Diagnosis and Prognostic Value of SPARC in Gastric Carcinoma: database mining for GCTA

Diqi Ying
Ningbo Women and Children's Hospital
Ding Li
Wenzhou Medical University https://orcid.org/0000-0002-9149-4362
Xiao Jin (nbjinxiao@163.com)
https://orcid.org/0000-0003-1964-9769

Research article

Keywords: Diagnosis, Prognostic, SPARC, Gastric Carcinoma

Posted Date: October 23rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-39818/v2

Abstract

Gastric carcinoma (GC) remains high incidence and mortality both in developed and developing countries. SPARC is extracellular non-structural matrix glycoprotein. Previous studies were closely associated with bone disease. However, the role of SPARC in GC remains largely unclear. In our study, we explored the diagnosis, prognosis and pathway enrichments value of SPARC in GC. Here, with the data from The Cancer Genome Atlas (TCGA), we used receiver operating characteristic (ROC) curve analysis to estimate the diagnosis value of the SPARC expression, Univariate and multivariate analysis to the prognosis, Gene set enrichment analysis (GSEA) to the signal pathway enrichments. As a result, SPARC expression was significantly higher in the GC tissue samples. Those with high SPARC expression of GC patients were worse prognosis. GSEA shows the gene sets related signal pathways including transforming growth factor (TGF) beta signaling pathway, pathways in cancer, Wnt signaling pathway, Mitogen-activated protein kinase (MAPK) signaling pathway etc. In brief, those results suggest that SPARC can serve as a potential biomarker for GC in diagnosis and prognosis.

1. Introduction

Gastric cancer (GC) is one of the five most common deadly cancers in the world [1]. In the developed country, gastric cancer has become one of the most mortality cancers among adults[2].GC is also the 5 most commonly diagnosed cancers which incidence and mortality rates were corresponding with aga among Chinese people[3]. In 2020, 27,600 new cases of gastric cancer and 11,010 deaths are estimated, with a 5-year survival rate of only 32% (2010–2016)[4]. Nowadays, early surgical treatment can obtain a 5-year survival rate of 85–95% in of stage I GC patients, so surgery is still the preferred treatment for gastric cancer[5]. However, the cost-effectiveness of colonoscopy and serology-based preventive screening and the inconspicuous early symptoms limit the early diagnosis of GC [6, 7], which makes the worse prognosis of GC. Thus, early and accurate diagnosis of gastric cancer biomarkers would help improve the prognosis of these patients.

Secreted protein acidic and cysteine rich (SPARC) ,also known as NOT or BM-40, is extracellular non-structural matrix glycoprotein which was first isolated and purified in human and foetal bovine bone[8]. As for the function of this protein, studies have shown there were certain relationships between the SPARC expression and tumorigenesis. A study indicated that SPARC conclusively shown to promote pancreatic cancer proliferation[9]. A prognostic report indicated that SPARC mRNA expression was a negative predictor of pathological complete response (pCR) following neoadjuvant nab-paclitaxel (nab-PTX) therapy regardless of breast cancer subtype[10]. Emerging shreds of evidence have manifested that SPARC expression may a potential therapeutic target or a potential clinical marker for the survival of GC[11–13]. However, whether the SPARC also plays a photodynmic dignosis and prognosis role in GC remains totally unclear.

Thus, the recent study aimed to evaluate diagnosis and prognosis of SPARC expression in human GC based on data obtained from TCGA. GSEA was performed to further understand the biological pathways involved in the SPARC regulatory network related to GC pathogenesis.

2. Materials And Methods

2.1. RNA-sequencing data collection

The gene expression data (407 cases with 32 normal samples and HTSeq-FPKM for workflow type:) and corresponding clinical information were downloaded from TCGA Genomic Data Commons (GDC) data portal (https://portal.gdc.cancer.gov/repository). RNA-Seq gene expression data and clinical data for 375 patients were retained and further analyzed (Table 1).

2.2 Gene set enrichment analysis (GSEA)

GSEA is a computational method to determine whether a priori defined set of genes shows a statistically significant consistent difference between two biological states that is intended to detect changes in the expression of modest but functionally coordinated genomes[14, 15]. In our study, datasets and phenotype marked files were generated and uploaded into GSEA software. GESA analysis was carried out to demonstrate the significant survival difference observed between high- and low- SPARC groups in GC patient obtained from TCGA. Gene set permutations were performed 1000 times for each. The nominal p-value (NOM p-val) < 0.05 and false discover rate q-value(FDR q-val) < 0.5 were set to sort the pathways enriched in each phenotype.

2.3. Statistical analysis

Relationship between clinical pathologic features and were conducted with the Wilcoxon signed-rank test and logistic regression. Clinicopathologic characteristics associated with overall survival in TCGA patients were used Cox regression and the Kaplan-Meier method. Wilson method and percentage results were used in receiver operating characteristic (ROC) curve analysis which fulfilled with survivalROC package. Univariate logistic regression was used to revealed SPARC expression was associated with clinicopathologic characteristics. Univariate and multivariate Cox analysis was used to compare the influence of SPARC expression on survival along with other clinical characteristics (age, stage, grade, distant metastasis status, lymph node status etc). The median value was set to cutoff the value of SPARC expression into two groups. All statistical analyses were conducted by R (v.3.6.3).

3. Result

3.1. Patient characteristics

TCGA data with 407 cases' gene expression of gastric was downloaded from in May 2020. The clinical information of 375 tumor cases was shown in Table 1. The main proportion was in the 70-79y's group (32.61%), followed by the 60-69y's (29.38%), 50-59y's (23.18%),>80 y's (7.82%) and < 49y's (7.01%).

Clinical Stage I classification had 53 cases (15.06%), stage II in 111(31.53%), stage III in 150 (42.61%) and stage IV in 38(10.80%). T1 disease of Tumor size was found in 19 patients (5.18%), T2 in 80(21.80%), T3 in 168(45.78%) and T4 in 100(27.17%). Most tumors (31.09%, N = 111) were of N0 classification, 27.17%(97) of N1, 21.73%(75) of N2 and 20.73%(74) of N3.The G1 cases of grade classification accounts for 3.76% (10), G2 for 51.50% (137) and G3 for 44.74%(119).The positive of metastasis(M) was Twenty-five of 355(7.04%) cases. The gender composition was 241(64.27%) males and 134 females (35.73%).

3.2. Association with SPARC expression and the value of diagnosis

Then, a total of 407 samples with SPARC expression data were analyzed from TCGA. As shown in Fig. 1A, increased expression of SPARC correlated significantly with the tumor type(p = 2.017e-12). There were also significant differences in the expression of SPARC in 27 paired groups of tumor tissues and adjacent tissues (p = 1.197e-05, Fig. 1B). To assess the diagnostic efficacy of SPARC, receiver operating characteristic (ROC) curve was used the expression data from 375 tumor samples and 32 normal samples. The area under the ROC curve was 0.874[95% confidence interval (CI), 0.8216–0.9021; Fig. 1D].

3.3. Associations between SPARC expression and clinicopathology parameters

Clinicopathology data of 375 GC patients from TCGA were generally analyzed which including gender, grade(G) classification, metastasis(M) stage, tumor(T) size, lymphatic node(N) metastasis, stage classification and age at diagnosis(age). As shown in Fig. 2(A-G), increased expression of SPARC was notably associated with T size (p = 8.184e-04, Fig. 2D) and G classification (p = 0.023, Fig. 2B).

Univariate logistic regression revealed that SPARC expression as a categorical dependent variable was associated with poor prognostic clinicopathologic characteristics (Table 2). Increased SPARC expression in GC as significantly associated with T3 vs. T1 classification (OR = 3.007, p = 0.042) and T4 vs. T1 classification (OR = 3.157, p = 0.039).

3.4. Survival outcomes and multivariate analysis

According to the Kaplan-Meier survival analysis, those with high SPARC expression of the 375 patients were worse prognosis (Fig. 1C, P = 0.009).The univariate Cox analysis revealed that SPARC-high correlated significantly with poor OS[hazard ratio(HR) = 1.300, 95% CI = 1.090–1.543,p = 0.003]. Other clinicopathologic variables associated with poor survival include age, advanced stage, TNM classification (Table 3).

At multivariate Cox analysis, high SPARC expression remained independently associated with overall survival (HR = 1.260, 95%CI = 1.040–1.526, p = 0.018), as well as age (HR = 1.354, 95%CI = 1.121–1.635, p = 0.002) among GC patient (Fig. 3).

3.5. GSEA identifies SPARC-related signal pathways

To identify signal pathways which are differentially activated in GC, we used GSEA comparing SPARC expression data which divided by the median expression level. GSEA revealed significant differences (NOM p-val = 0.05 and FDR q-val = 0.05) in enrichment of MsigDB collection(c2.cp.kegg.v7.1.symbols.gmt). Table 4 has showed the 20 items of GSEA analysis. As is shown in Fig. 4, gene sets related to transforming growth factor(TGF) beta signaling pathway, pathways in cancer, Wnt signaling pathway, Mitogen-activated protein kinase(MAPK) signaling pathway, focal adhesion, cell adhesion molecules cams, melanogenesis and small cell lung cancer, which were related to the tumor-associated.

4. Discussion

GC has long been one of the world's major cancers and remains one of the major causes of malignant disease morbidity and mortality[16]. Evidence have proved that SPARC has a crucial function in the process of tumorigenesis, but the bioinformation according to the TCGA data in GC are still firstly performed in this study.

According to our study, SPARC expression was significantly higher in the GC tissue samples compared to the control samples or the paired adjunct samples. Which suggested that the up-regulation of SPARC expression may be related to the development of GC.

Moreover, the clinical diagnosis and prognostic value of the SPARC expression were examined in our study of GC patients. At the beginning, we found that SPARC expression was significantly associated with clinical grade and T classification. Second, Kaplan–Meier curves for OS revealed that high expression of SPARC was associated with poor outcomes in GC patients. The area under the ROC curve showed the up-expression of SPARC in value of diagnosis. Further, univariate logistic analysis indicated the SPARC expression had relation with T classification. Univariate and multivariate Cox analysis showed the SPARC expression may be a potential independent marker for poor prognosis in GC patients. The multivariate Cox analysis revealed age was an independent risk factor or OS in GC patient. In general, these findings suggested that high expression of SPARC could indicate a factor of diagnosis and poor prognosis for GC patients. Which also might be a pivotal target gene involved in the process of GC cell growth and metastasis.

In this study, we observe that SPARC high expression phenotype was associated with TGF beta signaling pathway, pathways in cancer, Wnt signaling pathway, Mitogen-activated protein kinase (MAPK) signaling pathway, focal adhesion, cell adhesion molecules cams, melanogenesis and small cell lung cancer. TGF beta signaling pathway is instrumental in mammalian development which has pivotal role in many mechanisms of breast cancer[17], lung cancer[18] and other cancer[19–21]. Wnt signaling pathway is required for adult tissue maintenance, and perturbations in Wnt signaling promote human cancer[22, 23].MAPK signaling pathway activated during the differentiation of myogenic cell lines[24]. Which is

essential for human melanoma cells[25] and prostate cancer[26].Focal adhesion-dependent activation of these pathways has been involved in a diverse array of cellular processes and was a potential target in cancer therapy[27, 28].

However, prediction of protein expression according to mRNA was useful but far from perfect[29].In this report, the correlation between SPARC mRNA expression and SPARC protein expression has not been verified. We will conduct further research through experiments and local clinical information in the future.

Declarations

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Please contact authors for data requests.

Consent for publication

Not applicable.

Acknowledgements

Thanks for the anonymous reviewers for their valuable comments and suggestions that helped improve the quality of our manuscript.

References

1. F. Bray, J. Ferlay, I. Soerjomataram, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018, 68(6): 394-424.
2. P. A. Ganz. Current US Cancer Statistics: Alarming Trends in Young Adults? J Natl Cancer Inst 2019, 111(12): 1241-1242.
3. W. Chen, R. Zheng, P. D. Baade, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016, 66(2): 115-132.
4. Epidemiology Surveillance, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER Research Data, 9 Registries, Nov 2019 Sub (1975-2017) - Linked To County Attributes - Time Dependent (1990-2017) Income/Rurality, 1969-2017 Counties, National

Cancer Institute, DCCPS, Surveillance Research Program, released April 2020, based on the November 2019 submission.

5. Lei Yang, Rongshou Zheng, Ning Wang, et al. Incidence and mortality of stomach cancer in China, 2014. Chinese Journal of Cancer Reserach(English version), 030(003): 291-298.
6. P. Rawla and A. Barsouk. Epidemiology of gastric cancer: global trends, risk factors and prevention. Prz Gastroenterol 2019, 14(1): 26-38.
7. M. Venerito, A. Link, T. Rokkas, et al. Review: Gastric cancer-Clinical aspects. Helicobacter 2019, 24 Suppl 1: e12643.
8. John D. Termine, Hynda K. Kleinman, S.William Whitson, et al. Osteonectin, a bone-specific protein linking mineral to collagen. Cell 1981, 26: 99-105.
9. Amanda Munasinghe, Khalisha Malik, Fatemia Mohamedi, et al. Fibronectin acts as a molecular switch to determine SPARC function in pancreatic cancer. Cancer Letters 2020, 477: 88-96.
10. Y. Nakazawa, S. Nakazawa, S. Kurozumi, et al. The pathological complete response and secreted protein acidic and rich in cysteine expression in patients with breast cancer receiving neoadjuvant nab-paclitaxel chemotherapy. Oncol Lett 2020, 19(4): 2705-2712.
11. J. Feng and L. Tang. SPARC in Tumor Pathophysiology and as a Potential Therapeutic Target. Curr Pharm Des 2014, 20(39): 6182-6190.
12. Z. Wang, B. Hao, Y. Yang, et al. Prognostic role of SPARC expression in gastric cancer: a meta-analysis. Arch Med Sci 2014, 10(5): 863-869.
13. J. L. Zhang, G. W. Chen, Y. C. Liu, et al. Secreted protein acidic and rich in cysteine (SPARC) suppresses angiogenesis by down-regulating the expression of VEGF and MMP-7 in gastric cancer. PLoS One 2012, 7(9): e44618.
14. V. K. Mootha, C. M. Lindgren, K. F. Eriksson, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003, 34(3): 267-273.
15. A. Subramanian, P. Tamayo, V. K. Mootha, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005, 102(43): 15545-15550.
16. C. de Martel, D. Forman, and M. Plummer. Gastric cancer: epidemiology and risk factors. Gastroenterol Clin North Am 2013, 42(2): 219-240.
17. S. Mukherjee, M. J. Choi, S. W. Kim, et al. Secreted protein acidic and rich in cysteine (SPARC) regulates thermogenesis in white and brown adipocytes. Mol Cell Endocrinol 2020, 506: 110757.
18. X. Qin, X. Y. Wang, J. W. Fei, et al. MiR-20a promotes lung tumorigenesis by targeting RUNX3 via TGF-beta signaling pathway. J Biol Regul Homeost Agents 2020, 34(2).
19. Y. Shen, S. Dong, J. Liu, et al. Identification of Potential Biomarkers for Thyroid Cancer Using Bioinformatics Strategy: A Study Based on GEO Datasets. Biomed Res Int 2020, 2020: 9710421.
20. Q. Hong, S. Wang, S. Liu, et al. LRG1 May Accelerate the Progression of ccRCC via the TGF-beta Pathway. Biomed Res Int 2020, 2020: 1285068.

21. L. H. Katz, Y. Li, J. S. Chen, et al. Targeting TGF-beta signaling in cancer. Expert Opin Ther Targets 2013, 17(7): 743-760.
22. C. Y. Logan and R. Nusse. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol 2004, 20: 781-810.
23. B. Lustig and J. Behrens. The Wnt signaling pathway and its role in tumor development. J Cancer Res Clin Oncol 2003, 129(4): 199-221.
24. A. Keren, Y. Tamir, and E. Bengal. The p38 MAPK signaling pathway: a major regulator of skeletal muscle development. Mol Cell Endocrinol 2006, 252(1-2): 224-230.
25. H. Sumimoto, F. Imabayashi, T. Iwata, et al. The BRAF-MAPK signaling pathway is essential for cancer-immune evasion in human melanoma cells. J Exp Med 2006, 203(7): 1651-1656.
26. K. H. Shen, S. H. Hung, L. T. Yin, et al. Acacetin, a flavonoid, inhibits the invasion and migration of human prostate cancer DU145 cells via inactivation of the p38 MAPK signaling pathway. Mol Cell Biochem 2010, 333(1-2): 279-291.
27. David DSchlaepfer, Christof RHauck, and David JSieg. Signaling through focal adhesion kinase. Progress in Biophysics and Molecular Biology 1999, 71(3-4): 435-478.
28. M. J. van Nimwegen and B. van de Water. Focal adhesion kinase: a potential target in cancer therapy. Biochem Pharmacol 2007, 73(5): 597-609.
29. Y. Guo, P. Xiao, S. Lei, et al. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. Acta Biochim Biophys Sin (Shanghai) 2008, 40(5): 426-436.

Tables

Table 1. TCGA gastric cancer patient characteristics

Clinical characteristics		Total(375)	%
Age at diagnosis(y)	<49	26	7.01
	50-59	86	23.18
	60-69	109	29.38
	70-79	121	32.61
	>80	29	7.82
Stage classification	I	53	15.06
	II	111	31.53
	III	150	42.61
	IV	38	10.80
Tumor(T) size	T1	19	5.18
	T2	80	21.80
	T3	168	45.78
	T4	100	27.25
Lymphatic node(N) metastasis	N0	111	31.09
	N1	97	27.17
	N2	75	21.01
	N3	74	20.73
Metastasis(M) stage	M0	330	92.96
	M1	25	7.04
Grade(G) classification	G1	10	3.76
	G2	137	51.50
	G3	119	44.74
Gender	Male	241	64.27
	Female	134	35.73

Table 2. Logistic regression of SPARC expression[#] and clinical pathological characteristics

Clinical characteristics	Total(N)	OR	95%CI	P-value
Age at diagnosis(y)				
50-59 vs. <49	112	0.679	0.277-1.637	0.388
60-69 vs. <49	135	0.975	0.408-2.301	0.953
70-79 vs. <49	147	0.816	0.344-1.909	0.382
>80 vs. <49	55	1.055	0.362-3.075	0.921
Stage classification				
II vs. I	164	1.480	0.768-2.880	0.244
III vs. I	203	1.270	0.678-2.403	0.457
IV vs. I	91	1.449	0.628-3.374	0.385
Tumor(T) classification				
T2 vs. T1	99	2.533	0.878-8.444	0.101
T3 vs. T1	186	3.007	1.096-9.646	**0.042**
T4 vs. T1	117	3.157	1.115-10.375	**0.039**
Node(N) classification				
N1 vs. N0	207	0.859	0.497-1.482	0.584
N2 vs. N0	185	0.843	0.468-1.516	0.569
N3 vs. N0	184	0.914	0.506-1.647	0.764
Metastasis(M) classification				
M1 vs. M0	355	1.097	0.483-2.507	0.824
Grade(G) classification				
G2 vs. G1	145	1.135	0.310-4.607	0.850
G3 vs. G1	228	1.818	0.505-7.280	0.365
Gender				
Male vs. Female	375	1.374	0.900-2.104	0.142

Categorical dependent variable, greater or less than the median expression level.

SPARC, secreted protein acidic and cysteine rich

OR, odds ration

CI, confidence interval

Bold values indicate P<0.05

Table 3. Univariate and multivariate analysis of the relationship between SPARC expression[*] and overall survival among gastric patients

Parameter	Univariate analysis			Multivariate analysis		
	HR	95%CI	p-value	HR	95%CI	p-value
Age	1.264	1.055-1.514	0.011	1.354	1.121-1.635	**0.002**
Gender	1.484	0.980-2.247	0.062			
Grade	1.368	0.947-1.977	0.095			
Stage	1.535	1.221-1.931	**0.0002**	1.511	0.953-2.393	0.079
T	1.298	1.023-1.645	**0.032**	1.001	0.718-1.394	0.997
M	2.048	1.096-3.827	**0.025**	1.590	0.691-3.656	0.280
N	1.313	1.041-1.658	**0.022**	0.990	0.690-1.414	0.947
SPARC[#]	1.300	1.090-1.543	**0.003**	1.260	1.040-1.526	**0.018**

* Categorical dependent variable, greater or less than the median expression level.

Value=log2(value+1)

SPARC, secreted protein acidic and cysteine rich

HR, hazard ratio

CI, confidence interval

Bold values indicate P<0.05

Table 4 Gene sets (GS) enriched in high phenotype

GS follow link to MSigDB	NES	NOM -val	FDR q-val
FOCAL ADHESION	2.392	<0.001	<0.001
ECM RECETOR INTERACTION	2.347	<0.001	<0.001
TGF BETA SIGNALING ATHWAY	2.183	<0.001	0.001
REGULATION OF ACTIN CYTOSKELETON	2.144	<0.001	0.001
LEUKOCYTE TRANSENDOTHELIAL MIGRATION	2.129	<0.001	0.002
COMLEMENT AND COAGULATION CASCADES	2.087	<0.001	0.002
CYTOKINE CYTOKINE RECETOR INTERACTION	2.017	<0.001	0.006
LYSOSOME	1.983	0.004	0.007
CELL ADHESION MOLECULES CAMS	1.984	<0.001	0.007
ATHWAYS IN CANCER	1.918	<0.001	0.011
MELANOGENESIS	1.89	0.006	0.014
HYERTROHIC CARDIOMYOATHY HCM	1.866	0.002	0.018
SMALL CELL LUNG CANCER	1.848	<0.001	0.02
DILATED CARDIOMYOATHY	1.838	0.002	0.021
WNT SIGNALING ATHWAY	1.811	0.004	0.026
GLIOMA	1.794	0.008	0.029
MAK SIGNALING ATHWAY	1.779	0.002	0.032
ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOATHY ARVC	1.741	0.014	0.042
RION DISEASES	1.727	0.012	0.044
MELANOMA	1.716	0.014	0.046

GS, Gene sets; ES, enrichment score; NES, normalized ES; NOM p-val, normalized p-value. FDR q-val, false discovery rate q-value.

Figures

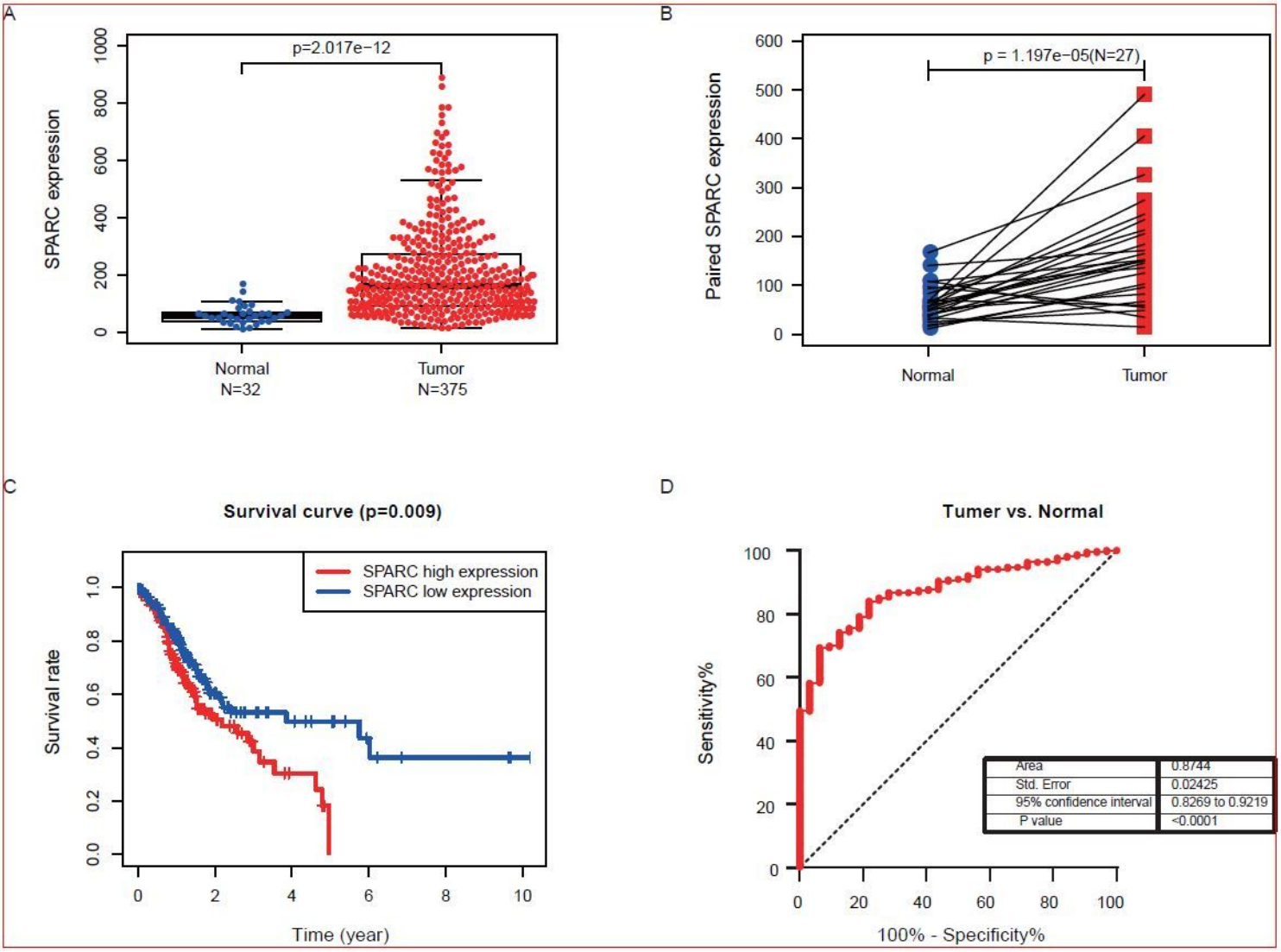

Figure 1

The mRNA expression level, clinical diagnosis and prognosis prediction of SPARC A. According to the TCGA cohort, tumor patients had higher levels of SPARC than the normal(p=2.017e-12). B. The expression level of SPARC in tumor tissues was higher than that in adjacent tissues (p=1.197e-05). C. SPARC expression and overall survival in gastric cancer patients in TCGA cohort(p=0.009). D. Receiver operating characteristic (ROC) curve SPARC expression in normal gastric tissue and tumor (AUC=0.874, 95%CI=0.827-0.922, p<0.0001).

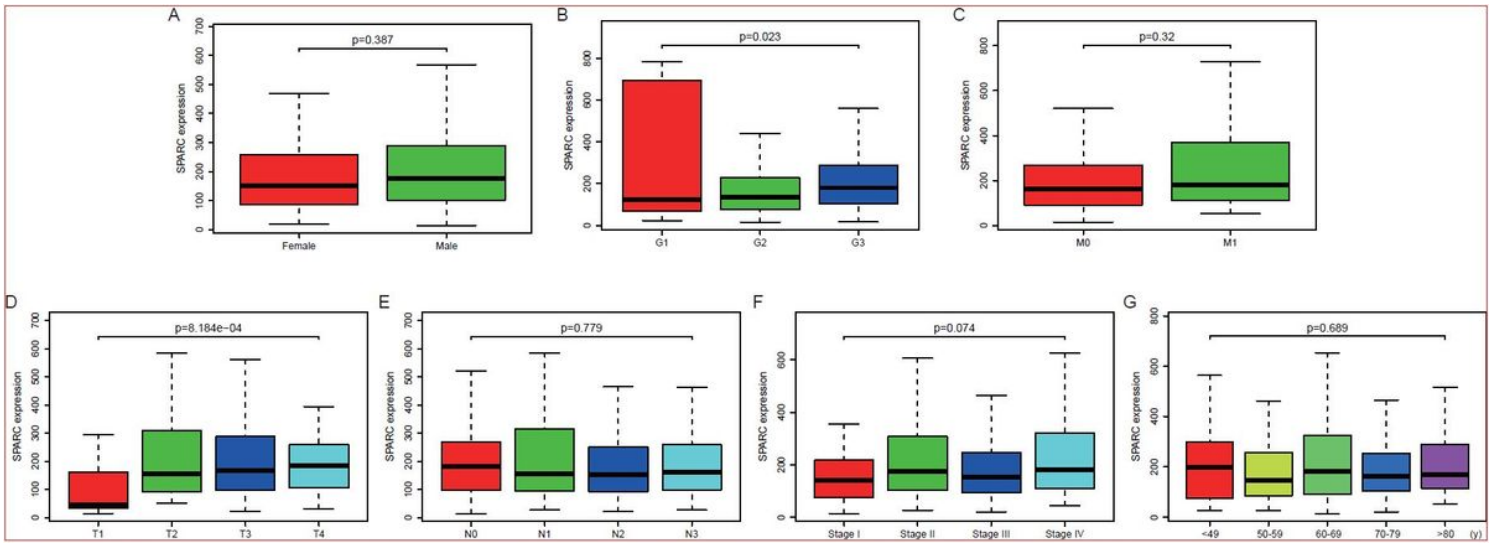

Figure 2

Association with SPARC expression and clinicopathologic characteristics, A: Gender, B: Grade(G) classification, C: Metastasis(M) stage, D: Tumor(T) size, E: Lymphatic node(N) metastasis. F: Stage classification, G: Age.

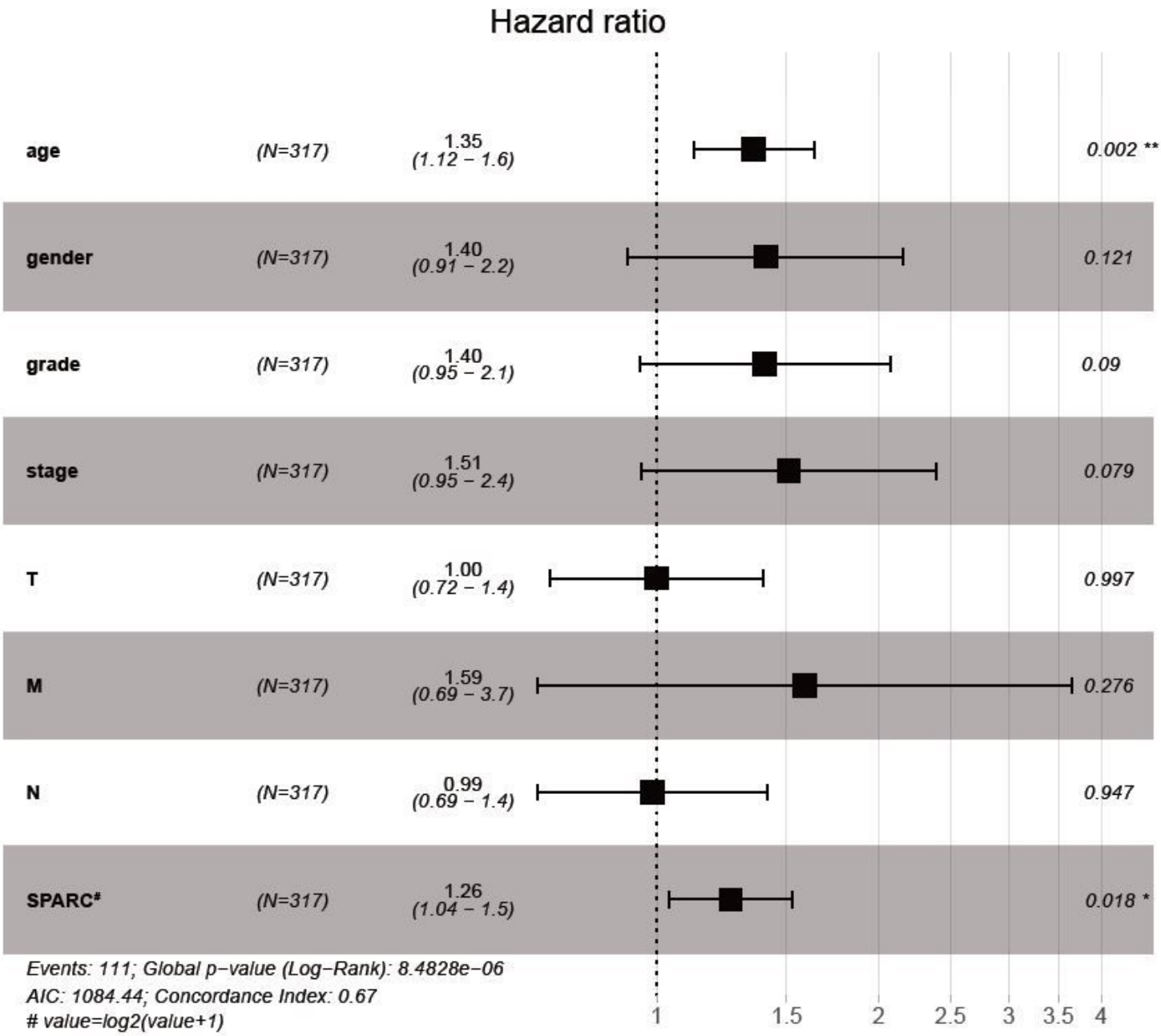

Figure 3

Multivariate analysis of the relationship between SPARC expression and overall survival among gastric patients

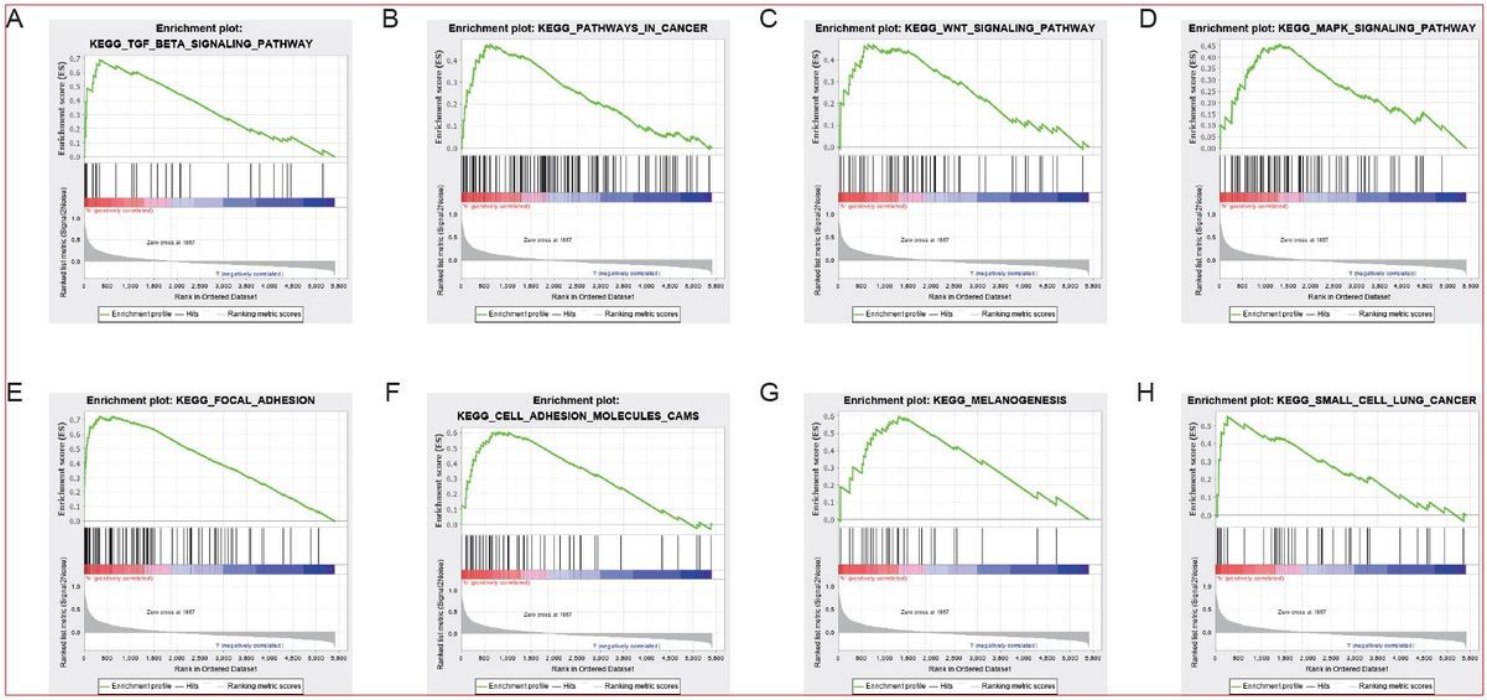

Figure 4

Enrichment plots from gene set enrichment analysis (GSEA).