Expression Pattern and Prognostic Value of p53-Pathway in Gastric Cancer

Wei Li
The First Affiliated Hospital of USTC
Changhong shi (ashi_248@163.com)
Guangzhou Medical University

Research

Keywords: Gastric cancer, p53-pathway, Prognostic signature

Posted Date: November 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1070589/v1

Abstract

Background: Gastric cancer is one of the most common cancer across the world. Increasing evidence suggest that p53-pathway plays a critical role in the initiation, progress and therapy of gastric cancer. However, the prognostic value of p53-pathway in gastric cancer is not fully understood.

Methods: A total of 67 p53-pathway-related genes and corresponding clinical information of 415 gastric cancer patients were downloaded from The Cancer Genome Atlas (TCGA) database. The consensus clustering algorithm were performed to analyze the expression level of all p53-pathway related genes. Based on the differentially expressed genes between different subgroups, Cox regression analysis were used to construct an independent prognostic signature, which was further validated in two external datasets.

Results: 3 subgroups of patients with significant different survival outcome were identified according to the expression level of p53-pathway-related genes. Comparison of gene expression between those subgroups identified 12 differentially expressed genes, 3 (THBS1, SERPINE1 and GADD45B) of which were significantly associated with the overall survival outcome. We further constructed a 3-gene signature as independence prognosis signature with promising performance in survival prediction of gastric cancer.

Conclusions: This study provides a potential prognostic signature for predicting prognosis of gastric cancer patients, which may benefit individual therapy of gastric cancer.

Background

Gastric cancer is the 5th most prevalent cancer in the world and the 3rd most lethal among all cancers [1]. Even though the 5-year survival rate of patients at early stage could reach 90-97%, the 5-year survival rate of patients at late stage is less than 30% [2]. Gastric cancer is a heterogeneous disease, and its outcomes can vary significantly even for patients with similar clinic features and therapeutic schedule [3, 4], indicating that conventional clinical characteristics are insufficient for prognostication. Therefore, developing novel prognostic signatures with significant prediction ability is critical to improve the therapeutic effect of gastric cancer.

Gastric cancer is a complex process that involves coordinate action of functionally associated genes from cancer related pathway. The p53-pathway is one of famous cancer related pathway, which controls cell-cycle progression, apoptosis and DNA repair, and plays an important role in cancer initiation, development and therapy of cancer [5–7]. The p53 pathway are frequently genetically altered in cancers. It has been reported that mutations of the p53 gene or other components of p53-pathway (e. g. Mdm2 and Mdmx) are detected in almost all cancer patients [8–12]. Prognostic value of p53 gene expression in gastric cancer has been intensively studied [13, 14]. However, some studies suggest that patients with low p53 expression tends to have better survival outcome, while other studies show that p53 expression has no relation with survival outcome [15, 16]. These contradictory results suggest that single p53 gene is not sufficient to serve as an effective prognostic signature for gastric cancer.

We note that previous studies often ignore the joint action between multiple genes involved in p53-pathway. Therefore, we investigate the prognostic value of the whole p53-pathway in this study. According to the expression level of p53-pathway related genes, we identified 3 subgroups of gastric cancer with significant different survival outcome. Based on these results, we further developed a novel prognostic signature with promising performance in predicting the survival outcome for patients with gastric cancer.

Methods

Samples and database

The discovery cohort that contains 415 gastric cancer patients were download from the TCGA database (https://portal.gdc.cancer.gov/). Two independent validation cohorts (GSE66229 and GSE84437) were download from Gene Expression Omnibus database. For datasets from the TCGA database, mRNA expression was quantified with fragment per kilobase of exon per million reads mapped (FPKM). For all expression datasets from the GEO database, background correction and quartile normalization were performed by applying the robust multi-array average algorithm. The average value of gene symbol with multiple probes was calculated as expression level.

Analysis of expression pattern of p53-pathway

A panel of genes involved in p53-pathway was retrieved from the KEGG pathway database (https://www.kegg.jp/kegg/). The consensus clustering algorithm were performed to analyze the expression level of all p53-pathway related genes, and divided the patients into three clusters [17]. The “edgeR” algorithm was performed to identify the differentially expressed genes (DEG) (FDR<0.05, |log2(foldchange)|>1) between two clusters [18].

Construction and validation of prognostic signature

Univariate Cox regression analysis were used to evaluate the association between the p53-pathway-related DEGs with the survival outcome. Multivariate Cox regression analysis were performed on these candidate genes to construct an independent prognostic signature. Age, gender, stage and risk score were used to construct nomogram using the “rms” and “survival” package in R. The calibration curve was drawn to assess the consistency between actual and predicted survival. The ROC curve was used to compare the performance of nomogram with the traditional TNM system using “timeROC” package in R [20]. The R package “clusterprofiler” was used to perform the Gene Ontology (GO) and KEGG enrichment analysis [21]. The co-expression network was constructed using R package “WGCNA”[19].

Statistical analysis

All the statistical analysis were performed in the R environments. Data were analyzed with standard statistical tests as appropriate. For identifying independent risk factor for survival, multivariate Cox regression analysis were performed to adjust covariates.

Results

Expression pattern of p53-pathway correlate with survival

We retrieved 67 genes involved in p53-pathway from the KEGG pathway database (pathway: map04115). Quantile method [22] was used to remove genes with low expression level and 62 genes remained. Finally, the expression level of the 62 genes were log transformed for further analysis (Fig. 1A).

Performing consensus clustering analysis in the discovery cohort, we found that all patients were clustered into 3 subgroups (Fig. 1A). The clinic characteristics of patients in 3 subgroups are listed in Table 1. The results showed that the age, TP53 mutation and histological diagnosis were significantly different between 3 clusters. The frequency of TP53 mutation in cluster 2 (55.4%) is much higher than that in cluster 1 (13.2%) and cluster 2 (31.8%). Comparing the survival curve between 3 clusters, we found that cluster 1 has the best prognosis outcome, while cluster 2 and cluster 3 showed comparable prognosis (Fig. 1B). There is a significant difference in prognosis outcome between cluster 1 and cluster 3 (log rank p=0.031) (Fig. 1C).

Comparing the expression level of genes between cluster 1 and cluster 3, we identified a total of 12 deferentially expressed genes (DEGs) (Fig. 1D), 10 of which was over-expressed in cluster 3. GO enrichment analysis showed that these DEGs were mainly related to cell cycle arrest (Fig. 1E).

Table 1
The clinical characteristics of each cluster

Variables	Cluster 1 (n=106)	Cluster 2 (n=177)	Cluster 3 (n=132)	p.value
Gender:				0.083
Female	47 (44.3%)	58 (32.8%)	42 (31.8%)	
Male	59 (55.7%)	119 (67.2%)	90 (68.2%)	
Age:				<0.001
<65	31 (29.5%)	69 (40.4%)	71 (54.6%)	
>=65	74 (70.5%)	102 (59.6%)	59 (45.4%)	
Histological diagnosis				<0.001
Intestinal type	16 (15.1%)	44 (24.9%)	17 (12.9%)	
NOS type	46 (43.4%)	64 (36.2%)	46 (34.8%)	
Diffuse type	15 (14.2%)	15 (8.47%)	37 (28.0%)	
Tubular type	23 (21.7%)	40 (22.6%)	10 (7.58%)	
Else	6 (5.66%)	14 (7.91%)	22 (16.7%)	
TP53 mutation:				<0.001
No	92 (86.8%)	79 (44.6%)	97 (78.2%)	
Yes	14 (13.2%)	98 (55.4%)	27 (21.8%)	
TNM Stage:				0.666
I	14 (14.4%)	30 (17.6%)	13 (10.6%)	
II	30 (30.9%)	53 (31.2%)	40 (32.5%)	
III	44 (45.4%)	67 (39.4%)	58 (47.2%)	
IV	9 (9.28%)	20 (11.8%)	12 (9.76%)	
T stage:				0.041
T1	8 (7.92%)	10 (5.71%)	4 (3.08%)	
T2	16 (15.8%)	46 (26.3%)	26 (20.0%)	
T3	42 (41.6%)	83 (47.4%)	56 (43.1%)	
T4	35 (34.7%)	36 (20.6%)	44 (33.8%)	
N stage:				0.817
N0	33 (32.7%)	54 (32.0%)	36 (28.6%)	

Variables	Cluster 1 (n=106)	Cluster 2 (n=177)	Cluster 3 (n=132)	p.value
N1	32 (31.7%)	43 (25.4%)	37 (29.4%)	
N2	16 (15.8%)	38 (22.5%)	25 (19.8%)	
N3	20 (19.8%)	34 (20.1%)	28 (22.2%)	
M stage				0.840
M0	94 (93.1%)	155 (93.9%)	118 (92.2%)	
M1	7 (6.93%)	10 (6.06%)	10 (7.81%)	

Construction and validation of p53-pathway-related prognostic signature

To evaluate the prognostic value of 12 p53-pathway-related DEGs, we first performed univariate Cox regression analysis and found that 3 of 12 genes were significant associated with the survival outcome (Fig. 2A), including SERPINE1 (p<0.001), THBS1 (p=0.001) and GADD45B (p=0.004). Then, the multivariate Cox regression model was employed to construct a prognostic signature as following: $\text{risk score} = 0.251 * \text{SERPINE1} + 0.028 * \text{THBS1} + 0.103 * \text{GADD45B}$. The distribution of risk score, survival status and expression profile of signature gene were showed in Fig. 2E.

Using the median risk score as the critical value, all patients were classified into high and low risk group. As shown in Fig. 2B, the patients in the low-risk group tends to live longer than those in the high-risk group (log rank test, p<0.001). The prognosis signature was further validated within two patients' subgroups with early stage (stage I/II) and late stage (stage III/IV) cancer, respectively. In both subgroups, patients were divided into the low-risk group with better survival and the high-risk group with worse outcomes (Fig. 2C and D). The prognosis signature was also validated by 2 external cohort, including GSE84437 (Fig. 3D) and GSE66229 (Fig. 3E). From the figures, we observed that two external cohort were consistently classified into the high-risk group with worse survivals and low-risk group with better outcome. Multivariate analysis also identified the prognostic signature as an independent prognostic factor (adjusted p<0.001), which is independent of the tumor stage and age (Table 2).

Table 2
multivariable analysis of prognostic signature in the prediction of gastric cancer survival

Variables	HR (95% CI)	*P* value
Age (Year) (>=65 vs <65)	1.030 (1.013,1.047)	<0.001
TNM stage		
Stage II vs. Stage I	1.276 (0.659, 2.470)	0.047
Stage III vs. Stage I	1.978 (1.054, 3.712)	0.034
Stage IV vs. Stage I	3.427 (1.778, 6.607)	<0.001
Risk score (high vs. Low)	2.172 (1.544, 3.056)	<0.001

Finally, we constructed a nomogram by integrating the p53-pathway-based risk score and the well-known prognostic factors including age and TNM stage (Fig. 3A). The calibration curve showed that the nomogram performed well, compared with an ideal model (Fig. 3B). The AUC of the ROC curve of the nomogram for 5 years reached 0.71, higher than that of TNM stage (AUC = 0.66), indicating better performance of our prognosis signature (Fig. 3C).

Functional analysis of the p53-pathway-related gene

To explore the potential target gene that regulated by three genes of the prognostic signature, we construct a co-expression network of the p53-pathway using the WGCNA package in R. The co-expression network showed that the three genes may regulated 10 genes involved in p53-pathway (Fig. 4A), including CDK1, CCNE2, CHEK1 and so on. GO (Fig. 4B) and KEGG analysis (Fig. 4C) showed that the differential expressed genes identified by comparing high risk and low risk group were enriched in several important pathways, including Extracellular matrix organization, ECM-receptor interaction, PI3K-Akt signaling pathway and so on.

Discussion

The p53 pathway plays a critical role in the tumorigenesis and progression of gastric cancer [23, 24]. The prognostic role of p53 expression in gastric cancer has been intensively studied and the conclusion are controversial [13], suggesting that single p53 gene may be not sufficient to serve as a prognostic signature. Through this study of expression of 62 genes involved in p53 signaling pathway, we demonstrated that the gastric cancer patients can be divided into 3 subgroups with significant different survival outcome. Among them, cluster 1 manifested the worst survival while cluster 3 manifest the best survival. Comparison of gene expression between cluster 1 and cluster 3 identified 12 differential expressed gene, 3 of which (THBS1, SERPINE1 and GADD45B) were significant associated with the survival outcome. Finally, we constructed a 3-gene prognostic signature which performed well in both discovery, validation and external cohort .

The role of the signature genes identified in this study have been previously reported in multiple types of cancers. THBS1 is a subunit of a disulfide-linked homotrimer protein that mediates cell-to-cell interaction, which plays important roles in angiogenesis and tumorigenesis [26]. Our study indicated that upregulated THBS1 was associated with poor prognosis in gastric cancer. SERPINE1 functions as the principal inhibitor of tPA and uPA, which promotes the metastasis of tumors [27, 28]. Our study showed that upregulated SERPINE1 was associated with a poor prognosis in gastric cancer. GADD45B is a member of the DNA damage-inducible gene family, which is associated with cell growth control and apoptosis [30]. Previous studies have reported that the over expression of GADD45B may inhibit cell growth in a variety of human tumor cell lines [31,32], including hepatocellular carcinoma cells, prostate cancer cells and others. Our study suggested that patients with high GADD45B tends to have shorter survival time.

Conclusion

In summary, in this study we revealed the gene expression profile of p53-pathway related genes and developed a novel prognostic signature in gastric cancer. Although the impact and function of distinctive genes need to be further experimental investigation, we believe that these results will facilitate the exploration of novel therapies.

Abbreviations

DEGs: Differentially expressed genes; GC: Gastric cancer; OS: Overall survival;

Declarations

Acknowledgements

None.

Authors' contributions

WL and CS participated in the design of this study, and they both performed the statistical analysis and drafted the manuscript. All authors read and approved the final manuscript.

Funding

This work is funded by the National Nature Science Foundation of China (No.12001129 to Changhong Shi)

Availability of data and materials

The datasets supporting the conclusion of this article are included within the article.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publishers Note

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7–34.
2. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet (London, England). 2020;396(10251):635–648.
3. Shah MA, Ajani JA. Gastric cancer-an enigmatic and heterogeneous disease. JAMA. 2010;303(17):1753–1754.
4. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet (London, England). 2020;396(10251):635–648.
5. Joerger AC, Fersht AR. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. Annu Rev Biochem. 2016; 85:375–404.
6. Vazquez A, Bond EE, Levine AJ, Bond GL. The genetics of the p53 pathway, apoptosis and cancer therapy. Nat Rev Drug Discov. 2008;7(12):979–987.
7. Stracquadanio G, Wang X, Wallace MD, et al. The importance of p53 pathway genetics in inherited and somatic cancer genomes. Nat Rev Cancer. 2016;16(4):251–265.
8. Hwang HJ, Nam SK, Park H, et al. Prediction of TP53 mutations by p53 immunohistochemistry and their prognostic significance in gastric cancer. J Pathol Transl Med. 2020;54(5):378–386.
9. Bykov VJN, Eriksson SE, Bianchi J, Wiman KG. Targeting mutant p53 for efficient cancer therapy. Nat Rev Cancer. 2018;18(2):89–102.
10. Stracquadanio G, Wang X, Wallace MD, et al. The importance of p53 pathway genetics in inherited and somatic cancer genomes. Nat Rev Cancer. 2016;16(4):251–265.
11. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202–209.
12. Ohtsuka J, Oshima H, Ezawa I, Abe R, Oshima M, Ohki R. Functional loss of p53 cooperates with the in vivo microenvironment to promote malignant progression of gastric cancers. *Sci Rep*.

2018;8(1):2291.

13. Nie K, Zheng Z, Wen Y, et al. Construction and validation of a TP53-associated immune prognostic model for gastric cancer. *Genomics*. 2020;112(6):4788–4795.
14. Hwang HJ, Nam SK, Park H, et al. Prediction of TP53 mutations by p53 immunohistochemistry and their prognostic significance in gastric cancer. *J Pathol Transl Med*. 2020;54(5):378–386.
15. Wei K, Jiang L, Wei Y, et al. The prognostic significance of p53 expression in gastric cancer: a meta-analysis. *J Cancer Res Clin Oncol*. 2015;141(4):735–748.
16. Gonçalves AR, Carneiro AJV, Martins I, et al. Prognostic significance of p53 protein expression in early gastric cancer. *Pathol Oncol Res*. 2011;17(2):349–355.
17. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–140.
18. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–140.
19. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9..
20. Seiler M, Huang CC, Szalma S, Bhanot G. ConsensusCluster: a software tool for unsupervised cluster discovery in numerical data. OMICS. 2010;14(1):109–113.
21. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000;56(2):337–344.
22. Conesa A, Madrigal P, Tarazona S, et al. A survey of best practices for RNA-seq data analysis. Genome Biol. 2016; 17:13.
23. Joerger AC, Fersht AR. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. Annu Rev Biochem. 2016; 85:375–404.
24. Pietsch EC, Humbey O, Murphy ME. Polymorphisms in the p53 pathway. Oncogene. 2006;25(11):1602.
25. Daubon T, Léon C, Clarke K, et al. Deciphering the complex role of thrombospondin-1 in glioblastoma development. *Nat Commun*. 2019;10(1):1146.
26. Qi C, Lei L, Hu J, Wang G, Liu J, Ou S. Thrombospondin-1 is a prognostic biomarker and is correlated with tumor immune microenvironment in glioblastoma. *Oncol Lett*. 2021;21(1):22.
27. Arroyo-Solera I, Pavón MÁ, León X, et al. Effect of serpinE1 over-expression on the primary tumor and lymph node, and lung metastases in head and neck squamous cell carcinoma. *Head Neck*. 2019;41(2):429–439.
28. Li L, Zhu Z, Zhao Y, et al. FN1, SPARC, and SERPINE1 are highly expressed and significantly related to a poor prognosis of gastric adenocarcinoma revealed by microarray and bioinformatics. *Sci Rep*. 2019;9(1):7827.
29. Gong L, Cai L, Li G, Cai J, Yi X. GADD45B Facilitates Metastasis of Ovarian Cancer Through Epithelial-Mesenchymal Transition. *Onco Targets Ther*. 2021; 14:255–269.

30. Zhao Z, Gao Y, Guan X, et al. GADD45B as a Prognostic and Predictive Biomarker in Stage II Colorectal Cancer. *Genes (Basel)*. 2018;9(7).

31. Wang L, Xiao X, Li D, et al. Abnormal expression of GADD45B in human colorectal carcinoma. *J Transl Med*. 2012;10(1):1–10.

Figures

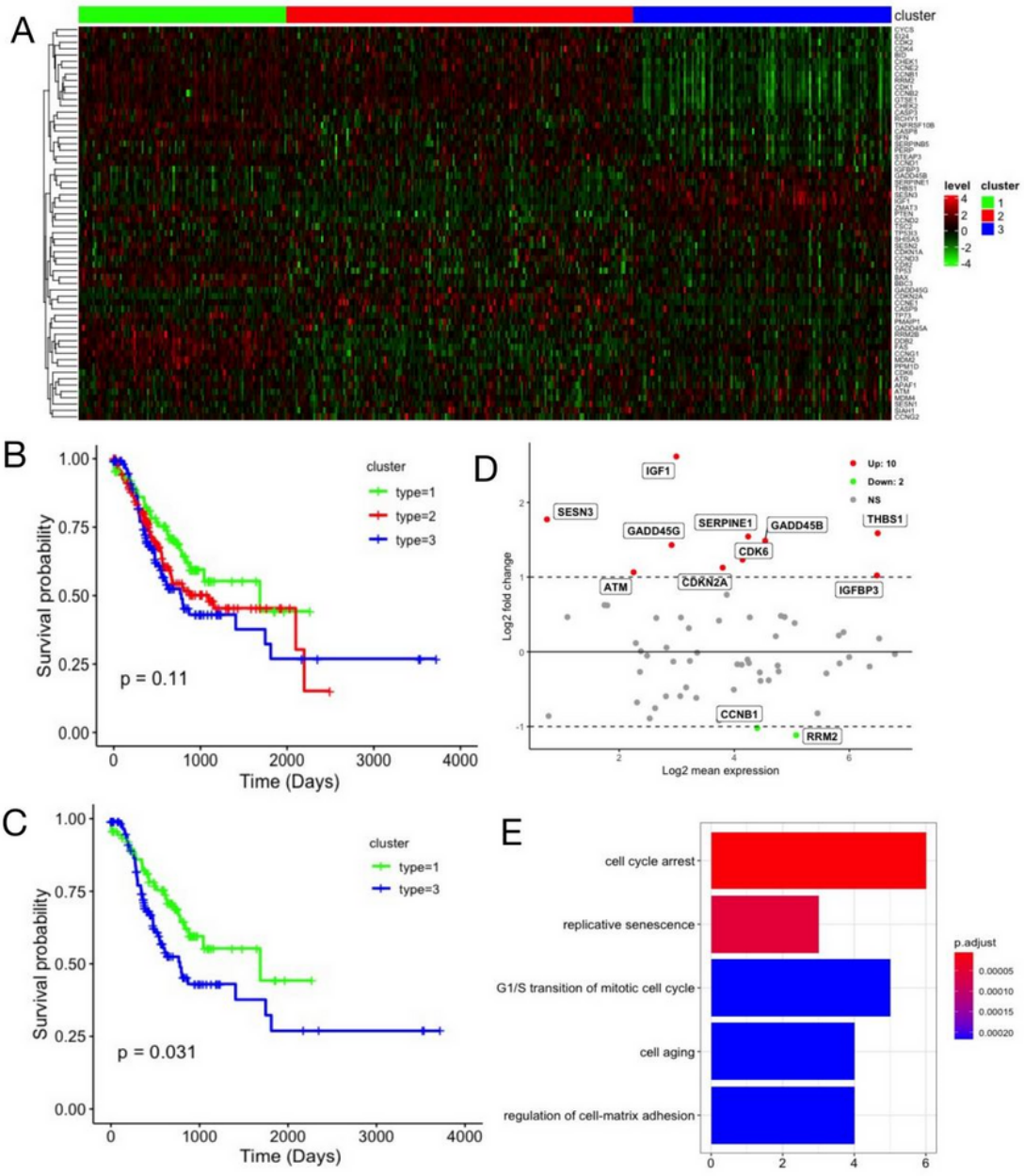

Figure 1

(A) The gastric cancer patients were divided into 3 clusters based on the mRNA expression; (B) Survival difference between 3 clusters; (C) Survival difference between cluster 1 and cluster 3; (D)Differentially expressed genes between cluster 1 and cluster 3; (E) The GO enrichment analysis of the p53-pathway-related risk DEGs.

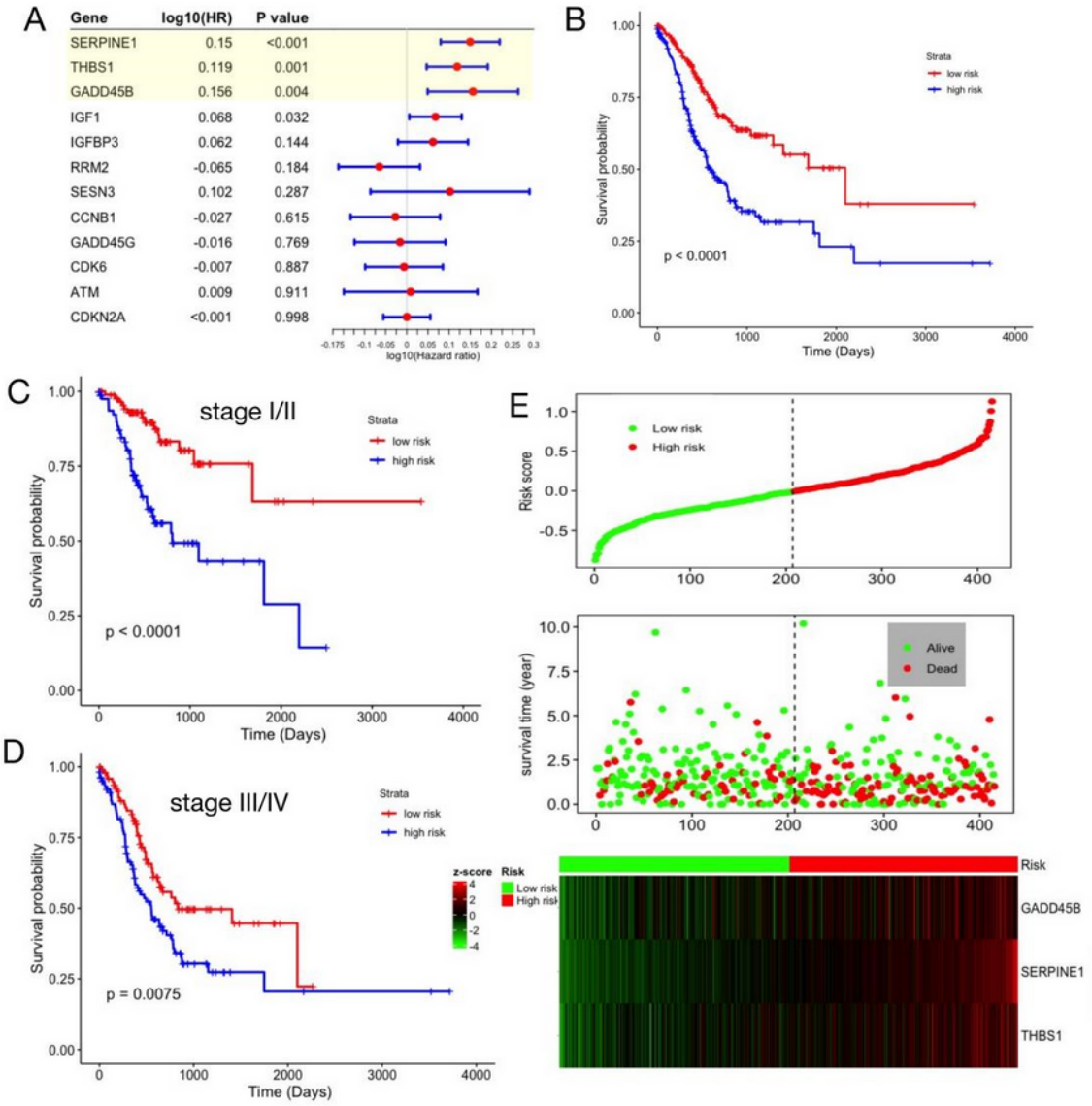

Figure 2

(A) Forest plot of hazard ratio for 12 p53-pathway-related risk DEGs; (B)(C)(D) Kaplan-Meier plot of overall survival for patients in low risk and high-risk group by P53-pathway-related signature in the whole cohort, early-stage subgroup and late-stage subgroup; (E)Distribution of risk score, survival status and expression profile of signature gene.

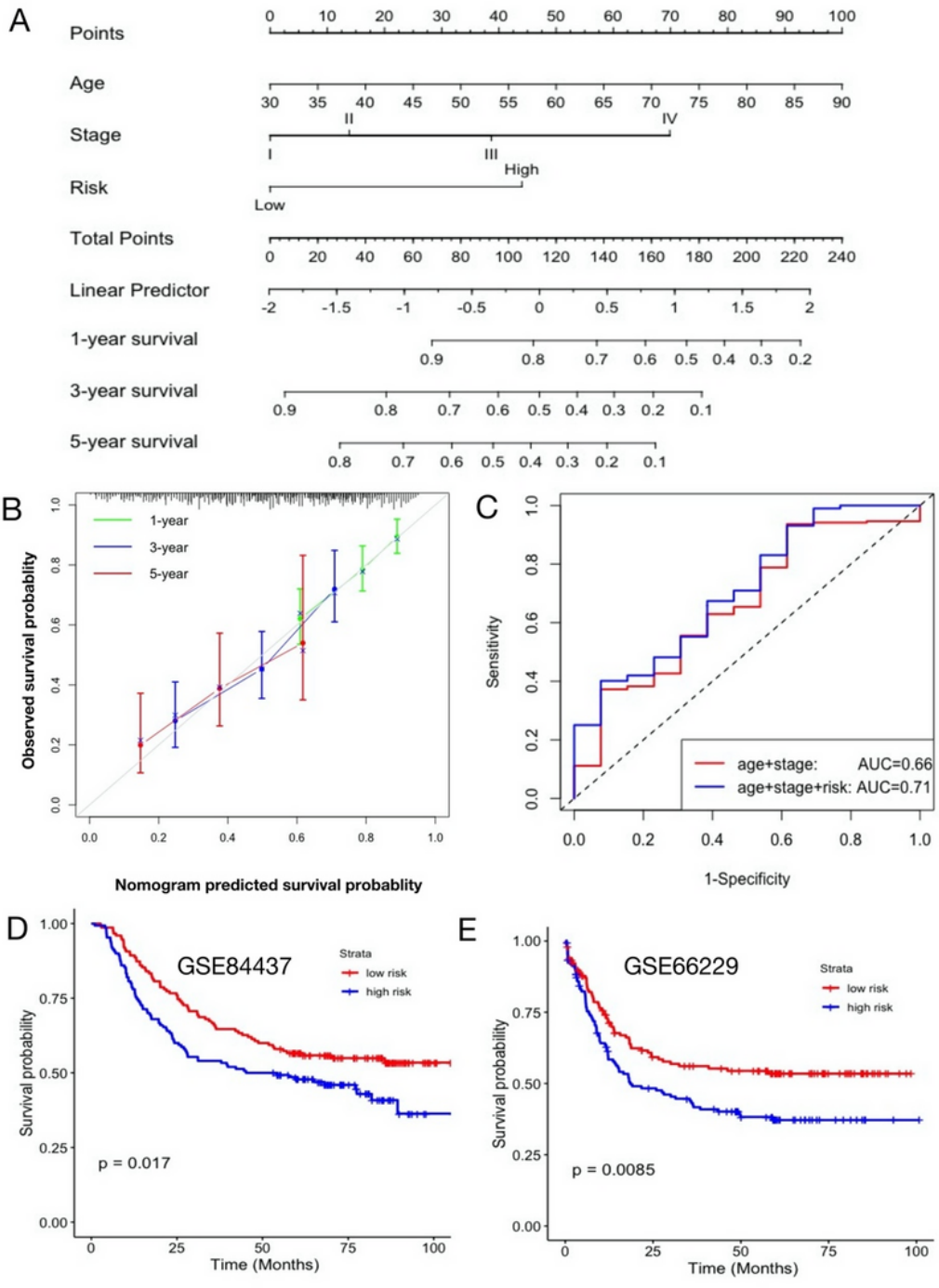

Figure 3

(A) Nomogram developed using discovery cohort to predict 1-, 3- and 5-year overall survival probability; (B) The calibration plot of nomogram in terms of agreement between predicted and observed outcome; (C) ROC curve of the nomogram for 5-year survival prediction; (D)(E) Kaplan-Meier plot of overall survival for patients in low-risk and high-risk group divided by P53-pathway-related signature in the GSE84437 and GSE66229 datasets.

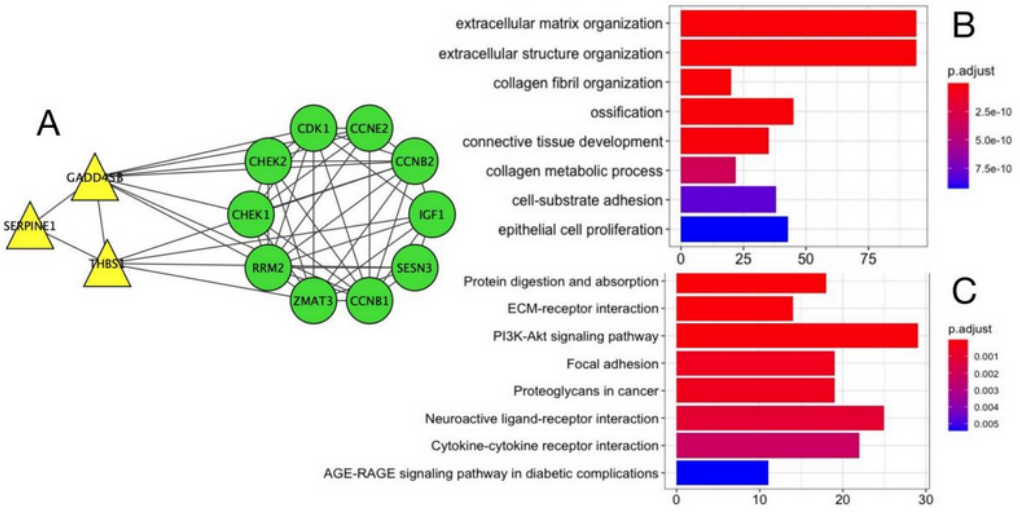

Figure 4

(A) co-expression network of the three genes of the prognostic signature; (B) GO analysis of DEGs between low and high-risk group; (C) KEGG analysis of DEGs between low and high-risk group.