Identification of Immune Infiltration-related Gene Signature Associated with Prognosis and Immune Features in Stomach Adenocarcinoma

Ya Yang (yangya426@126.com)
The First Affiliated Hospital of Zhengzhou University https://orcid.org/0000-0002-0394-1979
Xintan Zhang
The First Affiliated Hospital of Zhengzhou University
Tingxuan Li
The First Affiliated Hospital of Zhengzhou University
Yue Zhang
The First Affiliated Hospital of Zhengzhou University
Xiaoxiao Zuo
The First Affiliated Hospital of Zhengzhou University

Primary research

Keywords: Stomach adenocarcinoma, bioinformatics, Immune infiltrated genes, prognosis

Posted Date: November 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1057544/v1

Abstract

Background: Immune infiltrated genes (IIGs) have been identified to associated with the prognosis of various cancers, but their expression and prognostic significance remain largely unclear in stomach adenocarcinoma (STAD).

Methods: Gene expression profiles and clinical data of STAD patients were downloaded from The Cancer Genome Atlas (TCGA) as a training dataset (n = 375) and Gene Expression Omnibus (GEO) databases as a validation dataset (n = 300). Construction of high and low immune cell infiltration groups was performed by single sample gene set enrichment analysis (ssGSEA) and evaluated by ESTIMATE algorithm-derived immune scores. The overlapping differentially expressed genes (DEGs) in tumor *vs.* normal and Immunity-H *vs.* Immunity-L were selected as differentially expressed immune infiltrated genes (DEIIGs), which were used to construct DEIIG prognostic signature and its performance was validated using validation dataset. Moreover, the association between clinical data and immune features were explored. Furthermore, ADH4 and ANGPT2 were selected for analyzing their expression and prognostic values in STAD patients.

Results: A total of 191 overlapping DEGs, including 6 lnRNAs and 185 mRNA were identified. Consecutively, 9 DEIIG prognostic signature (LINC00843, ADH4, ANGPT2, APOA1, ASLC2, GFRA1, KIAA1549L, MTTP and PROC) were identified as risk signature and Kaplan-Meier curve and ROC curve verified its performance in TCGA and GEO datasets. Total five clinical outcomes (age, pathologic T, radiotherapy, tumor recurrence and prognostic score model status) were identified to be associated with the survival prognosis of STAD patients. The TIMER algorithm revealed that B cell, T cell CD4+, neutrophil, macrophage and myeloid dendritic cell were positively correlated with STAD prognosis, while CD8+ was negatively correlated with STAD prognosis. Additionally, we validated that higher ADH4 and lower ANGPT2 predicted better survival prognosis in STAD patients.

Conclusion: We constructed and verified a robust signature of nine DEIIG prognostic signature for the prediction of STAD patient survival.

Background

Stomach adenocarcinoma (STAD), as the most common pathology type of gastric cancer, is the fifth prevalent malignancies and leading causes among all malignancies, with estimated more than 100 million new cases and nearly 80 million deaths in 2018 [1, 2]. Despite the clinical outcomes have been improved under surgical therapies, chemotherapy and systemic treatments for STAD patients at early stage, most of 50% patients are identified as advanced stage, thereby causing less than 30% five-year survival rate [3–5]. Thus, it is great of importance to identify novel prognostic biomarkers and therapeutic targets for STAT patients.

As a well-recognized heterogeneous cancer, STAD is not only composed of cancer cells, but also by non-cancer cells, including endothelial cells, macrophages, stromal and immune cells [6]. Among these non-

cancer cells, the tumor infiltrating immune cells (TIICs) have been reported to be closely associated with the clinical outcomes and response to immunotherapy for their crucial roles in pro- and anti-tumorigenic processes [7]. Increasing evidence indicates that immune cell infiltration plays a vital role in the prognosis of cancer, including breast cancer [8], colon cancer [9] and bladder cancer [10]. For instance, tumor-associated lymphocytes (TALs), primarily T cells, can regulate the proliferation and migration of cancer cells by releasing soluble cytokines, as well as participate in activating angiogenesis and host defense mechanism [11]. The infiltration of tumor-associated macrophages (TAMs) into tumor tissue has been reported to be significantly correlated with tumor vascularity, the depth of tumor invasion, lymph node status and clinical stages [12, 13]. As one of the primary effector cells of anticancer immunity, CD8+ T cells is identified as a potential prognostic indicator of gastric cancer [14]. Recent studies indicated that the alter gene expression levels exerts anti-tumor effects through regulating an immune suppression mechanism in TIICs and is correlated with favorable prognosis as follows: Chaudhary et al. [15] reported for the first time that neuropilin 1 (NRP1) is upregulated on tumor-infiltrating lymphocytes (TILs) and can be induced on peripheral blood mononuclear cells (PBMCs) from colorectal cancer liver metastases. Wang et al. [16] previously identified that SUPV3L1 and SLC22A17 as hub genes affect immune cell infiltration, result in the different prognosis of gastric cancer. In addition, immune infiltration revealed a significant correlation between JAK3/TYK2 expression and the abundance of immune cells as well as immune biomarker expression in STAD [17]. Nevertheless, the association among gene expression levels, tumor infiltrating immune and survival prognosis remains largely unclear.

In the present study, we evaluated the immune cell infiltration in STAD tumor samples obtained from TCGA database based on single sample gene set enrichment analysis (ssGSEA) algorithm and distinguished the high immune infiltration group from the low infiltration group. On the basis of immune grouping, Cox regression analysis and LASSO algorithm were combined to screen the prognostic marker RNAs factors of STAD, and the survival prediction model was constructed and verified based on the prognostic marker RNAs.

Materials And Methods

Data acquisition

Gene expression profiles (lncRNAs and mRNAs) and corresponding clinical information from primary STAD tumors, uploaded up to the 20th October 2020, were obtained from The Cancer Genome Atlas (TCGA: https://cancergenome.nih.gov/) and Gene Expression Omnibus (GEO: https://www.ncbi.nlm.nih.gov/geo/). For TCGA datasets, total 407 samples containing 375 STAD samples and 32 normal samples were selected as training group. With the sample screening criteria (clinical follow-up information was retained and included samples at least 200), the gene expression assay GSE62254 (GP570, Affymetrix Human Genome U133 Plus 2.0 Array) as external validation dataset, including 300 STAD samples and their corresponding clinical information was retrieved from GEO database. Overlapping lncRNAs and mRNAs from these two datasets were selected for further analysis.

The overall study design and the different samples that were included at every stage of the analysis were illustrated as a flowchart in Figure 1.

Single-sample immune infiltration level analysis

The immune cell infiltration levels of STAD tumor samples were quantified by single sample gene set enrichment analysis (ssGSEA) in R3.6.1 package Gene Set Variation Analysis (GSVA) Version 1.36.3 (http://www.bioconductor.org/packages/release/bioc/html/GSVA.html) [18]. The ccGSEA employed gene signatures expressed by immune cell populations to individual tumor samples. Subsequently, 375 STAD samples were divided into high immunity infiltration (Immunity-H) and low immunity infiltration (Immunity-L) groups according to the results from ccGSEA data. Moreover, the reasonableness of immune infiltration grouping was validated using ESTIMATE method [19] and CIBERSORT algorithm [20].

Identification of differentially expressed immune infiltrated genes (DEIIGs)

The samples of TCGA dataset were divided into two groups according to sample source (tumor *vs.* normal) and obtained immunity group (Immunity-H *vs.* Immunity-L). Differentially expressed genes (DEGs) between tumor and normal or between Immunity-H and Immunity-L groups were identified using limma package of R3.6.1 Version 3.34.7 [21] with the cut-off value of FDR (false discovery rate) < 0.05 and log2 |fold change (FC)| > 1. These DEGs were visualized in a volcano plot in R. The overlapping DEGs in tumor *vs.* normal and Immunity-H *vs.* Immunity-L were selected as differentially expressed immune infiltrated genes (DEIIGs), which were visualized using Venn diagram.

Functional enrichment analysis of DEIIGs

Then, DEIIGs were analyzed by the Database for Annotation, Visualization and Integrated Discovery (DAVID) Version 6.8 bioinformatics tool (https://david.ncifcrf.gov/) [22, 23]. Gene Ontology (GO) biology process and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were then performed to annotate the potential functions for DEIIGs with the cut-off value of FDR < 0.05.

Construction of the DEIIG prognostic signature

Univariate Cox analysis in R3.6.1 survival package [24] was used to determine the association between the expression level of DEIIGs and patient's overall survival (OS) with the threshold of log-rank *p* value < 0.05. After filtration of prognostic DEIIGs, independent prognosis related DEIIGs were further screened via a multivariate Cox regression model with *p*-value < 0.05 as the cut-off criterion. Lasso-penalized Cox regression analysis [25] was performed to further reduce the number of independent prognosis related

DEIIGs with the optimal lambda using 1000-time cross-validation likelihood based on penalized package Version 0.9.50 [26]. According to the best lambda value, a prognostic gene signature of STAD patients was constructed with the following formula: Prognostic score (PS) = $\sum \beta DEIIGs \times Exp\ DEIIGs$. Here, β_{DEIIGs} represent the regression coefficients (β) derived from the Lasso Cox regression model and Exp_{DEIIGs} represent the expression levels of signature DEIIGs in training dataset.

Evaluation of the DEIIG prognostic signature

Taking the median PS as the cutoff point, we divided the samples in training dataset into high-risk group (PS > median value) and low-risk group (PS < median value). Kaplan–Meier (KM) survival curves analysis was used to analyze the OS between the high-risk and low-risk groups. The accuracy and sensitivity of survival prediction based on the PS were verified by receiver operating characteristic (ROC) curve analysis and determined by the value of area under the curve (AUC). Meanwhile, the expression levels of signature DEIIGs in validation dataset GSE62254 were extracted and the PS was calculated according to the formula as above. Similarly, Kaplan-Meier curve and ROC curve analysis were performed to evaluate the predictive ability of the signature.

Identification of independent prognostic parameters of STAD

Next, univariate and multivariate Cox regression analyses Survival package (Version2.41–1, http://bioconductor.org/packages/survivalr/) [24] were performed in the TCGA dataset on the DEIIG prognostic signature and clinicopathological parameters including age, gender, pathologic-M, pathologic-N, pathologic-T, pathologic-stage, neoplasm histologic grade, radiation therapy and PS model status. Log-rank p value < 0.05 was considered statistically significant. Parameters with log-rank p value < 0.05 based on the univariate analysis were further included in the multivariate Cox regression analysis to obtain independent prognostic parameters. Visual presentation of independent prognostic parameters was performed with forestplot Version 1.10 in R3.6.1 language [27]. Subsequently, we constructed the clinical prognosis models based on these independent prognostic parameters alone, which were compared with PS prognostic model by drawing ROC curves with the quantitative indicator AUROC (0.5-1) [28].

Correlation of PS with tumor-infiltrating immune cells (TICs)'s proportion

Total six types of TICs, including B cell, T cell CD4+, T cell CD8+, neutrophil, macrophage and myeloid dendritic cell were retrieved from Tumor Immune Estimation Resource (TIMER: https://cistrome.shinyapps.io/timer/) [29] as a web server for comprehensive analysis of TICs. These six kinds of TICs abundance distribution from 351 tumor samples in the training cohort was estimated by

CIBERSORT calculation method. The correlation between the TICs' proportion and PS was calculated using Spearman coefficient test.

Patients and specimens

Total 59 paired tumor tissues and matched adjacent tissues were collected from STAD patients from July 2017 to October 2020. All participating patients gave their written informed consent and did not receive adjuvant chemotherapy or radiotherapy prior to surgery. Basic clinical information, including sex, age, tumor size and lymph node metastasis. Follow-up information for all participants was obtained every three months by telephone or via a postal questionnaire. During the follow-up period, overall survival was measured from diagnosis to death or to the last follow-up (at five years). This study was approved by the Ethical Committee of The First Affiliated Hospital of Zhengzhou University (Henan, China).

RNA extraction and quantitative real time PCR

Total RNA sample was extracted from tissue specimens using TRIzol reagent (Thermo Fisher Scientific, Waltham, MA, USA) and reverse transcription of mRNA was performed with PrimerScript RT reagent Kit (Takara Biotechnology Co. Ltd., Dalian, China) according to the manufactures' instructions. SYBR Green quantitative PCR reaction was carried out in triplicate in a 20 μL reaction volume containing 2× PCR Master Mix (Applied Biosystems) with the cycling conditions as follows: 95 minutes or five minutes, followed by 40 cycles of 95 °C for 20 s and 60 °C for 30 s. Relative expression levels of ADH4 and ANGPT2 mRNA were calculated by the 2−ΔΔCq method [30].

Statistical analysis

Statistical analysis was performed using SPSS 22.0 (SPSS Inc.; Chicago, IL, USA). Continuous variables were analyzed using Student's t-test for paired samples. The association between gene expression levels and categorical variables were analyzed by the chi-square test. The relationship between gene expression levels and overall survival was analyzed through the Kaplan–Meier method, which was evaluated by the log-rank test. The univariate regression model was used to analyze the effects of individual variables on survival, and the multivariate Cox regression model was used to confirm the independent impact factors associated with survival. A p value < 0.05 was accepted as statistically significant.

Results

Groups and evaluation of tumor-infiltrating immune

The immune cell infiltration status was assessed by applying the ssGSEA approach to each tumor sample of TCGA STAD cohort. As shown in Figure 2, 375 tumor samples were distinctly divided into two

clusters, including Immunity-H (n = 192) and Immunity-L (n = 183) groups based on the landscape of 28 immune cell subpopulations infiltrations in STAD. Detailed results from ssGSEA were presented in Table S1. Next, we calculated the stromal score, immune score and estimate score using ESTIMATE method (Table S2). As depicted in Figure 3A, there were significant differences in stromal score, immune score and estimate score between Immunity-H and Immunity-L groups, of which corresponding scores in Immunity-H group were notably higher than those in Immunity-L group. Moreover, the results from CIBERSORT algorithm on immune cell type showed that the fraction of some important immune cell subtypes varied distinctly between Immunity-H and Immunity-L groups (Figure 3B), which were summarized in Table S3. Collectively, the Immunity-H and L groupings obtained based on ssGSEA evaluation can be used for subsequent analysis.

Identification of DEIIGs

We screened the DEGs between tumor and normal samples, and Immunity-H and Immunity-L samples in TCGA dataset. The volcano plots were drawn to visualize DEGs between tumor and normal group, and Immunity-H and Immunity-L group (Figure 4A). A total of 894 DEGs between tumor and normal group, and 592 DEGs between Immunity-H and Immunity-L group were screened, which were listed in Table S4. Moreover, the number of overlapping DEIIGs was 191, including 6 lncRNAs and 185 mRNAs (Figure 4B, Table S5).

GO function and KEGG pathway analyses

GO function and KEGG pathway enrichment analyses were performed for 185 DEIIGs. These DEIIGs were significantly enriched in 11 biological processes, such as retinoid metabolic process, cell-cell signaling, collagen catabolic process, potassium ion transport and potassium ion transmembrane transport (Table 1, Figure 5). The KEGG pathway analyses showed that the DEIIGs were mainly concentrated in transcriptional misregulation in cancer, vitamin digestion and absorption, fat digestion and absorption, protein digestion and absorption, and gastric acid secretion (Table 1, Figure 5).

Construction of the DEIIG prognostic signature

Univariate Cox regression analysis was performed for the 191 DEIIGs, of which 32 DEIIGs showed significant prognostic potential (log-rank p value < 0.05). Next, total 13 independent prognostic DEIIGs were further screened via a multivariate Cox regression model. After that, we performed the LASSO Cox regression analysis to reduce the number of independent prognosis related DEIIGs with the optimal lambda and finally obtained nine prognostic DEIIG signatures with corresponding coefficients for further study (Table 2).

Evaluation of the prognostic performance of DEIIG signature

According to the risk coefficient of each gene and the gene expression level, the PS of each patient in training dataset was calculated, which is a linear combination of the expression level of each gene weighted by its multivariate LASSO regression coefficient. The samples in training dataset were assigned into high-risk and low risk groups with the median PS as the cutoff value. The survival analysis indicated that the survival rate was remarkably lower in the high-risk group as opposed to low-risk group (*p*-value < 0.001, HR = 2.230, 95% CI = 1.583-3.142); whereas, the ROC curve analysis showed acceptable discrimination with AUC of 0.824, and high sensitivity and specificity in training dataset (Figure 6A). In addition, the external dataset GSE62254 was used to validate the prediction performance of the nine prognostic DEIIG signature. With the aforementioned formula, we calculated individual PS and classified the patients in validation dataset into high-risk and low-risk groups. Consistently, a significant separation was shown in the KM survival curve in validation dataset (*p*-value < 0.05, HR = 1.405, 95% CI = 1.020-1.935) and ROC curve analysis demonstrated accepted discrimination with an AUC of 0.766 (Figure 6B). In general, the nine prognostic DEIIG signature performed well at predicting OS of STAD.

Identification of independent prognostic parameters of STAD

Total 351 patients from the TCGA STAD dataset for which complete clinical information was provided, including age, gender, pathologic-M, pathologic-N, pathologic-T, pathologic-stage, neoplasm histologic grade and radiation therapy were included in the univariate and multivariate Cox regression analyses (Table S6). As shown in Table 3, univariate analysis revealed that age (*p* = 6.91E-03), pathologic M (*p* = 1.13E-02), pathologic N (*p* = 1.71E-03), pathologic T (*p* = 8.09E-03), pathologic stage (*p* = 1.68E-05), radiation therapy (*p* = 2.29E-04), recurrence (*p* = 3.44E-06) and PS model status (*p* = 2.56E-06) were significantly correlated with overall survival of STAD patients. Multivariate analysis further screened that age (*p* = 4.17E-03), pathologic T (*p* = 3.19E-02), radiation therapy (*p* = 1.05E-02), recurrence (*p* = 6.33E-04) and PS model status (*p* = 2.72E-03) were independent risk factors of overall survival. The results from forest map clearly described that age, pathologic T, tumor recurrence and PS model status were tumor risk factors, while radiotherapy was tumor protective factor (Figure 7). Subsequently, we performed ROC analyses to assess how these independent risk factors could behave in predicting prognosis. As shown in Figure 8, the AUC of PS model status performed on overall survival in the training cohort was 0.824, which was superior to those of age (0.545), pathologic T (0.537), radiotherapy (0.544) and recurrence (0.640), which may be the best performance in predicting overall survival.

Correlation of PS with the proportion of TICs

Based on the expression levels of TCGA STAD samples, we used TIMER to analyze the proportion of six kinds of TICs (Table S7). Combining the results of correlation analysis (Figure 9), B cell, T cell CD4+, neutrophil, macrophage and myeloid dendritic cell were positively correlated with PS, whereas T cell CD8+ was negatively correlated with PS. Thus, the significant infiltration with these TICs may potential act as one of the critical factors that the nine DEIIG signature holds to influence the outcome of STAD pronounced.

Validation of on DEIIG signature in clinical specimens

As described above, we have identified nine-DEIIG prognostic signature baed on the TCGA database. To further verify our findings, 59 cases of STAD specimens were collected and performed with quantitative real time PCR. As expected, ADH4 was downregulated in tumor tissues compared with adjacent tissues (Figure 10A). Clinical analysis further demonstrated that decreased ADH4 was associated with TNM stage, lymph node metastasis (Table 4), and represented an independent risk factor for overall survival (Table 5, Figure 10B). Conversely, ANGPT2 was upregulated in tumor tissues compared with adjacent tissues (Figure 10C), which was correlated with TNM stage (Table 6) and worse prognosis (Table 7, Figure 10D).

Discussion

As our best knowledge, gene expression and immune cell infiltration play a key role in the prognosis of tumors [31, 32]. Nevertheless, the association among gene expression levels, tumor infiltrating immune and survival prognosis remains largely unclear. Here, we used integrative bioinformatics to screen immune cell infiltration related genes based on the landscape of 28 immune cell subpopulations infiltrations in STAD derived from TCGA database. A total of 191 DEIIGs, including 6 lncRNAs and 185 mRNAs were obtained and used to apply for construction of the DEIIG prognostic signature. Total nine prognostic DEIIG signature (LINC00843, ADH4, ANGPT2, APOA1, ASLC2, GFRA1, KIAA1549L, MTTP and PROC) was identified to be associated with tumor cell immune infiltration. Alcohol dehydrogenases (ADHs), including class I (ADH1A, ADH1B, and ADH1C), class II (ADH4), class III (ADH5), class IV (ADH6), and class V (ADH7) [33], are huge family of dehydrogenase enzymes and associated with the prognosis of various cancers [34, 35]. A recent study by Wang et al. [36] identified that ADH4 was one of downregulated innate immunity genes in oral immune homeostasis. The presence of an ANGPT2-rich environment was associated with impairment of preexisting T-cell responses against tumor-associated antigens (TAA) and poor prognosis in patients with NSCLC [37]. In addition, APOA1 [38], GFRA1 [39], KIAA1549L [40] and MTTP [41] were all reported to be directly or indirectly associated with immune cell infiltration in cancer prognosis. There is little information concerning LINC00843, ASLC2 and PROC in immune infiltration related tumor prognosis, which need to be further explored.

Next, we evaluated the prognostic performance of DEIIG signature in TCGA and GEO datasets. Survival curves and time-dependent ROC and AUC analyses indicated that the nine prognostic DEIIG signatures

have powerful predictive capacity for STAD. Moreover, the AUC of PS model status performed on overall survival in the training cohort was 0.824, which was superior to those of age (0.545), pathologic T (0.537), radiotherapy (0.544) and recurrence (0.640), which may be the best performance in predicting overall survival. Consistent with the analysis of multivariate prognostic modules, the hazard ratio (HR) value of risk score based on the nine DEIIGs was higher among the factors in the forest map. These outcomes further confirmed that the nine-DEIIG signature was the most effective signature for prognostic assessment of STAD patients when compared with other clinical features. Similar to our study, Wang et al. [42] collected clinical data of STAD patients from TCGA database and established a stromal-immune-score-based gene signature and risk stratification. Yang et al. [43] collected RNA-seq data of immune infiltrated-related genes (IRGs) of 372 STAD patients from TCGA database and established a 10 prognostic gene prognostic model. Wu et al. [44] integrated clinical data to identify seven hub IRGs and establish the IRG prognostic model associated with STAD. Compared with previous studies, our study used updated data from TCGA and included 375 STAD patients. We used different validation dataset from GEO database and identified different nine prognostic DEIIG signature. Moreover, these DEIIGs were significantly correlated with the clinical outcomes (age, pathologic T, radiation therapy and recurrence) of STAD patients.

Our study also clarified the correlation between the three useful prognostic indicators and six types of tumor-infiltrating immune cells using TIMER. The results showed that nine prognostic DEIIG signature was positively correlated with B cell, T cell CD4+, neutrophil, macrophage and myeloid dendritic cell, but was negatively correlated with T cell CD8+. In fact, CD8+ T cells is one of the primary effector cells of anticancer immunity, which has been identified as a potential prognostic indicator of gastric cancer [14]. Consistently, previous study suggested that CD8 T cells with APOA1 as an alternative cellular vaccine for highly-active antiretroviral therapy [45]. ANGPT2 is a well-studied potential prognostic marker in B cell related chronic lymphocytic leukemia [46]. Furthermore, our validation experiments further demonstrated that both ADH4 and ANGPT2 were aberrantly expressed in STAD tissues and correlated with poor prognosis in STAD patients. Therefore, the identified nine prognostic DEIIG signature may also exert a vital function in immunotherapy of STAD. In addition, there were some limitations in our study as follows: We performed analysis at mRNA and non-coding level but not protein level. Furthermore, lacking of in vitro and in vivo experiments used for validating our results.

Conclusion

In summary, we screened nine DEIIGs (LINC00843, ADH4, ANGPT2, APOA1, ASLC2, GFRA1, KIAA1549L, MTTP and PROC) with marked prognostic capability for STAD. These DEIIGs were further confirmed as independent prognostic factors associated with OS of STAD patients. The findings might provide a new perspective that will help to find potential novel targets for STAD immunotherapy.

Abbreviations

STAD, stomach adenocarcinoma; ssGSEA, single sample gene set enrichment analysis; IIRGs, immune infiltrated-related genes; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; GSVA, Gene Set Variation Analysis; DEIIGs, differentially expressed immune infiltrated genes; DAVID, Database for Annotation, Visualization and Integrated Discovery; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PS, prognostic score; TIMER, Tumor Immune Estimation Resource

Declarations

Ethics approval and consent to participate

This study was approved by the Committee on the Ethics of Ethical Committee of The First Affiliated Hospital of Zhengzhou University (Henan, China). All of the experiments were performed in accordance with the Declaration of Helsinki. All volunteers who donated tissues have provided their written informed consent.

Authors' contributions

Ya Yang and Xintan Zhang wrote the main text of the article and designed the experiments. Ya Yang and Tingxuan Li were responsible for data analysis work. Wang Yue Zhang prepared Figures and revised the figure style. Xiaoxiao Zuo revised the manuscript. All the authors approved the manuscript. All authors contributed to the article and approved the submitted version.

Availability of data and materials

The analyzed data sets generated during the study are available from the corresponding author on reasonable request.

Acknowledgements

Not applicable.

Consent for publication

Not applicable

Funding

This work was supported by the Medical Science and Technology Co-construction Project of Henan Province.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries**. *CA: a cancer journal for clinicians* 2018, **68**(6):394–424.
2. Yakirevich E, Resnick MB: **Pathology of gastric cancer and its precursor lesions**. *Gastroenterology clinics of North America* 2013, **42**(2):261–284.
3. Haruhisa, Suzuki, Ichiro, Oda, Seiichiro, Abe, Masau, Sekiguchi, Genki, Mori: **High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection**. *Gastric cancer: official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* 2016.
4. Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang XS, Bannon F, Ahn JV, Johnson CJ, Bonaventure A *et al*: **Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2)**. *Lancet* 2015, **385**(9972):977–1010.
5. Katai H, Ishikawa T, Akazawa K, Isobe Y, Miyashiro I, Oda I, Tsujitani S, Ono H, Tanabe S, Fukagawa T *et al*: **Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001-2007)**. *Gastric Cancer* 2018, **21**(1):144–154.
6. Kim JW, Nam KH, Ahn SH, Park DJ, Kim HH, Kim SH, Chang H, Lee JO, Kim YJ, Lee HS *et al*: **Prognostic implications of immunosuppressive protein expression in tumors as well as immune cell infiltration within the tumor microenvironment in gastric cancer**. *Gastric Cancer* 2016, **19**(1):42–52.
7. Chen DS, Mellman I: **Elements of cancer immunity and the cancer-immune set point**. *Nature* 2017, **541**(7637):321–330.
8. Desmedt C, Salgado R, Fornili M, Pruneri G, Van den Eynden G, Zoppoli G, Rothe F, Buisseret L, Garaud S, Willard-Gallo K *et al*: **Immune Infiltration in Invasive Lobular Breast Cancer**. *Journal of the National Cancer Institute* 2018, **110**(7):768–776.
9. Zhou R, Zhang J, Zeng D, Sun H, Rong X, Shi M, Bin J, Liao Y, Liao W: **Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I–III colon cancer**. *Cancer Immunology, Immunotherapy* 2019.
10. Efstathiou JA, Mouw KW, Gibb EA, Liu Y, Wu CL, Drumm MR, da Costa JB, du Plessis M, Wang NQ, Davicioni E *et al*: **Impact of Immune and Stromal Infiltration on Outcomes Following Bladder-Sparing Trimodality Therapy for Muscle-Invasive Bladder Cancer**. *European urology* 2019, **76**(1):59–68.

11. Feng Y, Dai Y, Gong Z, Cheng JN, Zhu B: **Association between angiogenesis and cytotoxic signatures in the tumor microenvironment of gastric cancer**. *OncoTargets and therapy* 2018, **Volume 11**:2725–2733.
12. Zhang H, Wang X, Shen Z, Xu J, Qin J, Sun Y: **Infiltration of diametrically polarized macrophages predicts overall survival of patients with gastric cancer after surgical resection**. *Gastric Cancer* 2015, **18**(4):740–750.
13. Ishigami S, Natsugoe S, Tokuda K, Nakajo A, Okumura H, Matsumoto M, Miyazono F, Hokita S, Aikou T: **Tumor-associated macrophage (TAM) infiltration in gastric cancer**. *Anticancer research* 2003, **23**(5A):4079–4083.
14. Li F, Sun Y, Huang J, Xu W, Liu J, Yuan Z: **CD4/CD8 + T cells, DC subsets, Foxp3, and IDO expression are predictive indictors of gastric cancer prognosis**. *Cancer medicine* 2019, **8**(17):7330–7344.
15. Chaudhary B, Elkord E: **Novel expression of Neuropilin 1 on human tumor-infiltrating lymphocytes in colorectal cancer liver metastases**. *Expert opinion on therapeutic targets* 2015, **19**(2):147–161.
16. Wang M, Li Z, Peng Y, Fang J, Fang T, Wu J, Zhou J: **Identification of immune cells and mRNA associated with prognosis of gastric cancer**. *BMC cancer* 2020, **20**(1):206.
17. Meng L, Ding L, Yu Y, Li W, Huang T: **JAK3 and TYK2 Serve as Prognostic Biomarkers and Are Associated with Immune Infiltration in Stomach Adenocarcinoma**. *BioMed Research International* 2020, **2020**:1–15.
18. Li BL, Wan XP: **Prognostic significance of immune landscape in tumour microenvironment of endometrial cancer**. *Journal of cellular and molecular medicine* 2020, **24**(14):7767–7777.
19. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA *et al*: **Inferring tumour purity and stromal and immune cell admixture from expression data**. *Nature communications* 2013, **4**:2612.
20. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA: **Profiling Tumor Infiltrating Immune Cells with CIBERSORT**. *Methods in molecular biology* 2018, **1711**:243–259.
21. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies**. *Nucleic acids research* 2015, **43**(7):e47.
22. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources**. *Nature protocols* 2009, **4**(1):44–57.
23. Huang da W, Sherman BT, Lempicki RA: **Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists**. *Nucleic acids research* 2009, **37**(1):1–13.
24. Wang P, Wang Y, Hang B, Zou X, Mao JH: **A novel gene expression-based prognostic scoring system to predict survival in gastric cancer**. *Oncotarget* 2016, **7**(34):55343–55351.
25. Tibshirani R: **The lasso method for variable selection in the Cox model**. *Statistics in medicine* 1997, **16**(4):385–395.
26. Goeman JJ: **L1 penalized estimation in the Cox proportional hazards model**. *Biometrical journal Biometrische Zeitschrift* 2010, **52**(1):70–84.

27. Liu XF, Gao ZM, Wang RY, Wang PL, Li K, Gao S: **Comparison of Billroth I, Billroth II, and Roux-en-Y reconstructions after distal gastrectomy according to functional recovery: a meta-analysis**. *European review for medical and pharmacological sciences* 2019, **23**(17):7532-7542.
28. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M: **pROC: an open-source package for R and S+ to analyze and compare ROC curves**. *BMC bioinformatics* 2011, **12**:77.
29. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS: **TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells**. *Cancer research* 2017, **77**(21):e108-e110.
30. A KJL, B TDS: **Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2 ΔΔ C T Method**. *Methods* 2001, **25**(4):402–408.
31. Pagès F, Galon J, Dieu-Nosjean MC, Tartour E, Fridman WH: **Immune infiltration in human tumors: a prognostic factor that should not be ignored**. *Oncogene* 2010, **29**(8):1093–1102.
32. Kang JY, Gil M, Kim KE: **Neuropilin1 Expression Acts as a Prognostic Marker in Stomach Adenocarcinoma by Predicting the Infiltration of Treg Cells and M2 Macrophages**. *Journal of Clinical Medicine* 2020, **9**(5):1430.
33. Edenberg HJ, McClintick JN: **Alcohol Dehydrogenases, Aldehyde Dehydrogenases, and Alcohol Use Disorders: A Critical Review**. *Alcoholism, clinical and experimental research* 2018, **42**(12):2281–2297.
34. Wang P, Zhang L, Huang C, Huang P, Zhang J: **Distinct Prognostic Values of Alcohol Dehydrogenase Family Members for Non-Small Cell Lung Cancer**. *Medical science monitor: international medical journal of experimental and clinical research* 2018, **24**:3578–3590.
35. Guo E, Wei H, Liao X, Xu Y, Li S, Zeng X: **Prognostic value of alcohol dehydrogenase mRNA expression in gastric cancer**. *Oncology letters* 2018, **15**(4):5505–5516.
36. Wang Y, Anderson EP, Tatakis DN: **Whole transcriptome analysis of smoker palatal mucosa identifies multiple downregulated innate immunity genes**. *Journal of periodontology* 2020, **91**(6):756–766.
37. Lauret Marie Joseph E, Laheurte C, Jary M, Boullerot L, Asgarov K, Gravelin E, Bouard A, Rangan L, Dosset M, Borg C *et al*: **Immunoregulation and Clinical Implications of ANGPT2/TIE2(+) M-MDSC Signature in Non-Small Cell Lung Cancer**. *Cancer immunology research* 2020, **8**(2):268–279.
38. Li M, Wang Z, Zhu L, Shui Y, Zhang S, Guo W: **Down-regulation of RBP4 indicates a poor prognosis and correlates with immune cell infiltration in hepatocellular carcinoma**. *Bioscience reports* 2021, **41**(4).
39. Esseghir S, Todd SK, Hunt T, Poulsom R, Plaza-Menacho I, Reis-Filho JS, Isacke CM: **A role for glial cell derived neurotrophic factor induced expression by inflammatory cytokines and RET/GFR alpha 1 receptor up-regulation in breast cancer**. *Cancer research* 2007, **67**(24):11732–11741.
40. Janjanam VD, Mukherjee N, Lockett GA, Rezwan FI, Kurukulaaratchy R, Mitchell F, Zhang H, Arshad H, Holloway JW, Karmaus W: **Tetanus vaccination is associated with differential DNA-methylation: Reduces the risk of asthma in adolescence**. *Vaccine* 2016, **34**(51):6493–6501.
41. Dominguez JA, Xie Y, Dunne WM, Yoseph BP, Burd EM, Coopersmith CM, Davidson NO: **Intestine-specific Mttp deletion decreases mortality and prevents sepsis-induced intestinal injury in a murine**

model of Pseudomonas aeruginosa pneumonia. *PloS one* 2012, **7**(11):e49159.

42. Wang H, Wu X, Chen Y: **Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer**. *Frontiers in oncology* 2019, **9**:1212.
43. Yang W, Lai Z, Li Y, Mu J, Yang M, Xie J, Xu J: **Immune signature profiling identified prognostic factors for gastric cancer**. *Chinese journal of cancer research = Chung-kuo yen cheng yen chiu* 2019, **31**(3):463–470.
44. Wu M, Xia Y, Wang Y, Fan F, Li X, Song J, Ding J: **Development and validation of an immune-related gene prognostic model for stomach adenocarcinoma**. *Bioscience reports* 2020, **40**(10).
45. Ledru E, Christeff N, Patey O, de Truchis P, Melchior JC, Gougeon ML: **Alteration of tumor necrosis factor-alpha T-cell homeostasis following potent antiretroviral therapy: contribution to the development of human immunodeficiency virus-associated lipodystrophy syndrome**. *Blood* 2000, **95**(10):3191–3198.
46. Kopparapu PK, Miranda C, Fogelstrand L, Mishra K, Andersson PO, Kanduri C, Kanduri M: **MCPH1 maintains long-term epigenetic silencing of ANGPT2 in chronic lymphocytic leukemia**. *The FEBS journal* 2015, **282**(10):1939–1952.

Tables

Table 1

GO function and KEGG pathway analysis of DEIIGs

Category	Term	Count	P-Value	FDR	Genes
Biology Process	GO:0001523~retinoid metabolic process	8	2.65E-06	1.47E-03	ADH4, RDH12, RBP2, APOC3, APOA1, APOA4, LRAT, APOB
	GO:0030574~collagen catabolic process	7	4.55E-05	1.26E-02	MMP12, MMP7, COL11A1, MMP3, COL4A6, COL10A1, MMP10
	GO:0042158~lipoprotein biosynthetic process	4	8.19E-05	1.50E-02	MTTP, APOA1, APOA4, APOB
	GO:0033344~cholesterol efflux	5	1.09E-04	1.50E-02	ABCG8, APOC3, APOA1, APOA4, APOB
	GO:0006813~potassium ion transport	7	1.83E-04	2.02E-02	ATP4B, ABCC8, KCNJ13, KCNMB2, KCNMB3, KCNA5, ATP1A2
	GO:0071805~potassium ion transmembrane transport	8	2.31E-04	2.12E-02	KCNE2, KCNB1, ABCC8, KCNH8, KCNMB2, KCNMB3, KCNA5, TRPM5
	GO:0007267~cell-cell signaling	11	2.68E-04	2.12E-02	IL11, BMP3, VIPR2, SST, CD80, ASIP, CCL3, TNFSF9, ADRB2, INHA, MLN
	GO:0007586~digestion	6	4.39E-04	2.90E-02	CCKAR, CCKBR, SST, GKN1, MLNR, PGA3
	GO:0042632~cholesterol homeostasis	6	4.72E-04	2.90E-02	ABCG8, MTTP, APOC3, APOA1, APOA4, APOB
	GO:0042157~lipoprotein metabolic process	5	5.72E-04	3.14E-02	ALB, APOC3, APOA1, APOA4, APOB
	GO:0001508~action potential	4	6.24E-04	3.14E-02	KCNB1, KCNMB2, KCNMB3, AKAP6
KEGG Pathway	hsa04977: Vitamin digestion and absorption	5	1.34E-04	1.03E-02	RBP2, APOA1, APOA4, LRAT, APOB
	hsa04974: Protein digestion and absorption	7	7.19E-04	1.85E-02	COL11A1, KCNJ13, COL4A6, COL10A1, ATP1A2, PGA3, SLC8A2
	hsa04975: Fat digestion and absorption	5	7.93E-06	6.42E-04	ABCG8, MTTP, APOA1, APOA4, APOB
	hsa04060: Cytokine-	9	6.15E-	4.98E-	IL11, CXCL8, CSF2,

	cytokine receptor interaction		05	03	TNFRSF13B, OSM, LIF, CCL3, TNFRSF17, TNFSF9
	hsa04971: Gastric acid secretion	5	7.64E-05	6.19E-03	ATP4B, KCNE2, CCKBR, SST, ATP1A2
	hsa00830: Retinol metabolism	4	2.73E-04	2.21E-02	ADH4, CYP2B6, RDH12, LRAT
	hsa04022: cGMP-PKG signaling pathway	6	2.80E-04	2.27E-02	KCNMB2, KCNMB3, CACNA1D, ATP1A2, ADRB2, SLC8A2
	hsa05202: Transcriptional misregulation in cancer	6	3.41E-04	2.76E-02	CXCL8, CSF2, ZBTB16, MMP3, FCGR1A, RXRG
	hsa00500: Starch and sucrose metabolism	3	3.84E-04	3.11E-02	TREH, PYGM, GBA3
	hsa04020: Calcium signaling pathway	6	4.32E-04	3.50E-02	CCKAR, CCKBR, CACNA1D, ADRB2, RYR3, SLC8A2
	hsa04512: ECM-receptor interaction	4	5.69E-04	4.61E-02	COL11A1, CHAD, COL4A6, SPP1

Table 2

Information about the nine-DEIIR signature

Symbol	Type	Multi-variate Cox regression analysis			LASSO coef
		HR	95%CI	P value	
LINC00843	lncRNA	0.618	0.369-0.933	6.59E-02	-0.3194
ADH4	mRNA	1.166	1.034-1.314	1.22E-02	0.1550
ANGPT2		1.386	1.103-1.742	5.05E-03	0.2739
APOA1		0.799	0.889-0.979	4.67E-02	-0.0269
ASCL2		0.894	0.814-0.981	1.85E-02	-0.1004
GFRA1		1.217	1.034-1.435	1.86E-02	0.1002
KIAA1549L		1.433	1.093-1.878	9.21E-03	0.3762
MTTP		1.119	1.005-1.323	4.19E-02	0.1256
PROC		1.265	1.016-1.574	3.57E-02	0.1655

Table 3

The independent prognostic clinical factors according to univariate and multivariate Cox regression analyses

Clinical characteristics	TCGA (N=351)	Uni-variables cox			Multi-variables cox		
		HR	95%CI	P	HR	95%CI	P
Age (years, mean ± SD)	65.49 ± 10.58	1.022	1.006-1.039	6.91E-03	1.033	1.010-1.056	4.165E-03
Gender (Male/Female)	226/125	1.271	0.891-1.813	1.84E-01	-	-	-
Pathologic_M (M0/M1/-)	313/23/15	2.063	1.164-3.659	1.13E-02	1.138	0.423-3.062	7.986E-01
Pathologic_N (N0/N1/N2/N3/-)	104/94/73/70/10	1.329	1.144-1.544	1.71E-04	1.157	0.875-1.530	3.061E-01
Pathologic_T (T1/T2/T3/T4/-)	16/75/163/93/4	1.318	1.070-1.624	8.09E-03	1.472	1.034-2.095	3.187E-02
Pathologic_stage (I/II/III/IV/-)	47/109/147/35/13	1.551	1.264-1.903	1.68E-05	1.144	0.699-1.872	5.902E-01
Neoplasm histologic grade (G1/G2/G3/-)	9/127/206/9	1.366	0.991-1.883	5.63E-02	-	-	-
Radiation therapy (Yes/No/-)	62/265/24	0.429	0.262-0.703	2.29E-04	0.422	0.218-0.817	1.052E-02
H.pylori infection(Yes/No/-)	18/142/191	0.642	0.276-1.494	3.00E-01	-	-	-
Barretts esophagus (Yes/No/-)	14/191/146	0.941	0.344-2.573	9.05E-01	-	-	-
Recurrence (Yes/No/-)	59/229/63	2.517	1.681-3.767	3.44E-06	2.348	1.439-3.832	6.330E-04
PS model status (High/Low)	175/176	2.23	1.583-3.142	2.56E-06	2.059	1.284-3.301	2.722E-03
Death (Yes/No)	144/207	-	-	-	-	-	-
Overall survival time (months, mean ± SD)	20.37 ± 18.34	-	-	-	-	-	-

HR hazard ratio, CI confidence interval, SD standard deviation

Table 4

Association of ADH4 expression with clinicopathological features of GC patients

Variables	Cases (n = 59)	ADH4 expression		*P* value
		Low (n = 30)	High (n = 29)	(chi-square test)
Sex				0.506
Male	32	15	17	
Female	27	15	12	
Age				0.083
< 60	25	16	9	
≥ 60	34	14	20	
Tumor size (cm)				0.145
< 5	38	22	16	
≥ 5	21	8	13	
TNM stage				0.041*
I + II	40	24	16	
III + IV	19	6	13	
Lymph node metastasis				0.005*
Yes	23	17	6	
No	36	13	23	
Differentiation				0.195
Well/moderate	45	25	20	
Poor	14	5	9	

Note: *Statistically significant; **Abbreviations:** GC, gastric cancer; TNM, tumor-node-metastasis

Table 5

Univariate and multivariate analysis for overall survival in GC patients

	Univariate analysis		Multivariate analysis	
Characteristics	**HR (95% CI)**	**P value**	**HR (95% CI)**	**P value**
Sex	0.782 (0.489-1.252)	0.305	NA	NA
Age (years)	0.864 (0.588-1.269)	0.455	NA	NA
Tumor size (cm)	0.749 (0.508-1.105)	0.006	NA	NA
TNM stage	1.994 (1.398-2.846)	1.12E-04	1.904 (1.062-3.412)	4.58E-02
Lymph node metastasis	1.438 (1.181-1.751)	2.51E-04	1.025 (0.751-1.400)	8.75E-01
Differentiation	1.425 (0.719-2.823)	3.07E-01	NA	NA
ADH4 expression	2.456 (1.651-3.654)	4.71E-03	2.205 (1.415-3.435)	4.73E-02
Abbreviations: GC, gastric cancer; HR: hazard ratio; CI: confidence interval; NA, not analyzed				

Table 6

Association of ANGPT2 expression with clinicopathological features of GC patients

Variables	Cases (n = 59)	ANGPT2 expression		*P* value
		High (n = 30)	Low (n = 29)	(chi-square test)
Sex				0.235
Male	32	14	18	
Female	27	16	11	
Age				0.367
< 60	25	11	14	
≥ 60	34	19	15	
Tumor size (cm)				0.207
< 5	38	17	21	
≥ 5	21	13	8	
TNM stage				0.003*
I + II	40	15	25	
III + IV	19	15	4	
Lymph node metastasis				0.486
Yes	23	13	10	
No	36	17	19	
Differentiation				0.942
Well/moderate	45	23	22	
Poor	14	7	7	

Note: *Statistically significant; **Abbreviations:** GC, gastric cancer; TNM, tumor-node-metastasis

Table 7

Univariate and multivariate analysis for overall survival in GC patients

	Univariate analysis		Multivariate analysis	
Characteristics	**HR (95% CI)**	**P value**	**HR (95% CI)**	**P value**
Sex	0.989 (0.971-1.006)	1.937E-01	NA	NA
Age (years)	0.909 (0.537-1.540)	7.226E-01	NA	NA
Tumor size (cm)	1.594 (1.007-2.523)	4.370E-02	1.225 (0.744-2.019)	4.25E-01
TNM stage	2.227 (1.431-3.467)	2.620E-03	1.730 (1.524-3.141)	7.19E-03
Lymph node metastasis	1.811 (1.394-2.353)	6.230E-03	1.052 (0.309-1.371)	2.59E-01
Differentiation	0.885 (0.503-1.559)	6.710E-01	NA	NA
ANGPT2 expression	0.830 (1.746-4.589)	9.86E-03	2.508 (1.472-4.276)	7.27E-02
Abbreviations: GC, gastric cancer; HR: hazard ratio; CI: confidence interval; NA, not analyzed				

Figures

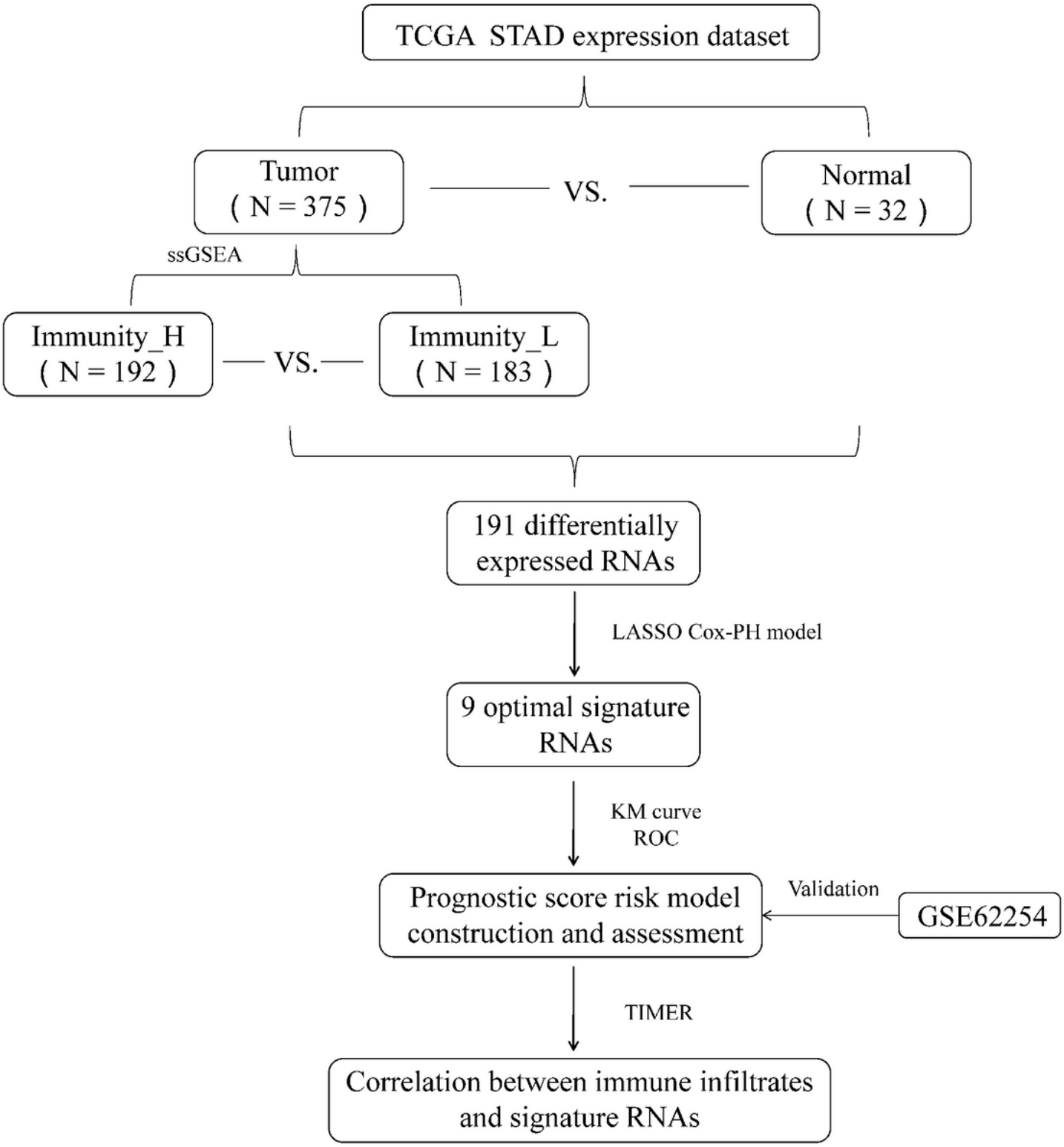

Figure 1

Flowchart detailing the overall study design and samples at each stage of analysis.

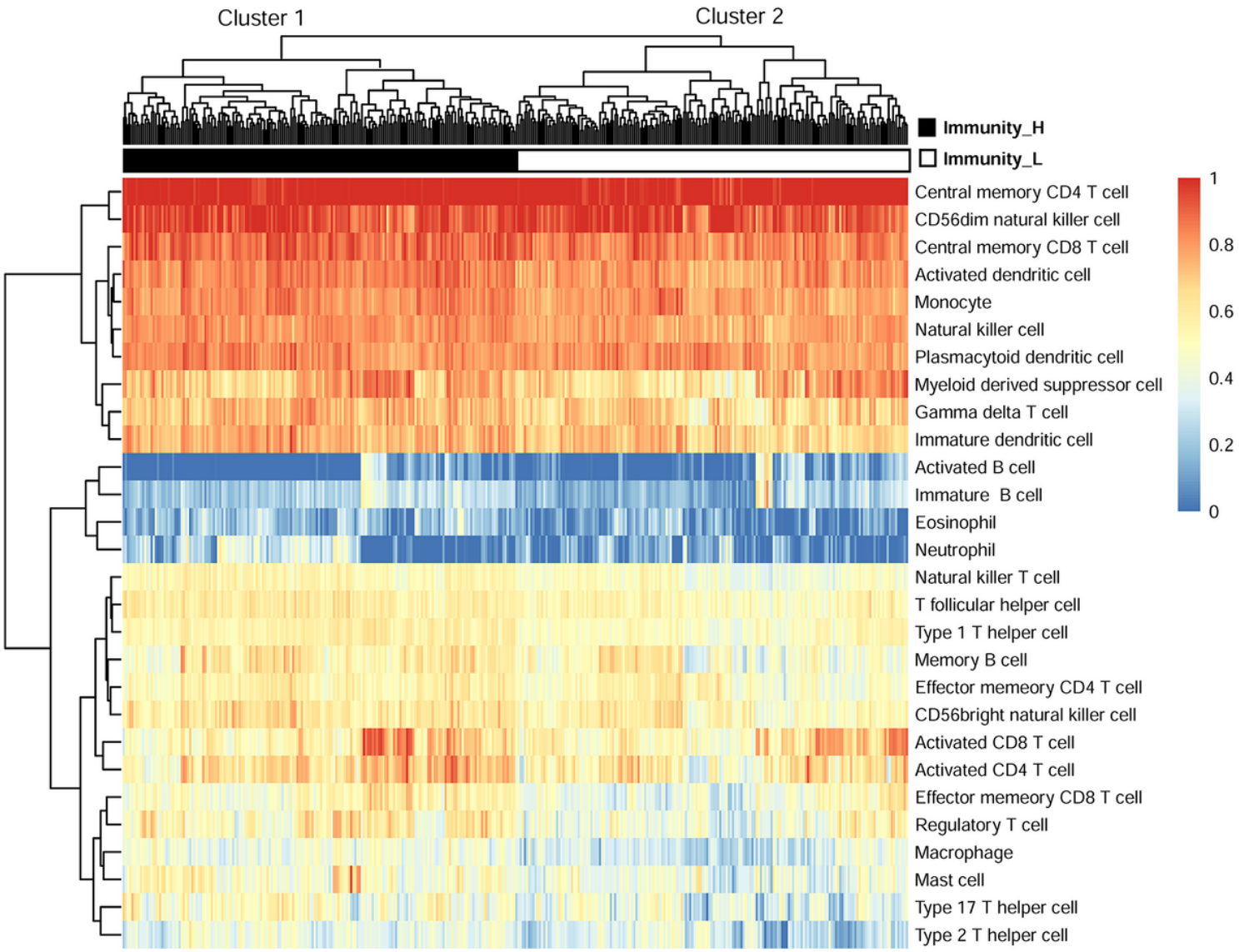

Figure 2

Heat map of the 28 immune cell proportions.

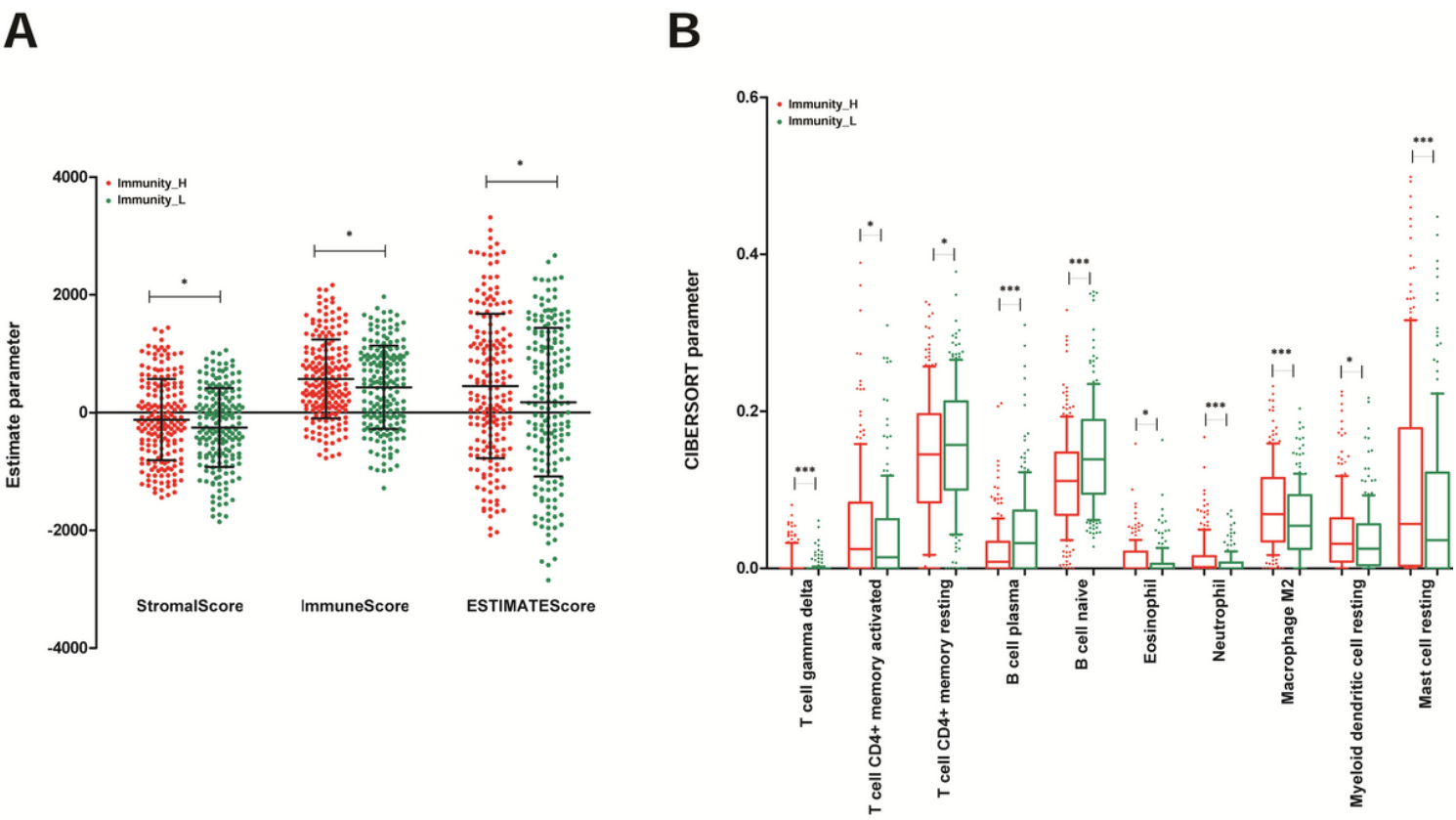

Figure 3

Evaluation of tumor-infiltrating immune group. (A) Results distribution of stromal score, immune score and estimate score based on the ESTIMATE method for Immunity-H and -L groups. Red and green dots represent samples from Immunity-H and -L groups, respectively. (B) The landscape of immune infiltration in STAD and difference of immune infiltration between Immunity-H and -L groups based on CIBERSORT algorithm.

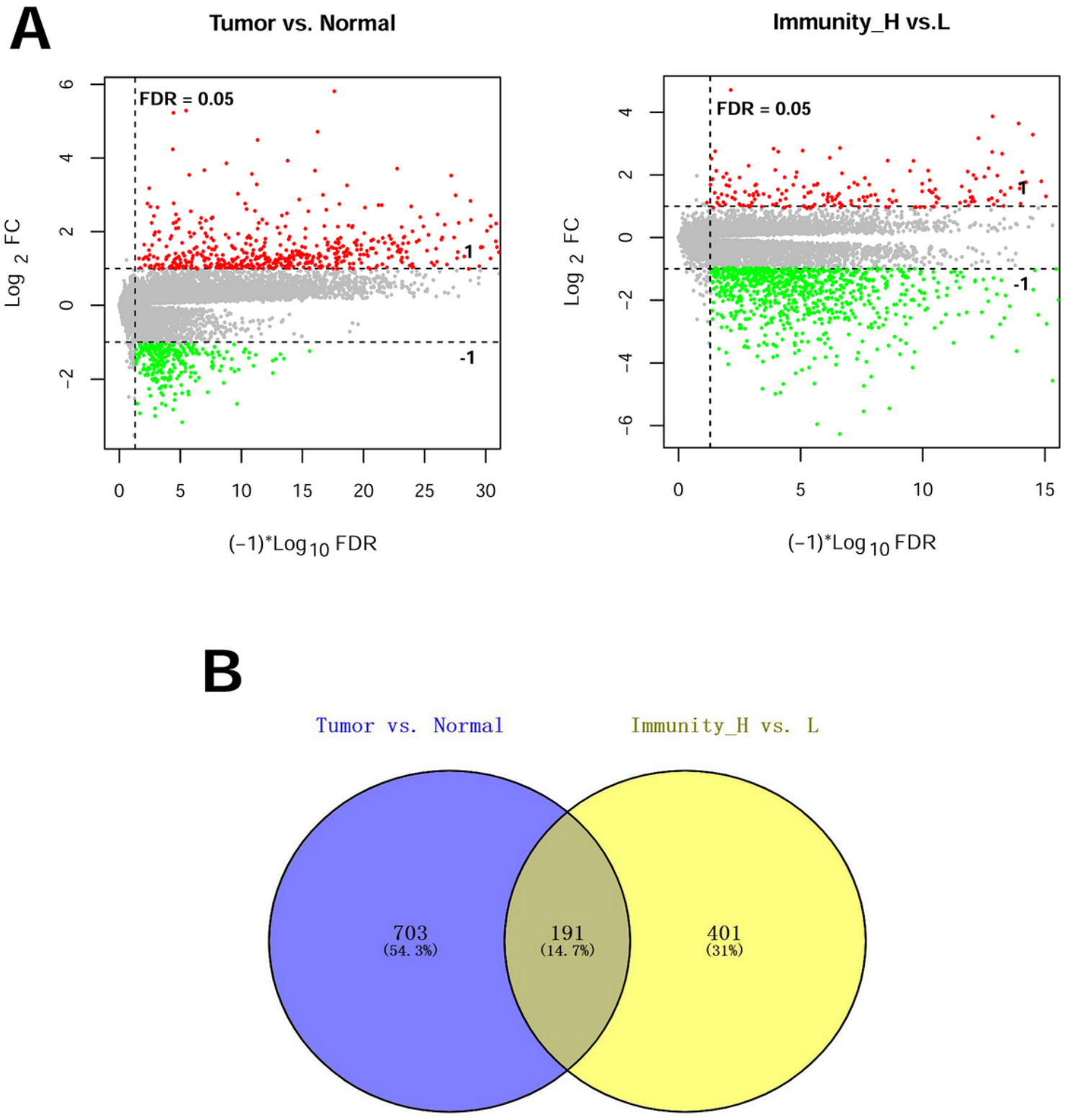

Figure 4

Identification of DEIIGs. (A) Effect size (log2FC)-log10 (FDR) volcano plots of DEGs in tumor vs. normal and Immunity-H vs. Immunity-L. Red and green dots indicate significant upregulated and downregulated RNAs, respectively. Horizontal dashed lines indicate FDR < 0.05 and two vertical dashed lines represent |log2FC| > 1. (B) The overlapping DEIIGs were visualized by Venn diagram.

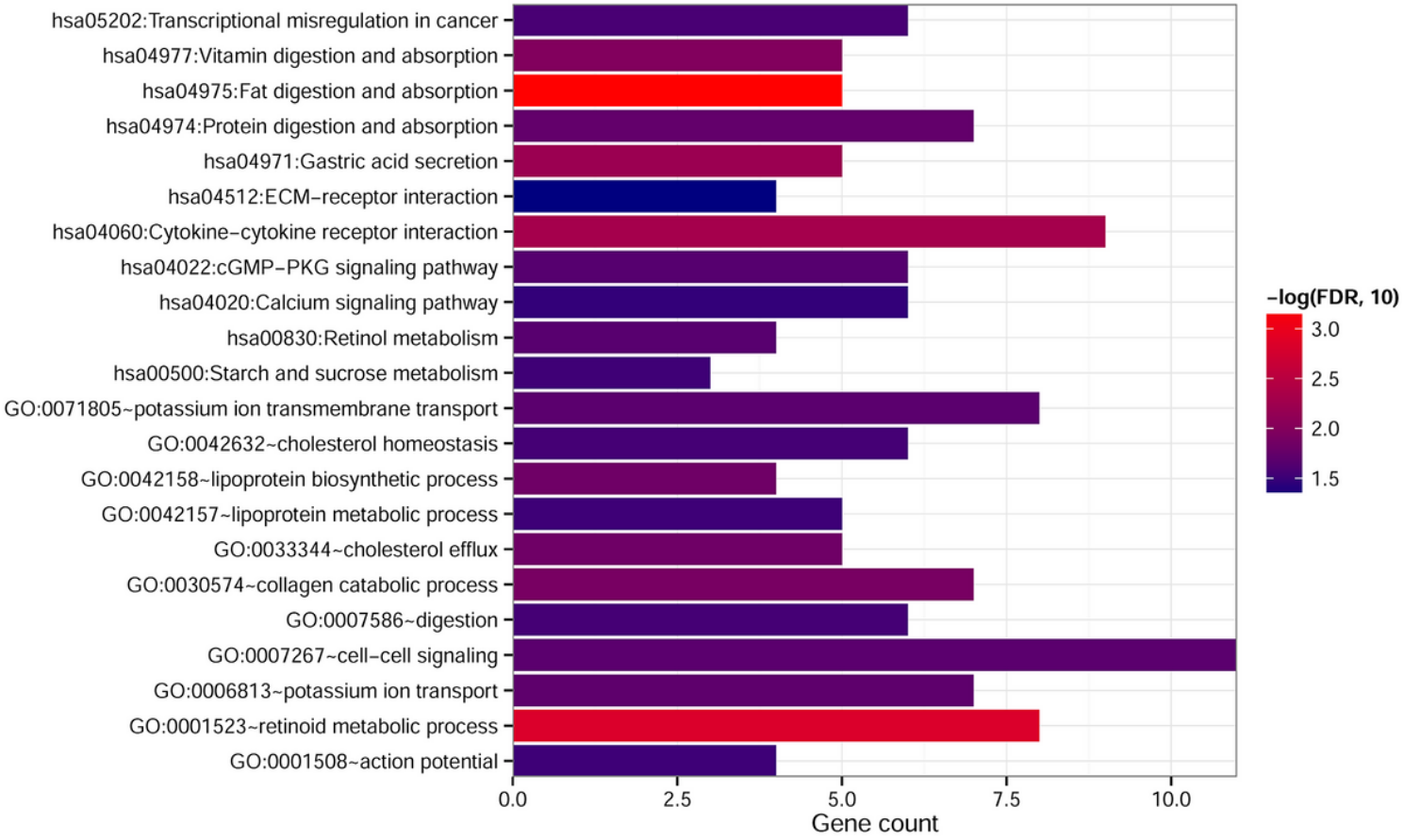

Figure 5

The GO and KEGG function enrichment analyses of overlapping 185 DEIIGs.

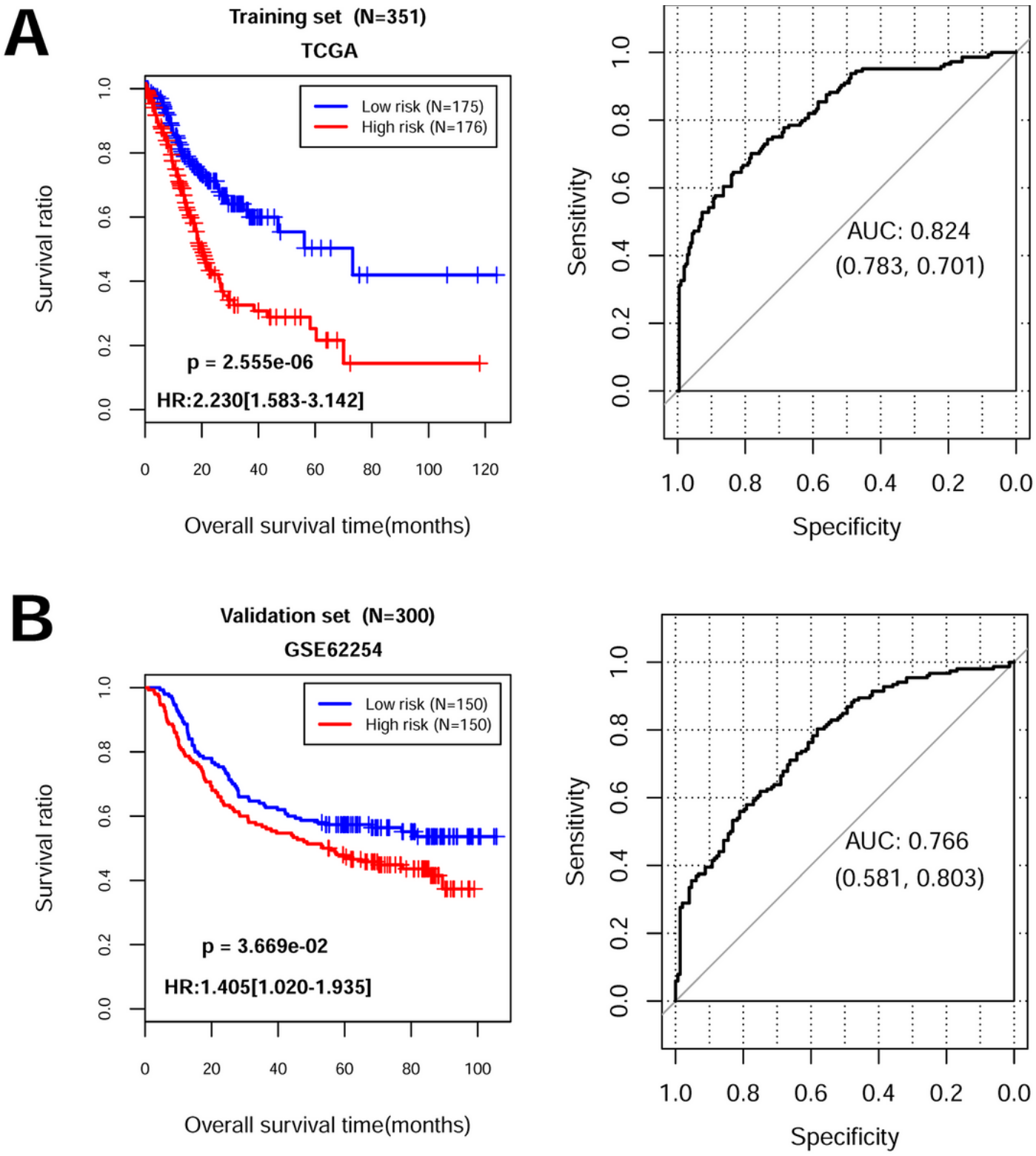

Figure 6

Evaluation of the performance of the nine prognostic DEIIG signature. (A) TCGA and (B) GSE62254; Left panel showed the Kaplan-Meier survival curves of the nine gene signature. Patients were stratified into high-risk and low-risk groups with the median PS as the cutoff value. Right panel showed the ROC curves for overall survival prediction for the nine gene signature.

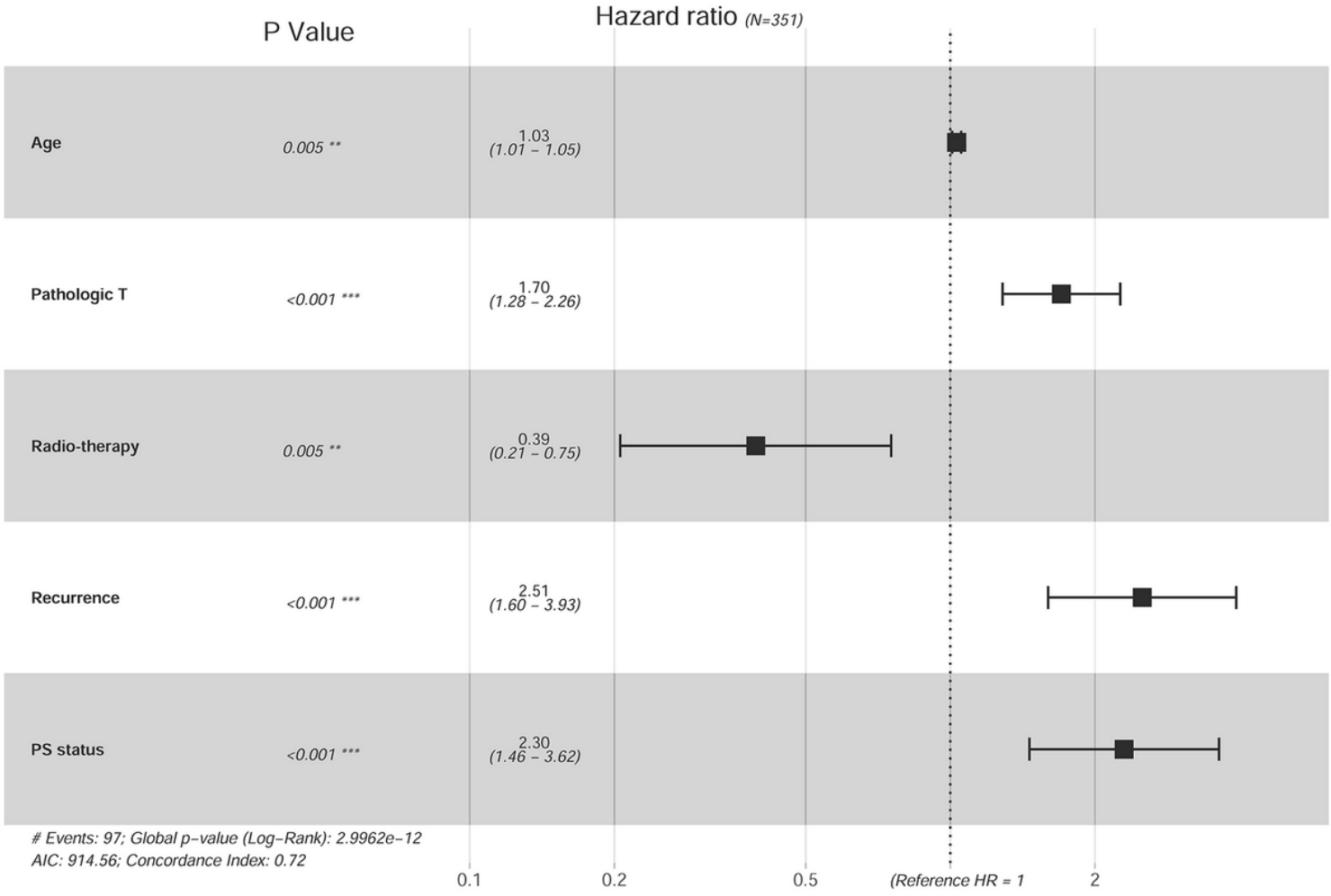

Figure 7

Prognostic value of the nine DEIIGs in STAD patients based on forest plots. Clinical features (age, pathologic T, radiotherapy, recurrence) and PS status were analyzed to assess the hazard ratio for STAD patients.

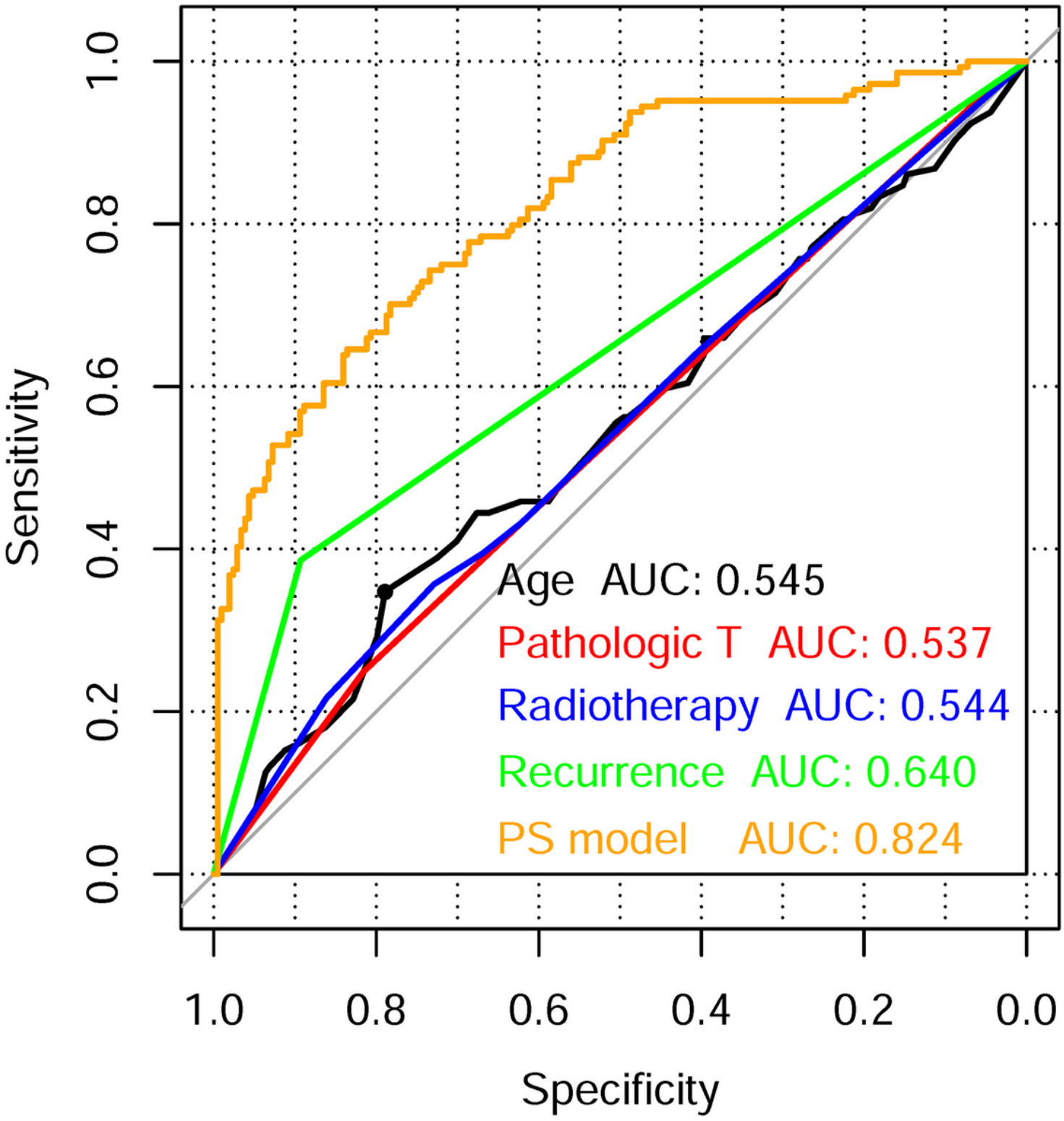

Figure 8

ROC analysis of the sensitivity and specificity of the prognosis prediction by the age, pathologic T, radiotherapy, recurrence and PS model in training cohort. AUC, area under the ROC curve.

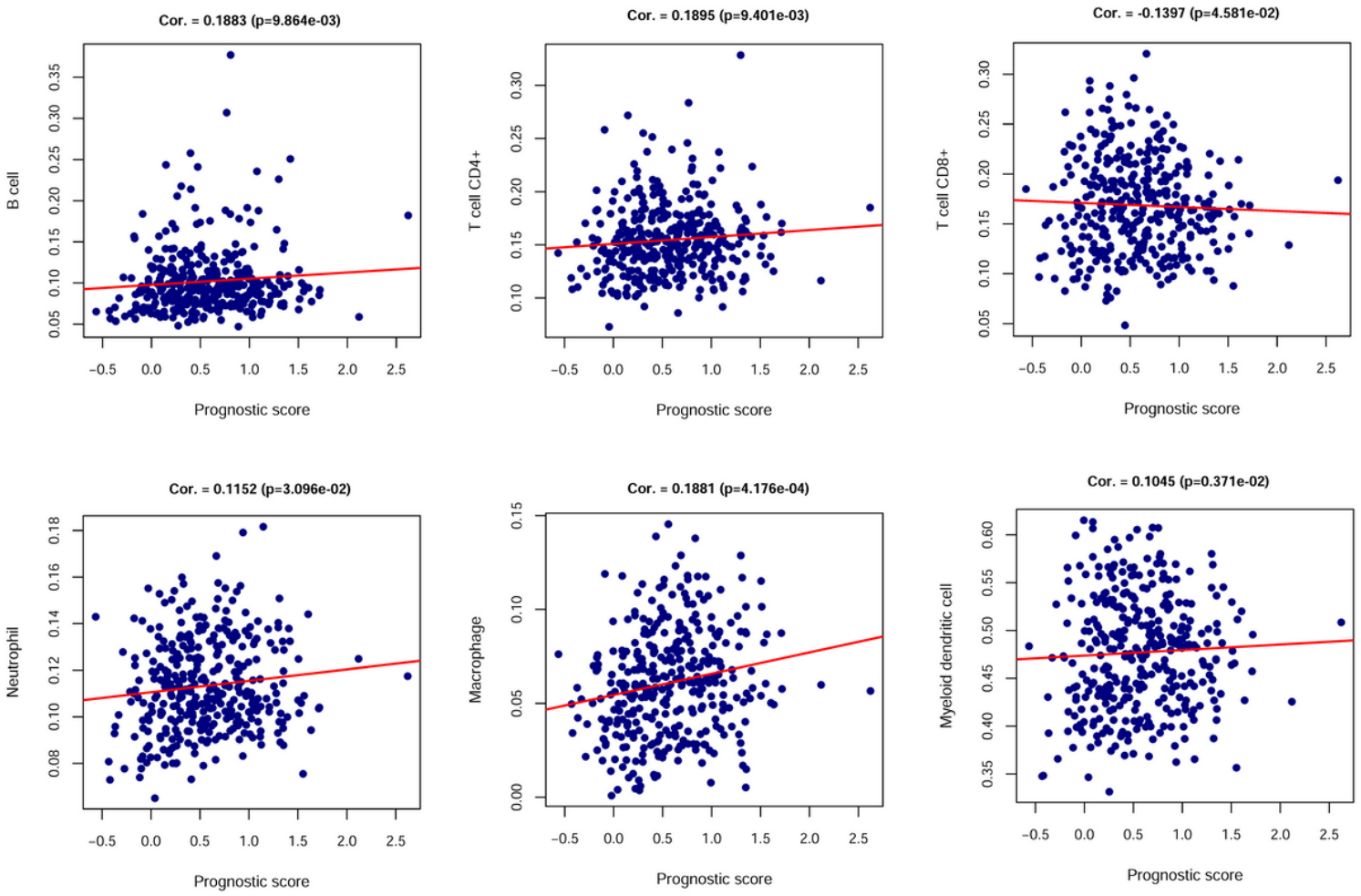

Figure 9

Correlation between six types of TICs proportion and nine-DEIIG signature prognostic score in the training cohort. Only significantly correlated TICs were plotted. The red line in each plot was fitted linear model indicating the proportion tropism of the immune cell along with prognostic score. The blue dots around the red line represents the 95% confidence interval. The Spearman coefficient was used for the correlation test.

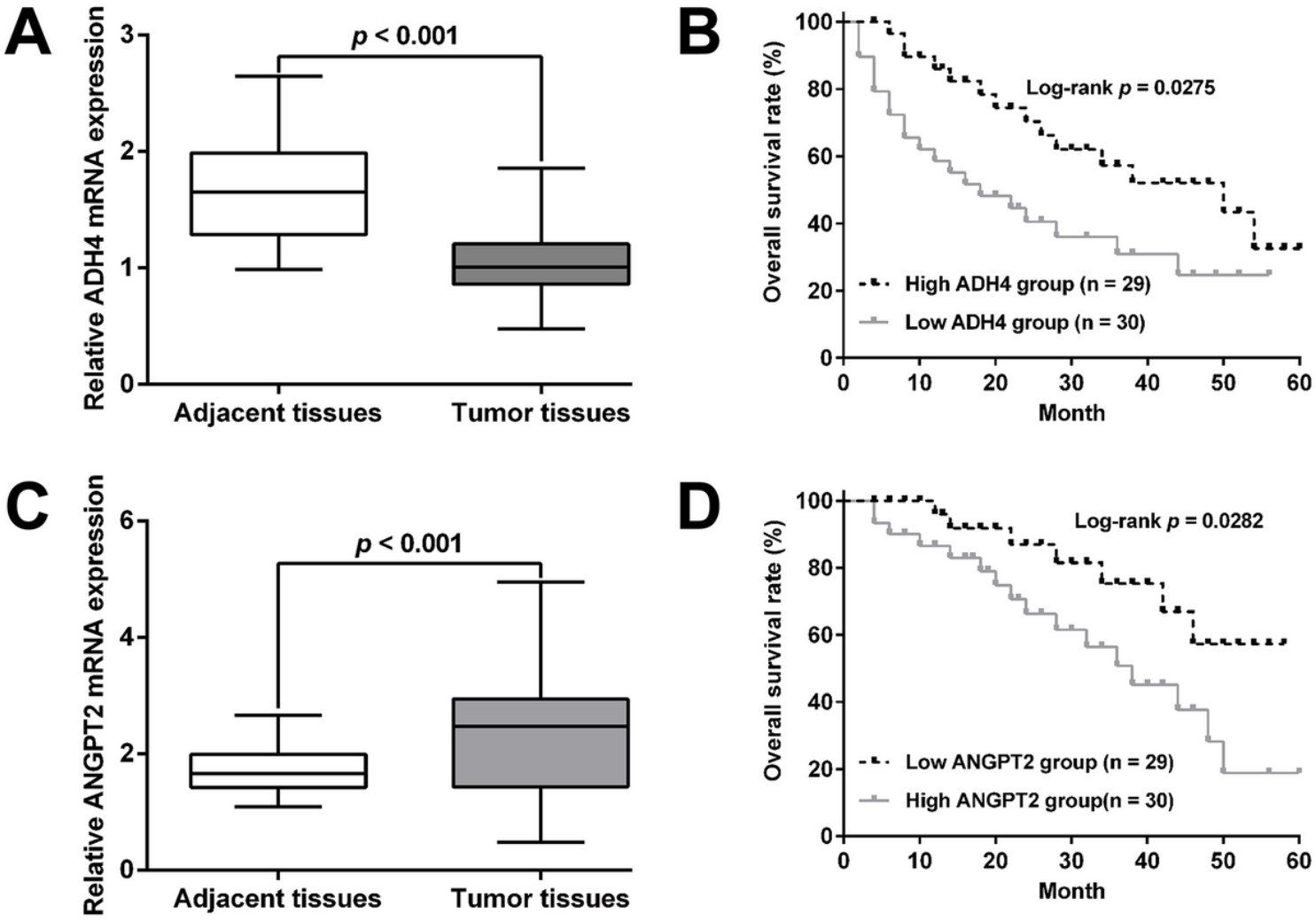

Figure 10

The expression levels and prognostic value of (A-B) ADH4 and (C-D) ANGPT2 in STAD patients. Left panel represented gene expression levels between tumor tissues and adjacent tissues. Right panel represented the Kaplan-Meier survival curves of corresponding genes.