Identification of An Immune Signature Predicting Prognosis Risk of Patients in Gastric Cancer

wei dang (dangwei@sjtu.edu.cn)
shanghai jiaotong university
Cai Chen
shanghai jiaotong university
Zhang Ruiqiao
shanghai jiaotong university
Fan Qingquan
shanghai jiaotong university
Song Xiaoling
shanghai jiaotong university
Wang Weimin
The Third People's Hospital of Chongming District
Gu Jun
shanghai jiaotong university

Research

Keywords: gastric cancer, prognosis, prediction, Immunotherapy, immune, somatic mutation

Posted Date: January 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-139109/v1

Abstract

Background

Gastric cancer is a common but lethal cancer owing to deficient in effective treatment. Substantial evidences have proved that immune infiltration plays a key role in progression of gastric cancer. This study aimed to establish a signature based on immune related genes that can predict clinical outcomes and therapeutic efficacy.

Methods

The expression data from The Cancer Genome Atlas database and 4617 immune related genes from previously published 160 immune gene sets were collected for development and validation of the signature. Cox proportional hazard regression model was used to construct the signature. The reliability and forecasting ability were evaluated by two independent datasets from GEO.

Results

A gene model consisting of 47 immune related genes was used as our signature. Risk scores were calculated based on the coefficient and the expression level of each gene in this model. The low risk score group had an obviously favorable prognosis than the other group in all cohorts. Both of univariate and multivariate analysis suggested that our immune gene signature was an independent prognostic factor. Single sample gene Set Enrichment Analysis (ssGSEA) revealed that high risk score was associated with high Th17 cell infiltration, low mast cell and pro- angiogenesis immune cell infiltration. More importantly, patients with high risk score presented high tumor mutation burden (TMB), which is an essential element for predicting therapeutic efficacy of immune check point inhibitor.

Conclusion

This signature is a promising tool to predict prognosis and screen out population who can get benefit from immune check point inhibitor.

Introduction

According to GLOBOCAN 2018 data, gastric cancer has become the 5th most frequently diagnosed tumor type and the 3rd most lethal cancer worldwide[1]. Given the heterogeneity of advanced stomach cancer, the prognosis of patients is difficult to predict and the treatment effects of various conventional means, including surgery, chemotherapy and radiotherapy, are rare. The classic prognosis predicting model is the eighth American Joint Committee on Cancer tumor, node and metastasis (TNM) classification[2]. Although

this model provides a quick estimate, but the specificity and sensitivity are limited. Han et al. established a prognosis predicting model based on age, sex, location, depth of invasion, number of metastatic lymph nodes and number of examined lymph nodes, which performs better in predicting long-term overall survival than previous models[3]. As the immune check point inhibitors achieved great success in the treatment of melanoma, tumor immunology has attracted much attention world widely. Thus, it is indispensable to take immune cells and immune related pathways into account when developing a prognosis predicting model.

The center of therapy has transferred from the tumor itself to the microenvironment. Evidence revealed that the most significant components of the tumor microenvironment are tumor-related stromal cells and infiltrating immune cells[4]. Several research have demonstrated that intra-tumoral IL17-producing cells, tumor-infiltrating lymphocytes, tumor-activated neutrophils and CD8 T cell are closely connected to tumorigenesis of gastric cancer[5–8]. Recently, Thorsson et al. identified six immune subtypes characterized by diversities of immune signature among 33 different cancer type according to 160 immune gene sets[9]. Profoundly understanding the expression level of immune related gene in stomach would provide clinicians with more precise prognostic information.

Therefore, it's crucial to develop an immune signature in the light of a comprehensive immune gene sets that can reflect the immune microenvironment of individuals and help to improve the ability of prognosis forecasting.

In this study, we concentrated efforts on developing an immune signature to estimate the prognosis and screen out population who can get benefit from immunotherapy based on well established immune gene sets. The expression data of 862 gastric cancer patients from The Cancer Genome Atlas (TCGA) database and the Gene Expression Omnibus (GEO) database were collected for analysis. Next, several analyses are employed to evaluate the performance of our model. Furthermore, we demonstrated that this immune signature had the ability to predict the response to immune checkpoint inhibitors.

Methods

Cohort Datasets and Immune Related Genes

TCGA RNA-seq datasets and clinical data of gastric cancer were collected through the Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/) as training set. Microarray data and corresponding clinical data of total 500 patients were download from the GEO database (http://www.ncbi.nlm.nih.gov/geo). Only the cohorts with more than 50 gastric cancer sample and overall survival time were retained, resulting in a total of 862 patients. The clinical features of the included cohorts were summarized in Table 1. The immune-related features of cancer were collected from previously published work[9], which was available at https://gdc.cancer.gov/about-data/publications/panimmune. Only the immune genes expressed among all the samples will be retained

for building our model. Since our data were obtained from public databases and we entirely complied by the publishing guidelines provided by TCGA and GEO, there were no necessity for ethical approvals.

Differentially Expressed Gene (DEG) Analysis

The DEseq2 package was applied to DEG analysis based on HTSeq-Counts data from TCGA database[10]. The DEGs were defined as genes with a log2 fold change > 1 and P-value <0.05. The Venn Diagram package was used for generation of differential immune related gene[11]. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were conducted by the cluster Profiler package in R to investigate the potential biologic functions of differentially expressed immune related genes[12-14]. Functional categories with an adjusted P value less than 0.05 were deemed as significant pathways.

Establishment and validation of the immune signature for gastric cancer

The expression data and survival data of TCGA database were used to develop the immune signature. Univariate analysis and log Rank test were applied to find immune gene associated with prognosis. For the genes with prognostic capability, Lasso and multivariate Cox regression analysis was adopted to identify the best gene model for predicting prognosis of gastric cancer patients. Both forward and backward selection were utilized to achieve a series of gene signature, and the best performing gene signature was established according to the Akaike Information Criterion (AIC). Then, the risk scores were calculated based on the coefficients and the expression level of genes in this model. Consequently, the patients can be divided into low-risk group and high-risk group according to the median of risk score. The concordance (c)-index was used to confirm the forecasting ability of the signature in all the cohorts[15]. The receiver operating characteristic (ROC) curve was applied to asses the sensitivity and specificity of the model. General expression of immune related genes in our signature were showed with heatmaps in all the TCGA dataset and GEO datasets. The survival time and status of patients were displayed with dot plots. The distribution of low-risk group and high-risk group were showed with curve graphs.

Survival Analysis

The overall survival (OS) of the low-risk group and the high-risk group were graphically exhibited by the Kaplan-Meier (K-M) survival curves. The univariate and multivariate analyses of survival were operated for both the immune signature and clinicopathologic features to identify the independent predictor of prognosis in gastric cancer. “survminer” package and “survival” package were applied to perform above analyses and generate several graphs.

Single Sample Gene Set Enrichment Analysis

The marker gene set for 25 most frequent immune cell types were available from Şenbabaoğlu Z et al.[16]. Enrichment scores of each immune cell in all patients were derived from the GSEA program.

Briefly, the infiltration levels of immune cell types were quantified by ssGSEA in the R package "GSVA"[17]. The difference of enrichment scores of 25 immune cells in low-risk group and high-risk group were investigated with student test. We also performed univariate analysis to find the immune cells associated with prognosis. The correlation between enrichment score of prognostic immune cell and risk score was figured out by Pearson correlation.

Somatic Mutation Load and Neoantigen Analysis

Aggregated somatic mutation data with Mutation Annotation Format (MAF) form were download from the GDC Portal (https://portal.gdc.cancer.gov/).

R package called "maftools" was used to extract the mutational information from download files and check the overview of mutation in gastric cancer [18]. Given the significant role of nonsynonymous mutations in generating neoantigen epitopes, we integrated the mutation data to figure out the number of each category of mutations. Since the tumor mutation burden (TMB) is the leading indicator of the therapeutic effect of immune checkpoint inhibitors, we explore the difference of TMB between low-risk group and high-risk group. The TMB is evaluated by the total number of nonsynonymous mutations, including nonsense mutation, missense mutation, splice site mutation, frameshift mutation and inframe mutation. The association of our signature with the single nucleotide polymorphism were also analyzed. Furthermore, the number of neoantigens was downloaded from previously published research [19] so that we can identify the relationship between the immune gene signature and the number of neoantigens in stomach cancer patients.

Statistical Analysis

Statistical comparison was performed with student's t test. The "ggplot2" and "ggpubr" R packages were utilized to generate bar plot and violin plots. "pheatmap" R package was used to generate heatmaps. Two-tailed p values less than 0.05 were thought to be statistically significant. All analyses were performed with R software version 3.6.2 (https://www.r-project.org/)

Results

Construction of immune signature

In order to identify gastric cancer specific immune related gene, we first perform differential gene analysis by DEseq2 R package. There are 4567 differentially expressed genes compared with adjacent normal tissue. Then, we check the immune gene list among all the three cohorts to make sure they express in all gastric cancer samples. A total of 608 immune-related genes were identified as differential expressed in gastric cancer tissues (Figure 1A). The univariate analysis was applied with all the 608 immune related genes for TCGA STAD datasets. There were 104 most remarkable genes filtered out with the criteria of a p-value less than 0.05. Multivariate Cox analysis were performed on the 104 immune related genes to generate the best gene signature in terms of predicting the overall survival of gastric cancer patients and

47 genes were eventually selected to develop a prognostic mode. The risk score of each patient was calculated based on the sum of the average expression of genes multiplied corresponding coefficients in multivariate Cox analysis, gene list and the coefficient values were showed in Table S1. Then all the 862 stomach cancer patients in three cohorts were classified into a low-risk group and a high-risk group in terms of the median risk score. The overview of survival time and expression heatmap of genes in our model were presented in Figure 1B-D.

Validation of Immune-related Prognostic Signature

To sustain our immune gene signature, we calculated the c-index for the forecast of overall survival. The c-index for TCGA dataset, GSE62254 and GSE15459 were 0.7754, 0.7082 and 0.7742 respectively (Figure 2A), which meant the high predictive veracity of the signature for survival. The area under the ROC curve for 5-year OS in TCGA dataset, GSE62254 and GSE15459 were 0.730, 0.733 and 0.798 respectively, indicating that this prognostic model presented a wonderful sensitivity and specificity (Figure 2B-D). Survival analysis presented that patients of high-risk showed dramatically poorer overall survival than those of low-risk in all the three datasets ($P < 0.001$; Figure 3A-C).

Independence of Immune-related Gene Signature from Other Clinical Features

A total of 862 gastric cancer patients with clinical information, including gender, age, tumor stage, TNM classification were selected for analysis. We further analyzed 300 patients of GSE62254 dataset with more detailed data, such as recurrence, MLH1 immunohistochemistry (IHC) and EBV in-situ hybridization (ISH). The univariate Cox analysis represented that stage III($p<0.05$), stage IV ($p<0.005$), M1 ($p=0.0157$), N1-N3 ($p<0.05$), MLH1 IHC ($p=0.0012$), recurrence ($p<0.0001$) and risk score ($p<0.0001$) were significantly associated with overall survival (Figure 4). Multivariate Cox analysis further indicated that our immune related gene signature could be independent predictor of patients' overall survival after regulated by other clinical factors in TCGA dataset [Hazard ratio (HR) =1.2865, 95% confidence intervals (95% CI) 1.2236-1.3526, $p<0.0001$], GSE62254 dataset (HR=1.3200, 95% CI 1.1965-1.45, $p<0.0001$) and GSE15459 dataset (HR=1.0240, 95% CI 1.0132-1.0350, $p<0.0001$), as demonstrated in Figure 5.

Association of Risk Score with Immune Cell Infiltration

As shown in Figure 6A, infiltration level of eosinophils, macrophages, mast cells, Natural killer cells (NK cells), Plasmacytoid dendritic cells (pDC), gamma delta T cells (Tgd cells) and angiogenesis related immune cells in high-risk group were higher than low-risk group, whereas infiltration level of T helper cells and Th2 cells in low-risk group were higher than high-risk group.

In addition, the relationships between immune cell infiltration and the risk score were analyzed by Pearson correlation. Eosinophils, macrophages, mast cells, Natural killer cells (NK cells), Plasmacytoid dendritic cells (pDC), gamma delta T cells (Tgd cells) and angiogenesis related immune cells were positive correlated with risk score. However, T helper cells and Th2 cells were negative correlated with risk score (Figure 6B).

Association of Risk Score with Tumor Mutation Burden and Neoantigen

Therapeutic efficiency of immune checkpoint inhibitor could be estimated with nonsynonymous mutation burden[20, 21]. Thus, we detected the association of our immune signature and each type of nonsynonymous mutation. Furthermore, we would like to explore whether our immune signature could be an predictor of reaction to immune checkpoint inhibitor in terms of the number of neoantigen. In general, patients with low risk score revealed higher nonsynonymous mutation load than those with high risk score (p=0.012, Figure7A). In addition, low-risk group patients had higher missense mutation (p=0.011, Fig. 7B), splice site mutation (p = 0.015, Fig. 7D) and Single Nucleotide Polymorphism (SNP) (p=0.012, Figure 7I) .We did not find this relationship in nonsense mutation (p = 0.052, Fig. 7C), frameshift mutation (Figure 7E and F), inframe deletion (p = 0.053, Figure 7G), inframe insertion (p=0.45, Figure 7H), total deletion mutation (p=0.075, Figure 7J), and total insertion mutation (p=0.37, Figure 7K). Moreover, we found there be a positive correlation between the signature and the number of neoantigens in TCGA STAD dataset (Figure 7L).

Discussion

The treatment and prognosis prediction of gastric cancer were still an exciting challenge for investigators and clinicians. Researches of tumor immunology and clinical trials of immune checkpoint inhibitor shade light on the critical role of immune-related gene in the development and

prognosis of stomach cancer[22–24]. In this study, we developed a reliable prognostic signature based on immune related gene sets in TCGA dataset and validated its efficacy in external GEO datasets. Our signature demonstrated the specific landscape of prognostic immune gene expression of gastric cancer and provided a predictor of prognosis and therapeutic effect of immune checkpoint inhibitors.

In the current investigation, 608 differentially expressed immune-related genes were filtered from 375 samples of gastric cancer and 32 normal tissues. This step could help us to screen out gastric cancer specific immune related genes. After univariate analysis and Cox hazard proportion regression analysis, 47 immune related genes were identified to constitute our signature. Since we screen out these 47genes from prognosis related gene, each gene could be a predictor of prognosis, and integration of these genes could develop a more robust model. Most genes of our signature, such as COL1A1, COL3A1, COL1A2, TIMP1, RGS5 and so on, had been proved that they played critical roles in development and prognosis of various types of cancer[25–30].This study demonstrated that our signature was closely correlated with overall survival of gastric cancer patients in all the three datasets. Concordance index and ROC curve were utilized to prove our model play well in predicting patients' OS in gastric cancer. Univariate and multivariate analysis revealed that risk score was positive related with high risk in TCGA dataset and GEO datasets and could be an independent predictor of prognosis. However, some clinical factor, including TNM classification and stage were not independent predictor of OS, which means that our signature performed better than classical prognostic model.

Furthermore, this research demonstrated that our signature was associated with several immune cell infiltration. The high-risk group mainly displayed enrichment of immunosuppressive cells, such as mast cell and angiogenesis related immune cells, which was consistent with other researches. Lv et al. had found mast cell degranulation could effectively promoted the proliferation and inhibited the apoptosis of gastric cancer cells in vitro and contributed to the growth and progression of gastric cancer tumors in vivo[31]. Angiogenesis is one of hallmarks of cancer, and stromal cells can produce various angiogenic factors, including vascular endothelial growth factor, interleukin-8, platelet-derived endothelial cell growth factor, and angiopoietin in gastric cancer[32, 33]. The low-risk group revealed enrichment of T cells to enhance the capacity of adaptive immunity[34].

More interesting, we found low-risk group was associated with higher nonsynonymous mutation, missense mutation, splice site mutation and Single Nucleotide Polymorphism than high-risk group. Given the significance of tumor mutation burden in forecasting the response to immunotherapy, we could infer there may be a relationship between our signature and response to immunotherapy. We also found that the low-risk group was positively correlated with number of neoantigens, which further demonstrated that our signature could accurately predict the therapeutic effect of immunotherapy. Considering tumor PD-L1 expression was regarded as candidate biomarkers of immunotherapy, we further estimated the relationship of our model and PD-L1 expression. However, our model was not associated with PD-L1 expression. Therefore, more efforts should be made to evaluate this model in gastric cancer patients treated with immune checkpoint inhibitors.

Much progression had made in the tumor immune microenvironment of gastric cancer. Jiang Y et al. had developed the SVM signature based on seven immune related features to predict response to chemotherapy in gastric cancer[35]. Wang H et al. had established the stromal-immune score-based gene signature as a prognosis stratification tool in gastric cancer[36]. Refolo MG et al. had utilized

an immune gene expression signature to divide the gastric cancer patients into different molecular classification[37]. The major strengths of this study were that we developed this signature based on all the known immune cell genes and genes participating in immune related pathways and the ability of predicting response to immune checkpoint inhibitors of this signature. Besides, most genes in our signature had been demonstrated to be key roles in development and progression of gastric cancer, which can further prove that our signature was reliable to predict the prognosis of gastric cancer patients.

Nonetheless, there were several shortcomings of this study should be acknowledged. First, our training cohort only contains 375 cases, which is too little to implement internal validation. Second, our signature was developed based on retrospective data. Thus, large number of gastric cancer samples with clinical information and expression data of 47 signature genes are needed to prove the efficiency of our model. Besides, the biological functions of our signature genes are required to be investigated by experimental studies. Finally, the ability of immunotherapeutic effect prediction of our immune signature is not well estimated due to lacking in patients treated with immune checkpoint blockades.

Conclusions

In summary, this study generates an immune-related signature that can not only predict the prognosis of gastric cancer patients but also forecast the response to immune checkpoint inhibitors. This signature can be utilized to identify the immune status of gastric cancer so that individualized therapeutic schedules will be made to achieve a better outcome for patients.

Declarations

Author Contributions

Jun Gu, Wang Weimin and Song Xiaoling designed the experiments and wrote the manuscript. Dang Wei and Cai Chen conducted the major parts of analyses. Zhang Ruiqiao and Fan Qingquan helped data collection.

Acknowledgments

Firstly, I wish to express my gratitude to all members of Shanghai Key Laboratory of Biliary Tract Disease Research laboratory for their assistance and discussion. Secondly, I would like to thank my families for their continuous support.

Ethics approval and informed consent

No animal or human studies are presented in this manuscript.

Consent for Publication

All the authors have agreed to publish this article in your journal if it is accepted.

Data availability

TCGA RNA-seq datasets were downloaded from the Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/). GSE62254 and GSE15459 were download from the GEO database (http://www.ncbi.nlm.nih.gov/geo).

Funding

Clinical Research Project of Shanghai Municipal Health Committee (201940244)

Competing interests

The authors declare that they have no competing interests for this work.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Clinicians AJJCACJf. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. 2018;
2. Ji X, Bu Z-D, Yan Y, et al. The 8th edition of the American Joint Committee on Cancer tumor-node-metastasis staging system for gastric cancer is superior to the 7th edition: results from a Chinese mono-institutional study of 1663 patients.
3. Han DS, Suh YS, Kong SH, et al. Nomogram Predicting Long-Term Survival After D2 Gastrectomy for Gastric Cancer. 30(31):3834-3840.
4. Wang M, Jingzhou Z, Lishen Z, et al. Role of tumor microenvironment in tumorigenesis. 8(5):761-773.
5. JT W, H L, H Z, et al. Intratumoral IL17-producing cells infiltration correlate with antitumor immune contexture and improved response to adjuvant chemotherapy in gastric cancer. 2019;30(2):266-273. doi:10.1093/annonc/mdy505
6. BW K, AN S, S Y, et al. Prognostic value of tumor-infiltrating lymphocytes in Epstein-Barr virus-associated gastric cancer. 2016;27(3):494-501. doi:10.1093/annonc/mdv610
7. TT W, YL Z, LS P, et al. Tumour-activated neutrophils in gastric cancer foster immune suppression and disease progression through GM-CSF-PD-L1 pathway. 2017;66(11):1900-1911. doi:10.1136/gutjnl-2016-313075
8. ED T, M Z, A M, et al. Patterns of PD-L1 expression and CD8 T cell infiltration in gastric adenocarcinomas and associated immune stroma. 2017;66(5):794-801. doi:10.1136/gutjnl-2015-310839
9. Thorsson V, Gibbs DL, Brown SD, Wolf D, Mariamidze AJI. The Immune Landscape of Cancer. 2018;48(4)
10. Love MI, Huber W, Anders SJGB. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 15(12):550.
11. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. 12(1):35-0.
12. research GOCJNa. The gene ontology (GO) project in 2006. 2006;34(suppl_1):D322-D326.
13. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima KJNar. KEGG: new perspectives on genomes, pathways, diseases and drugs. 2017;45(D1):D353-D361.
14. Yu G, Wang L-G, Han Y, He Q-YJOajoib. clusterProfiler: an R package for comparing biological themes among gene clusters. 2012;16(5):284-287.
15. Harrell FE, LEE KL, Medicine DBMJSi. MULTIVARIABLE PROGNOSTIC MODELS: ISSUES IN DEVELOPING MODELS, EVALUATING ASSUMPTIONS AND ADEQUACY, AND MEASURING AND REDUCING ERRORS. 15(4):361-387.
16. Şenbabaoğlu Y, Gejman RS, Winer AG, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger

RNA signatures. *Genome Biol*. 2016;17(1):231.

17. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7. doi:10.1186/1471-2105-14-7
18. A M, DC L, Y A, C P, research KHJG. Maftools: efficient and comprehensive analysis of somatic variants in cancer. 2018;28(11):1747-1756. doi:10.1101/gr.239244.118
19. MS R, SA S, CJ W, G G, Cell HNJ. Molecular and genetic properties of tumors associated with local immune cytolytic activity. 2015;160:48-61. doi:10.1016/j.cell.2014.12.033
20. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348(6230):124-128. doi:10.1126/science.aaa1348
21. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol*. 2016;17(12):e542-e551. doi:10.1016/S1470-2045(16)30406-5
22. R S, A Q, ALK T, et al. Transcriptional analysis of immune genes in Epstein-Barr virus-associated gastric cancer and association with clinical outcomes. 2018;21(6):1064-1070. doi:10.1007/s10120-018-0851-9
23. C L, H H, H L, et al. Tumour-associated macrophages-derived CXCL8 determines immune evasion through autonomous PD-L1 expression in gastric cancer. 2019;68(10):1764-1773. doi:10.1136/gutjnl-2018-316324
24. SY K, C P, HJ K, et al. Deregulation of immune response genes in patients with Epstein-Barr virus-associated gastric cancer and outcomes. 2015;148(1):137-147.e9. doi:10.1053/j.gastro.2014.09.020
25. Shi Y, Wang J, Xin Z, Duan Z, Wang G, Li FJPo. Transcription factors and microRNA-co-regulated genes in gastric cancer invasion in ex vivo. 2015;10(4)
26. Wang X-Q, Tang Z-X, Yu D, et al. Epithelial but not stromal expression of collagen alpha-1 (III) is a diagnostic and prognostic indicator of colorectal carcinoma. 2016;7(8):8823.
27. Zhuo C, Li X, Zhuang H, et al. Elevated THBS2, COL1A2, and SPP1 expression levels as predictors of gastric cancer prognosis. 2016;40(6):1316-1324.
28. Huang R, Gu W, Sun B, Gao LJMmr. Identification of COL4A1 as a potential gene conferring trastuzumab resistance in gastric cancer based on bioinformatics analysis. 2018;17(5):6387-6396.
29. Wang C-S, Wu T-L, Tsao K-C, Sun C-F. Serum TIMP-1 in gastric cancer patients: a potential prognostic biomarker. *Ann Clin Lab Sci*. 2006;36(1):23-30.
30. Wang J-H, Huang W-S, Hu C-R, Guan X-X, Zhou H-B, Chen L-BJWjogW. Relationship between RGS5 expression and differentiation and angiogenesis of gastric carcinoma. 2010;16(44):5642.
31. Lv Y-p, Peng L-s, Wang Q-h, et al. Degranulation of mast cells induced by gastric cancer-derived adrenomedullin prompts gastric cancer progression. 2018;9(10):1-12.
32. Hanahan D, Weinberg RAJc. Hallmarks of cancer: the next generation. 2011;144(5):646-674.
33. Kitadai YJJoo. Angiogenesis and lymphangiogenesis of gastric cancer. 2010;2010

34. Amedei A, Della Bella C, Silvestri E, Prisco D, D'Elios MMJC, Immunology D. T cells in gastric cancer: friends or foes. 2012;2012
35. Y J, J X, W H, et al. Tumor Immune Microenvironment and Chemosensitivity Signature for Predicting Response to Chemotherapy in Gastric Cancer. 2019;7(12):2065-2073. doi:10.1158/2326-6066.Cir-19-0311
36. H W, X W, oncology CYJFi. Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer. 2019;9:1212. doi:10.3389/fonc.2019.01212
37. MG R, C L, C M, MG C, biology DARJJol. Integrated immune gene expression signature and molecular classification in gastric cancer: New insights. 2020;doi:10.1002/jlb.4mr0120-221r

Tables

Features	TCGA (n,%)	GSE62254 (n,%)	GSE15459 (n,%)
Platform	Illumina Hiseq2000 RNA sequencing platform	Affymetrix Human Genome U133 Plus 2.0 Array	Affymetrix Human Genome U133 Plus 2.0 Array
Age (years)	65.44±10.82	63.00±11.55	64.37±13.23
Gender			
Male	186 (60.8%)	169 (66.0%)	125 (65.1%)
Female	120 (39.2%)	87 (34.0%)	67 (34.9%)
T stage			
T1	16 (5.2%)	NA	NA
T2	67 (21.9%)	NA	NA
T3	145 (47.4%)	NA	NA
T4	78 (25.5%)	NA	NA
N stage			
N0	98 (32.0%)	31 (12.1%)	NA
N1-3	208 (68.0%)	225 (87.9)	NA
M stage			
M0	283 (92.5%)	NA	NA
M1	23 (7.5%)	NA	NA
AJCC stage			
Stage I	45 (14.7%)	24 (9.4%)	31 (16.1%)
Stage II	97 (31.7%)	75 (29.3%)	29 (15.1%)
Stage III	129 (42.2%)	85 (33.2%)	72 (37.5%)
Stage IV	35 (11.4%)	72 (28.1%)	60 (31.3%)
Recurrence			
Yes	NA	116 (45.3%)	NA
No	NA	140 (54.7%)	NA
MLH1_IHC			
Positive	NA	201 (78.5%)	NA
Negative	NA	55 (21.5%)	NA
EBV_ISH			
Positive	NA	18 (7.0%)	NA
Negative	NA	238 (93.0%)	NA
Survival status			
Alive	182 (59.5%)	125 (48.8%)	97 (50.5%)
Dead	124 (40.5)	131 (51.2%)	95 (49.5%)

Note: age presented as mean±SD, *TCGA* The Cancer Genome Atlas

Figures

Figure 1

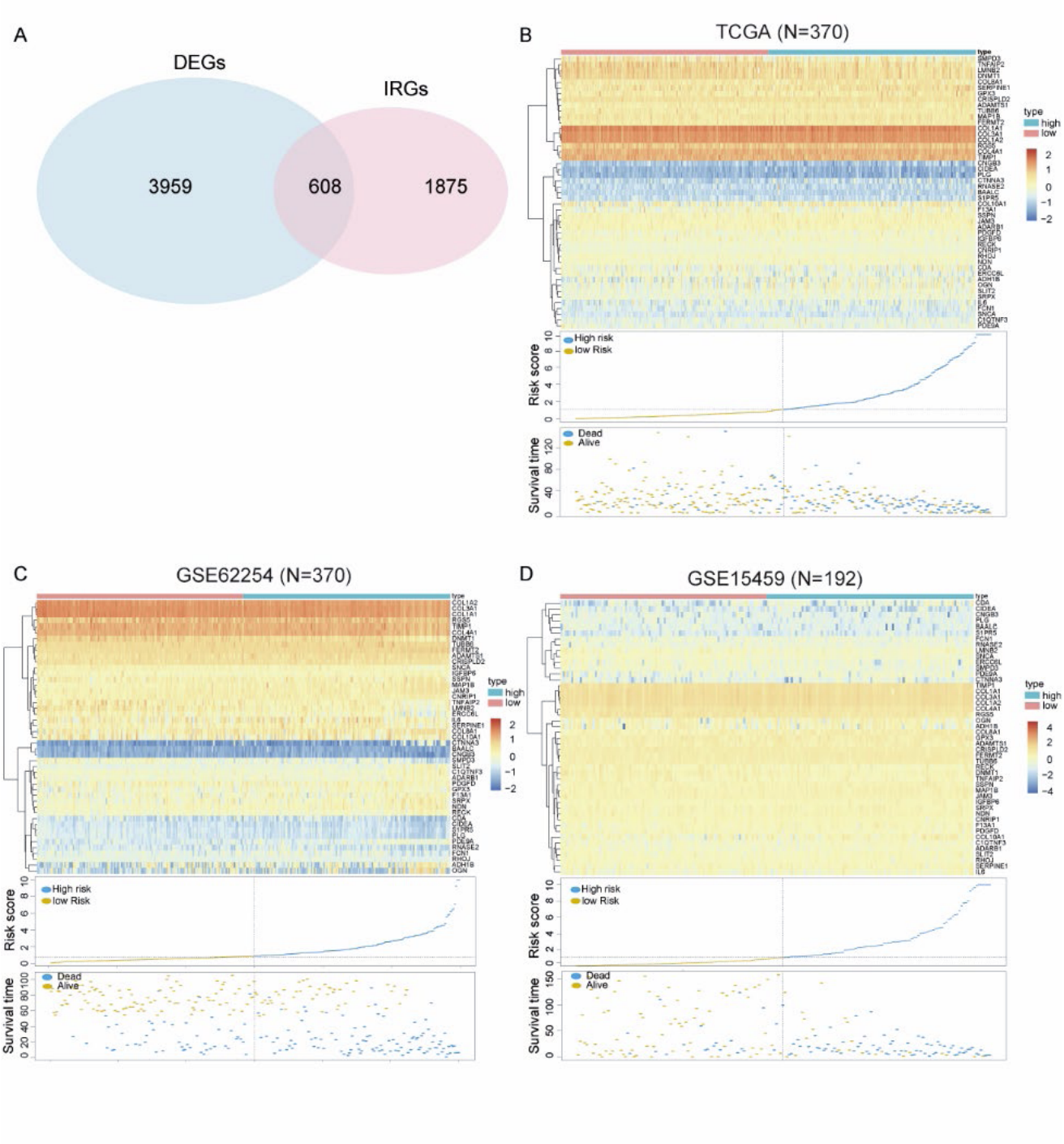

Figure 1

Overview of the immune-related signature. (A) The Venn diagram presented the differentially expressed immune-related genes in the TCGA dataset. (B-D) Heatmap of the signature consisting of 47 immune-related genes in the TCGA dataset, GSE62254 dataset, and GSE15459 dataset. The median risk score divided patients into the low-risk group and the high-risk group

Figure 2

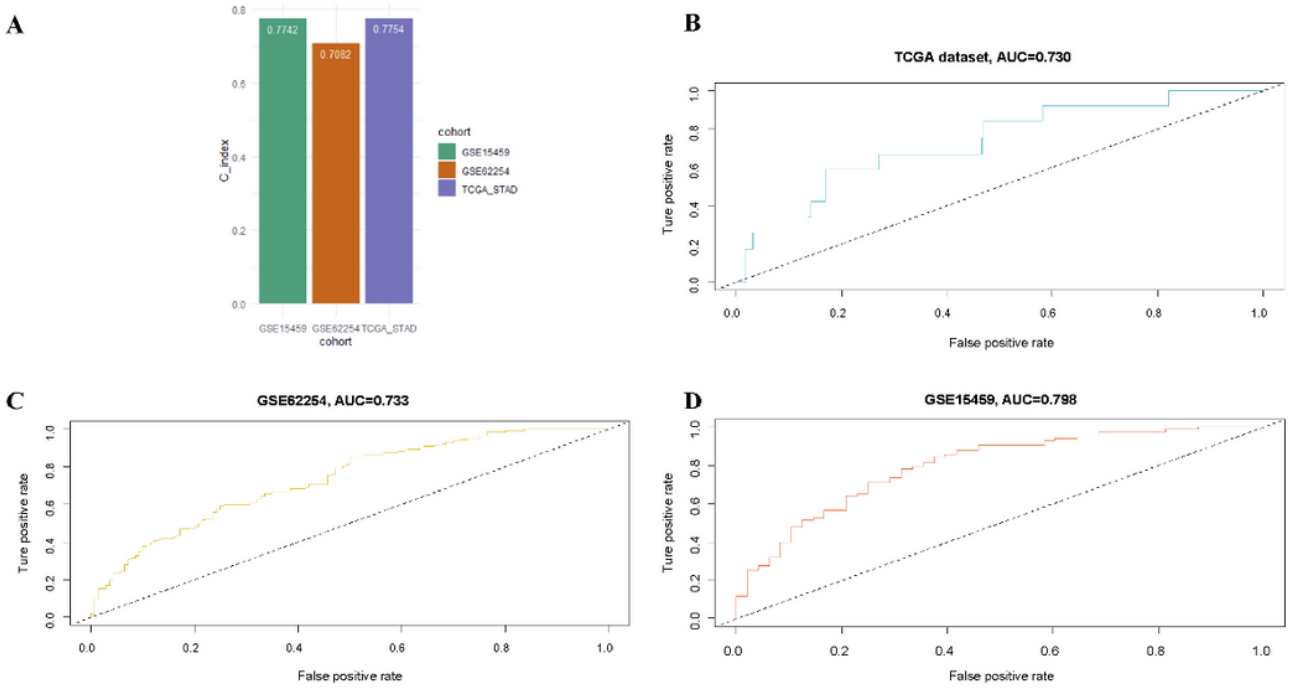

Figure 2

Validation of the immune-related signature. (A) The c-index of both training and external testing sets. The c-index for the TCGA dataset, GSE62254, and GSE15459 were 0.7754, 0.7082, and 0.7742, respectively. (B-D) ROC curves of three cohorts showing the excellent capability of the prediction model for overall survival rates of patients with gastric cancer. The area under the curve of the TCGA dataset, GSE62254, and GSE15459 were 0.730, 0.733, and 0.798, respectively

Figure 3

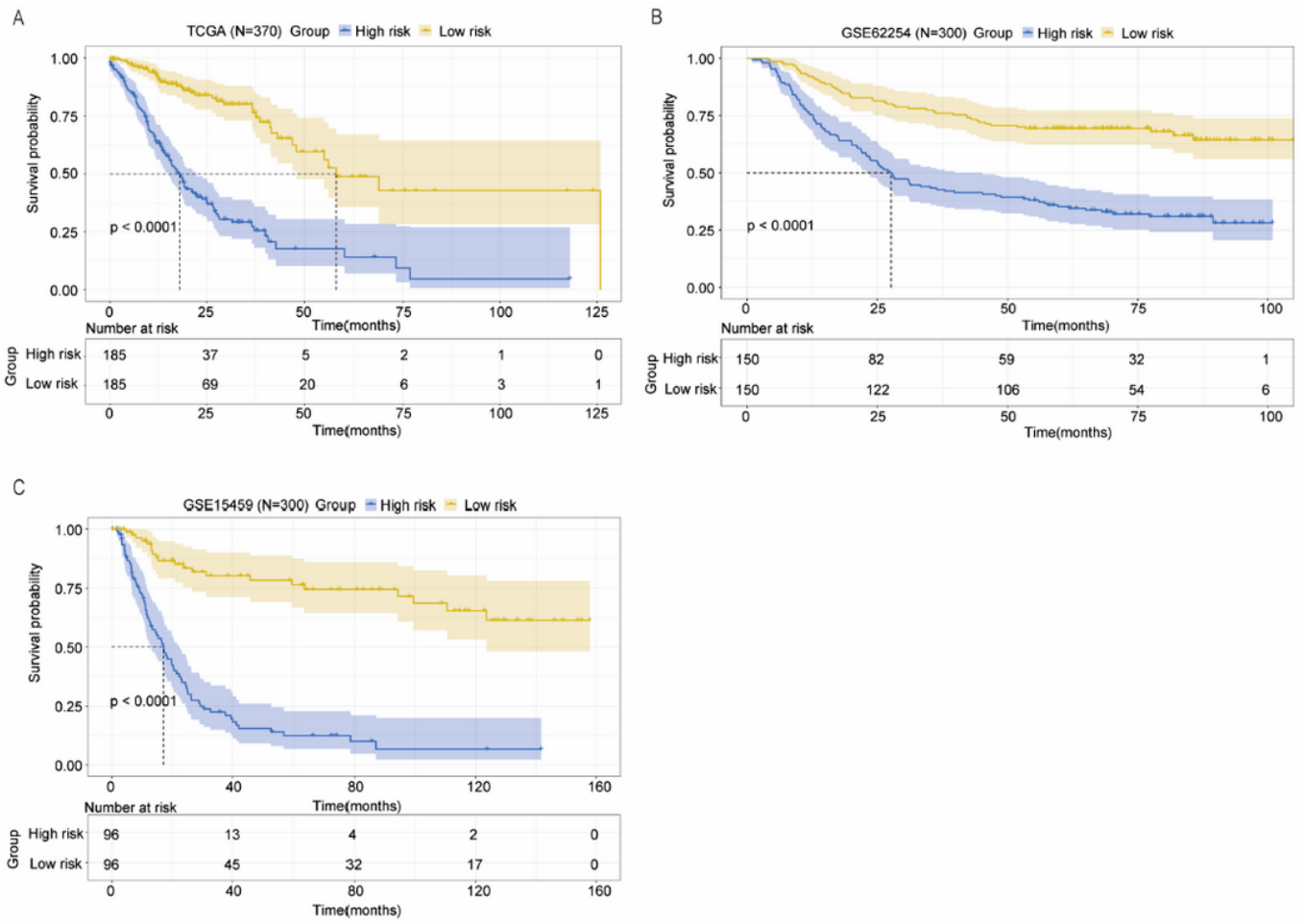

Figure 3

The survival analysis of the signature for three datasets. The high-risk group demonstrated poor overall survival than the low-risk group in the TCGA dataset, GSE62254, and GSE15459 (P<0.0001)

Figure 4

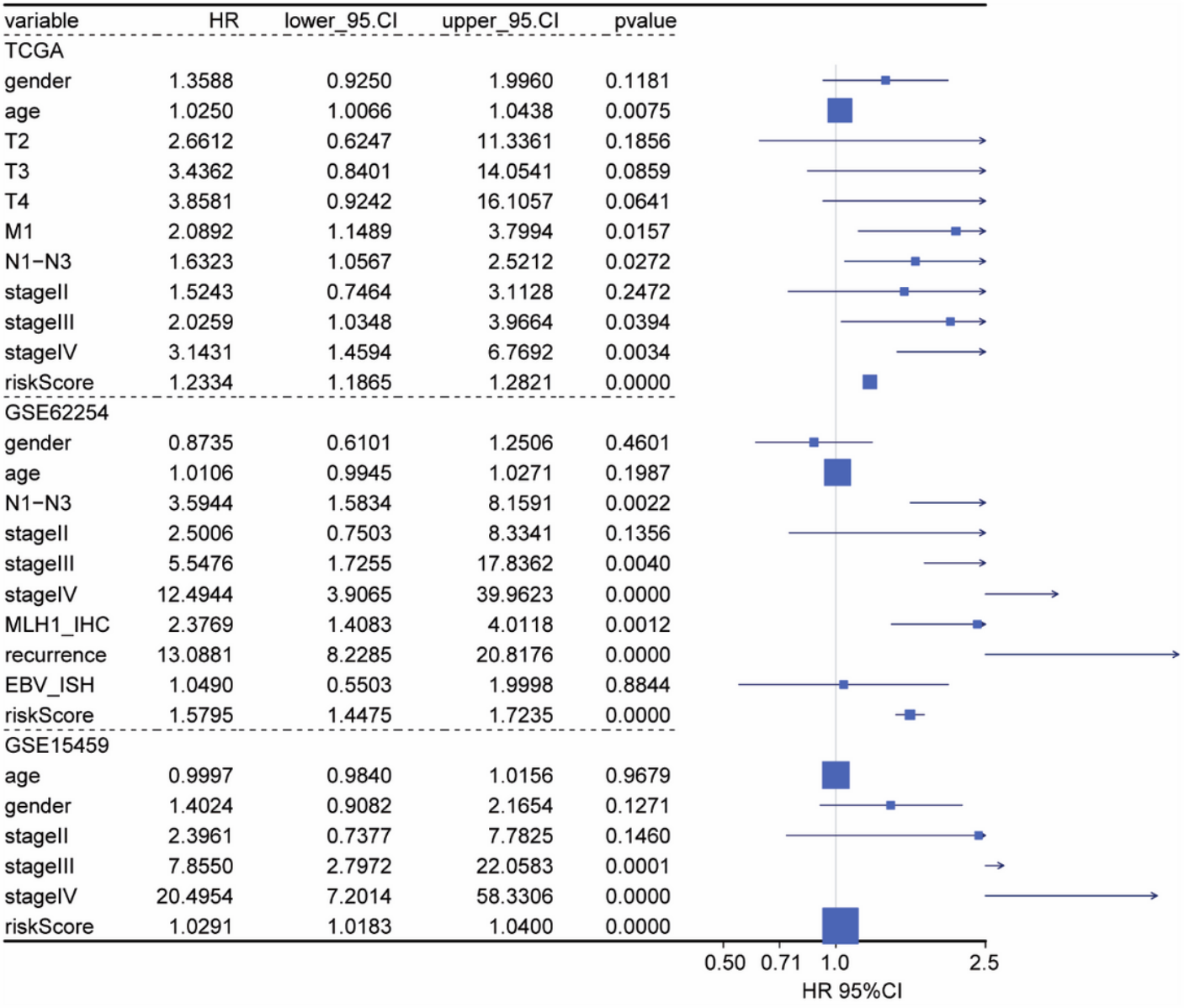

Figure 4

The Univariate Cox regression analysis of clinical features and the signature for training and testing sets. The HR of the signature in the TCGA cohort was 1.2334, with 95% CI from 1.1865 to 1.2821 (P<0.0001). The HR of the signature in the GSE62254 cohort was 1.5795, with 95% CI from 1.4475 to 1.7235 (P<0.0001). The HR of the signature in the GSE62254 cohort was 1.5795, with 95% CI from 1.4475 to 1.7235 (P<0.0001). TCGA The Cancer Genome Atlas, HR hazard ratio, 95% CI 95% confidence interval

Figure 5

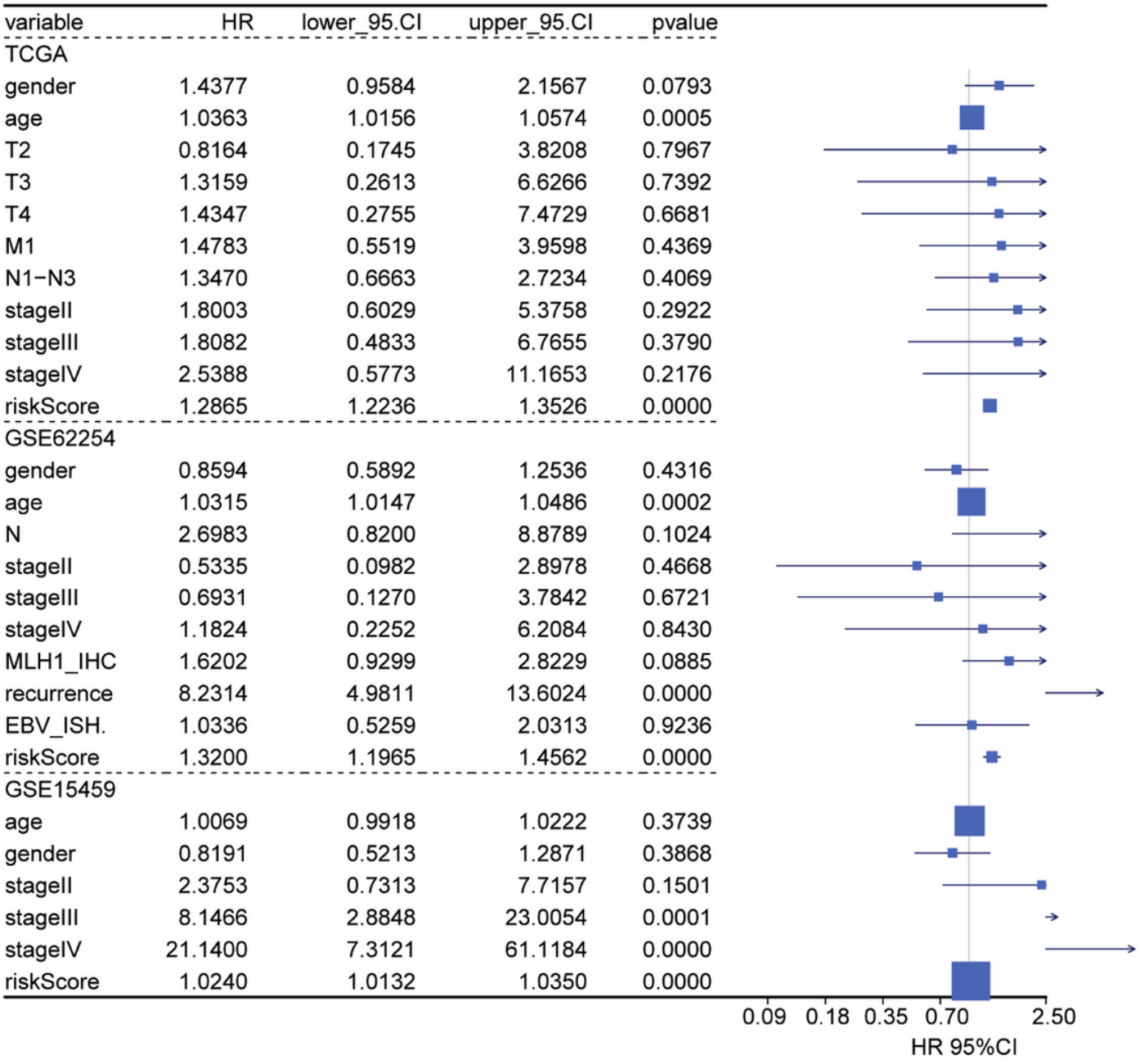

Figure 5

Multivariate Cox analysis revealing independently predictive capability of the signature for overall survival. The signature was able to independently predict patients' outcome in the TCGA cohort (HR=1.2865, 95% CI 1.2236-1.3526, P<0.0001), the GSE62254 cohort (HR=1.3200, 95% CI 1.1965-1.4562, P<0.0001), and the GSE15459 cohort (HR=1.0240, 95% CI 1.0132-1.0350, P<0.0001)

Figure 6

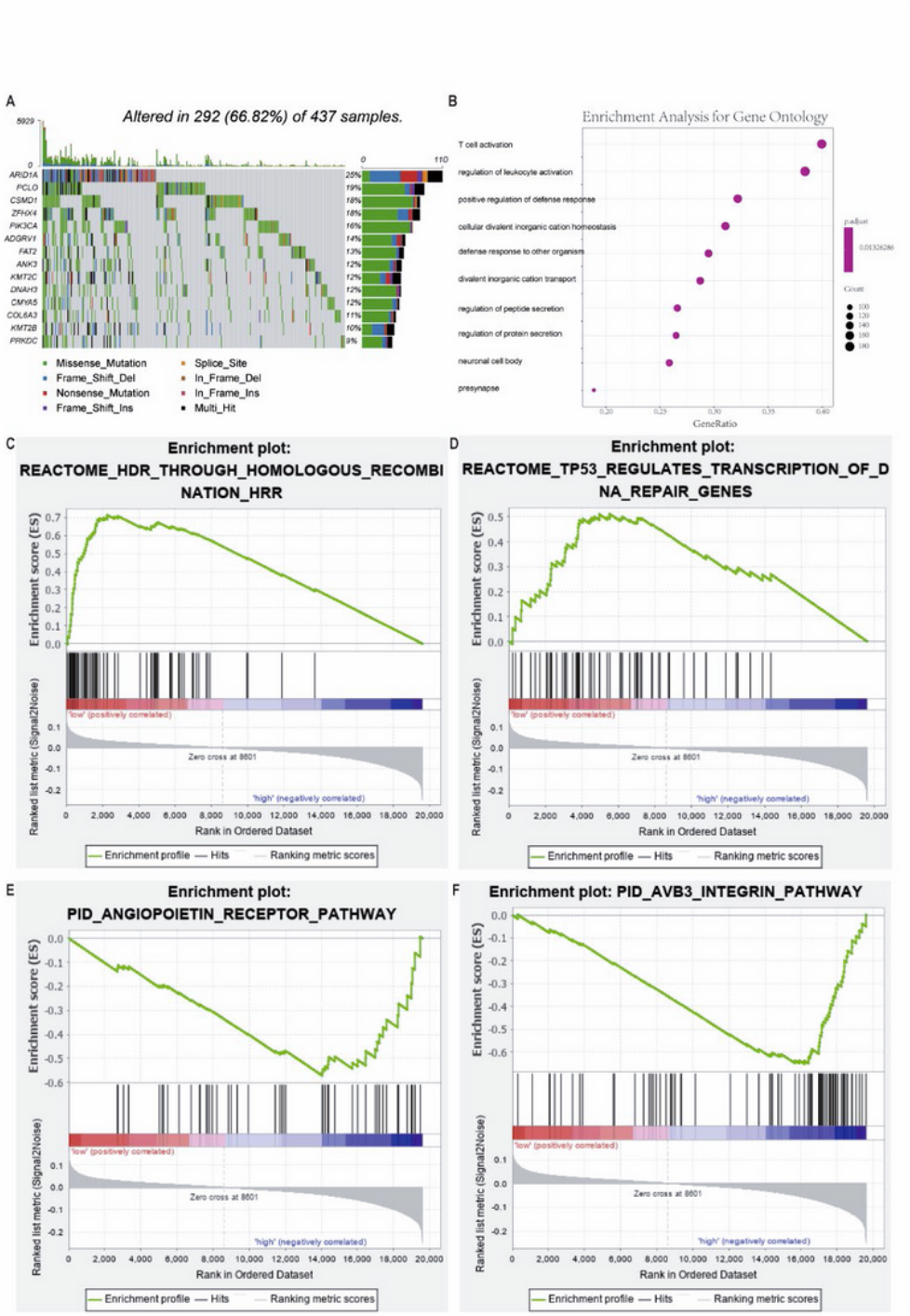

Figure 6

Genomic correlations with the signature in TCGA gastric cancer. (A) The oncoplot of 14 mutation genes which were associated with the signature. (b) The dot plot presented the result of enrichment analysis for Gene Ontology. (C-F) GSEA revealed that the most significant pathways correlated with the immune-related signature

Figure 7

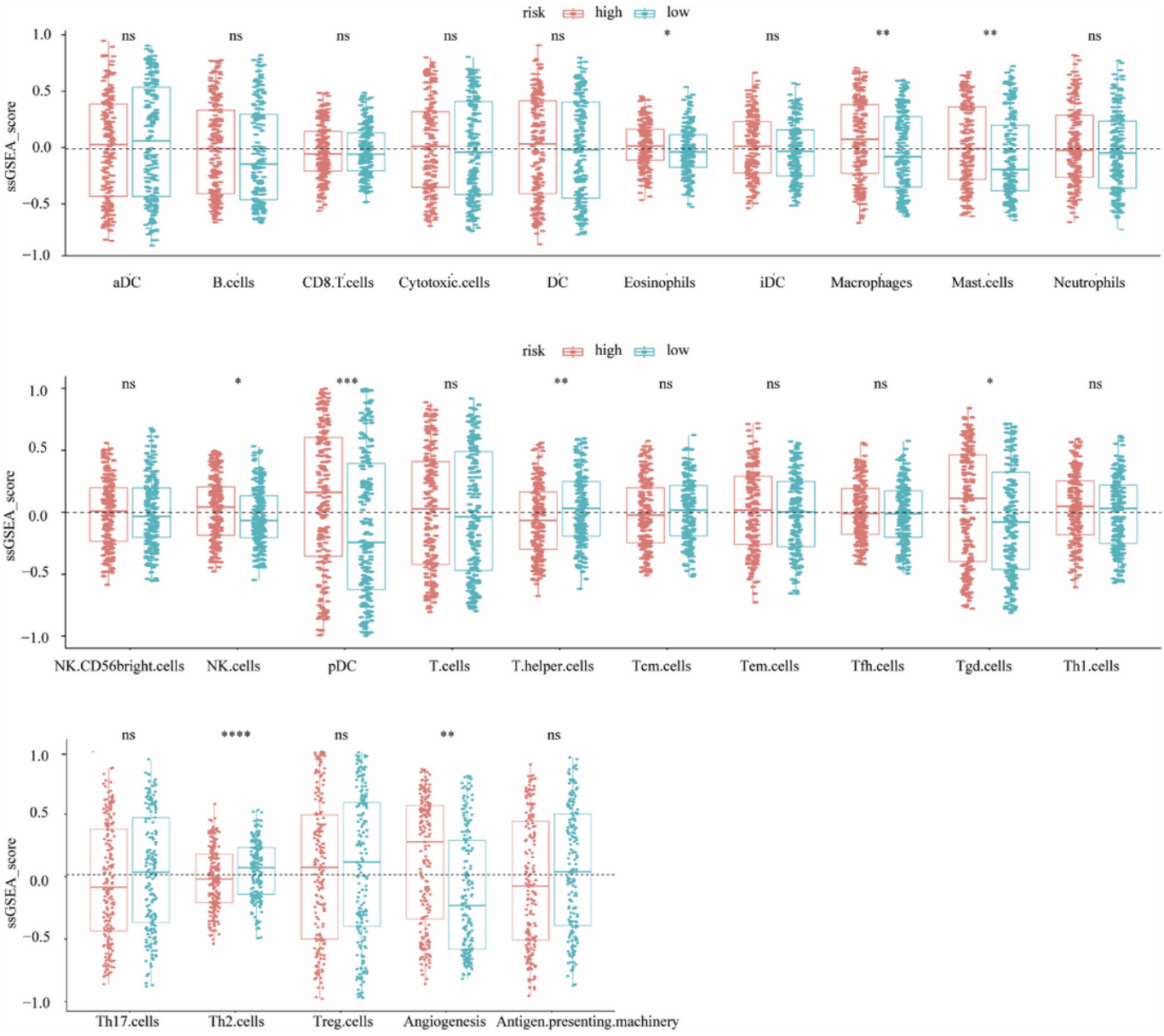

Figure 7

Single sample gene sets enrichment analysis revealed the infiltration level of 25 immune cell types between the high-risk group and the low-risk group

Figure 8

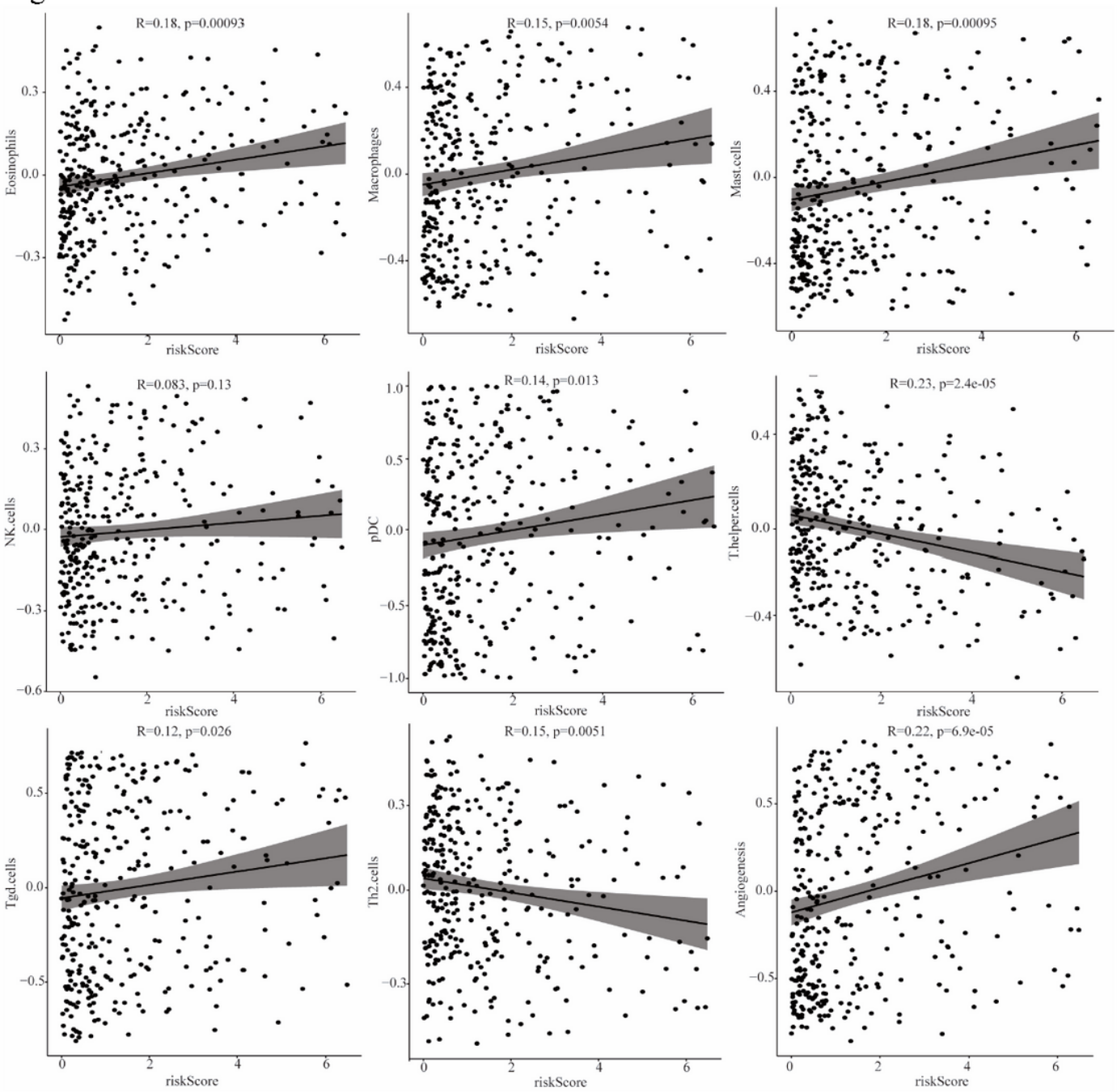

Figure 8

Relationship between the risk score and infiltration abundances of 25 types of immune cells

Figure 9

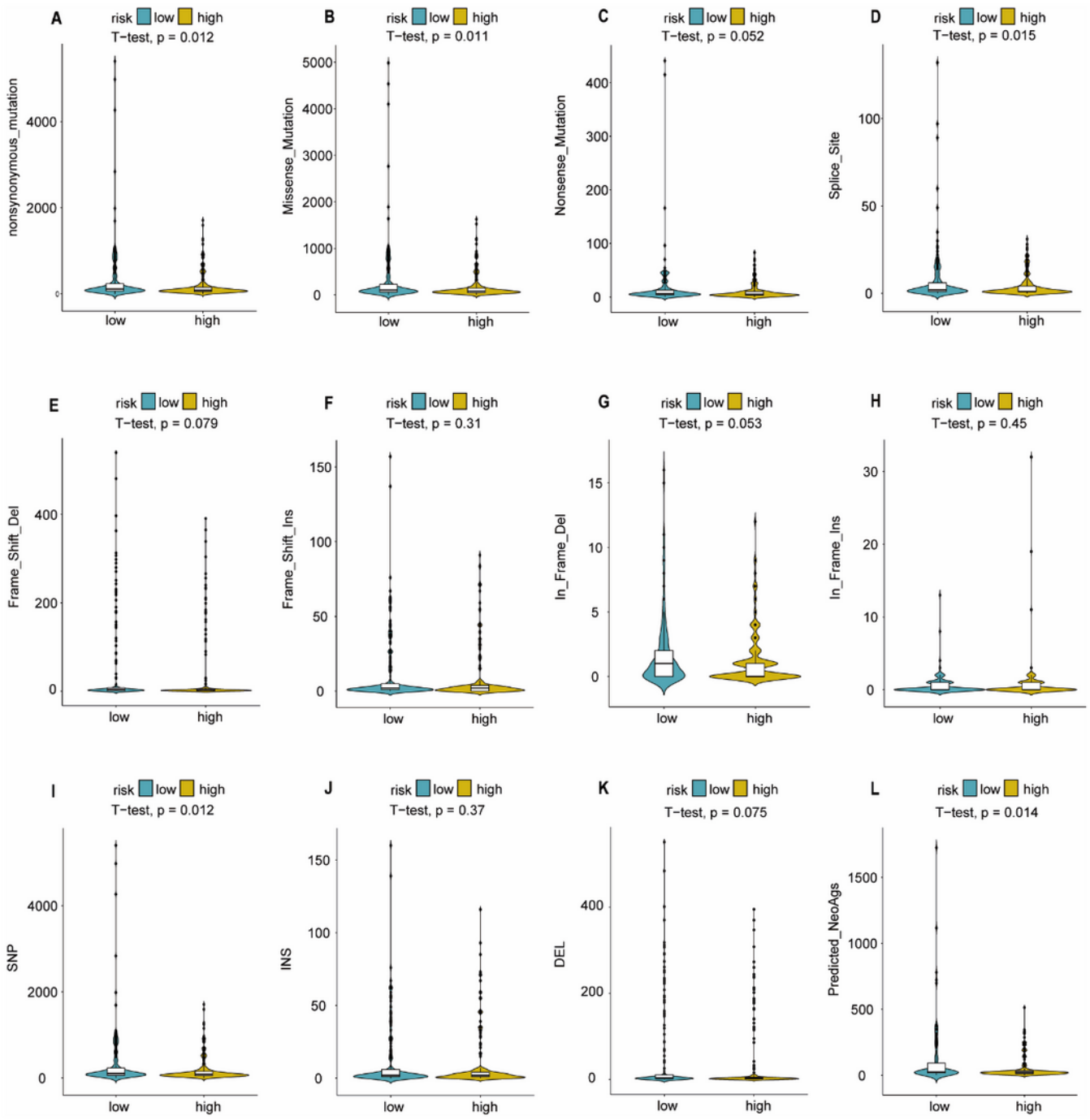

Figure 9

The relationship between the signature and nonsynonymous mutation burden, different nonsynonymous mutation types, and neoantigen. (A) The low-risk group had a higher nonsynonymous mutation burden (P=0.012). (B) The low-risk group was associated with a higher number of missense mutations (P=0.011). (C) There was no relationship between the risk score and the number of nonsense mutations. (D) The low-risk group was associated with a higher number of splice site mutations (P=0.011). (E) There

was no relationship between the risk score and the number of frameshift deletion. (F) There was no relationship between the risk score and the number of frameshift insertions. (G) There was no relationship between the risk score and the number of inframe deletion. (H) There was no relationship between the risk score and the number of inframe insertion. (I) The low-risk group was associated with a higher number of SNP (P=0.012). (J) There was no relationship between the risk score and the total number of insertion mutations. (K) There was no relationship between the risk score and the total number of deletion mutations. (L) The low-risk group was associated with a higher number of neoantigens (P=0.012). DEL deletion, INS insertion, SNP single nucleotide polymorphism, NeoAgs neoantigens

Supplementary Files

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

- TableS1.docx