scientific reports

OPEN

Bacterial lipopolysaccharide related genes signature as potential biomarker for prognosis and immune treatment in gastric cancer

Tianyi Yuan[1,4], Siming Zhang[2,4], Songnian He[1], Yijie Ma[1], Jianhong Chen[2✉] & Jue Gu[3✉]

The composition of microbial microenvironment is an important factor affecting the development of tumor diseases. However, due to the limitations of current technological levels, we are still unable to fully study and elucidate the depth and breadth of the impact of microorganisms on tumors, especially whether microorganisms have an impact on cancer. Therefore, the purpose of this study is to conduct in-depth research on the role and mechanism of prostate microbiome in gastric cancer (GC) based on the related genes of bacterial lipopolysaccharide (LPS) by using bioinformatics methods. Through comparison in the Toxin Genomics Database (CTD), we can find and screen out the bacterial LPS related genes. In the study, Venn plots and lasso analysis were used to obtain differentially expressed LPS related hub genes (LRHG). Afterwards, in order to establish a prognostic risk score model and column chart in LRHG features, we used univariate and multivariate Cox regression analysis for modeling and composition. In addition, we also conducted in-depth research on the clinical role of immunotherapy with TMB, MSI, KRAS mutants, and TIDE scores. We screened 9 LRHGs in the database. We constructed a prognostic risk score and column chart based on LRHG, indicating that low risk scores have a protective effect on patients. We particularly found that low risk scores are beneficial for immunotherapy through TIDE score evaluation. Based on LPS related hub genes, we established a LRHG signature, which can help predict immunotherapy and prognosis for GC patients. Bacterial lipopolysaccharide related genes can also be biomarkers to predict progression free survival in GC patients.

Gastric cancer (GC) is a general gastrointestinal malignancy in China. Global data show that GC ranks fifth in the incidence of all malignancies[1,2]. Environmental factors and lifestyle are the consequential causes of GC. In the past decades, the incidence of GC has gradually declined in developed countries, but it still threatens human health in developing countries. China is one of the countries with the highest incidence of GC, ranking the second in the incidence of tumors and the third in the mortality rate of malignant tumors[3]. There are 370,000 cases of death relevant to GC in China every year, even more than half of the global GC related deaths[3,4]. The backward technology of cancer screening leads to the diagnosis of most GC patients with advanced stage, accompanied by local invasion and distant metastasis, which is highly correlated with poor prognosis and mortality[5–7]. At present, gastrectomy is still the preferred treatment for patients with GC[8]. However, simple gastrectomy may not be suitable for patients with extensive invasion or lymphatic metastasis[9]. Although the pathogenesis of gastric cancer has been studied intensively by researchers around the world, the exact molecular mechanism remains unclear. Therefore, elucidating the mechanism of GC and finding effective markers of GC are of great significance for improving the prognosis of GC patients.

In addition to common infecting virus in GC development[10], intestinal microbial has become the focus of research on GC and environmental factors, and has been widely concerned[11,12]. Under normal circumstances, the

[1]Nantong Integrated Traditional Chinese and Western Medicine Hospital, Nantong, Jiangsu, China. [2]Affiliated Tumor Hospital of Nantong University, Nantong, Jiangsu, China. [3]Affiliated Hospital of Nantong University, Nantong, China. [4]These authors contributed equally: Tianyi Yuan and Siming Zhang. ✉email: chenhongjian2011@163.com; 511071555@163.com

gut microbiota as a real organ associated with coordinating health and wellness of our body and the coding genes of intestinal microorganisms are 100 times higher than those of human genes[13]. Studies have pointed out that the intestinal microbiota can promote the proliferation of gastrointestinal epithelial cells[14], ensure the normal energy and metabolism in the human body, and play an anti-inflammatory part[15,16]. Under normal circumstances, gut microbiota plays a significant part in maintaining the stability of flora and regulating the immune function of the body. After the imbalance of gastrointestinal microbiota, normal functions (immune function, metabolism, energy conversion) will be affected, resulting in the occurrence of various diseases[17–19]. The intestinal flora can affect local lesions by influencing immune factors and metabolites of the flora to have an impact on other parts of the body. With the deepening of the research on the gut microbiota, more and more studies have been conducted to diagnose the early stage of disease, treatment and prognosis by detecting microbiota markers[20–22]. As the main component of gram negative bacteria, lipopolysaccharide (LPS) plays an important role in the pathogenic components. Therefore, searching for LPS related hub genes (LRHG) has positive therapeutic significance for GC, and it is also very important to prevent the occurrence of GC.

In this study, we identified survival of GC associated with LPS relevant genes. After univariate and multivariate Cox regression analysis, we learned that LRHG could significantly predict the survival outcome of GC patients. This study examined potential molecular mechanisms of the prostate microbiome in GC. Besides, we used the chosen LRHG to construct a new prognostic model. In conclusion, our study provided a new LRHG model, and on this basis, we validated its ability in predicting the outcome of immunotherapy for GC.

Materials and methods

Data acquisition

TCGA (https://portal.gdc.cancer.gov/) provides GC mRNA expression and clinical data with 371 patients being enrolled. 433 patients from the GSE84437 dataset were screened as the external validation set. In the process of comparative Toxicogenomics Database (CTD, http://ctdbase.org/), 6555 genes related to LPS were found. The analysis process of our study is shown in in Fig. 1.

Gene-enrichment and functional annotation analysis

Tumor and adjacent normal's differentially expressed genes (DEGs) defined as $|\log FC| > 1.5$ and adjusted $P < 0.05$ were found. We utilize Gene ontology (GO) function enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG)[23] pathway enrichment analyses of DEGs by R package.

Construction of risk model based on LPS-related hub genes (LRHG)

Construct co-expression networks based on DEGs by utilizing WGCNA, which were then analyzed with clinical data using an appropriate soft-threshold power. We set the cut height to 0.3, set the minimum module size to 30, and calculate the dissimilarity results of the modules, some of which are merged. Finally, the significant positive and negative module (Blue and Turquoise) was selected for subsequent analysis. We selected prognostic genes associated with LPS based on modular genes using multivariate Cox analysis. The prediction of PPI network is input into prognostic genes in the database by using the online Search Tool for the Retrieval of Interacting Genes (STRING) database. Cytoscape tool can be used to visualize PPI networks. In the process of using plug-in cytoHubba, we comprehensively identified hub genes through Degree, DMNC, EPC, MCC and MNC scores. The ten genes with the highest degree of Degree, DMNC, EPC, MCC and MNC can be used as hub genes. Patients were divided into two groups according to LASSO cox analysis, namely, low-risk group and high-risk group. Principal Component Analysis (PCA) was applied to assess the clustering effect. To estimate the two groups' prognostic ability of the risk signature, Kaplan–Meier curves with the Log rank test were applied. To assess the performance of model predictions, the ROC curves for 1-, 3-, and 5-year survival were calculated.
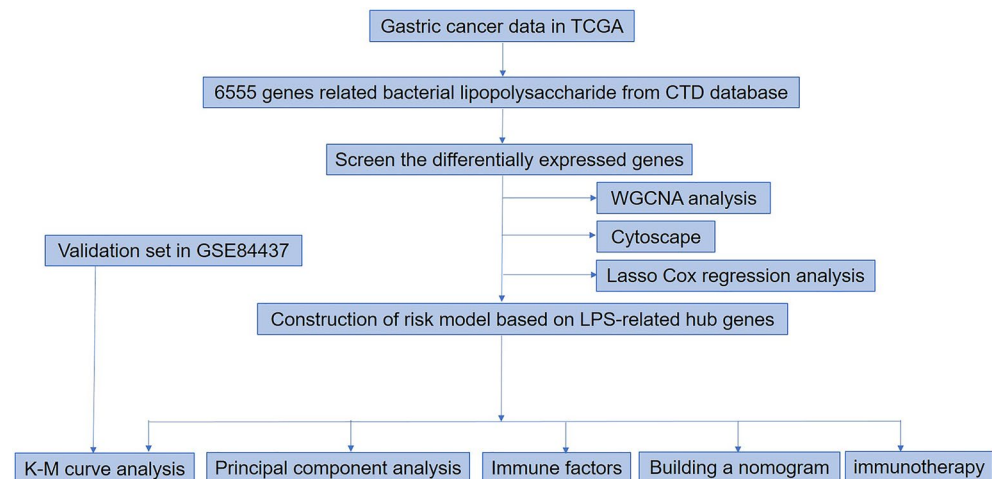

Figure 1. The flow chart of this study.

Construction of risk model nomogram based on LRHG

By establishing nomograms between risk scores and clinicopathological characteristics, we can predict 1 -, 3 -, and 5-year OS. According to the Hosmer lemeshow test, the consistency between the actual results and the predicted results can be explained by using the calculation correction curve. AUC and ROC curves were used in this study to evaluate the clinicopathological characteristics of prognosis.

Tumor-infiltrating immune cell analysis

After CIBERSORT computation, the immune response of 22 tumor infiltrating immune cells can be estimated. Through the use of CIBERSORT R software package, the relative scores of 22 immune cell infiltration status in TCGA samples were determined. After gene set variation analysis algorithm estimation, we calculated the infiltration score of 16 immune cells and the activity of 13 immune related pathways.

Molecular and immunological characterization of different LRHG groups and comprehensive analysis of immune treatment

As immunotherapy such as immune checkpoint inhibitor for digestive cancers is attracting increasing attention[24],we explored the LRHG groups in different immune subtypes using the Wilcoxon test, which are based on the immune subtype profile of each TCGA sample downloaded from the UCSC Xena. Then we evaluated the TMB and MSI scores of each GC patient in the TCGA cohort. We preliminarily evaluated the widely used cancer immunotherapy biomarkers and other published immune related markers to determine the KRAS mutation of LRHG in GC patients, and then compared them with our LRHG. We can predict the prognosis of tumor patients receiving immune treatment by calculating tide score online (http://tide.dfci.harvard.edu/)[25,26].

Statistical analysis

R 4.2.3 (https://www.R-project.org)[27] was used for the statistical analysis in this study. Statistics were deemed significant at $P < 0.05$.

Results

LPS-related genes and function analysis

In differential expression analysis, the DEGs were screened (Fig. 2A). The volcano plot showed the 853 DEGs. During the comparison between tumor samples and normal tissue samples, we found that 617 genes were up-regulated and 236 genes were down regulated (Fig. 2B). Figure 2C, D displayed the GO terms and KEGG pathways. Through functional enrichment analysis, we could see that the most relevant signal pathway of LPS-related genes was "cytokine cytokine receptor interaction". The most enriched term in biological process[28] was molecular function[16], and the enriched terms in cellular component (CC) were "extracellular matrix organization", "receptor ligand activity", and "collagen-containing extracellular matrix", respectively.

Through WGCNA analysis of candidate genes, we extracted LPS related-genes. Through the application of scale-free network, we get that the optimal soft-thresholding power is 5 (Fig. 3A). We established a dendrogram of 853 co-expressed genes identified by DEG in the module (Fig. 3B). On this basis, we identified seven modules according to the average linkage hierarchical clustering and the optimal soft-thresholding ability (Fig. 3C). By studying the Pearson correlation coefficient between each module and the sample characteristics, we concluded that the blue and Turquoise modules were closely related to GC, and selected one of them for further analysis and research. Multivariate Cox regression analysis of OS in modular genes allows us to select independent prognostic genes for follow-up study (Fig. 3D). According to the PPI network (Fig. 3E), we chose LPS related hub genes (LRHG) by Degree (Fig. 3F), DMNC (Fig. 3G), EPC (Fig. 3H), MCC (Fig. 3I) and MNC (Fig. 3J), and united these genes (Fig. 3K).

Survival outcomes in different LRHG groups

Firstly, we explored the association and found that DUSP1 with BRIP1, IGFBP1, NR4A3, CLSPN, PDK4 and ANGPT2 ($P < 0.05$); ADRB3 with ANKRD1, CLSPN and EZH2 ($P < 0.01$); PDK4 and BRIP1 ($P < 0.05$); AGT and CDC25A ($P < 0.01$); IGFBP1 and BRIP1 ($P < 0.05$); as well as CLSPN with ANGPT2 and BRIP1 ($P < 0.05$) had close connections (Fig. 4A). The mutant frequency of alteration between these genes displayed in Fig. 4B with DANTS1 exhibited the biggest mutation (3%), followed by ANGPT2, BRIP1, IGFBP1 and NR4A3 (2%), indicating important roles of these genes in GC development. To select independent prognostic genes, we used lasso Cox analysis to construct a prognostic index for all cancer samples (Fig. 4C), and established a risk score signature with the following formula:

$$\begin{aligned}
\text{LRHG signature} =\ & \text{ANGPT2} * 0.159 - \text{BRIP1} * 0.211 + \text{GPX3} * 0.091 + \text{IGFBP1} * 0.056 \\
& + \text{ANKRD1} * 0.214 + \text{RGS2} * 0.117 + \text{AGT} * 0.023 + \text{DUSP1} * 0.074 \\
& + \text{PON1} * 0.262.
\end{aligned}$$

In addition, we artificially divided people into two groups. PCA revealed clear boundaries between the two groups (Fig. 4D). Kaplan–Meier analysis further disclosed that the higher the risk score of TCGA GC patients, the shorter the survival time of patients (Fig. 4E). In the validation set GSE84437 similar phenomenon was also seen (Fig. 4F, G). These results revealed that constructed signature by prognostic genes of LRHG could stratify GC patients into two groups with different survival state. The inherent differences between the two groups deserve further exploration.
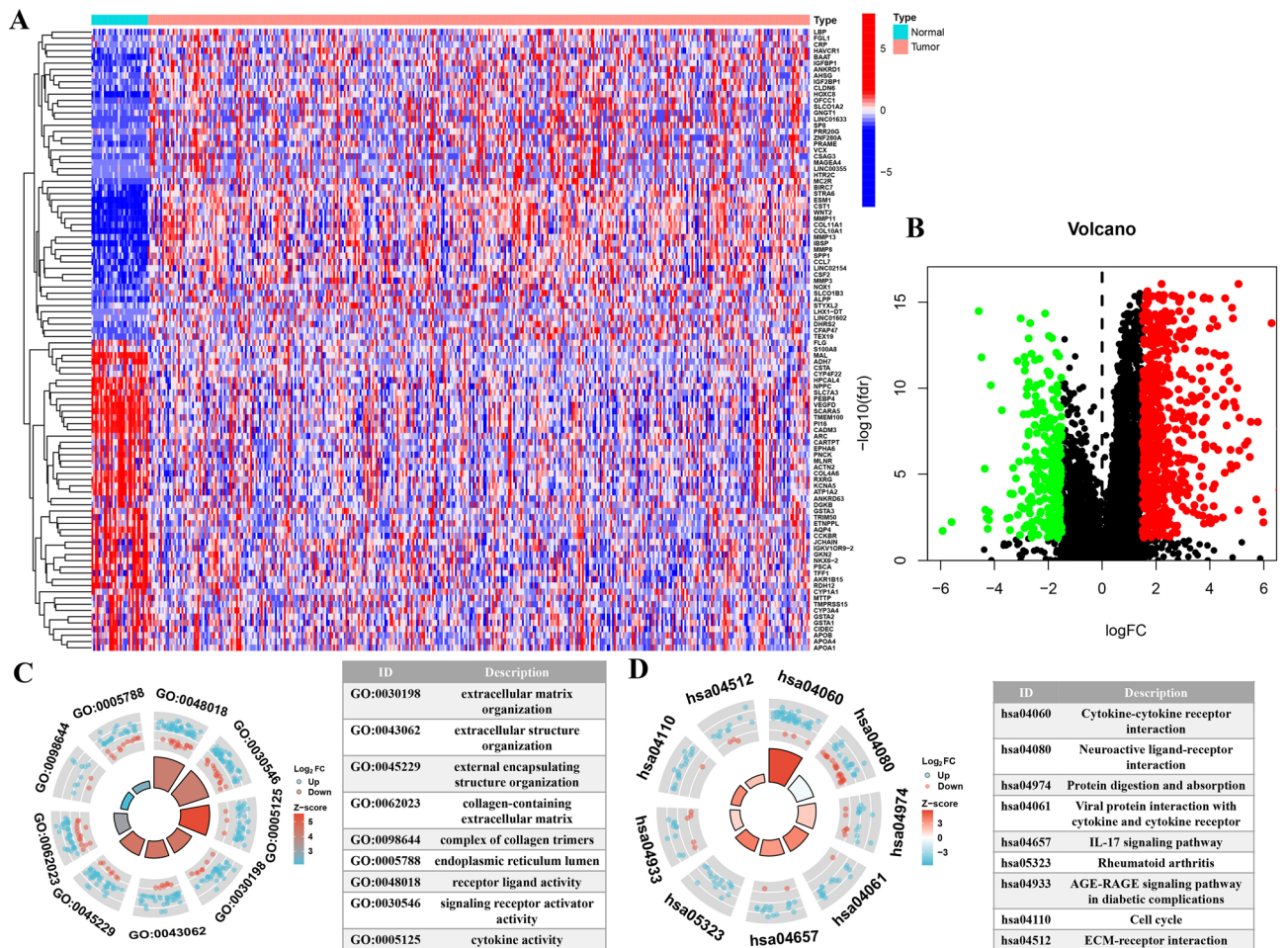
Figure 2. Screening for DEGs and enrichment analysis. (**A**) Heatmap of DEGs. (**B**) Volcano map of DEGs. (**C**) GO enrichment analysis for biological process, cellular component, molecular function respectively. (**D**) KEGG enrichment analysis.

Molecular characteristics of different LRHG groups

After Cox regression analysis, based on the univariate in Fig. 5A and the multivariate in Fig. 5B, we constructed the correlation between LRHG characteristics and clinical characteristics of GC patients. From ROC curve analysis, it can be concluded that LRHG signature and clinical features are able to predict the risk, age, gender, grade and Stage with AUCs of 0.690, 0.609, 0.559, 0.548 and 0.606 (Fig. 5C). ROC curve can also show that LRHG signature to predict the 1-, 3-, and 5-year with AUCs of 0.675, 0.655, and 0.690 (Fig. 5D). Through the analysis of the predictive nomogram, the overall survival rate of the whole cohort can be predicted relatively well compared with the ideal model (Fig. 6A, B). The ROC curve showed that LRHG signature nomogram and clinical feature to predict the risk, nomogram, age, gender, grade and Stage with AUCs of 0.683, 0.753, 0.602, 0.579, 0.535 and 0.617 (Fig. 6C). In order to construct the correlation between the LRHG characteristic nomogram results and clinical characteristics of GC patients, we used the univariate Cox regression analysis in Fig. 6D and the multivariate Cox regression analysis in Fig. 6E.

Immune characteristics of different LRHG groups

Recent studies have reported that LPS induces innate immune activation[29,30]. Therefore, we explored the composition of immune cells in different LRHG prognostic indicator groups, and also used Wilcoxon test for correlation analysis. This test compared the proportion of immune cells in different LRHG prognostic index groups (Fig. 7A). We found that T cells CD8, T cells CD4 memory activated, T cells follicular helper, monocytes (P < 0.001), T cells gamma delta, macrophages M1 and eosinophils (P < 0.05) suggested differences between different LRHG prognostic indicator groups (Fig. 7B). Then, some LRHG characteristics are used to define immune and molecular functions (Fig. 7C). Through the survival analysis, GC patients which have a lower score had a better outcome, with APC co inhibition (Fig. 7D), check point (Fig. 7E), cytolytic activity (Fig. 7F), inflammation promoting (Fig. 7G), MHC class I (Fig. 7H), T cell co inhibition (Fig. 7I), Type II IFN Response (Fig. 7J).
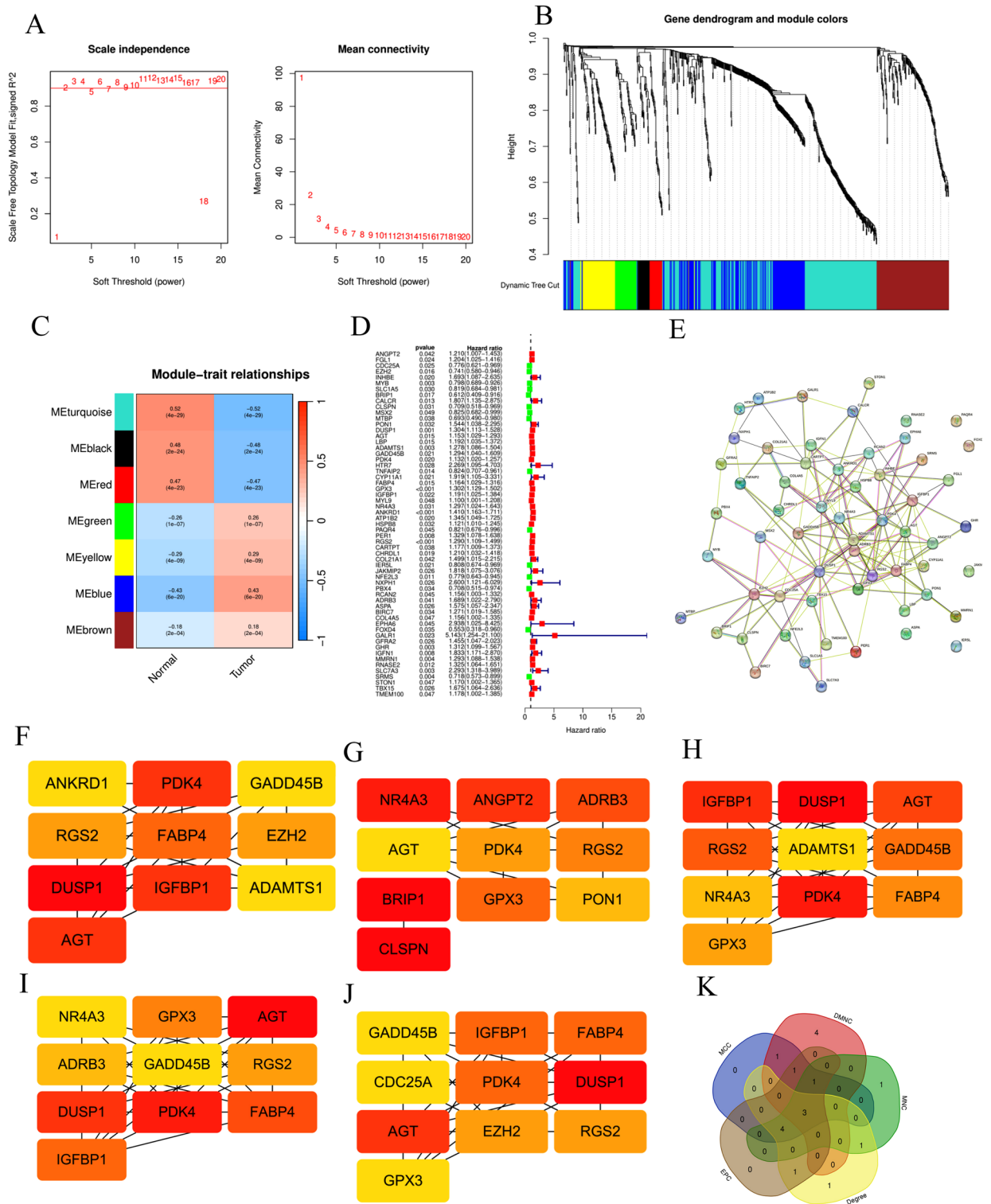
Figure 3. Analysis of LRHG modules by WGCNA. (**A**) Analysis of the scale-independence of various soft-thresholding powers. (**B**) Identification of co-expression modules. The branches of the tree diagram correspond to the seven different gene modules. (**C**) Correlation of gene modules with tissue type correlation scores. (**D**) Univariate Cox analysis of LRHG. (**E**) PPI network analysis of each gene in the blue and turquoise module. (**F**–**J**) Screen the top 10 genes by Degree, DMNC (**G**), EPC (**H**), MCC (**I**), MNC (**J**). (**K**) Veen map of Degree, DMNC, EPC, MCC and MNC.
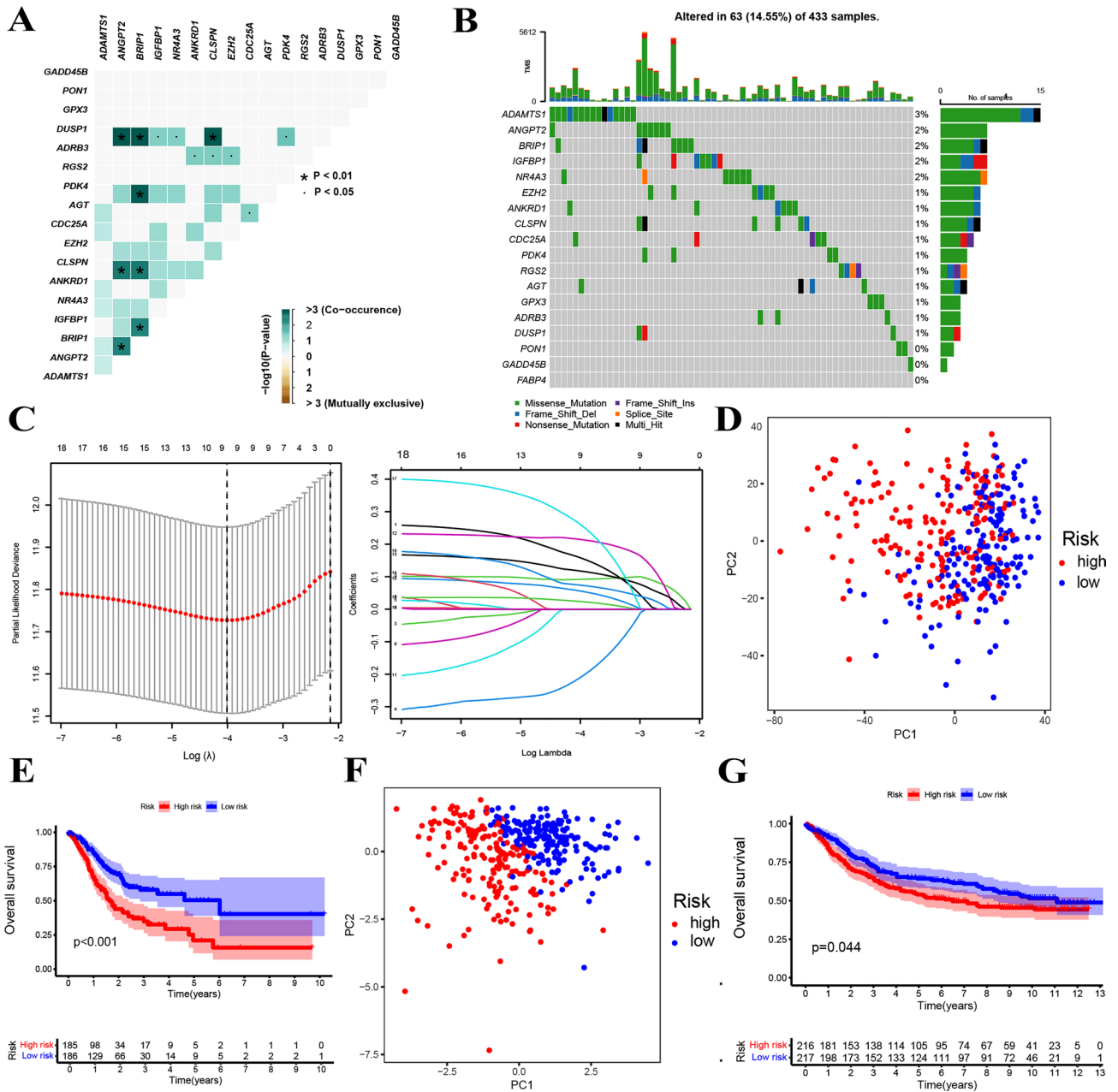
Figure 4. Identification of LRHG hub genes. (**A**) Correlation the LRHG genes. (**B**) Waterfall plot showing the genes mutation information. (**C**) The coefficient profile of prognostic genes by Lasso regression analysis. (**D**) PCA analysis with TCGA-STAD cohort. (**E**) Survival analysis by K–M curve for OS of TCGA-STAD patients. (**F–G**) Validation analysis in GSE84437dataset.

Benefits of immune treatment in different LRHG prognostic index groups

The TIDE algorithm can be used to evaluate the potential efficacy of immunotherapy in different groups of LRHG prognostic indicators. The higher the tide score, the higher the possibility of immune escape, which indicates that the patients are less likely to benefit from immunotherapy[31]. In our study, we found that there were differences in TIDE score dysfunction between the two groups, which confirmed that the group with low LRHG prognosis index may be more beneficial to immunotherapy (Fig. 8A). In addition, we found that the microsatellite instability (MSI) (Fig. 8B) and tumor mutation burden (TMB) (Fig. 8C) in the low LRHG prognostic index group were lower than those in the high LRHG prognostic index group. This finding suggests that patients with low LRHG prognostic index may benefit more from immunotherapy compared to patients with high LRHG prognostic index. In recent years, KRAS mutant subpopulations might also contribute to immune therapy failure[32]. For this purpose, we explore the LRHG prognostic index-low group had less KRAS mutant compared with the LRHG prognostic index-high group (Fig. 8D). Above all, LRHG signature shows the better result, which may benefit from TMB, MSI and KRAS mutant.
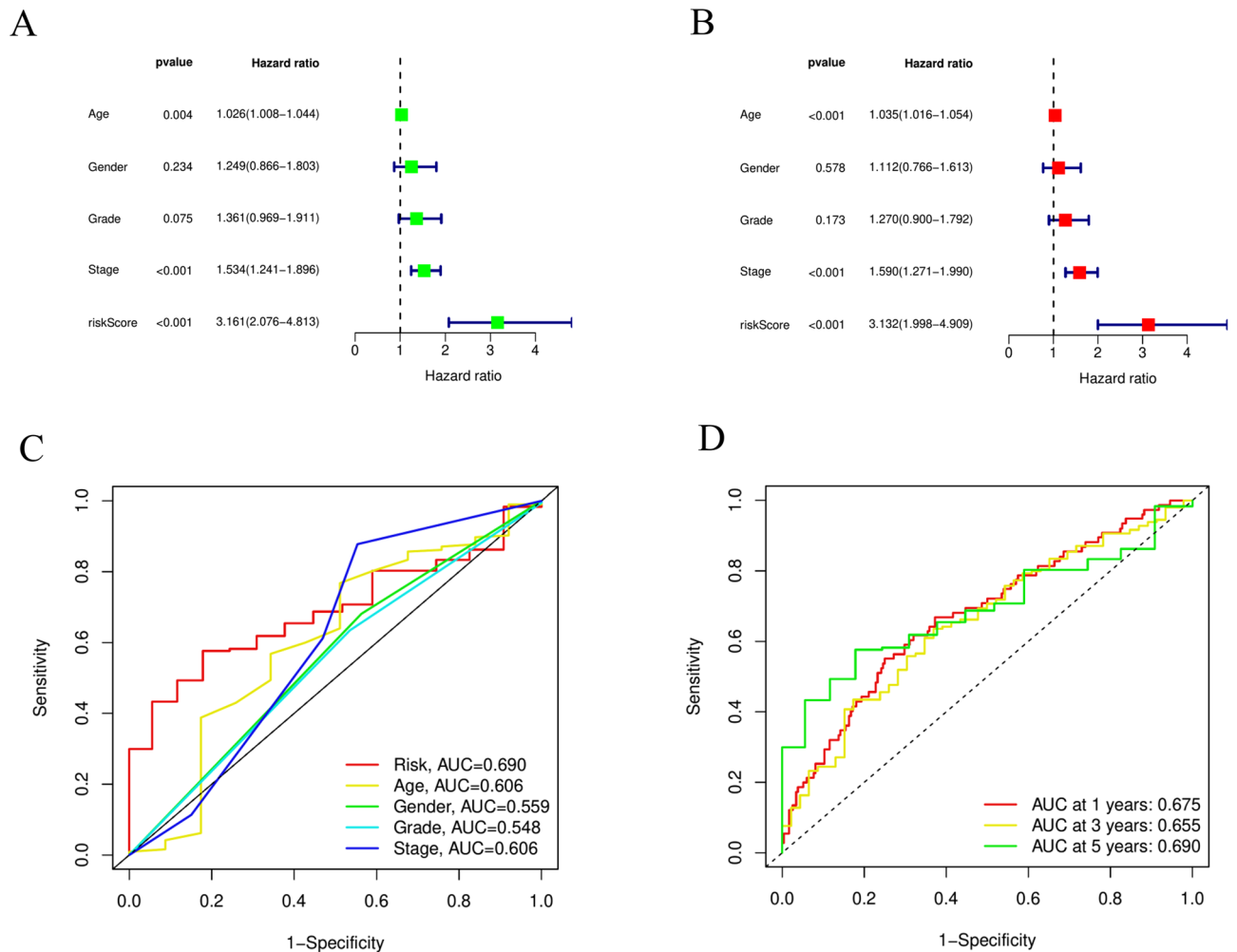
A

B

C

D

Figure 5. Establishing the association between the LRHG and clinical features. (**A**) Univariate and (**B**) multivariate analyses of the clinical features and risk score. (**C**) The risk score concordance indexes with clinical features in GC patients. (**D**) ROC curves in GC patients in 1-, 3-, and 5-year.

Discussion

GC is a common digestive tract tumor as a major cause of cancer-related mortality worldwide[28]. Most patients with gastric cancer are accompanied by local infiltration and distant metastasis. Currently, patients with advanced GC still have a 5-year survival rate of less than 5% after undergoing systemic treatment. So far, there is no accurate and comprehensive explanation of the exact pathogenesis of gastric cancer. To our knowledge, this is the first time that we applied widely used bioinformatics tools to explore the actions of LPS related genes on GC prognosis and immunotherapy.

Microorganisms are closely related to the development of GC, but due to the limitations of technology, the role of microorganisms in the development of GC has not been fully explained. At present, a small number of studies believe that the gastric microecology is basically unchanged in the progression of GC, and more studies believe that the gastric microorganisms is significantly changed in patients with GC. Therefore, one of the best methods may be to study and predict the role of gastric microbiome in gastric cancer patients through bioinformatics. We analyzed LPS related genes, which were obtained from the CTD database, and then combined with bioinformatics analysis to study the relevant molecular mechanisms of gastric microbiome involved in GC. Besides, we explored the immune treat in LRHG signature from TIDE score, TMB, MSI and common mutant KRAS.

For this study, the LRHG genes ANGPT2, BRIP1, GPX3, IGFBP1, ANKRD1, RGS2, AGT, DUSP1 and PON1 were obtained. In the event of bacterial LPS stimulation, silencing ANGPT2 could improve endoplasmic reticulum stress of intestinal epithelial cells via Notch signaling pathway[33]. Compared to the wild mice, RGS2 knockout mice show airways hyperreactivity and stiffer lungs in the LPS exposure[34]. Enhancing DUSP1 exerts cardioprotective effects by suppressing MAPK p38 and NF-κB pathway in LPS induced marked cardiac dysfunction model[35]. Paraoxonase 1 (PON1) was considered to have anti-inflammatory effect in previous studies[36]. Through promoting PON1 expression activated the ERK1/2 pathway, inhibiting the liver damage which is mediated by LPS[37]. However, the biological functions of LPS mediated stimulation of BRIP1, GPX3, IGFBP1, ANKRD1, and AGT are still unclear. The mechanism by which GPX3 restricts the development of colitis can be divided into two types, affecting the M2 macrophage subpopulation and promoting proliferation[38]. In addition, there are three
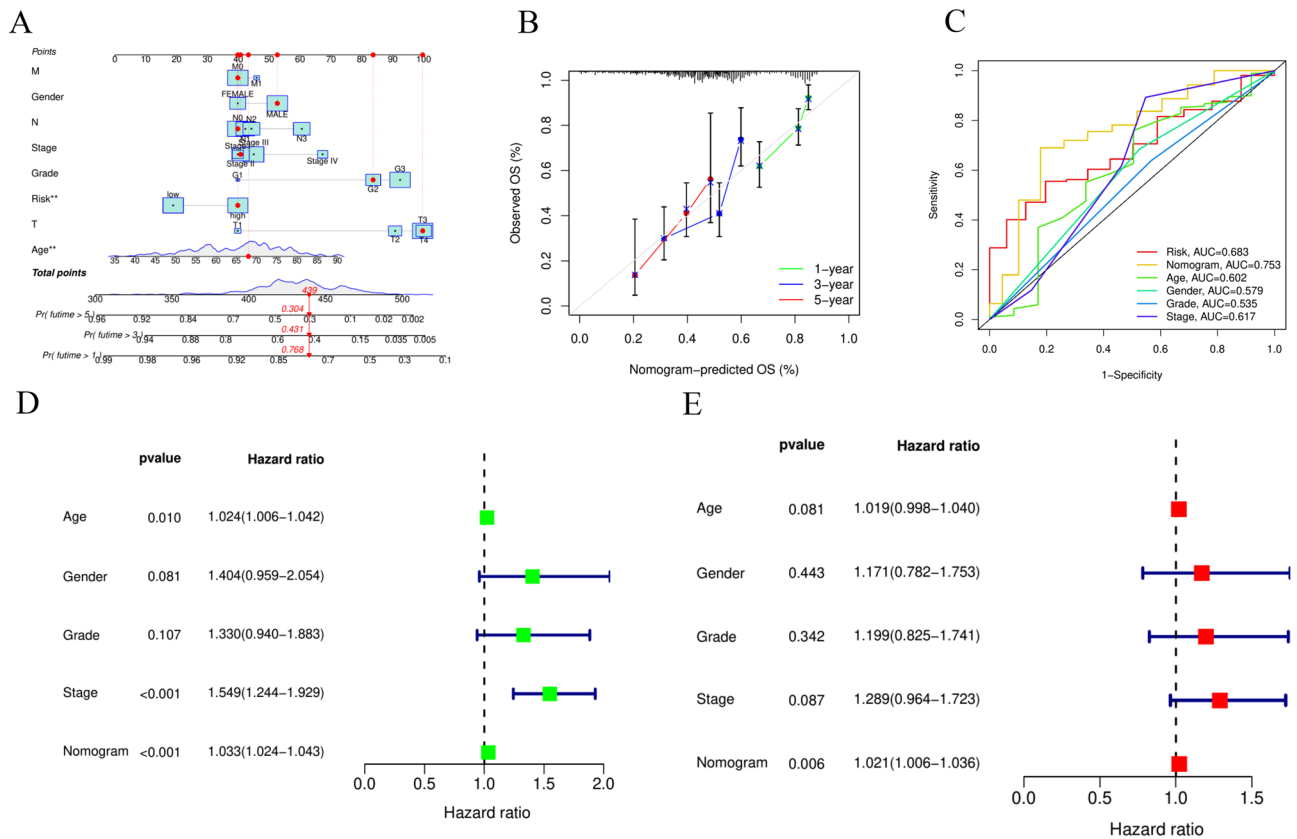
Figure 6. Establishing the nomogram. (**A**) The nomogram predicts the OS probability. (**B**) The calibration plot predicts the OS probability of the 1-, 3-, and 5-year. (**C**) ROC curves in GC patients with clinical features in GC patients. (**D**) Univariate and (**E**) multivariate analyses of the clinical features and nomogram.

ways for IGFBP1 to participate in tumor immunity, one is by mediating Cell surface receptor signaling pathway, and the others are by mediating cytokine production pathway or Monocyte signaling pathway[39]. In addition, AGT has the potential to serve as a biomarker for GC prognosis and immune infiltration[40].

The research and clinical application of immune cells have pointed out new directions for cancer treatment, but patients with GC have not benefited from it. The main reason is that during the onset of GC, there is a significant immunosuppressive microenvironment, which is caused by the presence of the infection factor Helicobacter pylori, disrupting the immune balance of the gastric mucosa[36]. Therefore, we also delved into the association with LRHG and the patterns of immune infiltration in GC. The results show that there is a significant relationship between LRHG and the degree of immune cell infiltration in the tumor in T cells CD8, T cells CD4 memory activated, T cells follicular helper, monocytes, T cells gamma delta, macrophages M1 and eosinophils, especially in T cells. This study reveals that T cell CD8 can be directly activated in a manner independent of MHC class I, which has potential value for the treatment of MHC class I deficient cancers[41].In recent studies, the completion of T cells cannot be separated from the crosstalk between T cells, intestinal cells, and other immune cells. T cells can also play an important role in the immune cell population of intestinal tissue[42].

We investigated mutations in related genes in different LRHG groups to further understand the immunological properties of the LRHG group. We are also exploring LRHG features related to the immune microenvironment
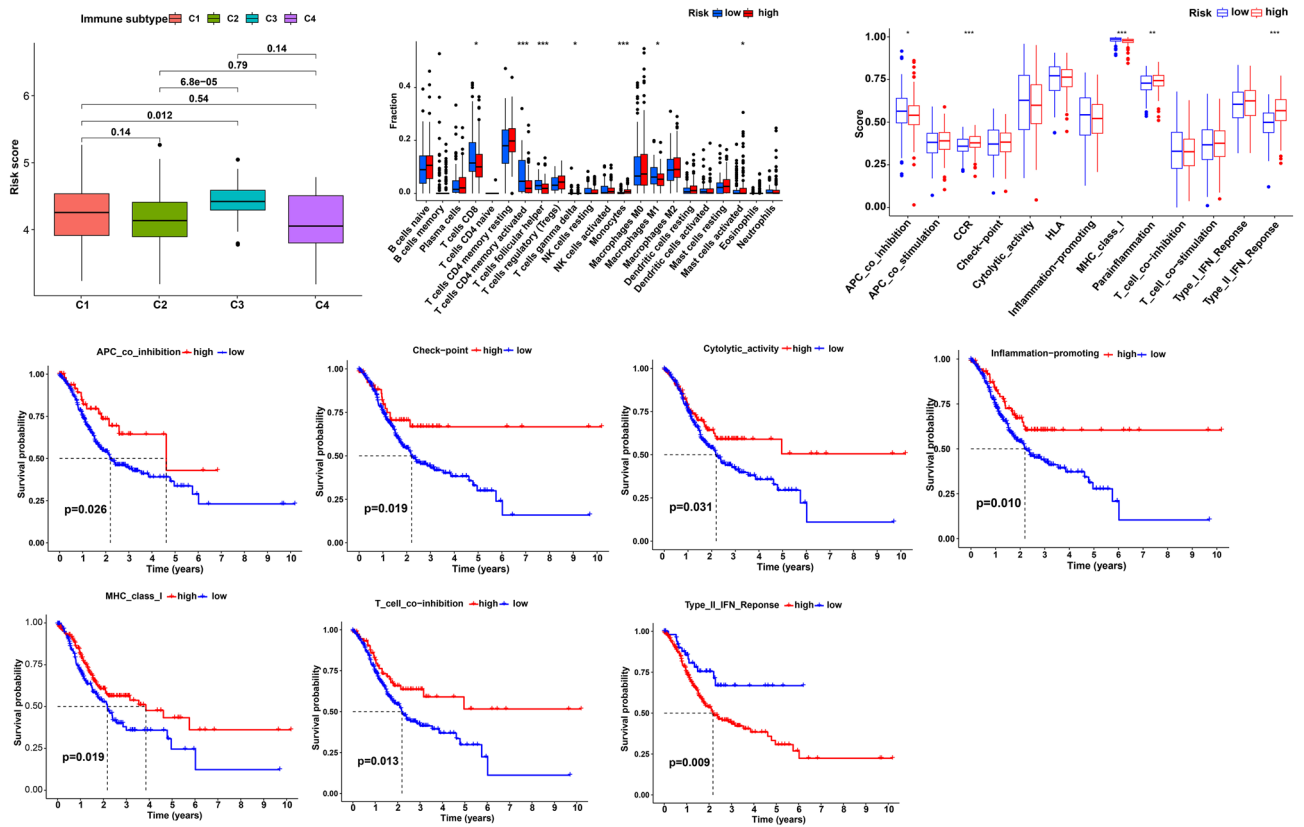
Figure 7. Immune cells infiltration between high-risk groups and low-risk groups. (**A**) LRHG signature analysis in different immune subtypes. (**B**) CIBERSORT showed the correlation between different groups. (**C**) Comparison of the enrichment scores of 13 immune-related pathways in LRHG group. (**D–J**) Kaplan–Meier survival analysis of the correlation of immune cell abundance ratios in the LRHG group.

in our research. Tumor Mutation Burden (TMB) is a potential predictive indicator for cancer immunotherapy, and research has also identified limitations that hinder TMB in clinical practice[43,44]. The proportion of LRHG low group in TMB is higher than that of LRHG high group, indicating a higher correlation between LRHG low group and higher reactivity of immune checkpoint inhibitors. In addition, microsatellite instability (MSI) is the first approved pan cancer biomarker that can be used to guide the treatment of immune checkpoint inhibitors[45]. We also explored MSI and LRHG signatures during the research process. The common mutation in GC is KRAS, which plays an important role in immunotherapy for solid tumors and colorectal cancer[46,47]. Whether there is a relationship between KRAS mutations and immunotherapy in GC patients remains unproven. In our study, the tidal fraction is used to predict the probability of LRHG signature model. The analysis methods used in this process are TMB and MSI analysis. TIDE score is a newly developed prediction method to predict the response of immunotherapy. TIDE score can be more accurate than TMB or PD-L1 expression and can be used to predict the clinical effectiveness of anti-pd1 and anti-CTLA4 treatment[48]. The high tide score in the LRHG high group suggested greater immune escape and poor outcome of high LRHG score, which may be closely related to TMB, MSI and KRAS mutations.

The limitations of this study must be acknowledged, although the findings of this study provide new innovations for elucidating the molecular mechanism of GC. First, due to technical limitations, the current animal and cell experiments have no ability to comprehensively detect the composition and structural characteristics of GC microorganisms. Second, on the basis of reviewing the expression data from public databases, we identified only genes associated with LPS from CTD. Last but not least, these LRHG can indeed control GC, but the basic mechanism of their control is still not clarified, such as upstream long non-coding RNA and microRNA[49] of these LRHG deserve to be studied. In the future, several advanced models such as integrating mathematical modeling[50], SWATH-MS-based network modeling[51], with experimental analysis as well as phase separation technique[52] could be applied to explore the mechanism of our screened 9 LRHGs. What's more, a new learning algorithm, graph convolutional network with graph attention network[53] will also be carried out to validate our finding. In short, we need conduct comprehensive research on their biological functions through more and more rigorous techniques, algorithm and experiments. More research in the future will enable the microbiome in GC to provide diagnostic tools for GC, microbiome based prevention and immunotherapy.
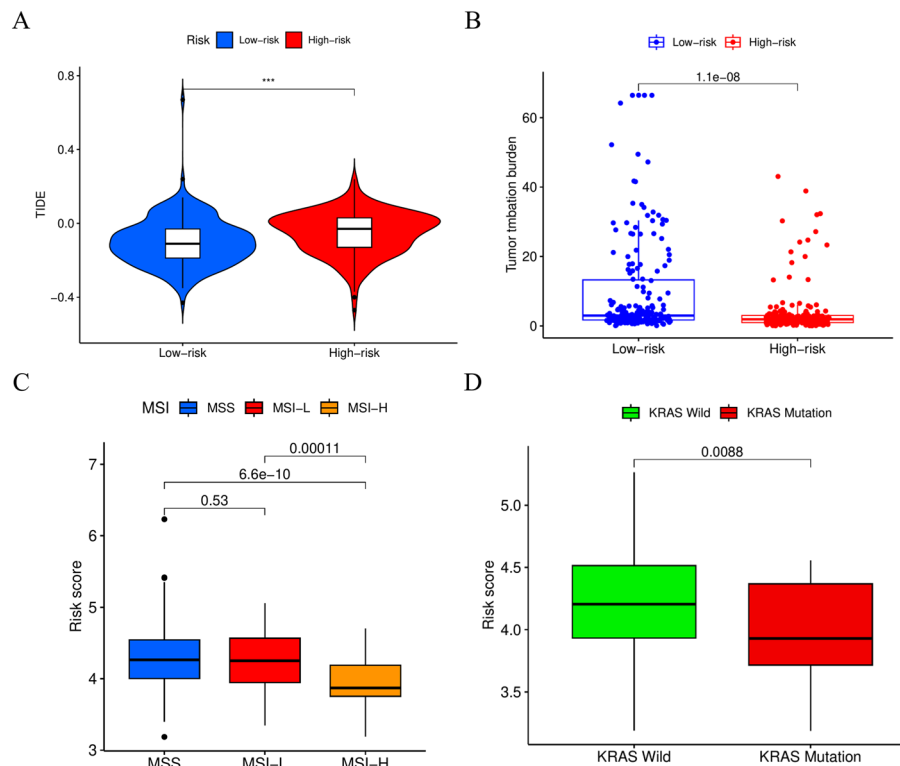
Figure 8. Immune response to immune therapy and the prognostic value. (**A**) TIDE score for different LRHG group. (**B**) TMB analysis, (**C**) MSI between the LRHG group. (**D**) KRAS mutant between in the LRHG high- and low-group.

Data availability

Publicly available datasets were analyzed in this study. These data can be found here: all relevant raw data used in the study can be accessed from TCGA (https://portal.gdc.cancer.gov/repository).

References

1. Siegel, R. L. *et al.* Cancer statistics, 2023. *CA Cancer J. Clin.* **73**(1), 17–48 (2023).
2. Guan, W. L., He, Y. & Xu, R. H. Gastric cancer treatment: Recent progress and future perspectives. *J. Hematol. Oncol.* **16**(1), 57 (2023).
3. Xia, C. *et al.* Cancer statistics in China and United States, 2022: Profiles, trends, and determinants. *Chin. Med. J. (Engl.)* **135**(5), 584–590 (2022).
4. Li, P. *et al.* Comparison of gastric cancer survival after R0 resection in the US and China. *J. Surg. Oncol.* **118**(6), 975–982 (2018).
5. Joshi, S. S. & Badgwell, B. D. Current treatment and recent progress in gastric cancer. *CA Cancer J. Clin.* **71**(3), 264–279 (2021).
6. Sasako, M. Progress in the treatment of gastric cancer in Japan over the last 50 years. *Ann. Gastroenterol. Surg.* **4**(1), 21–29 (2020).
7. Song, Z. *et al.* Progress in the treatment of advanced gastric cancer. *Tumour Biol.* **39**(7), 1010428317714626 (2017).
8. Li, M. *et al.* Surgery combined with molecular targeted therapy successfully treated giant esophageal gastrointestinal stromal tumor. *Oncologie* **24**(2), 349–356 (2022).
9. Liu, H. D., Wang, L. J. & Xu, Z. K. Diagnosis, prevention and treatment of post-operative rare complications after radical gastrectomy for gastric cancer. *Zhonghua Wei Chang Wai Ke Za Zhi* **26**(2), 138–143 (2023).
10. Jafari-Sales, A. *et al.* The presence of human papillomavirus and Epstein-Barr virus infection in gastric cancer: A systematic study. *Oncologie* **24**(3), 413–426 (2022).
11. Png, C. W. *et al.* Mucosal microbiome associates with progression to gastric cancer. *Theranostics* **12**(1), 48–58 (2022).
12. Bessede, E. & Megraud, F. Microbiota and gastric cancer. *Semin. Cancer Biol.* **86**(Pt 3), 11–17 (2022).
13. Di Iorio, B. R. *et al.* Urea and impairment of the gut-kidney axis in chronic kidney disease. *G Ital. Nefrol.* **34**, 2017 (2017).
14. Kelly, J. *et al.* Alterations of mucosa-attached microbiome and epithelial cell numbers in the cystic fibrosis small intestine with implications for intestinal disease. *Sci. Rep.* **12**(1), 6593 (2022).
15. Pomie, C. *et al.* CX3CR1 regulates gut microbiota and metabolism. A risk factor of type 2 diabetes. *Acta Diabetol.* **58**(8), 1035–1049 (2021).
16. Chen, X. *et al.* Maintenance of gastrointestinal glucose homeostasis by the gut–brain axis. *Curr. Protein Pept. Sci.* **18**(6), 541–547 (2017).
17. Ciccia, F. *et al.* The role of the gastrointestinal tract in the pathogenesis of rheumatic diseases. *Best Pract. Res. Clin. Rheumatol.* **30**(5), 889–900 (2016).
18. Choi, H. W., Lee, K. W. & Kim, Y. H. Microbiome in urological diseases: Axis crosstalk and bladder disorders. *Investig. Clin. Urol.* **64**(2), 126–139 (2023).
19. Masenga, S. K. *et al.* Recent advances in modulation of cardiovascular diseases by the gut microbiota. *J. Hum. Hypertens.* **36**(11), 952–959 (2022).

20. Aragones, G. *et al.* Gut microbiota-derived mediators as potential markers in nonalcoholic fatty liver disease. *Biomed. Res. Int.* **2019**, 8507583 (2019).
21. Hu, J. *et al.* Correlation between altered gut microbiota and elevated inflammation markers in patients with Crohn's disease. *Front. Immunol.* **13**, 947313 (2022).
22. Marizzoni, M. *et al.* A peripheral signature of Alzheimer's disease featuring microbiota–gut–brain axis markers. *Alzheimers Res. Ther.* **15**(1), 101 (2023).
23. Kanehisa, M. *et al.* KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**(D1), D545-d551 (2021).
24. Poureau, P.-G. & Metges, J.-P. Fundamentals of digestive cancers immunology, especially gastric and hepatocellular carcinomas (Fondamentaux de l'immunologie des Cancers Digestifs) (gastriques et hépatocellulaires). *Oncologie* **23**(1), 47–59 (2021).
25. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**(10), 1550–1558 (2018).
26. Blanche, P., Dartigues, J. F. & Jacqmin-Gadda, H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat. Med.* **32**(30), 5381–5397 (2013).
27. Null, R. C. T. R. *et al.* R: A language and environment for statistical computing. *Computing* **1**, 12–21 (2011).
28. Thrift, A. P., Wenker, T. N. & El-Serag, H. B. Global burden of gastric cancer: Epidemiological trends, risk factors, screening and prevention. *Nat. Rev. Clin. Oncol.* **20**(5), 338–349 (2023).
29. Yahfoufi, N. *et al.* Pubertal consumption of *R. badensis* subspecies *acadiensis* modulates LPS-induced immune responses and gut microbiome dysbiosis in a sex-specific manner. *Brain Behav. Immun.* **107**, 62–75 (2023).
30. Kong, X. N. *et al.* LPS-induced down-regulation of signal regulatory protein alpha contributes to innate immune activation in macrophages. *J. Exp. Med.* **204**(11), 2719–2731 (2007).
31. Li, D. *et al.* Identification and validation of EMT-related lncRNA prognostic signature for colorectal cancer. *Front. Genet.* **12**, 723802 (2021).
32. Queiros, P. *et al.* KRAS mutations in microsatellite instable gastric tumours: Impact of targeted treatment and intratumoural heterogeneity. *Virchows Arch.* **467**(4), 383–392 (2015).
33. Dai, L. *et al.* Angiopoietin-2 silence alleviates lipopolysaccharide-induced inflammation, barrier dysfunction and endoplasmic reticulum stress of intestinal epithelial cells by blocking Notch signaling pathway. *Bioengineered* **12**(1), 8116–8124 (2021).
34. George, T. *et al.* A bronchoprotective role for Rgs2 in a murine model of lipopolysaccharide-induced airways inflammation. *Allergy Asthma Clin. Immunol.* **14**, 40 (2018).
35. Xin, Y. *et al.* Inhibition of miR-101-3p protects against sepsis-induced myocardial injury by inhibiting MAPK and NF-kappaB pathway activation via the upregulation of DUSP1. *Int. J. Mol. Med.* **47**(3), 20 (2021).
36. Aharoni, S., Aviram, M. & Fuhrman, B. Paraoxonase 1 (PON1) reduces macrophage inflammatory responses. *Atherosclerosis* **228**(2), 353–361 (2013).
37. Ventro, G. *et al.* Molecular impact of omega 3 fatty acids on lipopolysaccharide-mediated liver damage. *J. Pediatr. Surg.* **51**(6), 1039–1043 (2016).
38. Barrett, C. W. *et al.* Tumor suppressor function of the plasma glutathione peroxidase gpx3 in colitis-associated carcinoma. *Cancer Res.* **73**(3), 1245–1255 (2013).
39. Xu, T. *et al.* MMP9 and IGFBP1 regulate tumor immune and drive tumor progression in clear cell renal cell carcinoma. *J. Cancer* **12**(8), 2243–2257 (2021).
40. Wu, F. *et al.* AGT may serve as a prognostic biomarker and correlated with immune infiltration in gastric cancer. *Int. J. Gen. Med.* **15**, 1865–1878 (2022).
41. Wang, X. *et al.* MHC class I-independent activation of virtual memory CD8 T cells induced by chemotherapeutic agent-treated cancer cells. *Cell Mol. Immunol.* **18**(3), 723–734 (2021).
42. Suzuki, T. *et al.* Gut gammadelta T cells as guardians, disruptors, and instigators of cancer. *Immunol. Rev.* **298**(1), 198–217 (2020).
43. Ravaioli, S. *et al.* Are we ready to use TMB in breast cancer clinical practice?. *Cancer Immunol. Immunother.* **69**(10), 1943–1945 (2020).
44. Jiao, X. D. *et al.* Tumor mutation burden in Chinese cancer patients and the underlying driving pathways of high tumor mutation burden across different cancer types. *Ann. Transl. Med.* **8**(14), 860 (2020).
45. Tian, J. *et al.* Genomic characteristics and prognosis of lung cancer patients with MSI-H: A cohort study. *Lung Cancer* **181**, 107255 (2023).
46. Liu, H. *et al.* Mutant KRAS drives immune evasion by sensitizing cytotoxic T-cells to activation-induced cell death in colorectal cancer. *Adv. Sci. (Weinh)* **10**(6), e2203757 (2023).
47. Zheng, W. *et al.* SHP2 inhibition mitigates adaptive resistance to MEK inhibitors in KRAS-mutant gastric cancer through the suppression of KSR1 activity. *Cancer Lett.* **555**, 216029 (2023).
48. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**(1), 218 (2016).
49. Wang, W. *et al.* Predicting the potential human lncRNA–miRNA interactions based on graph convolution network with conditional random field. *Brief Bioinform.* **23**(6), bbac463 (2022).
50. Li, X. *et al.* Caspase-1 and Gasdermin D afford the optimal targets with distinct switching strategies in NLRP1b inflammasome-induced cell death. *Research (Wash D C)* **2022**, 9838341 (2022).
51. Li, X. *et al.* RIP1-dependent linear and nonlinear recruitments of caspase-8 and RIP3 respectively to necrosome specify distinct cell death outcomes. *Protein Cell* **12**(11), 858–876 (2021).
52. Xu, F. *et al.* Specificity and competition of mRNAs dominate droplet pattern in protein phase separation. *Phys. Rev. Res.* **5**(2), 023159 (2023).
53. Sun, F., Sun, J. & Zhao, Q. A deep learning method for predicting metabolite-disease associations via graph neural network. *Brief Bioinform.* **23**(4), bbac266 (2022).

Author contributions

(I) Conception and design: T.Y. and S.Z.; (II) Administrative support: H.C. and J.G.; (III) Provision of study materials: T.Y.; (IV) Collection and assembly of data: S.H. and Y.M.; (V) Data analysis and interpretation: T.Y.; (VI) Manuscript writing: all authors; (VII) Final approval of manuscript: all authors.

Competing interests

The authors declare no competing interests. The results obtained in this study are partially based on data produced by the open source of TCGA dataset, so there are no ethical issues and other conflicts of interest.

Additional information

Correspondence and requests for materials should be addressed to J.C. or J.G.

Reprints and permissions information is available at www.nature.com/reprints.