An Immune-related Gene Signature of Predicting Lymph Node Metastasis and Responses to Immune Checkpoint Inhibitor Treatment in Gastric Cancer Patients

Ke-wei Wang
Jiangnan University
Mei-dan Wang
Jiangnan University
Hua Wang
Jiangnan University
Jian-feng Huang
Jiangnan University
Xiao-long Wu
Jiangnan University
Qin-fang Yuan
Jiangnan University
Ling-jun Li (lljzyy@hotmail.com)
Nanjing University of Chinese Medicine https://orcid.org/0000-0002-1615-8692

Research

Keywords: Lymph node metastasis, LNM, Gastric cancer, LGSGC, Immune, Checkpoints, Immunotherapy, Prognostic

Posted Date: December 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1099712/v1

Abstract

Background

Immune-related genes have been used as prognostic markers in multiple types of tumors. We aimed to develop an immune-related gene signature for predicting individual lymph node metastasis in gastric cancer (GC) patients, characterize the molecular and immune profiles of different risk patients and assess the potential value of this signature identifying patients with response to immune checkpoint inhibitor (ICI) treatment.

Methods

A total of 1338 GC patients from a training dataset, three external silico validation datasets and an external clinical dataset were included in this study. The microarray analysis was used to detect differentially expressed immune-related genes (DEIGs) between lymph node metastatic and non-lymph node metastatic gastric cancer tissues. Subsequently, we built a lymph node metastasis gene signature for gastric cancer (LGSGC), and then classified patients into low-risk and high-risk groups according to the LGSGC. Moreover, we implemented association analysis for this signature and the prognosis, molecular characteristics, immune profiles and the response of ICI treatment in different risk GC patients.

Results

he receiver operating characteristics (ROC) curve analysis (an area under curve [AUC] values of 0.85) showed that the LGSGC could distinguish lymph node metastatic patients from non-lymph node metastatic patients in the training dataset. Additionally, compared to low-risk group, high-risk group exhibited worse overall survival (hazard ratio [HR]=2.42) in the training dataset. Robust diagnostic and prognostic clinical ability of the LGSGC were successfully validated in four validation datasets. Next, the high-risk patients were characterized by active cancer and immune response-related pathways, high TP53, CSMD3 and FAT4 mutation rate, high infiltration of Neutrophils, M1 Macrophages, M0 Macrophages, M2 Macrophages, T cells gamma delta and T cells follicular helper, more abundant check point, more aggressive inflammation and Type I IFN response, and more benefit from ICI. On the contrary, low-risk patients were characterized by active cancer and tumor metabolism-related pathways, low TP53 mutation rate, high infiltration of Mast cells resting, NK cells resting, Plasma cells and T cells CD4 memory resting, and less benefit from ICI therapy. Of note, we also validated the LGSGC, which identified patients having response to ICI treatment with an AUC value of 0.71 in an advanced GC dataset and an AUC value of 0.64 in an IMvigor210 dataset.

Conclusions

The LGSGC is a reliable indicator to distinguish LNM in GC and could discriminate the prognosis, molecular characteristics, immune profiles and the response of ICI treatment in different risk groups. This signature may provide a reference for treatment decisions for different risk GC patients.

Introduction

Gastric cancer (GC) patients can be treated by radical operation. For patients without lymph node metastasis (LNM), the radical operation increased the incidence of surgical trauma and complications but the endoscopic sub-mucosal dissection (ESD) and the endoscopic mucosal resection (EMR) became the most commonly used treatments [1-4]. Therefore, accurate identification of LNM in GC, especially early GC patients, was of great guiding value for comprehensive treatment. It was found that the overall 5-year survival was over 90% among early-stage GC patients, 97% - 100% among GC patients without LNM and 80% - 85% among GC patients with LNM, respectively [5-7]. The LNM has become a crucial factor of poor prognosis of GC patients. However, big data analytics suggested that LNM only accounted for 16%-24% of early GC patients [4, 5]. Additionally, compared to preoperative imaging with limited accuracy to identify LNM in early GC, tracing sentinel lymph nodes widely used by clinicians during surgical procedure were not worthy of promotion and application because of lack of unified international standard [8]. Hence, there was an urgent to develop accurate indicators for individual lymph node metastatic risk assessment among GC patients, helping to optimize the treatment planning for patients.

In recent years, the immune checkpoint inhibitor (ICI) therapy has showed good curative effect of multiple types of cancers [9-14]. The main immune checkpoint molecules targeted by immunotherapy mainly included cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), programmed death-ligand 1 (PD-L1), programmed cell death-ligand 2 (PD-L2) and programmed cell death protein 1 (PD-1). The anti-PD1 therapy has been considered to be a promising treatment for GC patients with recurrence and metastasis [9, 15]. However, it was difficult to accurately identify patients with high response to ICI treatment [16]. Interestingly, the identification of LNM using immune related markers provided a new choice, and accurate lymph node metastatic risk assessment in GC patients could provide a reference for surgical treatment or immunotherapy [17-20]. So, it was necessary to accurately identify GC patients with LNM based on the molecular immune markers including tumor mutation burden (TMB), immune-related genes and tumor microenvironment (TME).

In this study, we tried to build a signature, consisted of immune-related genes from microarray gene data, for individual lymph node metastatic risk assessment among GC patients. Firstly, we identified immune-related genes set that were related to LNM in GC, and then constructed a final immune-related genes signature of detecting LNM and predicting prognosis. Secondly, we described the molecular and immune characteristics of both risk groups. Finally, we assessed the performance of the lymph node metastasis gene signature for gastric cancer (LGSGC) identifying patients with response to ICI treatment. The results show that LGSGC is a superior predictor of LNM in GC and provides a new inference for tailored surgical treatment and immunotherapy.

Methods

Patients and datasets

The study included a training dataset: the GSE15459 dataset (https://www.ncbi.nlm.nih.gov/), and three external silico validation datasets: the GSE62254 dataset (https://www.ncbi.nlm.nih.gov/), the GSE84437 dataset (https://www.ncbi.nlm.nih.gov/) and the TCGA dataset (https://portal.gdc.cancer.gov/), and a clinical dataset from affiliated hospital of Jiangnan University. Demographic characteristics and clinical-related variables were downloaded from above Websites and corresponding articles. Inclusion criteria for GC patients in affiliated hospital of Jiangnan University was complete clinical data and pathological tissue specimens. GC was staged based on the eighth edition American Joint Committee on Cancer (AJCC) staging system for GC. Altogether 107 patients GC selected in the clinical dataset underwent surgery between January 2019 and December 2020 in affiliated hospital of Jiangnan University, Wuxi, China. 107 formalin fixation and paraffin embedding (FFPE) samples for the clinical dataset were also selected. This study received the approval of the ethical committees of affiliated hospital of Jiangnan University. All experiments were implemented in full compliance with the Declaration of Helsinki. In this study each patient selected provided written informed consent.

We acquired the names of candidate immune-related genes by visiting to two databases: InnateDB (HTTPS://www.innateDBdb.com/) and ImmPort (HTTPS://www.immport.org/shared/home), and mutation information of candidate immune-related genes by visiting to the cBioPortal database. Furthermore, we acquired the data of somatic mutation for GC by visiting to the TCGA Datasets (https://portal.gdc.cancer.gov/). Four subtypes of GC, proliferative, metabolic, invasive and unstable were provided (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE15459) by Lei Z et al from Duke-National University of Singapore Graduate Medical School, Singapore. Two molecular subtypes of GC, epithelial phenotype (EP) and mesenchymal phenotype (MP) were provided (https://www.nature.com/articles/s41467-018-04179-8) by Oh SC et al from University of Texas, Houston, USA. Other four subtypes of GC, TP53 functional loss (MSS/TP53−) and intact TP53 activity (MSS/TP53+), epithelial-to-mesenchymal transition (EMT) and microsatellite instability (MSI) were provided (https://www.nature.com/articles/nm.3850#Sec19) by Cristescu R et al from Merck Research Laboratories, Boston, USA. To explore the efficacy of the LGSGC predicting response of cancer patients to the PD-1 inhibition, four datasets including a metastatic GC dataset, an IMvigor210 dataset, a bladder cancer dataset and a melanomas dataset were used. The metastatic GC dataset was provided https://www.nature.com/articles/s41591-018-0101-z) by Kim ST et al from Sungkyunkwan University School of Medicine, Seoul, Korea. The IMvigor210 dataset was provided (https://www.nature.com/articles/nature25501) by Mariathasan S et al from Lund University, Sweden. The bladder cancer dataset was provided (https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002309) by Snyder A et al from Memorial Sloan Kettering Cancer Center, New York, USA. The melanomas dataset was provided (https://www.sciencedirect.com/science/article/pii/S009286741630215X) by Jiang P et al from University of California, Los Angeles, USA.

RNA extraction and quantitative real-time polymerase chain reaction (qRT-PCR) assays

We used total RNA isolation kits to extract total RNA from the microdissected FFPE specimens. Nanodrop microspectrophotometer was used to detect the purity and concentration of total RNA. Based on 2 μg of total RNA the High-Capacity cDNA Reverse Transcription Kit (Invitrogen, Carlsbad, CA, USA) was use to synthesize complementary DNA (cDNA). The TB Green® Premix Ex Taq™ (Tli RNaseH Plus) was applied to qRT-PCR. The specific steps were described as follows: 2 μ L cDNA was taken for PCR amplification (10 min at 95°C and then 40 cycles of three steps consisting of 2 s at 95°C, 20 s at 60°C, and 10 s at 70°C). Then Fusion curves were collected. Each sample was repeated for 3 times and averaged. The internal-control gene was 5S. The relative quantification of mRNA was calculated by $2^{-\triangle\triangle CT}$ method.

Selection of candidate immune-related genes related to LNM in GC

Individual gene expression profiles of GC came from the GSE62254 dataset and the GSE84437 dataset. Differentially expressed genes (DEGs) between cancer tissues with LNM and cancer tissues without LNM were analyzed using the edgeR package in R. Next, those DEGs with the same direction of change were intersected with immune-related genes coming from ImmPort and InnateDB. Finally, we selected DEIGs for detecting LNM in GC. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and Gene ontology (GO) enrichment of candidate DEIGs were conducted using the Metascape software or the clusterProfiler package.

Selection of the 14 immune-related genes detecting LNM in GC

To make the prediction model robust and generalized, we used the GSE15459 dataset as the training dataset to build diagnostic model of LNM in GC based on above candidate immune-related genes. Moreover, to gain immune-related hub genes for detecting LNM in GC, the least absolute shrinkage and selection operator (LASSO) regression was conducted. Next, multivariate logistic regression analysis was conducted to develop the LGSGC according to the immune-related hub genes selected. Each sample gained a risk score which was calculated using the expression levels of immune-related hub genes and multivariate logistic regression coefficients. To assess the efficacy of the LGSGC detecting LNM in GC, ROC analysis using risk score of each sample as an independent variable was implemented in GC patients. The diagnostic power of the LGSGC detecting LNM in GC was validated in four external validation datasets: the GSE62254 dataset, the GSE84437 dataset, the TCGA dataset and the clinical dataset.

The LGSGC and survival of GC patients

According to the cutoff thresholds derived by the Youden index, risk scores of samples being continuous variables were regrouped into dichotomous variables (low-risk group and high-risk group). The Cox regression analysis or Landmark analysis while Kaplan-Meier (KM) survival curves were crossing was used to explore the prognostic potential of the LGSGC on different risk patients. We also used three external validation datasets, the GSE62254 dataset, the GSE84437 dataset and the TCGA dataset to validate the prognostic potential of the LGSGC.

Molecular and immune profiles and ICI therapy in different risk groups

Gene set enrichment analysis (GSEA) or single sample GSEA (ssGSEA) was implemented to identify the signaling pathways in which the gene sets were involved in different risk groups. We summarized, analyzed, annotated and visualized somatic variants of different risk groups by the Maftools package of R. To assess the relationship between the LGSGC and TMB and factors of the ICB therapy including CTLA-4, PD-L1, PD-L2 and PD-1, correlation analysis was conducted. We also assessed the performance of the LGSGC detecting LNM in GC by comparation with CTLA-4, PD-L1, PD-L2, PD-1 and TMB.

The LGSGC and 22 kinds of immune cells of TME

The immune profiles of GC patients were characterized. The gene expression profiles matrix as input, the CIBERSORT estimated abundance of 22 kinds of immune cells in GC. Moreover, to explore the effect of different-risk scores and clinicopathological factors on immune cells composition in GC, the relative proportions of 22 kinds of immune cells and the distribution of clinicopathological factors in both risk groups were analyzed.

Next, other four datasets, a metastatic GC dataset, a IMvigor210 dataset, a bladder cancer dataset and a melanomas dataset were used to assess the efficacy of the LGSGC predicting response of cancer patients to the PD-1 inhibition [21-25]. Based on the standardized expression of all genes, tumor immune dysfunction and exclusion (TIDE) scores were gained by visiting to the Dana Farber Cancer Institute & Harvard University(HTTP://tide.dfci.harvard.edu/).

Statistical analysis

In this study the following statistical methods were used: independent t-test for comparison of gene expression differences, univariate logistic regression and multivariate logistic regression for screening factors affecting LNM; the spearman correlation analysis for continuous variables; the Wilcoxon test for comparison of risk scores in different risk groups; ROC analysis for evaluating the effect of statistical models predicting LNM in GC; Cox regression analysis or landmark analysis while KM survival curves were crossing for comparing the survival of different-risk patients; the Chi-square test for categorical variables. The cutoff thresholds were derived by the Youden index from this 14-gene signature model. All statistical calculations were conducted in R (version 4.0.3). All P-values were two sided, and a value of less than 0.05 was deemed significant.

Results

Selection of DEIGs

Altogether 3841 genes in the GSE62254 dataset and 4772 genes in the GSE84437 were found to be differentially expressed between lymph node positive (LNP) patients and lymph node negative (LNN) patients (Figure S1A and B, Table S1 and 2). By intersecting above DEGs with the immune-related genes on the basis of InnateDB and ImmPort, 54 hub genes with the same direction of change were obtained

(Figure 1A). 54 hub genes were significantly associated with 338 GO terms and 40 KEGG pathways (Table S3 and 4) using functional enrichment analysis. We further showed the top six GO terms and 20 KEGG pathways in Figure S1C and S1D.

Construction of a lymph node metastasis gene signature for gastric cancer

To narrow down 54 differentially expressed immune-related hub genes affecting LNS of GC patients, a series of 14 genes were identified by executing LASSO logistic regression in the training dataset (Figure 1B). Further, the final model, a lymph node metastasis gene signature for gastric cancer (LGSGC), that consisted of the 14 immune-related genes by multivariate logistic regression was constructed (Figure 1C). Based on final model the formula of risk score: (1.191×expression of PMAIP1) -(0.777× expression of TRIM65) - (0.616× expression of RFXAP)+(0.647× expression of RB1) - (0.693× expression of CASP1)+ (0.747× expression of SECTM1) - (1.022×expression of UBE2V2)+(0.756× expression of UBE2W) + (0.779×expression of GNAI3) - (0.528×expression of DDX3X) - (0.821×expression of CMTM6)+ (0.751×expression of FSTL1) - (0.635×expression of SOCS2) - (0.734×expression of NLRP13). Compared to LNN patients, LNP patients had more risk score in the training dataset (p=3.2e−11), the GSE62254 dataset (p=1.3e−09), the GSE84437 dataset (p=2.22e−16), the TCGA dataset (p=2.9e−16) and the clinical dataset (p=2.3e−05) (Figure 2A1-E1). ROC analysis showed that the risk scores could distinguish LNP patients from LNN patients in the training dataset (AUC =0.85, 95% CI 0.79–0.91; Figure 2F).

To further explore the functional and molecular profiles of 14 immune-related genes identified, enrichment analysis was conducted by the Metascape software, showing that a total of 30 functional enrichment was involved in five related-apoptotic processes, four related-cell growth regulation and two related- autophagy, which were significantly correlated with the development and progression of cancer [26] (Figure 3A, Table S6, $P < 0.05$). Strikingly, only two pathways enriched by 14 immune-related genes included viral carcinogenesis and PID P53 downstream pathway, which were well known to play a key role in inducing tumor cell proliferation and migration (Figure 3B, Table S7, $P < 0.05$). Enrichment network showed that biological processes including apoptotic signaling pathway, protein polyubiquitination, homeostasis of number of cells, regulation of autophagy and positive regulation of proteolysis was generally shared among 14 genes (Figure 3C). For example, PMAIP1 played a centripetal role linking four enriched biological processes, and RB1, CASP1, UBE2V2 and DDX3X played a centripetal role linking three enriched biological processes, respectively.

Validation of the LGSGC

All patients in the several datasets (the GSE62254 dataset, the GSE84437 dataset, the TCGA dataset and the clinical dataset) obtained the risk scores that were calculated on the basis of the above formula. Expectedly, the LGSGC could distinguish LNP patients from LNN patients in the GSE62254 dataset (AUC =0.80, 95% CI: 0.73–0.88; Figure 2F), the GSE84437 dataset (AUC =0.82, 95% CI: 0.76–0.87), the TCGA dataset (AUC =0.77, 95% CI: 0.72–0.82) and the clinical dataset (AUC =0.76, 95% CI: 0.66–0.85). Next, in order to explore the efficacy performance of the LGSGC detecting LNS in early GC patients, ROC analysis was implemented in participants with pathological T1 and T2, showing that the LGSGC could distinguish

LNP patients from LNN patients in the training dataset (AUC=0.83, 95% CI: 0.72–0.94; Figure 2G), the GSE62254 dataset (AUC=0.79, 95% CI: 0.70–0.88), the GSE84437 dataset (AUC=0.83, 95% CI: 0.71–0.95), the TCGA dataset (AUC=0.78, 95% CI: 0.67–0.88) and the clinical dataset (AUC=0.79, 95% CI: 0.63–0.95).

Next, after adjusting for other clinicopathologic features such as subtypes, Union for international cancer control (UICC) stage, Lauren classification, gender and age, the LGSGC was an independent diagnostic biomarker of detecting LNM using multivariate logistic regression analysis in the training dataset (odds ratio [OR] =3.14, 95% CI: 1.93–5.90; Figure 1D). The LGSGC as an independent diagnostic factor of detecting LNM was validated in the GSE62254 dataset (OR=7.59, 95% CI: 3.07–23.76; Figure 2SA), the GSE84437 dataset (OR=6.37, 95% CI: 3.76–11.33; Figure 2SB), the TCGA dataset (OR=7.90, 95% CI: 3.95–17.75; Figure 2SC) and the clinical dataset (OR=2.13, 95% CI: 1.46–3.32; Figure 2SD). Furthermore, UICC stage, Neoplasm histology, T Stage, mesenchymal phenotype (MP) and epithelial phenotype (EP) were significantly associated with LNM in GC (Figure 1D and Figure 2S). Intriguingly, the combination model of the LGSGC, clinicopathologic factors such as subtypes, UICC stage, Neoplasm histology, T Stage, mesenchymal phenotype (MP) and epithelial phenotype (EP) had significantly superior diagnostic accuracy for LNM compared to the LGSGC in the GSE15459 dataset (AUC=0.92, 95% CI: 0.88–0.97; Figure 2H), the GSE62254 dataset (AUC=0.96, 95% CI: 0.92–1.00), the GSE84437 dataset (AUC=0.83, 95% CI: 0.78–0.88), the TCGA dataset (AUC=0.95, 95% CI: 0.93–0.98) and the clinical dataset (AUC=0.81, 95% CI: 0.72–0.89).

Association between the LGSGC and the survival of GC patients

LNM proved to have a significant association with poor survival of GC patients, we further estimated prognostic potential of the LGSGC for GC patients. Each patient was categorized into low- or high-risk group based on the cutoff thresholds, derived by the Youden index from the LGSGC model. Reassuringly, high-risk group exhibited worse prognosis than low-risk group in the training dataset (HR=2.42, 95% CI: 1.55–3.78; Figure 4A), consistent with the result in the GSE62254 dataset (HR=1.68, 95% CI: 1.17–2.40; Figure 4B) and the GSE84437 dataset (HR =1.51, 95% CI: 1.12–2.05; Figure 4C). The Landmark analysis showed that only those patients with survival more than 29 months from high-risk group exhibited worse prognosis (HR=4.31, 95% CI: 1.16–15.93; Figure 4D) in the TCGA dataset. Furthermore, high-risk group had shorter disease-free survival (DFS) compared with low-risk group (HR=1.78, 95% CI: 1.20–2.65; Figure 4SA) in the GSE62254 dataset and high-risk group with survival more than 44 months exhibited significantly shorter DFS compared with low-risk group (HR=6.65, 95% CI: 1.16-38.03; Figure 4SB) in the TCGA dataset. Because of limited dead samples (n=7), we did not conduct survival analysis in different risk groups from the clinical dataset.

Molecular characteristics of different risk groups

GSEA showed that the gene sets were significantly enriched in 18 pathways, which mainly included cancer and immune response-related pathways (Figure 5A and Table S5, $p < 0.05$) in high-risk group and 20 pathways which mainly included cancer and tumor metabolism-related pathways (Figure 5B and Table S5, $p < 0.05$) in low-risk group.

To understand biological immunological nature of different risk groups, we analyzed somatic mutations of GC patients from the TCGA dataset. It was found that missense mutation was the most frequent kind of mutations, followed by nonsense mutation and frameshift deletion in both risk groups (Figure 5C). The mutations of DNAH5, OBSCN and HMCN1 were very much in evidence in low-risk group while the mutations of CSMD3 and FAT4 were very much in evidence in high-risk group. Of note, the mutation rates of SYNE1, LRP1B, MUC16, TP53 and TTN were higher than 20% in both risk groups.

Immune characteristics of different risk groups

It was found that low-risk group had more abundant T cells CD4 memory resting, Plasma cells, NK cells resting, Mast cells resting and B cells naïve in the training dataset while high-risk group had more abundant T cells follicular helper, T cells gamma delta, M0 Macrophages, M1 Macrophages, M2 Macrophages and Neutrophils (Figure 6A). Next, we analyzed the relation between risk score and some clinicopathological factors such as subtypes, gender, Lauren-classification, UICC stage, alive time, alive status, age, T stage and LNS. It was found that the distribution of GC subtypes (proliferative, metabolic, invasive and unstable), UICC stages, survival time, lymph node status and alive status had significant differences in both risk groups (Figure 6B). The ssGSEA scores showed that low-risk group had more plentiful Mast cells, APC co inhibition, NK cells and B cells while high-risk group had more plentiful Type I IFN response, check point, APC co stimulation, Macrophages, Inflammation promoting, Para inflammation, MHC class I and Treg (Figure 6S1).

Immune cells can affect the survival in multiple types of tumors. So, we assessed the prognostic efficacy of certain immune cells in the training dataset. Figure 6S2 indicated those patients with higher scores of APC co inhibition, T cells CD4 memory activated, B cells naïve, B cells, MHC class I, iDCs, Tfh, NK cells, NK cells resting, Mast cells activated, TIL and Plasma cells had longer ($p < 0.05$) survival, while patients with higher scores of NK cells activated, M2 Macrophages, Dendritic cells resting, M1 Macrophages, T cells follicular helper, APC co stimulation, Para inflammation, T helper cells, Treg and Type I IFN response had shorter ($p < 0.05$) survival.

Association between LGSGC-GC subtypes and other molecular and immune-GC subtypes

Previous a study suggested that four subtypes of GC including proliferative, metabolic, invasive and unstable had significant differences of molecular and genetic characteristics and response to therapy [19]. Figure 7A showed a significant difference (p=0.014, chi-square test) in the distribution of four subtypes of GC between two risk groups. Low-risk group had more metabolic-subtypes samples and less unstable-subtypes samples while high-risk had less metabolic-subtypes samples and more unstable-subtypes samples. Next, a novel molecular classification of GC according to genomic alterations, recurrence pattern and prognosis was reported.

Figure 7B showed a significant difference (p=0.003, chi-square test) in the distribution of four subtypes of GC in two risk groups. Low-risk group had more EMT and MSI subtypes and less MSS/TP53+ subtypes while high-risk group had less EMT and MSI subtypes and more MSS/TP53+ subtypes in the TCGA

dataset. However, significantly different distribution of four subtypes of GC including MSS/TP53-, MSS/TP53+, EMT and MSI subtypes was not found in the GSE15459 dataset and the GSE62254 dataset (Figure 7S, p=0.071 and 0.752, chi-square test, respectively). Although Oh SC et al identified two molecular subtypes of GC, epithelial phenotype (EP) and mesenchymal phenotype (MP) which were corelated with LNM in GC (Figure 2SA), the distribution of two subtypes did not have a significant difference in two risk groups. (Figure 7C, p=0.366, chi-square test) [27].

Association between the LGSGC and cancer ICI treatment response

To assess the efficacy of the LGSGC identifying patients with response to ICI treatment, 43 patients with advanced GC administered pembrolizumab or nivolumab from a dataset study were selected [21, 22]. It was found that only 32.4% (11/43) of subjects achieved partial response and complete response to pembrolizumab or nivolumab. Next, according to the RNA-seq data provided by Noh MG, et al, each participant gained the risk score calculated by the 14 immune-related gene signature model [22]. Those patients who achieved response to PD-1 inhibition had higher risk scores than those who did not achieve response to PD-1 inhibition (Figure 8A, p=0.039). Strikingly, the LGSGC had significantly diagnostic accuracy for responders from non-responders to PD-1 inhibition with an AUC value of 0.71 (95% CI: 0.53-0.89; Figure 8B).

Other three datasets were also used to assess the efficacy of the LNMGC predicting response of cancer patients to PD-1 inhibition [23-25]. The standardized risk scores according to the LNMGC in both risk groups were significantly different in the IMvigor210 dataset (p=0.028, Figure 8C), the bladder cancer dataset (p=0.00014, Figure 8F) and the melanomas dataset (p=0.034, Figure 8H). Additionally, patients with high-risk scores exhibited worse prognosis compared with that with low-risk scores in the IMvigor210 dataset (n=326) (p = 0.013, Figure 8D) and in the bladder cancer dataset (n=25) (p = 0.029, Figure 8G). However, Figure 8I did not show significantly different survival of both risk groups in the melanoma dataset (p = 0.14, Figure 8I). The cause of this result might be statistical limitation in the melanomas dataset, which included only 27 patients, among whom eight dead patients from high-risk group and four dead patients from low-risk group. Of note, the LNM in GC had significantly diagnostic accuracy for selecting patients with an AUC value of 0.64 (95% CI: 0.56-0.72; Figure 8E), who had response to PD-1 inhibition, from the IMvigor210 dataset.

The cancer immunotherapy response for advanced GC patients was further assessed by TIDE score from the web application for response prediction using gene expression profiles [14]. The results showed that patients with high-risk scores achieved lower TIDE scores compared with patients with low-risk scores in the training dataset, suggesting that patients with low-risk scores were more likely to obtain immune evasion compared with patients high-risk scores (Figure 8J). Besides, MSI score and T cell exclusion score were significantly different in both groups, while T cell dysfunction score were not significantly different.

Association between the LGSGC and cancer ICI treatment-related gene expression

The ICB therapy that embraced CTLA-4, PD-L1, PD-L2, PD-1 and TMB blocking agents provided significant clinical benefits for GC patients. Next, we explored the relationship between the LGSGC and CTLA-4, PD-L1, PD-L2 and PD-1 expression in the GSE63354 dataset (Figure 8SA-D).

The relationship between the LGSGC and TMB was assessed in the TCGA dataset (Figure 8SE). It was found that risk score of the LGSGC yielded significant positive association with expression of CTLA-4, PD-L1 and PD-L2 (r = 0.18, *p*=0.0022; r = 0.25, *p*=1.5e-05; r=0.15, *p*=0.0077; Figure 8SA, B and D), but did not yield significant associations with the expression of PD-1 (r = 0.06, *p*= 0.30; Figure 8SC) and TMB (r = 0.049, *p*= 0.37; Figure 8SE). Meanwhile, the efficacy of the LGSGC detecting LNM of GC patients comparation with CTLA-4, PD-L1, PD-L2 and PD-1 was explored by ROC analysis (Figure 8SF). The results showed that the LGSGC had bigger AUC value of 0.80 (95% CI: 0.73-0.88) than known immunological biomarkers, CTLA-4 (AUC=0.56, 95% CI: 0.45-0.65), PD-1 (AUC=0.57, 95% CI: 0.48-0.67), PD-L2 (AUC=0.60, 95% CI: 0.51-0.70) and PD-L1 (AUC=0.57, 95% CI: 0.47-0.68). Moreover, the combination of the LGSGC and PD-L1 with AUC value of 0.84 (95% CI: 0.76-0.92) did not have significantly superior diagnostic accuracy for LNM compared with the LGSGC (Z = -1.8617, p-value = 0.06264, DeLong test; Figure 8SF).

Discussion

The ESD and the EMR are the two most commonly used treatments for early GC patients without LNM [4]. Meanwhile, it is extremely important to clarify lymph nodes status for GC patients treated by radical operation, providing guidance for the extensive dissection of lymph nodes. Patients with advanced metastatic GC may need excellent internal medical therapeutic schedule. The immunotherapy, which has been widely used in multiple types of tumors, can provide a personalized treatment basis for advanced GC patients [9, 20]. Therefore, there was an urgent need to develop a novel signature which not only detected the LNM in GC to provide a reference for radical operation but also predict the survival of GC patients to be good for personalized immunotherapy regimens.

The LNM was known as an important risk factor of poor prognosis of GC patients [3, 28-30]. There were also significant associations between clinical factors including tumor differentiation, tumor invasion depth, lymphatic infiltration and tumor size and the TNM. Unfortunately, these above clinical factors still did not accurately predict LNM. Interestingly, to some extent the occurrence of LNM was affected by immune-related genes and immune microenvironment [31-33]. At present, many immune-related genes have been found to be related to the LNM in GC. Nevertheless, the molecular and immune characteristics of LNM in GC were still unclear. So, we assessed the mechanism of immune-related genes used to build the prediction model of LNM and prognosis of GC patients. The GSE6228 dataset and the GSE84437 dataset were used to select 54 immune-related genes that could detect LNM. Subsequently, we established the LGSGC of LNM detection by LASSO logistic regression in the GSE15459 dataset. We further validated the diagnostic ability of the LGSGC predicting LNM using three external silico validation datasets and an external clinical dataset from the affiliated hospital of Jiangnan University. It was found that the LGSGC of LNM detection in each dataset with AUCs of more than 0.75 were superior than a preoperative model of predicting LNM based on 26 genes with AUCs of 0.69-0.74 [30]. Another study

showed that a nomogram of combination 23-gene signature and Lauren classification achieved an AUCs of 0.78-0.92[3]. Similarly, the LGSGC combination with clinicopathologic factors such as subtypes, UICC stage, Neoplasm histology, T Stage, MP and EP were more accuracy of predicting the LNM of GC patients than the LGSGC.

The detection of LNM in early GC patients was always difficult in the radical operation in GC. Previous a study suggested that a 15-gene classifier of predicting the LNM in GC patients from the training and the validation datasets with AUC values of 0.77 and 0.74, respectively [1]. In this result, the LGSGC for detection of LNM in early GC achieving AUC value of more 0.76 in each dataset, showing that the LGSGC was superior to these with currently used modalities in the clinical settings.

Expected, several processes and pathways such as related-apoptotic process, related-cell growth regulation, related-autophagy, viral carcinogenesis and PID P53 downstream pathway which were closely related to tumorigenesis and metastasis were enriched by the 14 immune-related genes. Several studies showed that autophagy played a major role in some physiological and pathological processes of multiple types of tumors [34-38]. Although autophagy-related signaling pathways including the P53 pathway, the AMPK/mTOR and PI3K/Akt/mTOR pathways might contribute to the development of GC cell invasion and metastasis, the role of autophagy remained unclear [39]. Even anticancer drugs for GC including autophagy inducers and autophagy inhibitors are being developed [40-42]. It is well known that the viral carcinogenesis mainly includes Epstein–Barr virus which can cause chronic atrophic gastritis, dysplasia, gastric intestinal metaplasia and adenocarcinoma. Recent studies have shown that GC patients with Epstein–Barr virus infection could benefit from immunotherapy [43]. In this study enrichment network showed that the PMAIP1, linking four enriched biological processes, could determine whether a cell committed to apoptosis and was correlated with development of multiple types of cancers including lung cancer cells[44, 45], metastatic melanoma[46] and bladder cancer cells[47]. Moreover, another study showed that PMAIP1 was correlated with unfavorable prognosis in pancreatic ductal adenocarcinoma [48]. Similarly, the expression of PMAIP1 was found to have a significant negative association with the prognosis survival of GC patients (HR=3.29, 95% CI 1.66–7.13).

At present, the research showed that immune-related genes had appreciable impacts on the survival of people with breast cancer [49], leukemia [13], lung cancer [12], ovarian cancer [50], colorectal cancer [51], hepatocellular carcinoma [52], melanomas [53], osteosarcomas [54], pancreatic cancer [55] and gastric cancers [30, 56], and even multiple combination-models on the basis of immune-related genes were accurate indicators of the survival of cancer patients [28, 57]. The LGSGC was a robust predictor of the LNM and the prognosis of GC patients, and could classify patients into high-risk or low-risk group. Moreover, high-risk group exhibited shorter survival compared with low-risk group. However, ROC analysis showed that AUCs of the LNMGC on overall survival of GC patients from each dataset with time were less than 0.7 expect for training dataset. On the contrary the five-year predictive AUCs of prognosis of GC patients in a study with ten immune-related gene signature and in another study with six immune-related gene signature were 0.85 [58]and 0.79[59] , respectively. The reason may be that our signature was constructed based on some genes which could predict LNM in GC. Although LNM can lead to metastasis

and poor prognosis of GC, its direct prediction of patient survival was limited [60]. Consistent with this result was also noted in several previous studies [5, 17, 29]. So, it was almost impossible to accurately predict the survival of GC patients solely according to genes. Other factors including immune cell infiltration, tumor microenvironment and immune detection that could affect the occurrence, development, invasion and metastasis of tumors should be added in prediction model of LNM and the survival of GC patients.

The TME played complex roles in tumorigenesis, progression, metastasis, recurrence, and drug resistance. For example, multiple immune cells including NK cells, B cells, macrophages and all T cell subsets were associated with the outcome of treatment for cancer patients [20]. Memory T cells and cytotoxic T cells were found to be related to the long-term survival in different tumor origins and tumor cell types, while the effects of other subsets of macrophages, NK cells, B cells and some helper T cells on prognosis were related to tumor types and stages. As a key component of humoral immunity, B lymphocytes including plasma cells, B cells memory and B cells naïve have been found to play a positive immunomodulatory role through various regulatory ways. B lymphocytes can not only actively regulate the anti-tumor immune process by acting as antigen presenting cells, secreting a variety of cytokines and producing anti-tumor antibodies, but also play a role in immunosuppressive regulation by inhibiting the proliferation of immune-activated T cells [61]. However, our study showed that low-risk group had more abundant plasma cells and B cells naïve than high-risk group, and patients with low proportions of three types of B cells had significantly worse prognosis than patients with high proportions. Similarly, previous research showed that tumor progression and recurrence could be inhibited by eliminating B lymphocytes [62]. Moreover, chemosensitivity of cancer patients could be improved by eliminating B lymphocytes. Consistent with this result was also noted in another study, which showed that B cells mainly infiltrated tissue cancer and were significantly correlated with longer survival of GC patients [63].

Several T cell subsets including activated CD4 memory T cells, CD8 cells, natural killer T cells, follicular helper T cells, T cells regulatory (Tregs) and activated NK cells could impact on the occurrence, progression and response to immunotherapy in multiple types of cancers [9, 15, 16, 64, 65]. In this study, high-risk group had more abundant T cells gamma delta and T cells follicular helper while low-risk group had more abundant T cells CD4 memory resting. We also found that different kinds of T cells had inconsistent effects on the survival of patients. For example, Tregs was significantly related to shorter survival of GC patients while T cells follicular helper (Tfh) was significantly associated with better prognosis of GC patients. Recent a study showed that the gastric tumor sites were infiltrated by Th17 and Treg, and the imbalance between Th17 and Treg might promote the metastasis of advanced GC [66]. Tumor associated-macrophages proved to play a complex role in the gastric carcinogenesis. Different macrophage subsets had an influence on the progression and prognosis of GC [18, 67-69]. Tumor-associated macrophages were divided into M1 and M2 phenotypes. M1 Macrophages were classically activated macrophage with strong antigen presentation ability and the ability to promote immunity and kill tumors. M2 Macrophages was a substitute activated macrophage, which had immunosuppression, promotes tumor matrix remodeling and angiogenesis, and was related to tumor progression [10, 70, 71]. In this study, high-risk group had more abundant M1 macrophages, M0

macrophages and M2 macrophages which were found to be significantly associated with longer survival of GC patients. The effect of macrophage phenotypes on tumors were actually controversial, even some researchers believed that M2 Macrophages could be reprogrammed into M1 macrophages to inhibit tumor growth and proliferation through therapeutic procedures [72, 73].

Moreover, high-risk patients had lower TIDE scores, lower T cell exclusion score and higher MSI score than low-risk patients in the training dataset. Lower TIDE scores, lower T cell exclusion score and higher MSI score have been found to be benefit from ICI therapy in various types of cancers [20, 60, 71, 72]. However, given that the overall response rate to PD-1 therapy was only 20% to 40%, it was so important to select effective immune-related biomarkers for GC immunotherapy [11, 74]. Interestingly, the LNMGC could identify GC patients who had respondence to PD-1 inhibition with more accuracy in advanced GC dataset [21]. At present, the MSI detection was a general approach to screen advanced GC patients who could benefit from ICI therapy. Although microsatellite instability had a high incidence of endometrial cancer and colorectal cancer, this incidence was only 11.68%~33.82% in GCs, failing to detect GC patients who could be sensitive to immunotherapy. In the study the LNMGC may provide a reference for classifying GC patients who were sensitive to immunotherapy besides MSI detection. The surrounding microenvironment of the tumor, closely correlated with the biological function of tumors, consisted of many types of cells such as endothelial cells, fibroblasts, lymphocytes and macrophages, and many soluble molecules including cytokines, growth factors, antibodies, proteases, chemokines, metabolites and various enzymes, and extracellular matrix. However, it barely grasped the status of the microenvironment in specific tissues because of lack of precise localization of the tissue space in tumor and immune microenvironment [75]. In this study the expression of the immune-related genes selected from specific GC tissues or microenvironment could reflect cell function or assess the status of the microenvironment, even predict the immunotherapy for GC patients.

In conclusion, the LGSGC is a reliable predictor to detect the LNM and the prognosis of GC patients. Moreover, the LGSGC could classify immune and molecular characteristics of GC patients in different risk groups and identify immunotherapy-positive patients. So, the LNMGC might be a potential indicator may guiding diagnose and treatment strategies for GC patients.

References

1. Izumi D, Gao F, Toden S, Sonohara F, Kanda M, Ishimoto T, Kodera Y, Wang X, Baba H, Goel A: **A genomewide transcriptomic approach identifies a novel gene expression signature for the detection of lymph node metastasis in patients with early stage gastric cancer**. *EBIOMEDICINE* 2019, **41**:268-275.
2. Cristescu R, Lee J, Nebozhyn M, Kim K, Ting JC, Wong SS, Liu J, Yue YG, Wang J, Yu K *et al*: **Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes**. *NAT MED* 2015, **21**(5):449-456.

3. Zhong X, Xuan F, Qian Y, Pan J, Wang S, Chen W, Lin T, Zhu H, Wang X, Wang G: **A genomic-clinicopathologic Nomogram for the preoperative prediction of lymph node metastasis in gastric cancer**. *BMC CANCER* 2021, **21**(1):455.
4. Abdelfatah MM, Othman MO: **Lymph node metastases in early gastric cancer, when the East and the West come to terms**. *Translational gastroenterology and hepatology* 2018, **3**:87.
5. Gotoda T, Yanagisawa A, Sasako M, Ono H, Nakanishi Y, Shimoda T, Kato Y: **Incidence of lymph node metastasis from early gastric cancer: estimation with a large number of cases at two large centers**. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* 2000, **3**(4):219-225.
6. Hahn KY, Park CH, Lee YK, Chung H, Park JC, Shin SK, Lee YC, Kim H, Cheong J, Hyung WJ *et al*: **Comparative study between endoscopic submucosal dissection and surgery in patients with early gastric cancer**. *Surgical endoscopy* 2017, **32**(1):73-86.
7. Hahn KY, Lee SK, Park CH, Chung H, Park JC, Shin SK, Lee YC: **Mo1065 Comparative Study Between Endoscopic Submucosal Dissection and Surgery in Patients With Early Gastric Cancer**. *GASTROINTEST ENDOSC* 2016, **83**(5):B448.
8. Pham H, Richardson AJ: **The Prognostic Importance of Lymph Node Metastasis in Patients With Gastric Cancer and Identifying Those Suitable for Middle Segmental Gastrectomy**. *JAMA network open* 2021, **4**(3):e211877.
9. Kawazoe A, Shitara K, Boku N, Yoshikawa T, Terashima M: **Current status of immunotherapy for advanced gastric cancer**. *JPN J CLIN ONCOL* 2021, **51**(1):20-27.
10. Wei Q, Xu Q, Yuan X, Li JJ, Chen L, Luo C, Zhu X, Ying JE: **Immunological impact of chemotherapy on the tumor microenvironment in gastric cancer**. *J SURG ONCOL* 2021, **123**(8):1708-1715.
11. Yarchoan M, Hopkins A, Jaffee EM: **Tumor Mutational Burden and Response Rate to PD-1 Inhibition**. *The New England journal of medicine* 2017, **377**(25):2500-2501.
12. Bodor JN, Boumber Y, Borghaei H: **Biomarkers for immune checkpoint inhibition in non–small cell lung cancer (NSCLC)**. *CANCER-AM CANCER SOC* 2020, **126**(2):260-270.
13. Vadakekolathu J, Minden MD, Hood T, Church SE, Reeder S, Altmann H, Sullivan AH, Viboch EJ, Patel T, Ibrahimova N *et al*: **Immune landscapes predict chemotherapy resistance and immunotherapy response in acute myeloid leukemia**. *SCI TRANSL MED* 2020, **12**(546).
14. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B *et al*: **Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response**. *NAT MED* 2018, **24**(10):1550-1558.
15. Zeng D, Wu J, Luo H, Li Y, Xiao J, Peng J, Ye Z, Zhou R, Yu Y, Wang G *et al*: **Tumor microenvironment evaluation promotes precise checkpoint immunotherapy of advanced gastric cancer**. *J IMMUNOTHER CANCER* 2021, **9**(8):e2467.
16. Walsh RJ, Tan DSP: **The Role of Immunotherapy in the Treatment of Advanced Cervical Cancer: Current Status and Future Perspectives**. *J CLIN MED* 2021, **10**(19):4523.

17. Kim SM, Lee H, Min BH, Kim JJ, An JY, Choi MG, Bae JM, Kim S, Sohn TS, Lee JH: **A prediction model for lymph node metastasis in early-stage gastric cancer: Toward tailored lymphadenectomy**. *J SURG ONCOL* 2019, **120**(4):670-675.
18. Peng L, Zhang J, Teng Y, Zhao Y, Wang T, Mao F, Lv Y, Cheng P, Li W, Chen N *et al*: **Tumor-Associated Monocytes/Macrophages Impair NK-Cell Function via TGFβ1 in Human Gastric Cancer**. *CANCER IMMUNOL RES* 2017, **5**(3):248.
19. Lei Z, Tan IB, Das K, Deng N, Zouridis H, Pattison S, Chua C, Feng Z, Guan YK, Ooi CH *et al*: **Identification of Molecular Subtypes of Gastric Cancer With Different Responses to PI3-Kinase Inhibitors and 5-Fluorouracil**. *Gastroenterology (New York, N.Y. 1943)* 2013, **145**(3):554-565.
20. Fridman WH, Pagès F, Sautès-Fridman C, Galon J: **The immune contexture in human tumours: impact on clinical outcome**. *Nature reviews. Cancer* 2012, **12**(4):298-306.
21. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JI, Kim K, Liu XQ, Sher X, Jung H, Lee M *et al*: **Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer**. *NAT MED* 2018, **24**(9):1449-1458.
22. Noh MG, Yoon Y, Kim G, Kim H, Lee E, Kim Y, Park C, Lee KH, Park H: **Practical prediction model of the clinical response to programmed death-ligand 1 inhibitors in advanced gastric cancer**. *EXP MOL MED* 2021, **53**(2):223-234.
23. Snyder A, Nathanson T, Funt SA, Ahuja A, Buros Novik J, Hellmann MD, Chang E, Aksoy BA, Al-Ahmadie H, Yusko E *et al*: **Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: An exploratory multi-omic analysis**. *PLOS MED* 2017, **14**(5):e1002309.
24. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel IEE, Koeppen H, Astarita JL, Cubas R *et al*: **TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells**. *Nature (London)* 2018, **554**(7693):544-548.
25. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G *et al*: **Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma**. *CELL* 2017, **168**(3):542.
26. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: **Metascape provides a biologist-oriented resource for the analysis of systems-level datasets**. *NAT COMMUN* 2019, **10**(1):1523.
27. Oh SC, Sohn BH, Cheong J, Kim S, Lee JE, Park KC, Lee SH, Park J, Park Y, Lee H *et al*: **Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype**. *NAT COMMUN* 2018, **9**(1):1714-1777.
28. Wang Q, Li P, Wu W: **A systematic analysis of immune genes and overall survival in cancer patients**. *BMC CANCER* 2019, **19**(1):1225.
29. Zhou C, Wang Y, Ye H, Yan S, Ji M, Liu P, Yang J: **Machine learning predicts lymph node metastasis of poorly differentiated-type intramucosal gastric cancer**. *SCI REP-UK* 2021, **11**(1):1300.

30. Yang Y, Zheng Y, Zhang H, Miao Y, Wu G, Zhou L, Wang H, Ji R, Guo Q, Chen Z *et al*: **An Immune-Related Gene Panel for Preoperative Lymph Node Status Evaluation in Advanced Gastric Cancer**. *BIOMED RES INT* 2020, **2020**:8450656.
31. Xu K, Wang R, Xie H, Hu L, Wang C, Xu J, Zhu C, Liu Y, Gao F, Li X *et al*: **Single-cell RNA sequencing reveals cell heterogeneity and transcriptome profile of breast cancer lymph node metastasis**. *Oncogenesis (New York, NY)* 2021, **10**(10):66.
32. Chen B, Zhang G, Lai J, Xiao W, Li X, Li C, Mok H, Li K, Wang Y, Cao L *et al*: **Genetic and immune characteristics of sentinel lymph node metastases and multiple lymph node metastases compared to their matched primary breast tumours**. *EBIOMEDICINE* 2021, **71**:103542.
33. Ulintz PJ, Greenson JK, Wu R, Fearon ER, Hardiman KM: **Lymph Node Metastases in Colon Cancer Are Polyclonal**. *CLIN CANCER RES* 2018, **24**(9):2214-2224.
34. Mah LY, Ryan KM: **Autophagy and cancer**. *CSH PERSPECT BIOL* 2012, **4**(1):a8821.
35. Huang T, Song X, Yang Y, Wan X, Alvarez AA, Sastry N, Feng H, Hu B, Cheng S: **Autophagy and Hallmarks of Cancer**. *Critical reviews in oncogenesis* 2018, **23**(5-6):247.
36. Galluzzi L, Pietrocola F, Bravo-San Pedro JM, Amaravadi RK, Baehrecke EH, Cecconi F, Codogno P, Debnath J, Gewirtz DA, Karantza V *et al*: **Autophagy in malignant transformation and cancer progression**. *The EMBO journal* 2015, **34**(7):856-880.
37. Chen J, Wu P, Xia R, Yang J, Huo X, Gu D, Tang C, De W, Yang F: **STAT3-induced lncRNA HAGLROS overexpression contributes to the malignant progression of gastric cancer cells via mTOR signal-mediated inhibition of autophagy**. *MOL CANCER* 2018, **17**(1):6.
38. Li Z, Wu Y, Zhang Y, Guan X, Niu M, He L, Zheng X, Xue X, Bo Y, Gao W *et al*: **circPARD3 drives malignant progression and chemoresistance of laryngeal squamous cell carcinoma by inhibiting autophagy through the PRKCI-Akt-mTOR pathway**. *MOL CANCER* 2020, **19**(1):166.
39. Cao Y, Luo Y, Zou J, Ouyang J, Cai Z, Zeng X, Ling H, Zeng T: **Autophagy and its role in gastric cancer**. *CLIN CHIM ACTA* 2019, **489**:10-20.
40. Shen Y, Li C, Yen C, Hsu C, Lin Y, Lin Z, Chen L, Su W, Chao Y, Yeh K *et al*: **Phase II Multicentered Study of Low-Dose Everolimus plus Cisplatin and Weekly 24-Hour Infusion of High-Dose 5-Fluorouracil and Leucovorin as First-Line Treatment for Patients with Advanced Gastric Cancer**. *Oncology* 2014, **87**(2):104-113.
41. Tang C, Yang L, Jiang X, Xu C, Wang M, Wang Q, Zhou Z, Xiang Z, Cui H: **Antibiotic drug tigecycline inhibited cell proliferation and induced autophagy in gastric cancer cells**. *BIOCHEM BIOPH RES CO* 2014, **446**(1):105-112.
42. Dong Z, Abbas MN, Kausar S, Yang J, Li L, Tan L, Cui H: **Biological Functions and Molecular Mechanisms of Antibiotic Tigecycline in the Treatment of Cancers**. *INT J MOL SCI* 2019, **20**(14):3577.
43. Machlowska J, Baj J, Sitarz M, Maciejewski R, Sitarz R: **Gastric Cancer: Epidemiology, Risk Factors, Classification, Genomic Characteristics and Treatment Strategies**. *INT J MOL SCI* 2020, **21**(11):4012.

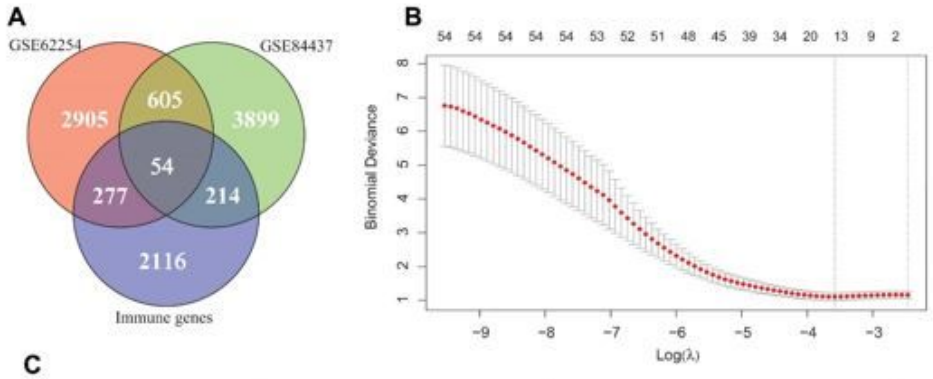
44. Do H, Kim D, Kang J, Son B, Seo D, Youn H, Youn B, Kim W: **TFAP2C increases cell proliferation by downregulating GADD45B and PMAIP1 in non-small cell lung cancer cells**. *BIOL RES* 2019, **52**(1):35.
45. Zhao X, Liu X, Su L: **Parthenolide induces apoptosis via TNFRSF10B and PMAIP1 pathways in human lung cancer cells**. *J EXP CLIN CANC RES* 2014, **33**(1):3.
46. Qiao S, Cabello CM, Lamore SD, Lesson JL, Wondrak GT: **d-Penicillamine targets metastatic melanoma cells with induction of the unfolded protein response (UPR) and Noxa (PMAIP1)-dependent mitochondrial apoptosis**. *Apoptosis (London)* 2012, **17**(10):1079-1094.
47. Cao Y, Sun J, Li M, Dong Y, Zhang Y, Yan J, Huang R, Yan X: **Inhibition of G9a by a small molecule inhibitor, UNC0642, induces apoptosis of human bladder cancer cells**. *ACTA PHARMACOL SIN* 2019, **40**(8):1076-1084.
48. Feng Z, Shi M, Li K, Ma Y, Jiang L, Chen H, Peng C: **Development and validation of a cancer stem cell-related signature for prognostic prediction in pancreatic ductal adenocarcinoma**. *J TRANSL MED* 2020, **18**(1):360.
49. Kwon MJ: **Emerging immune gene signatures as prognostic or predictive biomarkers in breast cancer**. *ARCH PHARM RES* 2019, **42**(11):947-961.
50. Shen S, Wang G, Zhang R, Zhao Y, Yu H, Wei Y, Chen F: **Development and validation of an immune gene-set based Prognostic signature in ovarian cancer**. *EBIOMEDICINE* 2019, **40**(C):318-326.
51. Wu J, Zhao Y, Zhang J, Wu Q, Wang W: **Development and validation of an immune-related gene pairs signature in colorectal cancer**. *ONCOIMMUNOLOGY* 2019, **8**(7):1596715.
52. Liu X, Niu X, Qiu Z: **A Five-Gene Signature Based on Stromal/Immune Scores in the Tumor Microenvironment and Its Clinical Implications for Liver Cancer**. *DNA CELL BIOL* 2020, **39**(9):1621-1638.
53. Gartrell RD, Marks DK, Rizk EM, Bogardus M, Gérard CL, Barker LW, Fu Y, Esancy CL, Li G, Ji J *et al*: **Validation of Melanoma Immune Profile (MIP), a Prognostic Immune Gene Prediction Score for Stage II-III Melanoma**. *CLIN CANCER RES* 2019, **25**(8):2494-2502.
54. Hong W, Yuan H, Gu Y, Liu M, Ji Y, Huang Z, Yang J, Ma L: **Immune-related prognosis biomarkers associated with osteosarcoma microenvironment**. *CANCER CELL INT* 2020, **20**(1):83.
55. Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, Jin X, Yin J, Chen L, Zhang Y *et al*: **Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers**. *NAT COMMUN* 2020, **11**(1):1000.
56. Wang H, Wu X, Chen Y: **Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer**. *FRONT ONCOL* 2019, **9**:1212.
57. Rhee J, Jung YC, Kim KR, Yoo J, Kim J, Lee Y, Ko YH, Lee HH, Cho BC, Kim T: **Impact of Tumor Purity on Immune Gene Expression and Clustering Analyses across Multiple Cancer Types**. *CANCER IMMUNOL RES* 2018, **6**(1):87.
58. Yang W, Lai Z, Li Y, Mu J, Yang M, Xie J, Xu J, Laboratory Of Malaria And Vector Research NIOA, Department Of Biochemistry And Molecular Biology SMUT, Department Of Oncology RHOW *et al*:

Immune signature profiling identified prognostic factors for gastric cancer. *CHINESE J CANCER RES* 2019, **31**(3):463-470.

59. Guan X, Xu Z, Chen R, Qin J, Cheng X: **Identification of an Immune Gene-Associated Prognostic Signature and Its Association With a Poor Prognosis in Gastric Cancer Patients**. *FRONT ONCOL* 2020, **10**:629909.
60. Li X, Liu S, Yan J, Peng L, Chen M, Yang J, Zhang G, Kazuhiko U, Uchiyama K: **The Characteristics, Prognosis, and Risk Factors of Lymph Node Metastasis in Early Gastric Cancer**. *GASTROENT RES PRACT* 2018, **2018**:6945743-6945747.
61. Zhang Y, Morgan R, Podack ER, Rosenblatt J: **B cell regulation of anti-tumor immune response**. *IMMUNOL RES* 2013, **57**(1-3):115-124.
62. Affara NI, Ruffell B, Medler TR, Gunderson AJ, Johansson M, Bornstein S, Bergsland E, Steinhoff M, Li Y, Gong Q *et al*: **B Cells Regulate Macrophage Phenotype and Response to Chemotherapy in Squamous Carcinomas**. *CANCER CELL* 2014, **25**(6):809-821.
63. Sakimura C, Tanaka H, Okuno T, Hiramatsu S, Muguruma K, Hirakawa K, Wanibuchi H, Ohira M: **B cells in tertiary lymphoid structures are associated with favorable prognosis in gastric cancer**. *The Journal of surgical research* 2017, **215**:74-82.
64. Lee N, Zakka LR, Mihm MC, Schatton T: **Tumour-infiltrating lymphocytes in melanoma prognosis and cancer immunotherapy**. *PATHOLOGY* 2016, **48**(2):177-187.
65. Kurozumi S, Fujii T, Matsumoto H, Inoue K, Kurosumi M, Horiguchi J, Kuwano H: **Significance of evaluating tumor-infiltrating lymphocytes (TILs) and programmed cell death-ligand 1 (PD-L1) expression in breast cancer**. *MED MOL MORPHOL* 2017, **50**(4):185-194.
66. Rezalotfi A, Ahmadian E, Aazami H, Solgi G, Ebrahimi M: **Gastric Cancer Stem Cells Effect on Th17/Treg Balance; A Bench to Beside Perspective**. *FRONT ONCOL* 2019, **9**:226.
67. O Reilly LA, Putoczki TL, Mielke LA, Low JT, Lin A, Preaudet A, Herold MJ, Yaprianto K, Tai L, Kueh A *et al*: **Loss of NF-κB1 Causes Gastric Cancer with Aberrant Inflammation and Expression of Immune Checkpoint Regulators in a STAT-1-Dependent Manner**. *Immunity (Cambridge, Mass.)* 2018, **48**(3):570-583.
68. Li W, Zhang X, Wu F, Zhou Y, Bao Z, Li H, Zheng P, Zhao S: **Gastric cancer-derived mesenchymal stromal cells trigger M2 macrophage polarization that promotes metastasis and EMT in gastric cancer**. *CELL DEATH DIS* 2019, **10**(12):916-918.
69. Eum HH, Kwon M, Ryu D, Jo A, Chung W, Kim N, Hong Y, Son D, Kim ST, Lee J *et al*: **Tumor-promoting macrophages prevail in malignant ascites of advanced gastric cancer**. *Experimental & molecular medicine* 2020, **52**(12):1976-1988.
70. Locati M, Curtale G, Mantovani A: **Diversity, Mechanisms, and Significance of Macrophage Plasticity**. *Annual review of pathology* 2020, **15**(1):123-147.
71. Oya Y, Hayakawa Y, Koike K: **Tumor microenvironment in gastric cancers**. *CANCER SCI* 2020, **111**(8):2696-2707.

72. Räihä MR, Puolakkainen PA: **Tumor-associated macrophages (TAMs) as biomarkers for gastric cancer: A review**. *Chronic diseases and translational medicine* 2018, **4**(3):156-163.
73. **P. Sarode, X. Zheng, A. Weigert, et al. LSC-2017-reprogramming of tumor associated macrophages by modulating Wnt/ß-catenin signaling in lung cancer Eur Respir J, 50 (2017), p. OA4859**.
74. Menzies AM, Amaria RN, Rozeman EA, Huang AC, Tetzlaff MT, van de Wiel BA, Lo S, Tarhini AA, Burton EM, Pennington TE *et al*: **Pathological response and survival with neoadjuvant therapy in melanoma: a pooled analysis from the International Neoadjuvant Melanoma Consortium (INMC)**. *NAT MED* 2021, **27**(2):301-309.
75. Han S, Huang K, Gu Z, Wu J: **Tumor immune microenvironment modulation-based drug delivery strategies for cancer immunotherapy**. *NANOSCALE* 2020, **12**(2):413-436.

Figures



C

Variables	N	Odds ratio	95%CI	p
PMAIP1	159		3.29 (1.66, 7.13)	0.001
TRIM65	159		0.46 (0.24, 0.81)	0.011
RFXAP	159		0.54 (0.29, 0.91)	0.031
RB1	159		1.91 (1.08, 3.61)	0.035
CASP1	159		0.50 (0.25, 0.94)	0.034
SECTM1	159		2.11 (1.17, 4.04)	0.017
UBE2V2	159		0.36 (0.15, 0.76)	0.010
UBE2W	159		2.13 (1.12, 4.34)	0.027
GNAI3	159		2.18 (1.16, 4.28)	0.018
DDX3X	159		0.59 (0.34, 0.99)	0.051
CMTM6	159		0.44 (0.20, 0.89)	0.031
FSTL1	159		2.12 (1.16, 4.08)	0.018
SOCS2	159		0.53 (0.28, 0.97)	0.049
NLRP13	159		0.48 (0.27, 0.83)	0.010

0.2 0.5 1 2 5

D

Variables		N	Odds ratio	95%CI	p
Riskscore		159		3.14 (1.93, 5.90)	< 0.001
Subtype	Invasive	43		Reference	
	Metabolic	34		8.54 (1.71, 53.77)	0.01
	Proliferative	57		1.70 (0.33, 9.32)	0.53
	Unstable	25		0.88 (0.14, 5.83)	0.89
Gender	Female	56		Reference	
	Male	103		2.09 (0.59, 7.73)	0.25
Lauren classification	Diffuse	66		Reference	
	Intestinal	80		0.58 (0.11, 2.77)	0.49
	Mixed	13		7.10 (0.32, 293.44)	0.27
Stage		159		3.58 (2.10, 6.66)	< 0.001
Age		159		0.99 (0.94, 1.04)	0.66

0.2 1 200

Figure 1

Selection of candidate immune-related genes associated with LNM of GC patients. (A) Diagram of immune-related genes with the same direction of change between LNP patients and LNN patients in the GSE62254 dataset and the GSE84437 dataset. (B) LASSO logistic regression for the 54-candidate differentially expressed immune-related genes (DEIGs) associated LNM in the training dataset (GSE15459). Two dashed lines are ploted at the optimal value using the minimum value criterion and the

one-standard error criterion, respectively. (C) Multivariate logistic regression analysis of 14 DEIGs associated with LNM in GC in the training dataset. (D) Multivariate logistic regression analysis of 14 DEIGs associated with LNM in GC adjusting for other clinicopathologic factors such as Lauren classification, gender, age and subtypes of GC in the training dataset. LNM, lymph node metastasis; LNN, lymph node negative; GC, gastric cancer; OR, odds ratio; LNP, lymph node positive; CI, confidence interval.

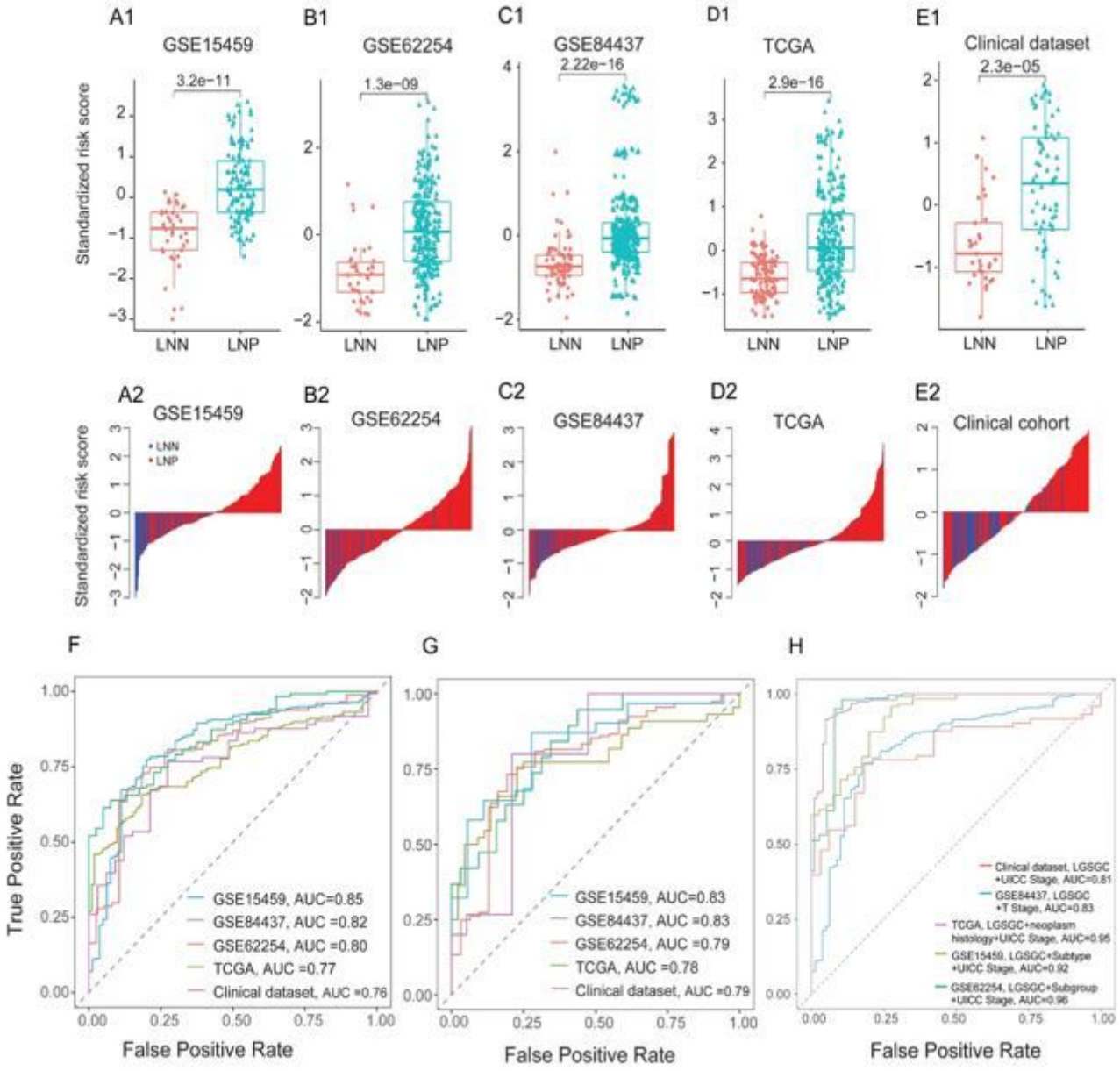


Figure 2

Training and validation of the LGSGC detecting LNM of GC patients. (A1-E5) Box plots show standardized risk scores between LNP patients and LNN patients. The boxes indicate the median ± 1 quartile. (A2-E2) Waterfall plot shows the standardized risk score according to lymph node status. ROC analysis shows the predictive value of the LGSGC identifying LNP from LNN in all GC patients (F) and in early GC patients (G). LGSGC, lymph node metastasis gene signature for gastric cancer.

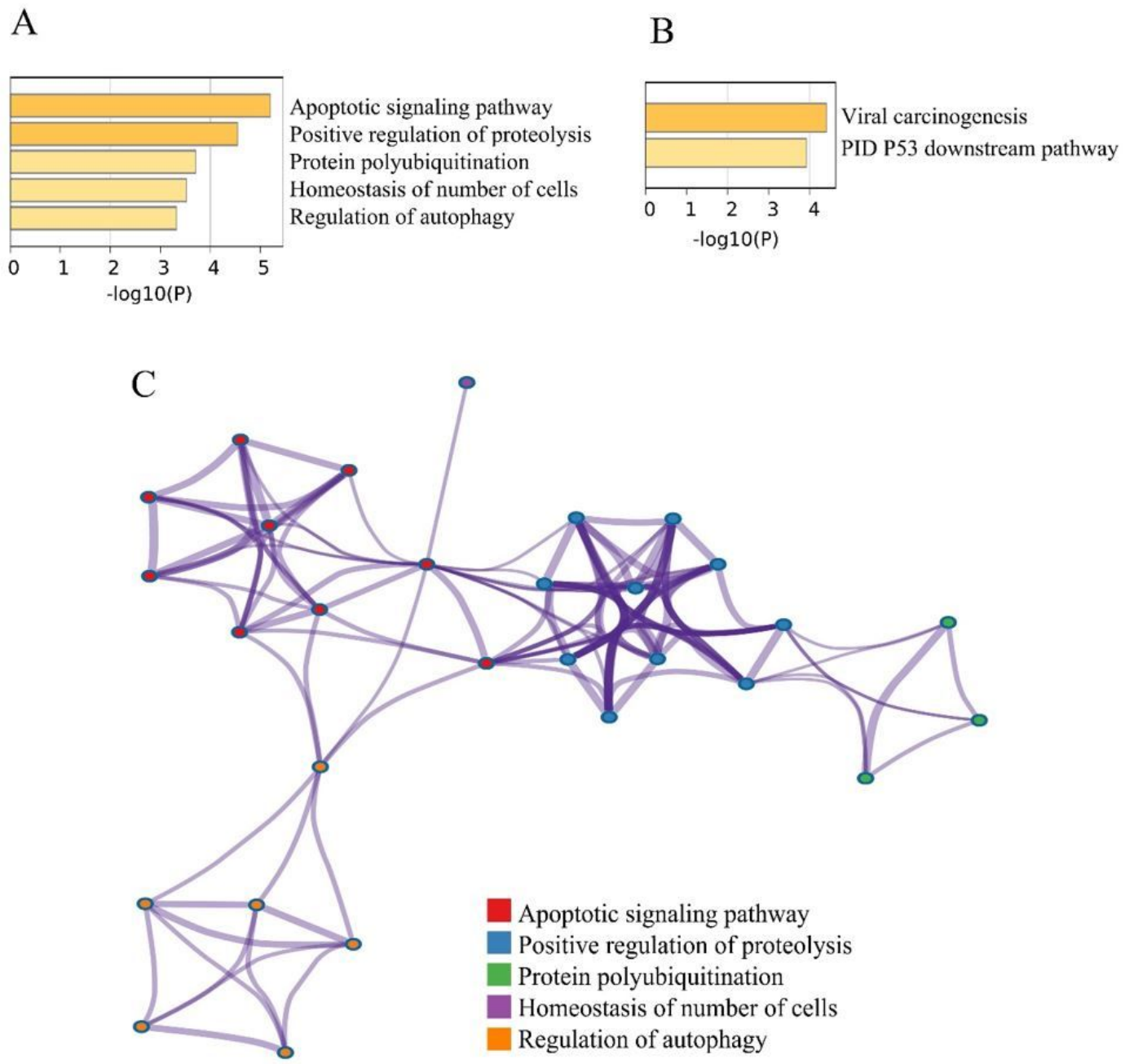


Figure 3

Gene function and pathway analysis for the 14 DEIGs. (A) Bar graph of the top five functional enrichment terms and (B) two pathways enrichment terms. (C) Network of the five functional enrichment terms, where nodes are represented by pie charts.

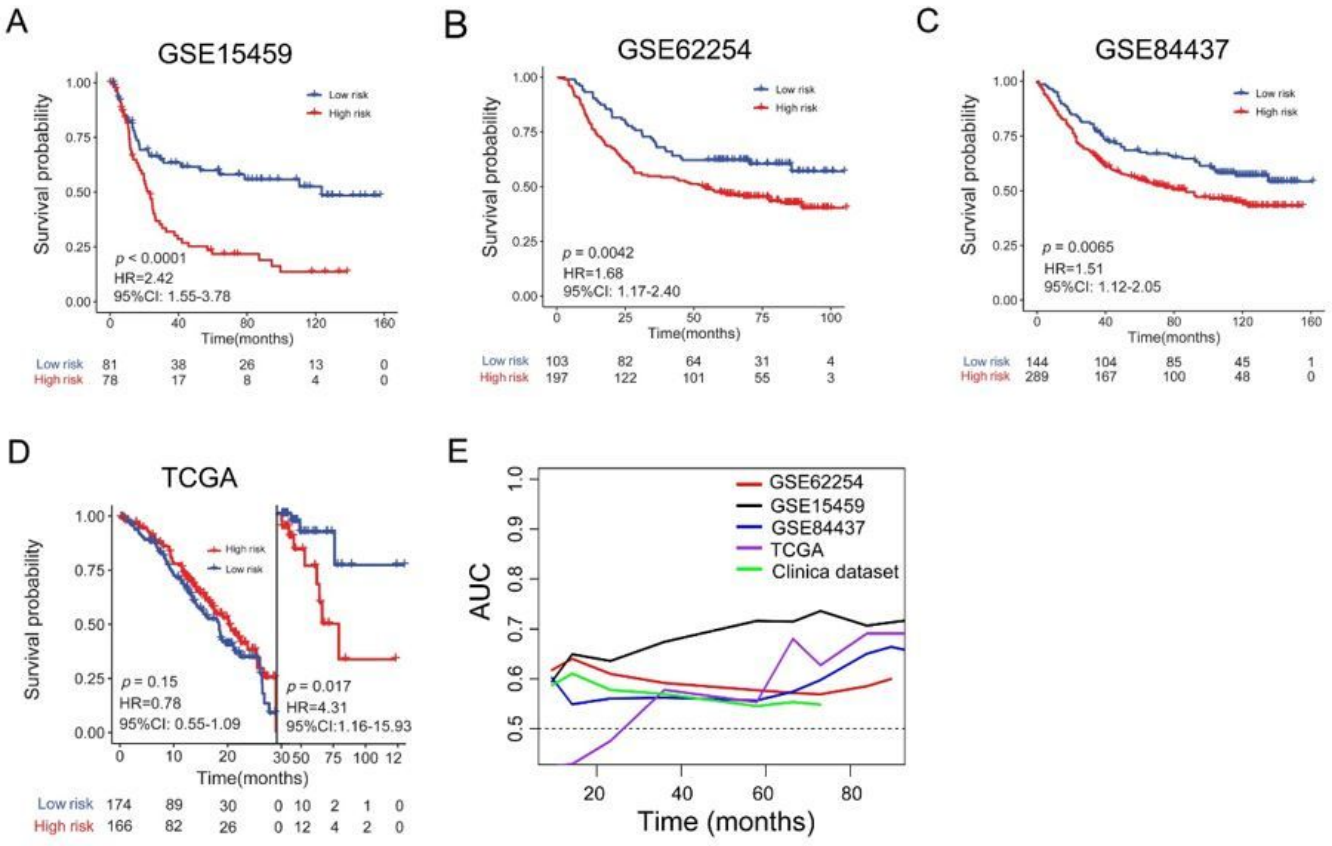


Figure 4

Association between the LGSGC and the survival of GC patients. (A-D) Kaplan–Meier plots of both risk groups estimated by the LGSGC in each dataset. (E) ROC analysis of the LNMGC on overall survival of GC patients in each dataset with time.

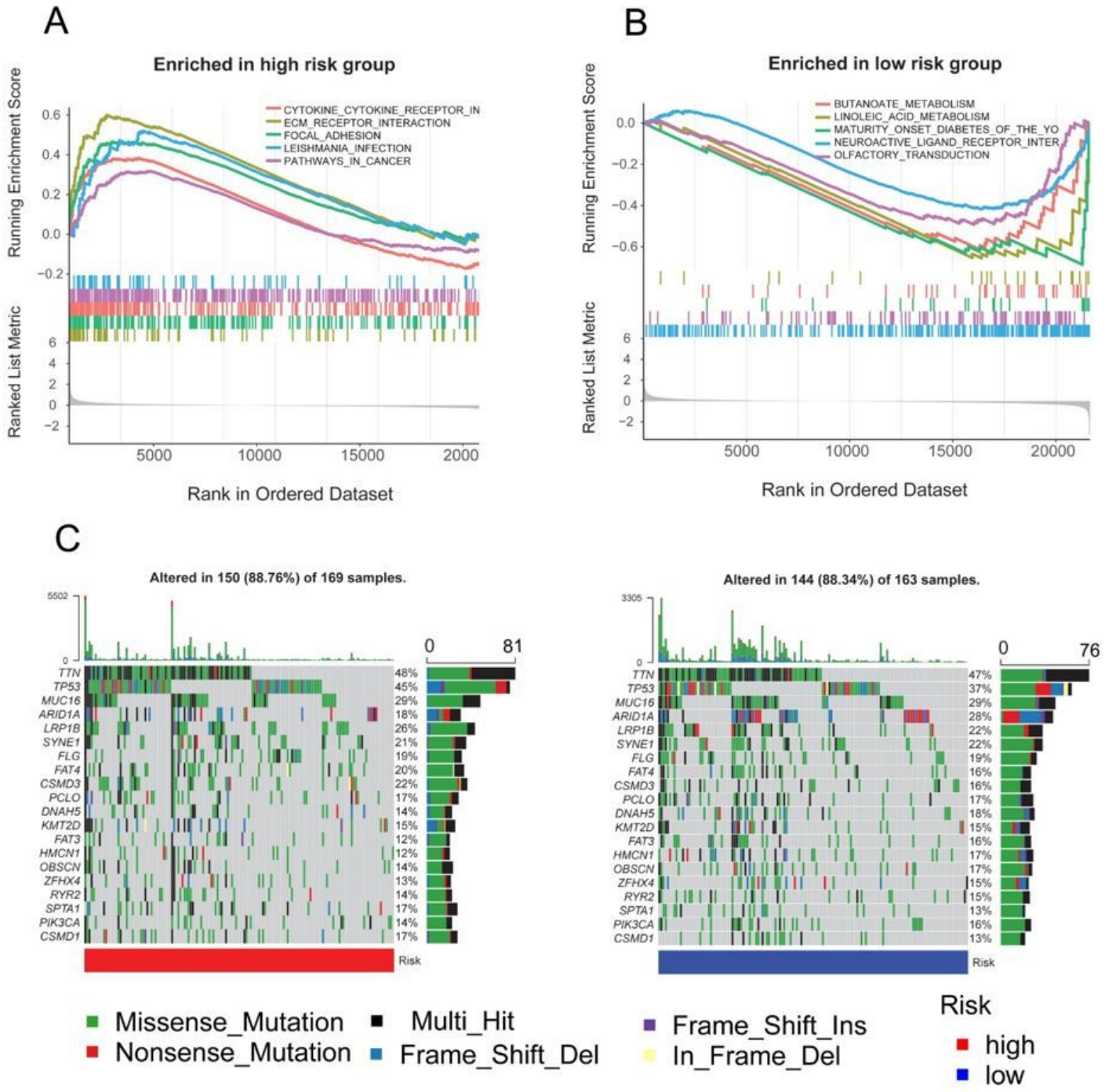


Figure 5

Molecular characteristics of different risk groups. (A) Enrichment plots of gene sets in high-risk group and (B) low-risk group from the training dataset ($p < 0.05$). (C) The mutation frequency of genes in both risk groups from the TCGA dataset. Columns represent different patients. The upper bar graph represents the total number of mutations, the left is the mutated gene symbols, and the right bar graph represents the percentage of mutations. The top 20 mutated genes are sorted by mutation rate.

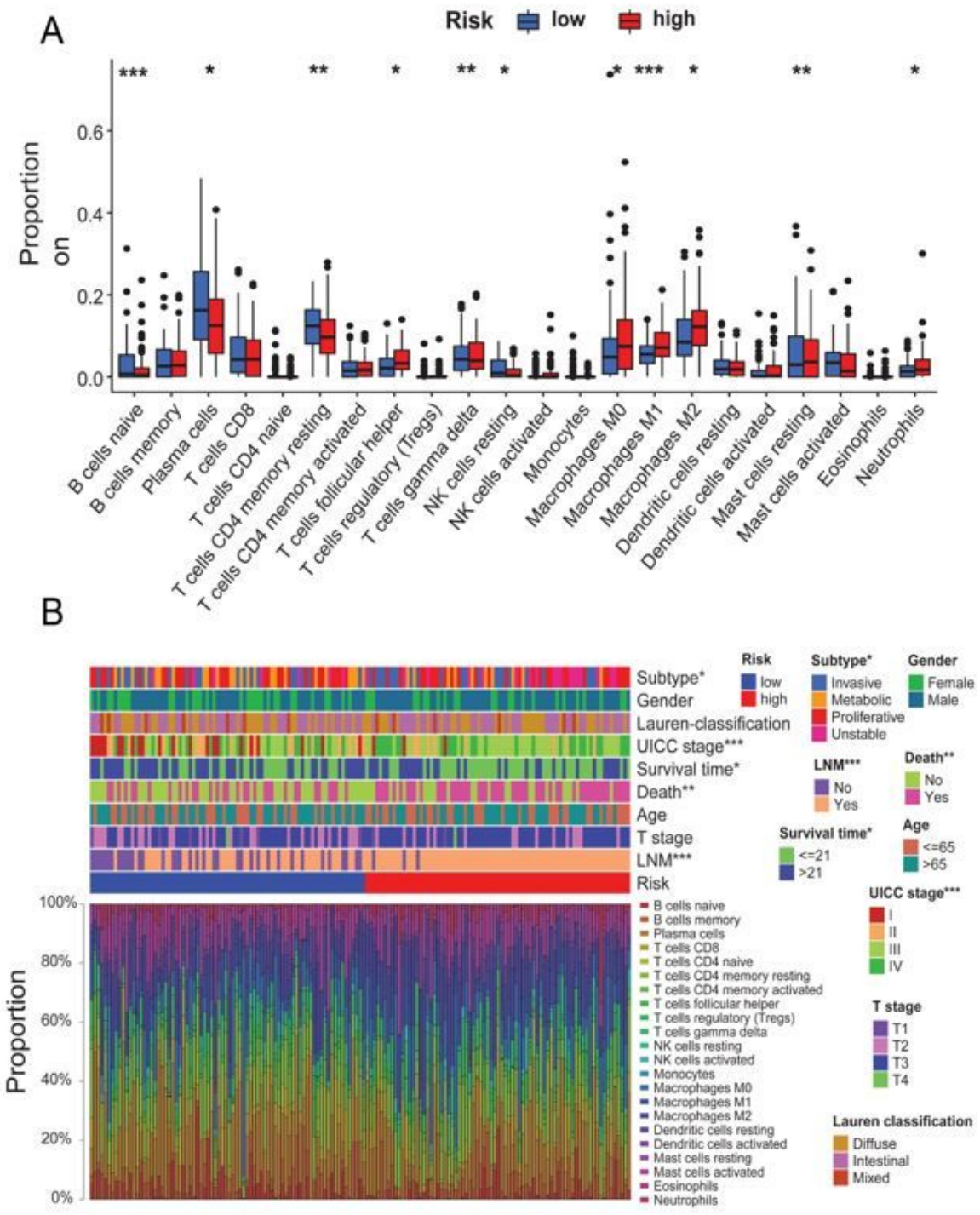


Figure 6

The landscape of clinicopathological factors and immune cell subsets in both risk groups from the training dataset. (A) Box plots of 22 immune cell types proportion. Blue and red colors represent low-risk group and high-risk group, respectively; the horizontal line in the middle of the box represents the median, with the 25th and 75th percentiles at the bottom and top, respectively. The Wilcoxon test was used to compare the differences in both risk groups with significant statistical *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$,

**** $p < 0.0001$. (B) Subtypes of GC, gender, Lauren-classification, UICC stage, alive time, alive status, age, T stage and LNS are annotated in both groups; Bar charts of proportions of 22 immune cell types for each GC patient.

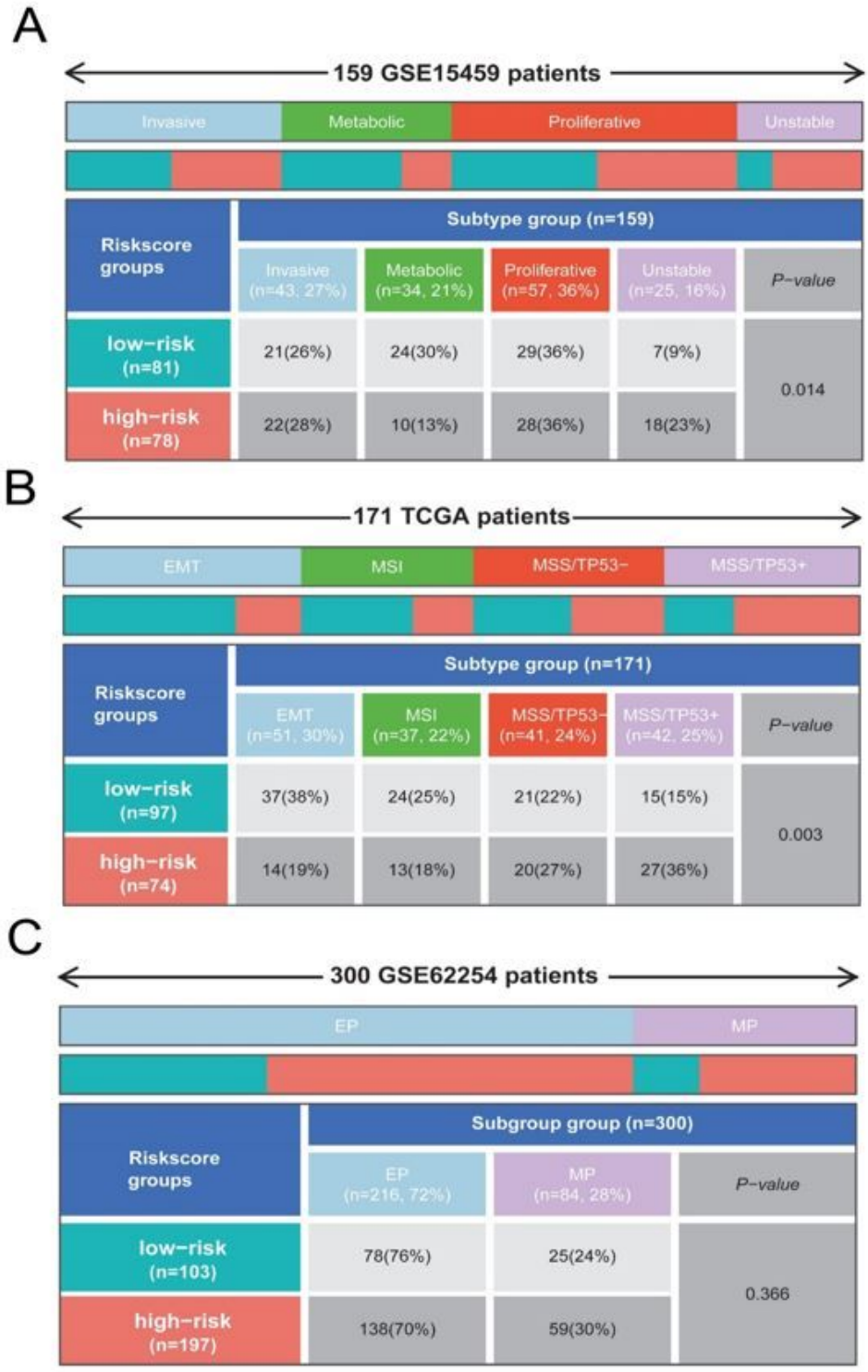


Figure 7

Heat maps and tables of molecular subtypes of GC in both risk groups. (A) The distribution of four subtypes (proliferative, metabolic, invasive and unstable) of GC in both risk groups from the training dataset. (B) The distribution of four subtypes (MSS/TP53−, MSS/TP53+, MSI and EMT) of GC in both groups from the TCGA dataset. (C) The distribution of two subtypes (MP and EP) of GC in groups from the GE62254 dataset. MSS/TP53−: TP53 functional loss, MSS/TP53+: intact TP53 activity, MSI: microsatellite instability, EMT: epithelial-to-mesenchymal transition, MP: mesenchymal phenotype, EP: epithelial phenotype. The difference of the molecular subtype's distribution of GC patients between in both risk groups was analyzed by the chi-square test.

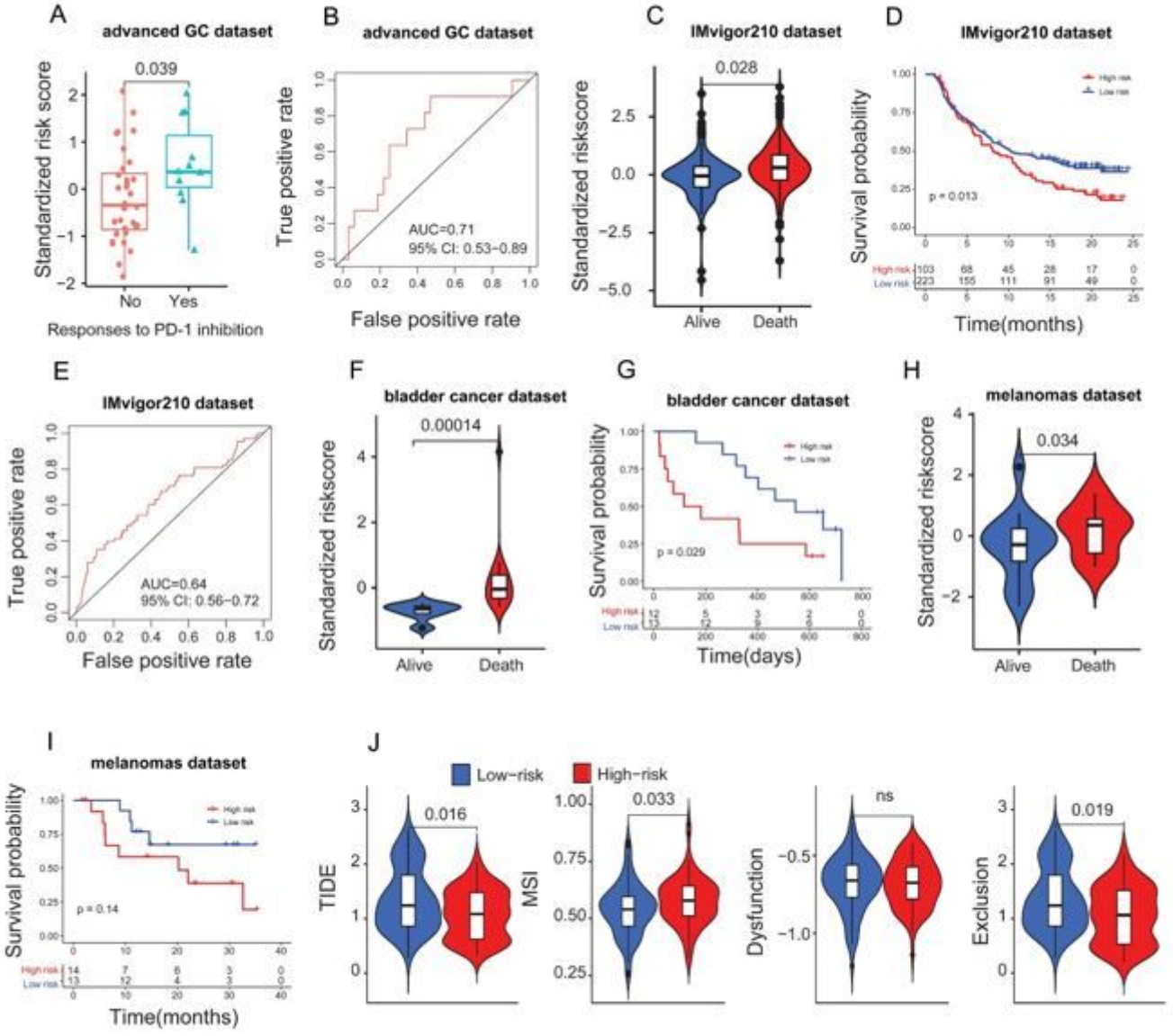


Figure 8

Assessment of the LNMGC predicting cancer immunotherapy response. (A) Box plots show standardized risk score in both groups, from the advanced GC dataset, who have or do not have response to PD-1 inhibition (*p*-value is determined by the Wilcoxon test). (B) ROC curve shows the predictive value of the LNMGC discriminating patients, from the advanced GC dataset, who have response to PD-1 inhibition. (C) The comparation of standardized risk score based on the LNMGC in patients with various types of cancers, from the IMvigor210 dataset, who are alive or dead (p-value is determined by the Wilcoxon test). (D) The Kaplan-Meier survival curves of both groups from the IMvigor210 dataset with PD-1 inhibition therapy (the log-rank test). (E) ROC curve shows the predictive value of the LNMGC discriminating patients, from the IMvigor210 dataset, who have response to PD-1 inhibition. (F) The comparation of standardized risk score based on the LNMGC in bladder cancer patients who are alive or dead (the Wilcoxon test). (G) The Kaplan-Meier survival curves of both groups-patients with PD-1 inhibition therapy in the bladder cancer dataset (the log-rank test). (H) The comparation of standardized risk score based on the LNMGC in melanomas patients who are alive or dead (the Wilcoxon test). (I) The Kaplan-Meier survival curves of both groups-patients with PD-1 inhibition therapy in the melanomas dataset (the log-rank test). (J) The comparation of the TIDE, MSI, and T cell dysfunction and exclusion score based on the TIDE module in both group-patients from the training dataset (the Wilcoxon test, ns: not significant). PD-1: programmed cell death 1; TIDE: Tumor Immune Dysfunction and Exclusion.

Supplementary Files

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

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- figs1.jpg
- TableS5.xlsx
- figs2.jpg
- TableS6.xlsx
- figs4.jpg
- TableS7.xlsx
- figs61.jpg
- figs62.jpg
- figs7.jpg
- figs8.jpg