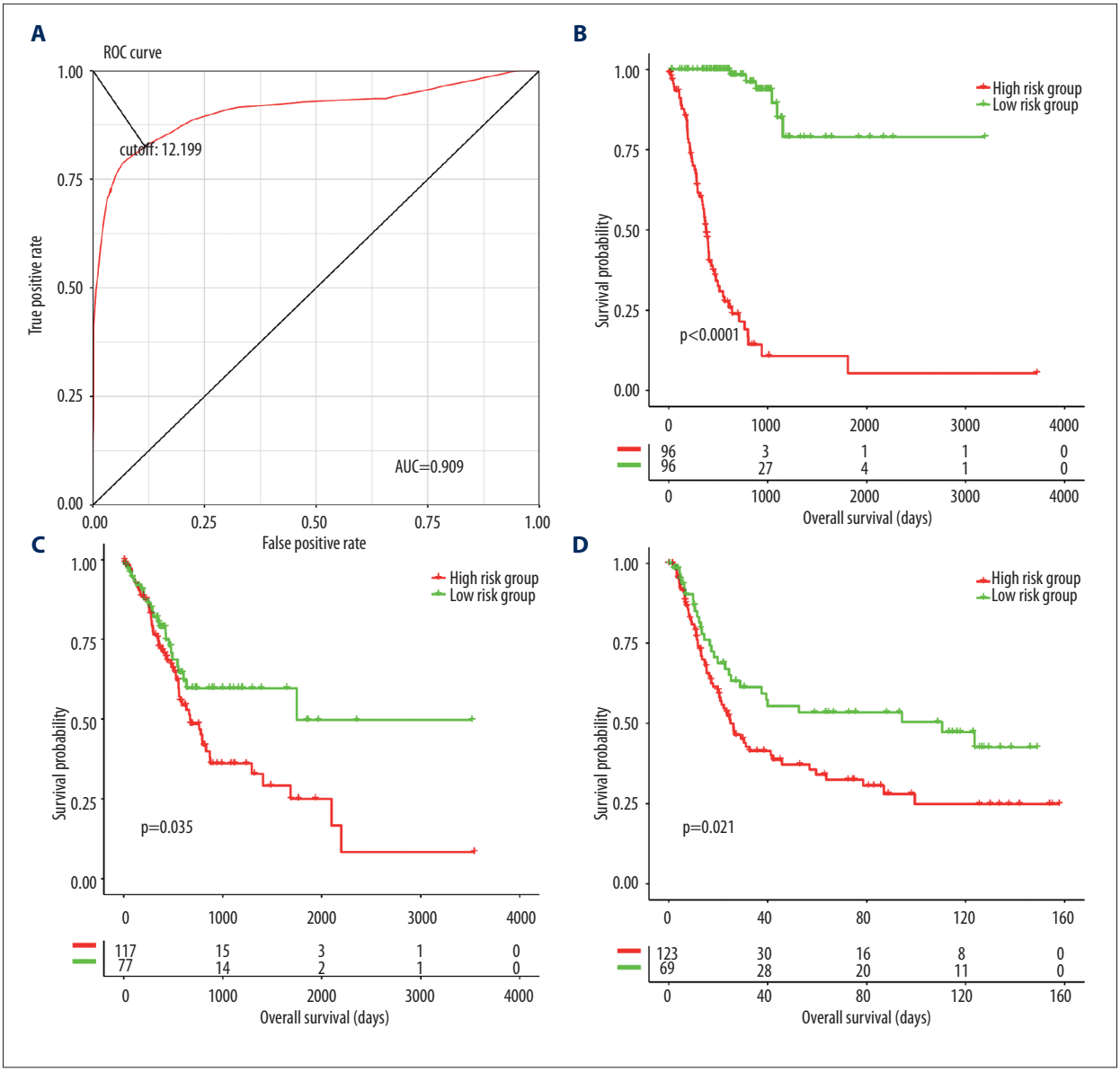

Figure 1. The prognosis role of the 5-gene combination in patients with gastric cancer. (**A**) Optimal cutoff to classify gastric cancer in to the 5-gene combination low-risk group and high-risk group. (**B**) The overall survival of gastric patients in the 5-gene combination low-risk group and 5-gene signature high-risk group in the training set. (**C**) The overall survival of gastric patients in the 5-gene combination low-risk group and 5-gene signature high-risk group in the test set. (**D**) The overall survival of gastric patients in the 5-gene combination low-risk group and 5-gene signature high-risk group in the validation set.

Discussions

Due to the low 5-year survival rate and high rate of recurrence of GC patients, as well as the fact that early diagnosis and early invention can improve the clinical outcomes of GC patients, identification and evaluation of novel biomarkers for patients with GC is of great significance [26-28]. In the present study, by using an RSF approach, we identified and validated a 5-gene combination (FRMD7, FLJ16779, PRR20A, SLC7A2, and SLC22A16) that could predict the survival of GC patients. Univariate and multivariable Cox proportional hazards regression analysis indicated that the 5-gene combination might be an independent prognostic factor in GC. We compared the prognostic performance of the 5-gene combination with another 9-gene based prognostic combination introduced by Wang et al. [22], which suggested that our 5-gene combination performed better. The results of GSEA indicated that the 5-gene combination might affect the OS of GC patients through E2F signaling, MYC signaling, and G2M checkpoint.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

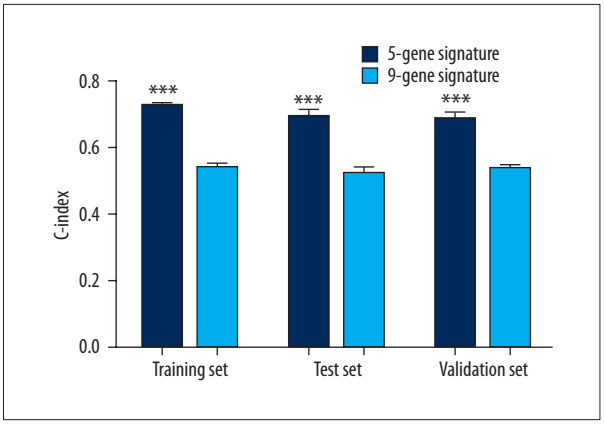

Figure 2. Comparison of the C-index of the 5-gene combination and the 9-gene signature in the training set, test, and validation set.

RSF, which applies a combination of decision trees for prediction and to rank variables according to their importance for survival data, is specifically designed for exploratory analysis of right-censored survival data of prospective cohorts where the outcome is a time-dependent variable, and to reduce the data dimension by picking out the most important variables correlated with the survival time of interest [29,30]. Thus, we applied this robust algorithm to identify the most survival-related genes, and translated them into a 5-gene combination to predict the OS of GC patients by the cumulative hazard function of each patient at each time point. Among the 5 genes, we noticed that SLC7A2 and SLC22A16 had been reported to be involved in the pathogenesis of several human cancers. Coburn et al. [31] and Tozlu et al. [32] demonstrated that SLC7A2 participates in the carcinogenesis of colon and breast cancer. Lal et al. indicated that the c.146A>G variation in SLC22A16 is involved in variations in the pharmacokinetics of doxorubicin and doxorubicinol patients with malignant tumors [33]. Wu et al. demonstrated that SLC22A16 is up-regulated in acute myeloid leukemia cells and it affect the proliferation and viability of acute myeloid leukemia cells [34]. Therefore, these studies further confirmed the robustness and utility of the 5-gene signature.

The results of GSEA suggested that E2F signaling, MYC signaling, and G2M checkpoint influence the 5-gene combination in predicting the survival of GC patients. Members of the E2F family, MYC signaling, and G2M checkpoint regulate various cellular functions related to cell cycle, apoptosis, and carcinogenesis [35–37].

Wang et al. [22] introduced a 9-gene signature (as mentioned above) to predict the survival of GC patients by robust likelihood-based modeling with 1000 iterations. We compared the prognostic performance of our 5-gene combination with theirs, and the results suggested that our 5-gene combination outperformed the 9-gene signature. Thus, we suggest that our 5-gene combination is a good supplement for use in the prognosis of patients with GC.

Although the 5-gene combination showed excellent performance, our study has the following defects. First, this study was an integration and reanalysis of existing published GC gene expression studies. Although it showed good performance in various aspects, it has not been verified by large-scale prospective studies. Second, molecular biology experiments were not been carried out to verify its specific mechanisms in GC cells. Therefore, our subsequent research will focus on verifying the conclusions of this study in terms of clinical application and molecular mechanisms.

Conclusions

We introduced a 5-gene combination that can predict the survival of GC patients and might be an independent prognostic factor in GC.

Conflicts of interest

The authors declare no conflicts of interest.

Table 1. Genes sets enriched in the 5-gene combination low-risk group.

Gene set	SIZE	ES	NES	NOM P value	FDR
E2F signaling	193	0.7650	1.9174	0.0038	0.0408
MYC signaling	58	0.8257	1.8950	<0.0001	0.0285
G2M checkpoint	194	0.6679	1.8672	0.0058	0.0282

ES – enrichment score; NES – normalized enrichment score; NOM P value – normal P value; FDR q value – false discovery rate

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Supplementary Tables

Supplementary Table 1. Characteristics of gastric patients in the training set and test.

No. of samples	Training set n=194	Test set n=194	P value	GSE15459 n=192
Median age in years (range)	66 (34–86)	68 (30–90)	P>0.05	66.55 (23.4–92.4)
Female (%)	69 (35.57)	67 (34.54)	P>0.05	67 (34.9)
Male (%)	125 (64.43)	127 (65.46)	P>0.05	125 (65.1)
Stage I (%)	24 (12.37)	27 (13.92)	P>0.05	31 (16.15)
Stage II (%)	58 (29.9)	63 (32.47)	P>0.05	29 (15.1)
Stage III (%)	91 (46.91)	74 (38.14)	P>0.05	72 (37.5)
stage IV (%)	19 (9.79)	19 (9.79)	P>0.05	60 (31.25)
Grade 1 (%)	4 (2.06)	6 (3.09)	P>0.05	NA
Grade 2 (%)	69 (35.57)	68 (35.05)	P>0.05	NA
Grade 3 (%)	118 (60.82)	114 (58.76)	P>0.05	NA
No. of deaths (%)	71 (36.6)	86 (44.33)	P>0.05	95 (49.48)

Supplementary Table 2. Genes that were associated with the overall survival of patients with gastric cancer patients.

Genes	Coefficients	HR	LCI	UCI	p Value	FDR
FLJ16779	0.3368	1.4005	1.2126	1.6175	<0.0001	0.0157
FRMD7	0.2597	1.2966	1.1537	1.4571	<0.0001	0.0441
CPNE8	0.3698	1.4474	1.2220	1.7145	<0.0001	0.0637
APOD	0.3599	1.4332	1.2159	1.6894	<0.0001	0.0608
PRR20A	0.2271	1.2549	1.1338	1.3890	<0.0001	0.0398
LOC113230	–0.3293	0.7195	0.6170	0.8389	<0.0001	0.0906
NRP1	0.3764	1.4570	1.2298	1.7262	<0.0001	0.0462
MAGED4B	0.3286	1.3890	1.1914	1.6193	<0.0001	0.0922
SLC22A16	0.3050	1.3566	1.1763	1.5644	<0.0001	0.0940
ZNF804B	0.2239	1.2510	1.1322	1.3822	<0.0001	0.0371
GABRG1	0.2633	1.3013	1.1667	1.4513	<0.0001	0.0077
TBX22	0.2275	1.2554	1.1301	1.3946	<0.0001	0.0767
PRTG	0.3190	1.3757	1.1851	1.5971	<0.0001	0.0948
SLC7A2	0.3209	1.3783	1.1874	1.5999	<0.0001	0.0840
CGB5	0.3143	1.3693	1.1959	1.5679	<0.0001	0.0184
CGB1	0.2726	1.3134	1.1671	1.4781	<0.0001	0.0207
SERPINE1	0.3661	1.4421	1.2303	1.6902	<0.0001	0.0213
PCDHB5	0.3158	1.3714	1.1866	1.5849	<0.0001	0.0647
GPX3	0.3591	1.4321	1.2136	1.6899	<0.0001	0.0719

HR – hazards ratio; LCI – lower limit of confidence interval; UCI – upper limit of confidence interval; FDR – false discovery rate.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

Supplementary Table 3. Univariate and multivariable Cox proportional hazards regression model of the overall survival of gastric cancer patients in the training set.

Variable	Univariate analysis				Multivariable analysis			
	HR	LCI	UCI	P value	HR	LCI	UCI	P value
5-gene combination	1.1136	1.0952	1.1324	<0.0001	1.1242	1.1033	1.1456	<0.0001
Age	1.0117	0.9890	1.0350	0.3158	1.0415	1.0147	1.0690	0.0022
Gender Male	1.5118	0.9007	2.5374	0.1178	0.8170	0.4615	1.4462	0.4878
Gender Female	Reference				Reference			
Pathologic stage	1.1777	1.0381	1.3362	0.0111	1.2341	1.0807	1.4092	0.0019
Grade	1.4874	0.9373	2.3604	0.0919	0.9076	0.5262	1.5654	0.7273

HR – hazards ratio; LCI – lower limit of confidence interval; UCI – upper limit of confidence interval; FDR – false discovery rate.

Supplementary Table 4. Univariate and multivariable Cox proportional hazards regression model of the overall survival of gastric cancer patients in the test set.

Variable	Univariate analysis				Multivariable analysis			
	HR	LCI	UCI	P value	HR	LCI	UCI	P value
5-gene combination	1.0341	1.0174	1.0511	0.0001	1.0232	1.0044	1.0424	0.0156
Age	1.0280	1.0068	1.0495	0.0093	1.0489	1.0227	1.0757	0.0002
Gender Male	0.9718	0.6202	1.5229	0.9007	1.0586	0.6543	1.7127	0.8166
Gender Female	Reference				Reference			
Pathologic stage	1.3168	1.1666	1.4863	<0.0001	1.3224	1.1648	1.5012	<0.0001
Grade	1.4126	0.9209	2.1669	0.1136	1.3735	0.8688	2.1714	0.1745

HR – hazards ratio; LCI – lower limit of confidence interval; UCI – upper limit of confidence interval; FDR – false discovery rate.

Supplementary Table 5. Univariate and multivariable Cox proportional hazards regression model of the overall survival of gastric cancer patients in the validation set.

Variable	Univariate analysis				Multivariable analysis			
	HR	LCI	UCI	P value	HR	LCI	UCI	P value
5-gene combination	1.0618	1.0059	1.1209	0.0299	1.0559	0.9995	1.1155	0.0522
Age	0.9944	0.9786	1.0105	0.4943	1.0044	0.9882	1.0209	0.5960
Gender Male	1.1431	0.7414	1.7623	0.5450	0.8669	0.5499	1.3666	0.5385
Gender Female	Reference				Reference			
Lauren classification Intestinal	0.7527	0.4956	1.1433	0.1828	0.7522	0.4781	1.1834	0.2180
Lauren classification Mixed	0.5549	0.2357	1.3067	0.1777	0.4796	0.2003	1.1481	0.0990
Lauren classification Diffuse	Reference				Reference			
Stage	1.9697	1.5647	2.4795	<0.0001	2.0460	1.6065	2.6057	<0.0001

HR – hazards ratio; LCI – lower limit of confidence interval; UCI – upper limit of confidence interval; FDR – false discovery rate.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]

References:

1. Lott PC, Carvajal-Carmona LG: Resolving gastric cancer aetiology: An update in genetic predisposition. Lancet Gastroenterol Hepatol, 2018; 3(12): 874–83
2. Tsang YH, Lamb A, Chen LF: New insights into the inactivation of gastric tumor suppressor RUNX3: The role of *H. pylori* infection. J Cell Biochem, 2011; 112(2): 381–86
3. Dos Santos MP, Sallas ML, Zapparoli D et al: Lack of association between IL-6 polymorphisms and haplotypes with gastric cancer. J Cell Biochem, 2019; 120(6): 9448–54
4. Rocken C: Molecular classification of gastric cancer. Expert Rev Mol Diagn, 2017; 17(3): 293–301
5. Zhang XY, Zhang PY: Gastric cancer: Somatic genetics as a guide to therapy. J Med Genet, 2017; 54(5): 305–12
6. Geng R, Li J: Apatinib for the treatment of gastric cancer. Expert Opin Pharmacother, 2015; 16(1): 117–22
7. Orditura M, Galizia G, Sforza V et al: Treatment of gastric cancer. World J Gastroenterol, 2014; 20(7): 1635–49
8. Aoyama T, Yoshikawa T: Adjuvant therapy for locally advanced gastric cancer. Surg Today, 2017; 47(11): 1295–302
9. Dietrich S, Floegel A, Troll M et al: Random Survival Forest in practice: A method for modelling complex metabolomics data in time to event analysis. Int J Epidemiol, 2016; 45(5): 1406–20
10. Mogensen UB, Ishwaran H, Gerds TA: Evaluating random forests for survival analysis using prediction error curves. J Stat Softw, 2012; 50(11): 1–23
11. Ooi CH, Ivanova T, Wu J et al: Oncogenic pathway combinations predict clinical prognosis in gastric cancer. PLoS Genet, 2009; 5(10): e1000676
12. Tao J, Deng NT, Ramnarayanan K et al: CD44-SLC1A2 gene fusions in gastric cancer. Sci Transl Med, 2011; 3(77): 77ra30
13. Cancer Genome Atlas Research Network: Comprehensive molecular characterization of gastric adenocarcinoma. Nature, 2014; 513(7517): 202–9
14. Cui J, Yin Y, Ma Q et al: Comprehensive characterization of the genomic alterations in human gastric cancer. Int J Cancer, 2015; 137(1): 86–95
15. Chia NY, Deng N, Das K et al: Regulatory crosstalk between lineage-survival oncogenes KLF5, GATA4 and GATA6 cooperatively promotes gastric cancer development. Gut, 2015; 64(5): 707–19
16. Lei Z, Tan IB, Das K et al: Identification of molecular subtypes of gastric cancer with different responses to PI3-kinase inhibitors and 5-fluorouracil. Gastroenterology, 2013; 145(3): 554–65
17. Ishwaran H, Gerds TA, Kogalur UB et al: Random survival forests for competing risks. Biostatistics, 2014; 15(4): 757–73
18. Ishwaran H, Kogalur UB: Consistency of random survival forests. Stat Probab Lett, 2010; 80(13–14): 1056–64
19. Wang Y, Xia ST, Tang Q et al: A novel consistent random forest framework: Bernoulli random forests. IEEE Trans Neural Netw Learn Syst, 2018; 29(8): 3510–23
20. Chen X, Ishwaran H: Pathway hunting by random survival forests. Bioinformatics, 2013; 29(1): 99–105
21. Kamarudin AN, Cox T, Kolamunnage-Dona R: Time-dependent ROC curve analysis in medical research: current methods and applications. BMC Med Res Methodol, 2017; 17(1): 53
22. Wang Z, Chen G, Wang Q et al: Identification and validation of a prognostic 9-genes expression signature for gastric cancer. Oncotarget, 2017; 8(43): 73826–36
23. Schroder MS, Culhane AC, Quackenbush J, Haibe-Kains B: Survcomp: An R/Bioconductor package for performance assessment and comparison of survival models. Bioinformatics, 2011; 27(22): 3206–8
24. Mootha VK, Lindgren CM, Eriksson KF et al: PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet, 2003; 34(3): 267–73
25. Subramanian A, Tamayo P, Mootha VK et al: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA, 2005; 102(43): 15545–50
26. Kim JS, Kang SH, Moon HS et al: Clinical outcome of doublet and triplet neoadjuvant chemotherapy for locally advanced gastric cancer. Korean J Gastroenterol, 2016; 68(5): 245–52
27. Punjachaipornpon T, Mahachai V, Vilaichone R: Severe manifestations and grave prognosis in young patients with gastric cancer in Thailand. Asian Pac J Cancer Prev, 2016; 17(7): 3427–29
28. Liang YX, Deng JY, Guo HH et al: Characteristics and prognosis of gastric cancer in patients aged >/=70 years. World J Gastroenterol, 2013; 19(39): 6568–78
29. Datema FR, Moya A, Krause P et al: Novel head and neck cancer survival analysis approach: Random survival forests versus Cox proportional hazards regression. Head Neck, 2012; 34(1): 50–58
30. Hsich E, Gorodeski EZ, Blackstone EH et al: Identifying important risk factors for survival in patient with systolic heart failure using random survival forests. Circ Cardiovasc Qual Outcomes, 2011; 4(1): 39–45
31. Coburn LA, Singh K, Asim M et al: Loss of solute carrier family 7 member 2 exacerbates inflammation-associated colon tumorigenesis. Oncogene, 2019; 38(7): 1067–79
32. Tozlu S, Girault I, Vacher S et al: Identification of novel genes that co-cluster with estrogen receptor alpha in breast tumor biopsy specimens, using a large-scale real-time reverse transcription-PCR approach. Endocr Relat Cancer, 2006; 13(4): 1109–20
33. Lal S, Wong ZW, Jada SR et al: Novel SLC22A16 polymorphisms and influence on doxorubicin pharmacokinetics in Asian breast cancer patients. Pharmacogenomics, 2007; 8(6): 567–75
34. Wu Y, Hurren R, MacLean N et al: Carnitine transporter CT2 (SLC22A16) is over-expressed in acute myeloid leukemia (AML) and target knockdown reduces growth and viability of AML cells. Apoptosis, 2015; 20(8): 1099–108
35. Evangelou K, Havaki S, Kotsinas A: E2F transcription factors and digestive system malignancies: How much do we know? World J Gastroenterol, 2014; 20(29): 10212–16
36. Stine ZE, Walton ZE, Altman BJ et al: MYC, metabolism, and cancer. Cancer Discov, 2015; 5(10): 1024–39
37. Lobrich M, Jeggo PA: The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. Nat Rev Cancer, 2007; 7(11): 861–69

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS]