A Stable Gene Set for Prediction of Prognosis and Efficacy of Chemotherapy in Gastric Cancer

Rui Wu
First Affiliated Hospital of Nanchang University
Sixuan Guo
Nanchang University
Shuhui Lai
Nanchang University
Guixing Pan
Shangrao Maternity and Child Care Hospital
Linyi Zhang
Nanchang University
Huanbing Liu (liuhuanbing6911@sina.com)
First Affiliated Hospital of Nanchang University

Research Article

Keywords: Gastric cancer, Molecular typing, Prognosis, Prediction of efficacy of chemotherapy, Immune infiltration

Posted Date: February 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-156785/v1

Abstract

Background

Gastric cancer is one of the leading causes of cancer-related death worldwide. How to eliminate gastric cancer is an urgent public health problem. Prediction of prognosis is critical to the development of clinical treatment regimens. The aim of this study was to establish a stable prognostic gene set to guide the clinical diagnosis and treatment of gastric cancer.

Methods

A public microarray dataset of TCGA providing clinical information was obtained. The selection operator regression method was used to reduce the dimensionality of stable prognostic genes identified via the bootstrap method and survival analysis.

Conclusion

We established two prognostic models, respectively designated as stable gene risk score of OS(SGRS-OS) and stable gene risk score of PFI(SGRS-PFI) consisting of 18 and 21 genes. With specific risk score formulae, the SGRS set possesses a strong ability to predict overall survival and progression-free interval through both univariate and multivariate analyses. Compared with the TNM stage, the SGRS set showed much higher predictive accuracy. Further analysis revealed that patients with higher SGRS exhibited worse chemotherapy outcomes. Our SGRS set may be an effective tool to predict survival and guide treatment in patients with gastric cancer.

Introduction

Gastric cancer (GC) ranks the 6th place in terms of cancer morbidity, and it is also the 5th cause of cancer deaths in the world (1). The overall survival (OS) rate of GC cannot be improved through surgery or neoadjuvant therapy (2). Gastric cancer is a kind of heterogeneous malignant tumor, whose primary or acquired drug resistance makes chemotherapy unable to completely destroy tumor cells, while insensitivity to chemotherapy is a common cause of tumor recurrence and metastasis (3). Therefore, the evaluation of the overall survival, progression-free interval (PFI) and chemotherapy effect of patients with gastric cancer can help optimize the treatment strategy. The development of clinical prediction model is a conventional method to predict prognosis, and the key of modeling lines in the selection of stable and effective variables.

Conventional clinicopathologic variables, such as depth of invasion (T Stage) or lymph node metastasis (N stage), are predominantly focused on cancer cells to predict prognosis. While these variables are valid and widely used, they do not provide sufficient prediction (4). Before this study, some articles have

proposed new factors in addition to clinical factors for predicting the prognosis of gastric cancer, but the area under the curve (AUC) of the prediction model was not high, which suggests that new, more effective predictors need to be discovered. (5)

DNA microarray technology or "gene chips," derived from large-scale sequencing methods, are increasingly used to produce much more data than represents the sequence itself. It sheds novel lights on the pathophysiology and classification of disease, gene function, as well as drug research (6). Using DNA microarray technology, we developed a reliable prognostic gene set in the hope of predicting overall survival, progression-free interval (PFI) and the chemotherapeutic effects on GC cases, thus laying solid foundation for treatment in clinic.

Materials And Methods

Transcriptome data acquisition and clinical information collection

The Cancer Genome Atlas (TCGA) provides a large, free reference database for cancer research through the collection of cancer-related omics data, which is publicly available at the Data Portal TCGA (https://cancergenome.nih.gov). We downloaded the expression matrix of gastric cancer patients and relevant clinical information from the TCGA database in September 2018. The clinical information includes overall survival, progression-free interval, AJCC pathologic tumor stage, histologic grade, gender and age.

Study population and clinicopathological variables

We used the “createDataPartition” package in R to divided the data set into training cohort and validation cohort according to the stage stratified sampling with a ratio of 7:3. In this study, we used two analysis endpoints: OS, the time interval from diagnosis to death; PFI, the time interval between the beginning of observation and tumor progression.

Stable prognostic gene identification and selection

In order to obtain stable prognostic genes, bootstrapping testing was used to test the stability of the initial genes. Seventy percent of patients were randomly selected from samples to assess the genetic impact on survival. After 1,000 iterations, genes enrolled into 70% resampled runs ($P<0{\cdot}05$ upon stability test) were selected to be the creditable prognostic genes. survival analysis was performed on all patients using the R software, and the genes with P value less than 10e-3 were screened for further study.

Gene set generation using LASSO Cox regression

LASSO regression is a statistical method that can not only select variables but also make regularization (7). In biological and medical research, it is also used to build prediction models in data sets with many interrelated independent variables (8). Therefore, LASSO regression has important statistical characteristics that help to assess the relationship between many biomarkers and clinical characteristics

(9). Using LASSO regression, select ten-fold cross validation, intercept the modeled optimal penalty parameter value, and finally generate the optimal genetic set for predicting prognosis. Based on the generated gene set, we used Cox analysis to obtain the risk score of OS and PFI with OS and PFI as endpoint variables, respectively.

Estimation of immune infiltration

Tumor is a kind of tissue with high heterogeneity, where the tumor microenvironment (TME) surrounds and interacts with the malignant cells, and the TME contains various immunocyte types. The dialectical relationship of cancer cells with immune microenvironment is of critical clinical significance; therefore, it is necessary to develop approaches to investigate the cell components in immune microenvironment (10). MCP-counter package from the R software might be used in this case, which using the gene expression matrix to produce the scores of immunocytes (T cells, CD8+ T cells, NK cells, monocytes, myeloid dendritic cells, endothelial cells, B lymphocytes, cytotoxic lymphocytes, fibroblasts and neutrophils). The MCP-counter estimates represented scores of individual samples because they are calculated independently from each sample (11). The MCP-counter package of R software was adopted for converting the mRNA data to non-tumor cell infiltrating levels within TME. Before the analysis by MCP-counter, the standard annotation file was used to make the gene expression profile.

Gene set variation analysis (GSVA)

GSVA calculates the enrichment fraction of the sample gene set according to the gene function inside and outside the gene set, which is a non-parametric, non-supervised competitive gene set test. Conceptually, such method may be interpreted to alter the gene expression data coordinate system from one gene to one gene set (12). To assess pathway variability in large heterogeneous populations with complex phenotypic characteristics, we applied RNA-seq data and GMT to GSVA and acquired the enrichment fraction of each sample.

Immunohistochemistry

Immunohistochemistry was obtained from THPA (http://www.proteinatlas.org/) (13). The expression levels of different expression genes, which chosen to build the OS and PFI models, were evaluated between normal stomach tissues and gastric cancer tissues from THPA.

Statistical analysis

The survival rate was calculated by the Kaplan-Meier method, while significance of difference was determined by log-rank test. Cox proportional hazard models with the stepwise method “LRforward” were used for single factor and multiple factor analysis. The lasonos' guide was used to construct and validate the nomogram (14). The accuracy of survival prediction of the prognostic model was evaluated by time-dependent ROC as well as the Harrell's concordance index (c-index). R package was employed for statistical analysis and P value were tested by double-tail. The truncation points of P values were statistically significant.

Results

Gastric cancer patients' characteristics and stable prognostic gene identification

The detailed characteristics of the patients in this study are as follows (Supplemental Table S1). In this study, 362 patients with clinical information in the TCGA data set were screened during modeling. The mean age at diagnosis was 67.0 years (range:30.0 - 90 years), 234 (64.6%) were males, and 128 (35.4%) were females. All patients screened had OS and PFI information. The mean survival days of OS was 603.7 days, and the mean survival days of PFI was 543.6 days. Through bootstrapping testing described in the materials, 1,446 genes were screened. After survival analysis, 425 of the 1,446 genes were screened and identified as stable prognostic genes. (Supplemental Table S2).

Construction of Immune infiltration subgroups using stable prognostic genes

Firstly, unsupervised clustering was adopted for classifying 362 cancer tissues to diverse molecular subtypes on the basis of those 425 stable prognostic genes. Thereafter, the R package "ConsensusClusterPlus" function was adopted for assessing cluster stability and selecting the best cluster number. At last, the Type1 and Type2 patient clusters were discovered (Fig. 1a-b). Differences in OS (Fig. 1c) and PFI (Fig. 1d) between the two groups were statistically significant as shown by Kaplan-Meier curve. In terms of clinical features, we conducted further studies and found significant differences between the two types in the grade. The patients of Type2 significantly had a more advanced grade compared with Type1(Supplemental Table S3). Subsequent cell infiltration analysis revealed significant differences in the number of stromal cells and immune cells in Type1 and Type2 patients, including neutrophils(t=-3.6) and endothelial cells(t=-13.3) (Fig. 1e). We investigated the relationship between immune scores stromal scores and OS and found that the higher the score of neutrophils (Fig. 1f) and endothelial cells (Fig. 1g), the poorer the survival of patients, while such results were contradictory with previous results. As can be seen from the violin plot, there are significantly fewer neutrophils and endothelial cells in Type1 than in Type2, which also confirms that Type1 has a better survival than Type2. Finally, our result of molecular typing was compared with other established molecular subtypes of gastric cancer. The results showed that Type1 patients were mainly concentrated in C1, C2, GI.CIN and GI.HM-indel subtypes, and patients of Type2 most in C1, C2, C3, GI.CIN as well as GI.GS subtypes (Fig. 1h, Supplemental Table S3).

Construction of prognostically relevant gene set

For developing a gene set with clinical effectiveness, LASSO Cox regression model was utilized to reduce the dimensionality of those 425 identified prognostic genes. Thereafter, all cases were classified as the training or the validation cohort to analyze the prognosis. Differences were not statistically significant in clinical features between both groups (Supplementary Table S1). Through the LASSO model, based on the information OS and PFI, we generated stable gene sets (Supplemental Figure S1a-d). The OS stable gene set contained 18 genes, and the PFI gene set contained 21 genes (Supplemental Table S4). Then, Cox analysis was performed on the two gene sets to establish two prognostic models respectively.

Finally, we acquired stable gene risk score of OS (SGRS-OS) and PFI (SGRS- PFI). All cases were classified as 2 groups based on SGRS- OS and SGRS- PFI, and the cutoff value calculated by the whole queue was adopted (0.14 for SGRS-OS and 1.44 for SGRS-PFI). In the training and validation sets, the Kaplan Meier curves showed that patients in the high SGRS-OS cohort had a worse prognosis. (Fig. 2a-b). In the ROC, SGRS-OS, which served as the continuous variable in both training and validation cohorts, displayed high predicting ability compared with the TNM classification system. Stage was a categorical variable, so SGRS-OS was converted into a four-categorical variable, for the sake of enhancing the comparability. Even as a categorical variable, the prediction accuracy of SGRS-OS remains good (Supplemental Figure S2a-b). Similar results were also found for the SGRS-PFI set with documented PFI information (Fig. 2c-d, Supplemental Figure S2c-d). The predictive ability of SGRS-OS and SGRS -PFI models was tested in each subgroup stratified by immune subtype, level, sex, stage and age in the whole cohort, respectively, and SGRS-OS and SGRS-PFI were analyzed as continuous variables. As observed from the forest plots, the greater values of the two models markedly identified cases with dismal prognostic outcomes in each subgroup (Fig. 2e-f).

Stable gene set predicts the efficacy of chemotherapy in gastric cancer

Relative to supportive care (15), systemic chemotherapy, which is associated with the advantages of survival as well as quality of life, is developed to be the standard therapeutic modality to manage the metastatic or unresectable GC (16). Therefore, the outcome of chemotherapy is crucial for survival in patients with gastric cancer. We screened the patients with chemotherapy information and combined the chemotherapy results with SGRS-OS to explore the relationship. We used SGRS-OS to predict the efficacy of chemotherapy in patients with gastric cancer and found that low SGRS-OS patients were associated with good chemotherapy outcomes, while high SGRS-OS patients tended to be associated with bad chemotherapy outcomes. The accuracy of ROC curve was plotted and indicating a passable accuracy (Fig. 3a). Therefore, we can use the SGRS set to predict the chemotherapy efficacy of patients, providing a strong reference for the survival of clinical patients. For developing a related quantitative approach to predict the mortality possibility in patients, 2 nomograms were established in the present work (Fig. 3b-c) by enrolling the prognostic factors and scores obtained from the stable gene set. As suggested by the calibration plots, those as-constructed nomograms had better performance than the ideal model (Fig. 3d-e).

Identification of SGRS-OS and SGRS-PFI related clinical characters and biological pathways

This study also examined the correlations between scores obtained from the stable gene set and clinical features/molecular subtypes (Fig. 4a–b). In terms of clinical features, SGRS-OS and SGRS- PFI were significantly increased in more advanced stage patients. In addition, grade also affects the scores of the stable gene set, while age and gender have less influence on the it. In terms of molecular typing, we observed that the SGRS for C3, C6, and Type2 were also higher than other types. In terms of the pathway, we found that both SGRS-OS and SGRS-PFI values were significantly correlated with base excision repair, DNA replication and RNA degradation (Fig. 4c). Therefore, the development of gastric cancer is closely

related to gene expression, which provides a strong basis for gene expression to predict the prognosis of gastric cancer.

Identification of CGB3 as a potential biological target

In order to further explore the function of genes chosen to module in the development of gastric cancer, differential expression analysis of modeling genes was performed using gastric cancer samples and normal samples. Nine differential expression genes (DEGs) were identified, of which 3 were down-regulated and 6 were up-regulated (Fig. 5a). For better validating the as-constructed stable signature, those 9 DEGs expression levels were compared in normal versus GC tissues derived from The Human Protein Atlas (THPA). It was suggested by immunohistochemical results that, CGB8 (ENSG00000213030.5) expression upregulated within GC tissues, confirming the difference in CGB8 level in normal versus GC tissues (Fig. 5b). Furthermore, ROC curve analysis was also performed for evaluating CGB8 sensitivity and specificity in diagnosing GC. ROC curves of CGB8 in TCGA database was displayed (Fig. 5c), showing good sensitivity and specificity with AUC of 0.700. In addition, survival analysis showed that CGB8 is a risk factor in the progression of gastric cancer (Fig. 5d). Of note, the expression and function of CGB8 in gastric cancers remain largely unknown. Therefore, we proposed CGB8 as a biological target and tried to discover its role in gastric cancer development.

Discussion

GC ranks the 6th place in terms of its morbidity within cancer globally, and it is also a major reason for cancer deaths. Although important advances have been made in the molecular mechanism, diagnosis, treatment selection and strategies of tumorigenesis, OS in gastric cancer patients still needs to be further improved (17). The great GC morbidity may be ascribed to the fact that, specific prognostic markers are lacking, which leads to the failure to timely adjust the clinical treatment plan of gastric cancer patients (18). Carbohydrate antigen (CA) 19 – 9, CA72-4, and carcinoembryonic antigen have been the extensively adopted GC biomarkers, yet they are not the best diagnostic and prognostic biomarkers for GC because of the limited specificity or sensitivity (19, 20). As a result, it is necessary to identify the novel prognostic biomarkers for GC.

The DNA microarray technique is the efficient biomedical approach at present, and it can be applied in various diagnostic fields (21). There have been many reports on predicting the prognosis of gastric cancer with single gene, but the accuracy of prediction results still needs to be improved (22). In addition, the prognostic value of Tumor-associated macrophages (TAM) density in gastric cancer patients has been analyzed. The results showed that compared with low-density TAM patients, the HR of OS and PFI of high-density TAM patients were 1.56 and 1.10 respectively, indicating that TAM density did not significantly predict adverse survival of gastric cancer patients, and TAM density was not an independent predictor of survival of gastric cancer patients (23). Our analysis of cell infiltration showed that there were also significant differences in the composition of stromal cells such as fibroblasts and endothelial cells in Type1 and Type2 patients in addition to immune cells. It can be seen that the number of stromal cells

is also an important factor in predicting gastric cancer, and it is one-sided and inaccurate to analyze the number of immune cells only. Sequencing all human genes is not practical in clinical diagnosis, but single gene prediction is not accurate enough, so we need to develop an effective gene group for prediction. We performed stability analysis and survival analysis on all genes of gastric cancer to screen out the stable prognostic genes. The results of the immune infiltration estimation showed that the genes were related to a variety of immune cells and stromal cells, which were in close connection with the tumor microenvironment, providing a comprehensive view of gastric cancer.

We combined the results of molecular typing with the result of infiltration analysis and found that neutrophils and endothelial cells were strongly associated with prognosis. An increase in neutrophils and endothelial cells often predicts a worse prognosis. Neutrophil levels have been shown to be a strong predictor of poor survival in gastric cancer patients. In patients with gastric cancer, accumulation of peripheral blood and invasive marginal neutrophils promotes disease progression and predicts poor survival (24). In addition, studies have shown that endothelial cells such as lymphatic endothelial cells and vascular endothelial cells can promote the metastasis or growth of gastric cancer (25, 26). The accuracy of our results is further verified.

We use LASSO Cox regression to screen the optimal combination of genes and establish two models, called SGRS-OS and SGRS- PFI. The two models contain 18 and 21 genes, respectively. We confirmed the clinical feasibility and high accuracy of the two models. In the future, the development of a kit to test this gene set could promote the clinical prognosis prediction of gastric cancer for the benefit of mankind.

Further studying of selected genes, we found that some genes such as PLA2R1, GPC3, AKR1B1 and SERPINE1B were closely related to the tumor microenvironment. Some reports found that PLA2R1 is expressed in neutrophils (27) and pulmonary macrophages (28). Additionally, PLA2R1 is able to enhance the tumor suppressing responses, such as apoptosis, senescence, or transformation suppression. PLA2R1 is down-regulated in a number of cancer types, which supports its tumor suppressor role, and its expression can be suppressed by c-MYC and HIF2α, the oncogenes (29). Additionally, GPC3, one of the tumor-associated antigens, elevated F4/80 + CD86 + macrophage (M1) percentage within tumor, in the meantime of inducing CD8 + T cell immune response specific to GPC3 (30). Fidarestat, an inhibitor of AKR1B1, can markedly suppress the inflammatory signals induced by growth factors, tumor necrosis factor-alpha (TNF-α), environmental allergens, and lipopolysachharide (LPS), and such signals may result in various inflammatory disorders. The inflammatory disorder animal model like cardiovascular disease (CVD), diabetes, metastasis, uveitis, cancer and asthma, inhibiting AKR1B1 evidently promotes disease occurrence (31). SERPINE1B is associated with B cell function (32). These genes related to the tumor microenvironment were selected and involved in modeling that greatly improving the accuracy of the model. Although these have been proved to be closely related to tumor, there are still few studies related to gastric cancer. Our findings provide new ideas and methods for searching for potential biological targets of gastric cancer

Generally, 80–90% GC cases are diagnosed at the advanced stage when the cancer cannot be resected or may relapse or metastasize after surgery (33, 34). Although molecular targeted therapy is promising for improving the survival of patients with advanced gastric cancer, due to the high heterogeneity of gastric cancer and the lack of targets, fewer patients receive appropriate molecular targeted therapy. Therefore, systemic chemotherapy is still the main treatment method for patients with advanced gastric cancer (35). Therefore, prediction of chemotherapy outcomes is crucial to the formulation of patient prognosis and improvement of patient survival. We found significant differences in the efficacy of chemotherapy in different patients with SGRS-OS. The chemotherapy efficacy of patients with low SGRS-OS was significantly better than that of patients with low SGRS-OS, suggesting a correlation between the two. The results of SGRS-OS can be used to predict the chemotherapy efficacy of patients with good accuracy. This method is expected to solve the problem that prognosis of gastric cancer is difficult to predict.

This study has some limitations. Firstly, the patient population was heterogeneous. Secondly, we used the patients in the TCGA dataset to model. Some modeling genes were not found in the patient expression matrix in the GEO dataset. Therefore, we did not use a validation set from the GEO database. Special attention should be paid when using the stable gene set to detect patients in other databases. Thirdly, the gene expression data were imported to the Cox regression model as categorical variables in this work. Therefore, more studies are needed to verify the optimal thresholds.

To sum up, our constructed stable gene set can stably predict patient survival and guide the treatment for GC patients and has a good prospect of clinical application.

Declarations

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the TCGA. The URL link and the accession number of the data used from the TCGA database is: https://xenabrowser.net/

Authors' contributions

Rui Wu and Sixuan Guo conceived and supervised the study. Sixuan Guo, Shuhui Lai, Guixing Pan and Linyi Zhang analyzed the data and data visualization. Huanbing Liu conceptualization and project administration. Sixuan Guo wrote the manuscript. All authors read and approved the final version of manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68: 394-424, 2018.
2. Kim YJ, Hwang KC and Kim SW: Lee YCPotential miRNA-target interactions for the screening of gastric carcinoma development in gastric adenoma/dysplasia. Int J Med Sci. 14: 610-616, 2018.
3. Zheng P, Chen L, Yuan X, Luo Q, Liu Y, Xie G, Ma Y and Shen L: Exosomal transfer of tumor-associated macrophage-derived miR-21 confers cisplatin resistance in gastric cancer cells. J Exp Clin Cancer Res 36: 53, 2017.
4. Zhou ZH, Ji CD, Zhu J, Xiao HL, Zhao HB, Cui YH and Bian XW: The prognostic value and pathobiological significance of Glasgow microenvironment score in gastric cancer. J Cancer Res Clin Oncol 143: 883-894, 2017.
5. Zhang D, Zhou S and Liu B: Identification and Validation of an Individualized EMT-Related Prognostic Risk Score Formula in Gastric Adenocarcinoma Patients. BioMed Res Int 7082408: 2020.
6. Grünblatt E: The benefits of microarrays as tools for studying neuropsychiatric disorders. Drugs Today (Barc) 40: 147-156, 2004.
7. Tibshirani R: Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B-Methodol. 58:267–88, 1996.
8. Tibshirani R: Regression shrinkage and selection via the lasso: A retrospective. Journal of the Royal Statistical Society Series B-Statistical Methodology. 73: 273–82, 2011.
9. Pripp AH and Stanišić M: Association between biomarkers and clinical characteristics in chronic subdural hematoma patients assessed with lasso regression. PLoS One 12: e0186838, 2017.
10. Petitprez F, Vano YA, Becht E, Giraldo NA, de Reyniès A, Sautès-Fridman C and Fridman WH: Transcriptomic analysis of the tumor microenvironment to guide prognosis and immunotherapies. Cancer Immunol Immunother 67: 981-988, 2018.
11. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, et al: Estimating the population abundance of tissue-infiltrating immune

and stromal cell populations using gene expression. Genome Biol 17: 218, 2016.

12. Hänzelmann S, Castelo R and Guinney J: GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 16: 14-17, 2013.
13. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, et al: Towards a knowledge-based Human Protein Atlas. Nat Biotechnol 28: 1248-1250, 2010.
14. Iasonos A, Schrag D, Raj GV and Panageas KS: How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 10: 1364-70, 2008.
15. Wagner AD, Grothe W, Haerting J, Kleber G, Grothey A and Fleig WE: Chemotherapy in advanced gastric cancer: A systematic review and meta-analysis based on aggregate data. J Clin Oncol 24: 2903-2909, 2006.
16. Kim SY, Yoon MJ, Park YI, Kim MJ, Nam BH and Park SR. Nomograms predicting survival of patients with unresectable or metastatic gastric cancer who receive combination cytotoxic chemotherapy as first-line treatment. Gastric Cancer 21: 453-463, 2018.
17. Van Cutsem E, Sagaert X, Topal B, Haustermans K and Prenen H: Gastric cancer. Lancet 388: 2654-2664, 2016.
18. Sawaki K, Kanda M and Kodera Y: Review of recent efforts to discover biomarkers for early detection, monitoring, prognosis, and prediction of treatment responses of patients with gastric cancer. Expert Rev Gastroenterol Hepatol 12: 657-670, 2018.
19. Emoto S, Ishigami H, Yamashita H, Yamaguchi H, Kaisaki S and Kitayama J: Clinical significance of CA125 and CA72-4 in gastric cancer with peritoneal dissemination. Gastric Cancer 15: 154-161, 2012.
20. Kanda M, Fujii T, Takami H, Suenaga M, Inokawa Y, Yamada S, Nakayama G, Sugimoto H, Koike M, Nomoto S, et al: Combination of the serum carbohydrate antigen 19-9 and carcinoembryonic antigen is a simple and accurate predictor of mortality in pancreatic cancer patients. Surg Today 44: 1692-1701, 2014.
21. Marzancola MG, Sedighi A and Li PC: DNA Microarray-Based Diagnostics. Methods Mol Biol 1368: 161-178, 2016.
22. Ren J, Kuang TH, Chen J, Yang JW and Liu YX. The diagnostic and prognostic values of microRNA-21 in patients with gastric cancer: a meta-analysis. Eur Rev Med Pharmacol Sci 21: 120-130, 2017.
23. Liu JY, Yang XJ, Geng XF, Huang CQ, Yu Y and Li Y: Prognostic significance of tumor-associated macrophages density in gastric cancer: A systemic review and meta-analysis. Minerva Med 107: 314-321, 2016.
24. Li TJ, Jiang YM, Hu YF, Huang L, Yu J, Zhao LY, Deng HJ, Mou TY, Liu H, Yang Y, et al: Interleukin-17-Producing Neutrophils Link Inflammatory Stimuli to Disease Progression by Promoting Angiogenesis in Gastric Cancer. Clin Cancer Res. 15: 1575-1585, 2017.
25. Wang Z, Wang Z, Li G, Wu H, Sun K, Chen J, Feng Y, Chen C, Cai S, Xu J, et al: CXCL1 from tumor-associated lymphatic endothelial cells drives gastric cancer cell into lymphatic system via activating

integrin β1/FAK/AKT signaling. Cancer Lett 385: 28-38, 2017.
26. Yang H, Zhang H, Ge S, Ning T, Bai M, Li J, Li S, Sun W, Deng T, Zhang L, et al: Exosome-Derived miR-130a Activates Angiogenesis in Gastric Cancer by Targeting C-MYB in Vascular Endothelial Cells. Mol Ther 26: 2466-2475, 2018.
27. Silliman CC, Moore EE, Zallen G, Gonzalez R, Johnson JL, Elzi DJ, Meng X, Hanasaki K, Ishizaki J, Arita H, et al: Presence of the M-type sPLA(2) receptor on neutrophils and its role in elastase release and adhesion. Am J Physiol Cell Physiol 283: C1102-C1113, 2002.
28. Gelb MH, Lambeau G, Marone G and Triggiani M: Activation of cytokine production by secreted phospholipase A2 in human lung macrophages expressing the M-type receptor. J Immunol 174: 464-474, 2005.
29. Bernard D and Vindrieux D: PLA2R1: Expression and function in cancer. Biochim Biophys Acta 1846: 40-44, 2014.
30. Luo C, Shibata K, Suzuki S, Kajiyama H, Senga T, Koya Y, Daimon M, Yamashita M and Kikkawa F: GPC3 expression in mouse ovarian cancer induces GPC3 specific T cell-mediated immune response through M1 macrophages and suppresses tumor growth. Oncol Rep 32: 913-921, 2014.
31. Srivastava SK, Yadav UC, Reddy AB, Saxena A, Tammali R, Shoeb M, Ansari NH, Bhatnagar A, Petrash MJ, Srivastava S, et al: Aldose reductase inhibition suppresses oxidative stress-induced inflammatory disorders. Chem Biol Interact 191: 330-338, 2011.
32. McGrath-Morrow SA, Ndeh R, Helmin KA, Khuder B, Rothblum-Oviatt C, Collaco JM, Wright J, Reyfman PA, Lederman HM and Singer BD: DNA methylation and gene expression signatures are associated with ataxia-telangiectasia phenotype. Sci Rep 10: 7479, 2020.
33. Van Cutsem E, Sagaert X, Topal B, Haustermans K and Prenen H: Gastric cancer. Lancet 388: 2654-2664, 2016.
34. Wagner AD, Syn NL, Moehler M, Grothe W, Yong WP, Tai BC, Ho J and Unverzagt S: Chemotherapy for advanced gastric cancer. Cochrane Database Syst Rev 8: CD004064, 2017.
35. Wadhwa R, Song S, Lee JS, Yao Y, Wei Q and Ajani JA: Gastric cancer-molecular and clinical dimensions. Nat Rev Clin Oncol 10: 643-655, 2013.

Figures

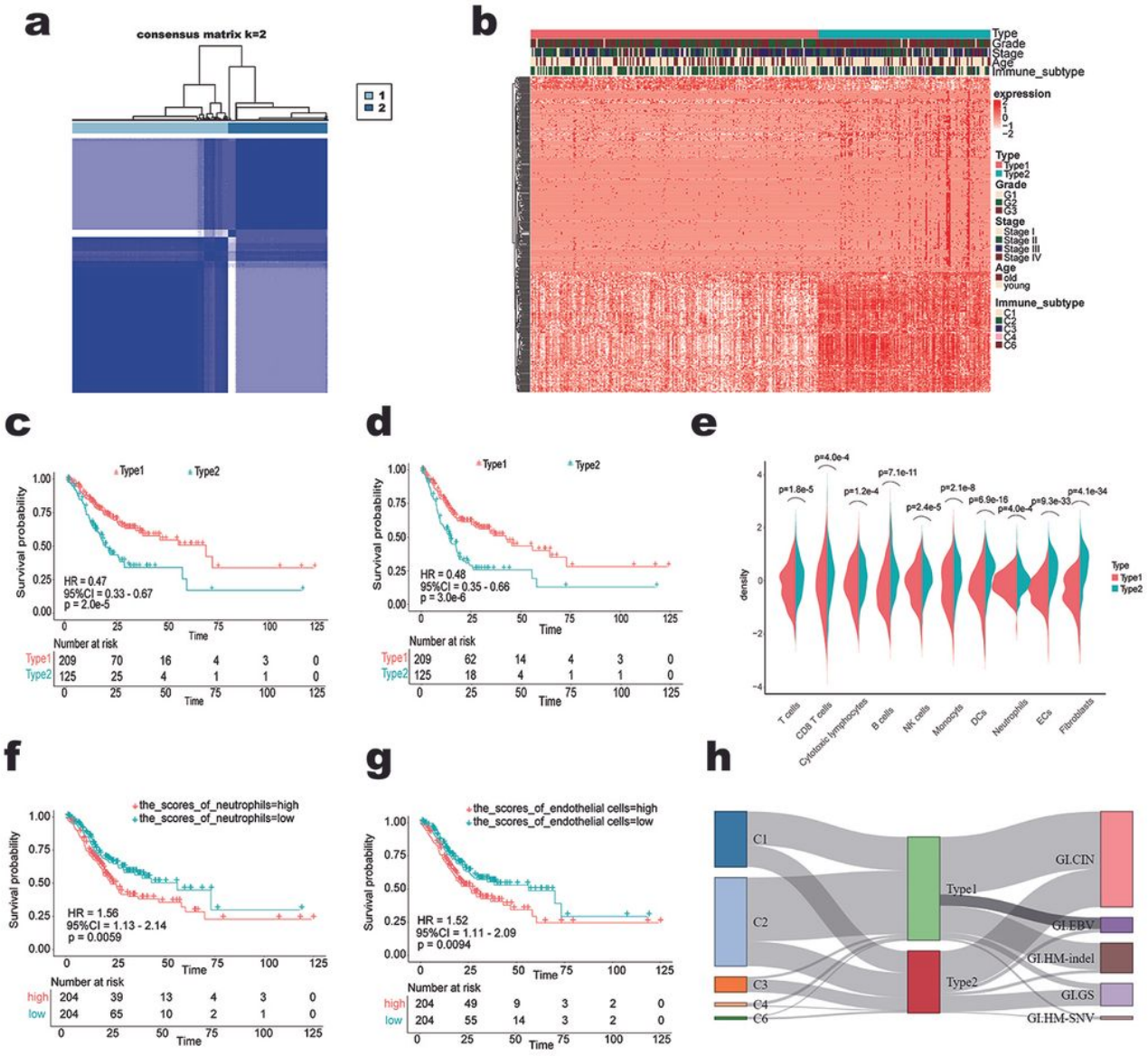

Figure 1

Consensus clustering of stable prognostic genes in gastric cancer. (a) Consensus matrices of gastric cancer patients for k = 2; (b) Gastric cancer cases are divided into two subtypes based on unsupervised analysis and hierarchical clustering of 425 stable prognostic genes. Clinical information (AJCC pathologic tumor stage, histologic grade, gender), immune subtype and type are indicated above the heatmap; (c–d) Differences in patient overall survival (c)and progression-free interval (d) with two clusters; (e) Violin plot of the comparison of immune and stromal cell infiltration between the two types; (f-g) Kaplan–Meier curves of overall survival according to the cell infiltrating scores of neutrophils (f) and endothelial cells (g)□ (h) Sankey chart displaying the distribution of the two types in C1–C6 subtypes and GI subtypes.

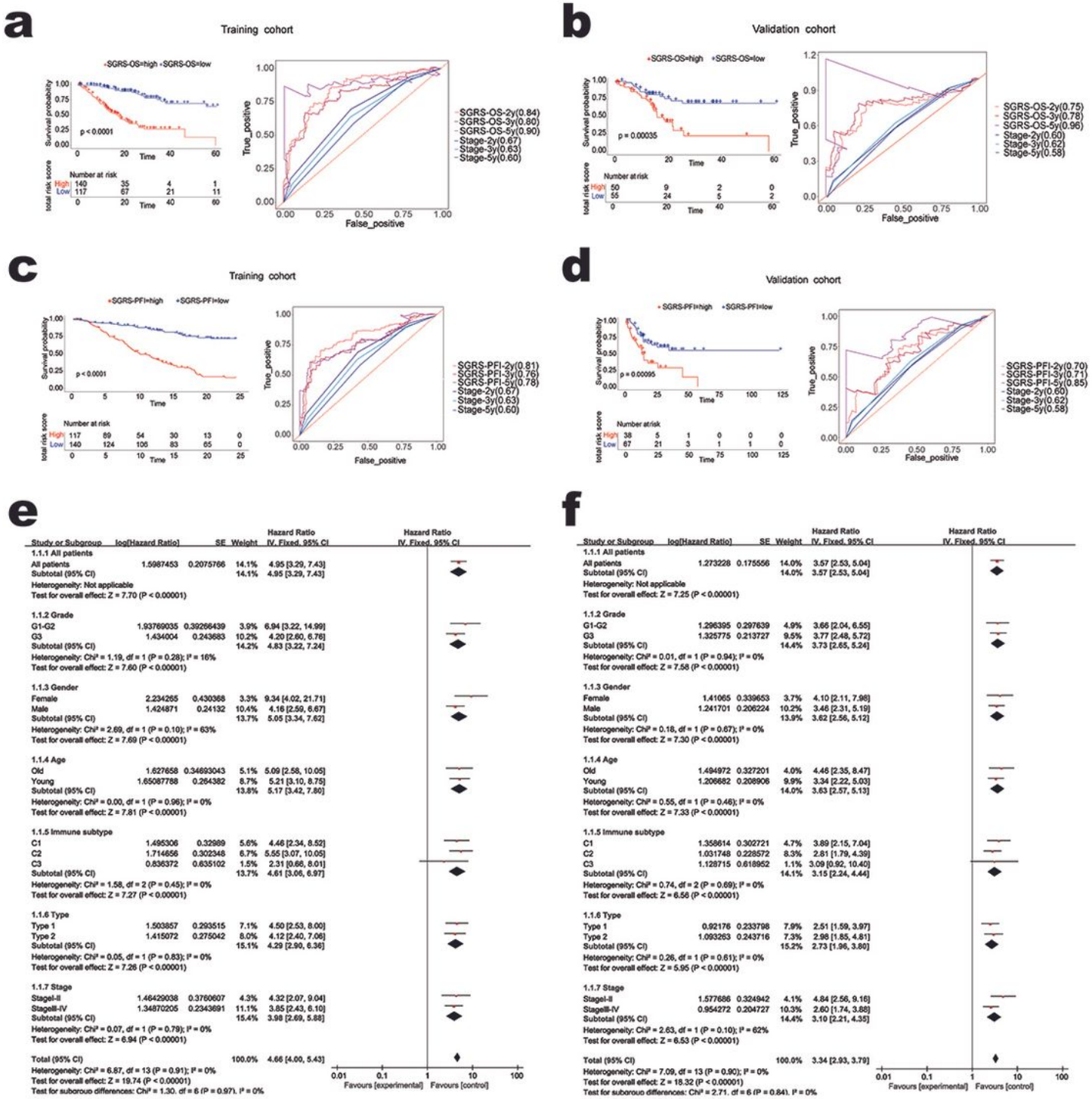

Figure 2

SGRS panel is a prognostic marker. (a–b) Kaplan–Meier curves (left) and ROC curves (right) of overall survival according to SGRS-OS groups in the training cohort (a) and validation cohort(b); (c-d) Kaplan–Meier curves(left) and ROC curves (right) of progression-free interval according to SGRS-PFI groups; (e-f) Forest plots of the associations between SGRS-OS and overall survival (e) and the associations between SGRS-PFI and progression-free interval (f) in various subgroups.

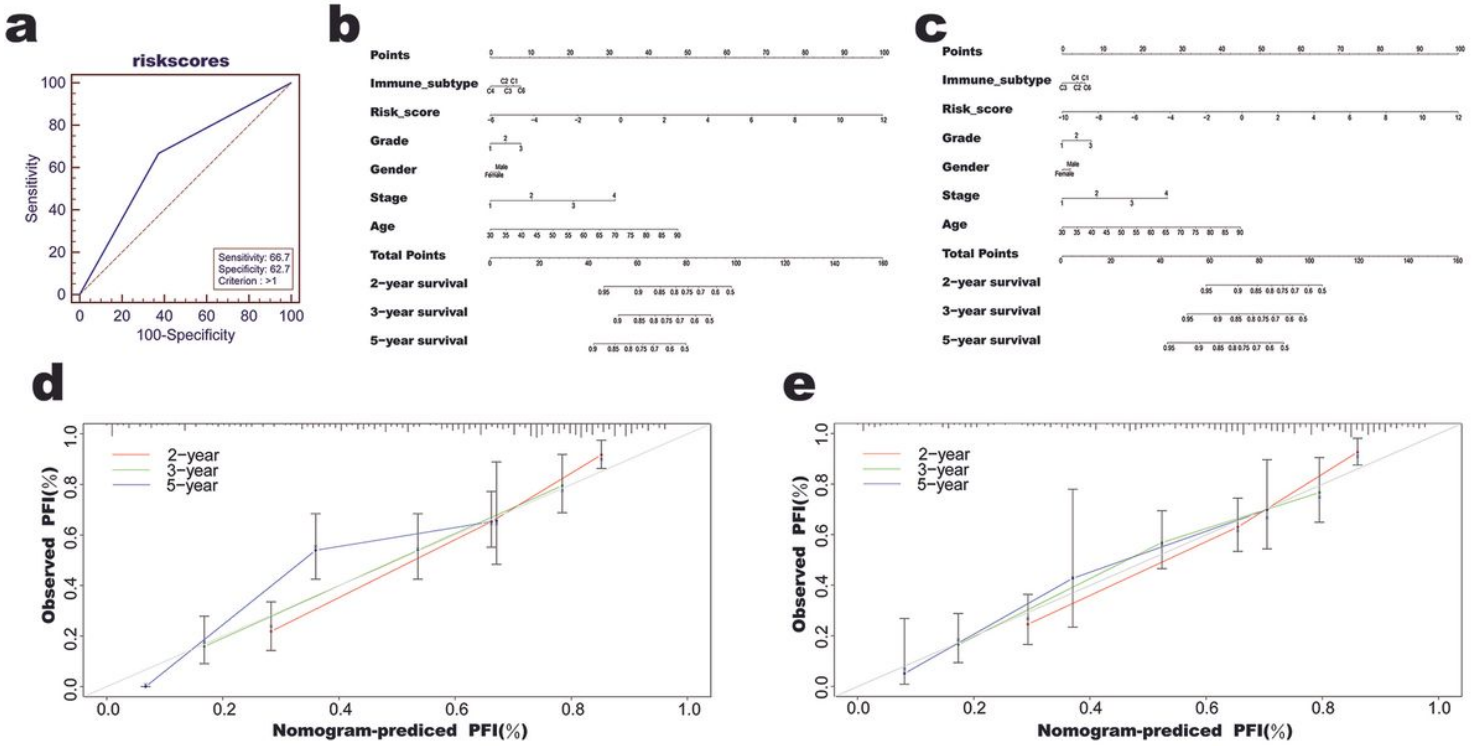

Figure 3

Prediction of the chemotherapy efficacy in gastric cancer (a) ROC curves of using SGRS-OS to predict the efficacy of chemotherapy; (b-c) Nomograms for predicting the probability of patient mortality based on SGRS-OS (b), SGRS-PFI(c) and clinical variables; (d-e) Plots depict the calibration of nomograms based on SGRS-OS (d) and SGRS-PFI (e) in terms of agreement between predicted and observed 2-year, 3-year, and 5-year outcomes. Nomogram performance is shown by the plot, relative to the 45-degree line, which represents the ideal prediction.

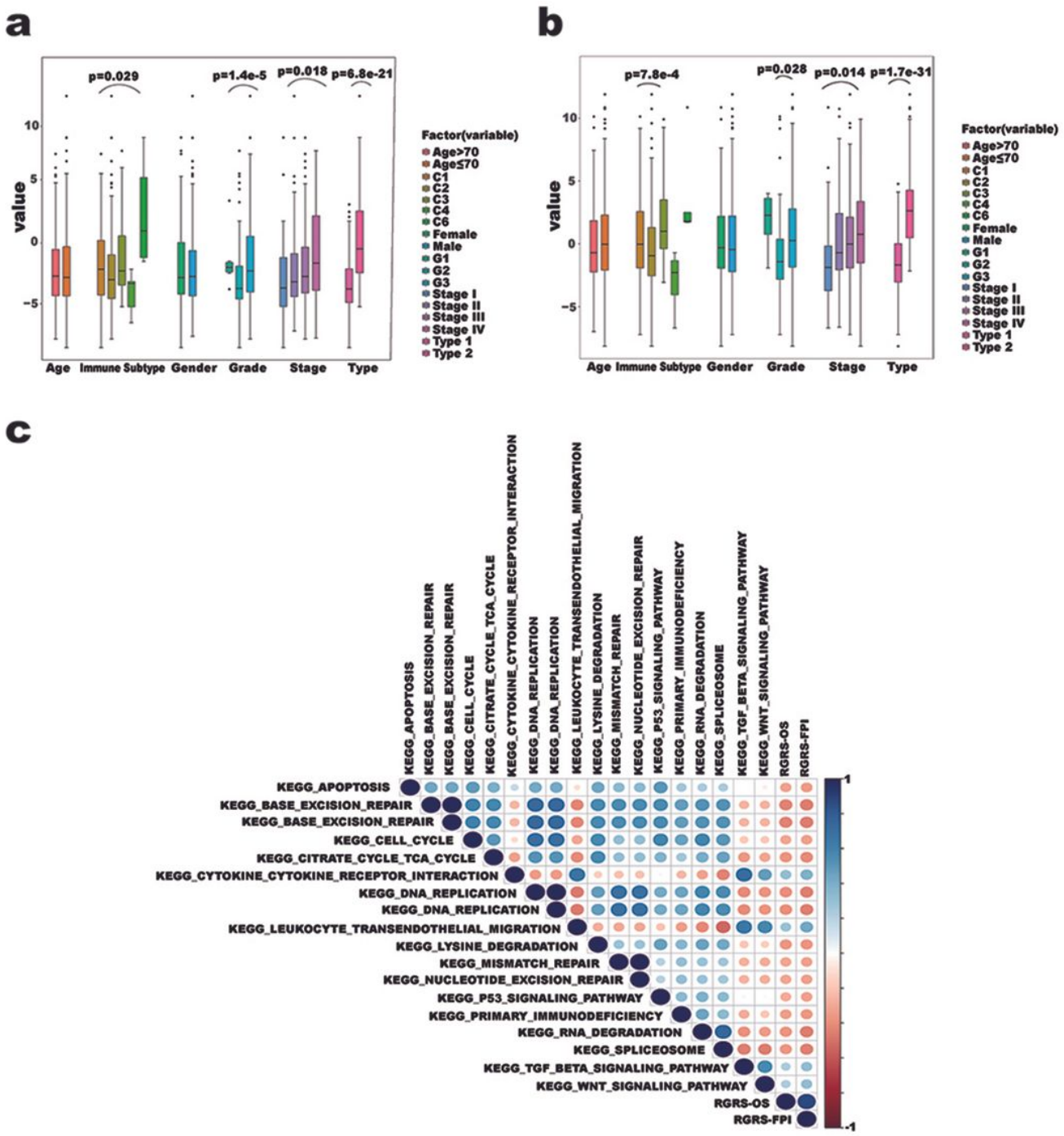

Figure 4

Clinical significance and biological function of SGRS panel. (a–b) SGRS-OS (a) and SGRS-PFI (b) values in different clinical subgroups. Boxes represent 25–75% of values, blacklines in boxes represent median values, whiskers represent 1.5 interquartile ranges, and black dots represent outliers; (c) Correlation matrix of SGRS-OS, SGRS-PFI values and the activation levels of biological process. Shading colours represents the value of corresponding correlation coefficients and size represents the p-values.

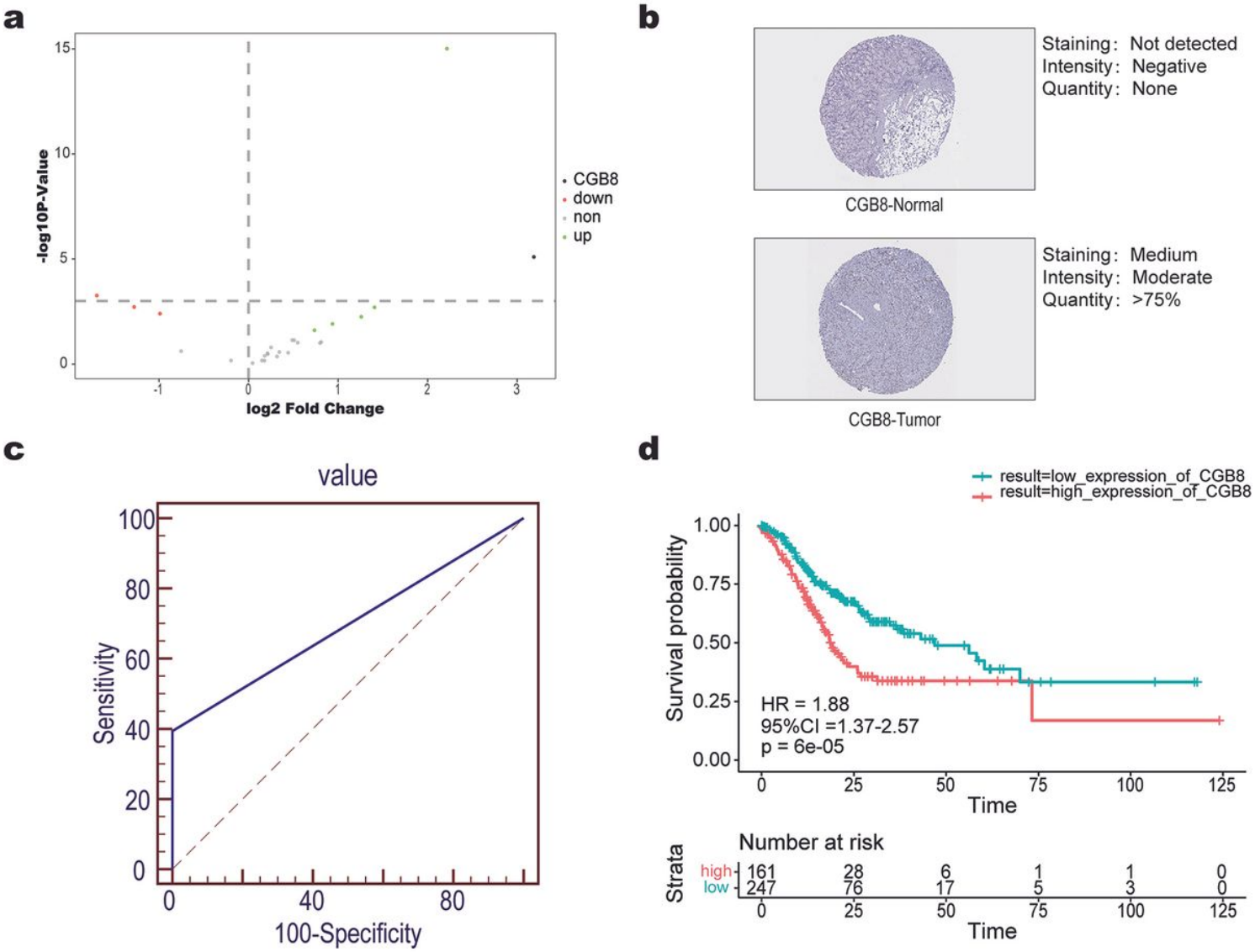

Figure 5

Identification of CGB3 as a potential biological target (a) Volcano plots of gene expression profiles in TCGA. Red/Green symbols classified the downregulated / upregulated genes according to the criteria: P-value < 0.05; (b) Immunohistochemistry from THPA was used to explore CGB8 between normal tissues and gastric cancer tissues; (c) ROC curve analysis of CGB8 in TCGA; (d) Kaplan–Meier curves of overall survival according to the expression of CGB8.

Supplementary Files

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

- FigureLegendsd1.docx
- FigureS1.tif
- FigureS2.tif

- TableS1.docx
- Tables2.xlsx
- TableS3.docx
- Tables4.xlsx