Research Paper

The tumor microenvironment in gastrointestinal adenocarcinomas revealed a prognostic and immunotherapeutic biomarker

Yong Zhang[1,2,3,4,*], Lu Li[1,2,3,*] , Feifei Chu[1,2,3], Lei Zhang[1,2,3], Li Zhang[1,2,3], Huili Wu[1,2,3], Kunkun Li[1,2,3]

[1]Department of Gastroenterology, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou 450007, China
[2]Medical Key Laboratory for Diagnosis and Treatment of Colorectal Cancer in Henan Province, Zhengzhou 450007, China
[3]Zhengzhou Key Laboratory for Diagnosis, Treatment and Research of Colorectal Cancer, Zhengzhou 450007, China
[4]Branch Center of Advanced Medical Research Center, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou 450007, China
*Equal contribution and co-first authors

Correspondence to: Kunkun Li, Huili Wu; **email:** likunkun770610@zzu.edu.cn, wuhuili660912@zzu.edu.cn

ABSTRACT

Accumulated evidence has elucidated that the tumor microenvironment (TME) is great of clinical significance in predicting survival outcomes and therapeutic efficacy. Nonetheless, few studies have investigated the prognostic and immunotherapeutic signature correlated with TME phenotypes in gastrointestinal adenocarcinomas (GIAC). Here, by estimating the TME pattern of immune infiltration and expression in over 1,000 GIAC patients, we revealed three TME subgroups and identified six key differential genes. To predict the TME phenotypes, TMEscore was established and validated to be an independent prognostic factor, where the high TMEscore was characterized by immune activation and response to immunotherapy and accompanied with favorable prognosis in GIAC. Furthermore, TMEscore was confirmed to predict prognosis and immunotherapeutic response in six datasets. In summary, depicting TME landscape of GIAC patients may be beneficial for interpreting survival and immunotherapeutic response, and provide new strategies for clinical treatment of GIAC.

INTRODUCTION

Gastrointestinal adenocarcinomas (GIAC) mainly consist of esophageal adenocarcinoma, gastric adenocarcinoma and colorectal adenocarcinoma and have the highest incidence and mortality across all kinds of malignant tumors. Approximately 1.4 million people die from GIAC each year worldwide [1]. Though GIAC has a similar epithelial tissue origin and gastrointestinal physiological environment, the clinical phenotypes and genetically molecular characteristics are quite distinct [2]. Genomic and transcriptomic analysis have further defined the heterogeneity of GIAC by identifying molecular subtypes. According to DNA mutation and copy number alteration and DNA methylation pattern, GIAC patients can categorize into five subtypes, namely, EpsteinBarr virus(EBV), chromosomal instability (CIN), microsatellite instability (MSI), hypermutated-SNV (HM-SNV) and genome stable (GS) tumors, or other four subtypes, i.e., CIMP-H, CIMP-L, EBV-CIMP and non-CIMP [3]. Recent studies suggest that nearly all GIAC tumors appear to present with four subtypes that are either characterized by canonical epithelial origin, extensive immune infiltration, metabolic dysregulation or mesenchymal gene expression signatures [4]. Although these studies have greatly enhanced the understanding of the tumor underlying mechanism, effective molecular markers to predict survival and guide treatment in

specific subtypes are still lacking. Thus, it is quite essential to identify tumor subgroups, exploit prognostic and therapeutic response biomarker at the pan-digestive tract tumor level.

Calculated evidence suggests that tumor initiation and progression are not only governed by the genetic changes of cancer cells but also by tumor microenvironment (TME) factors [5]. Transformed cancer cells admixed with immune cells and stromal cellular elements form complicated TME, which significantly influences therapeutic response and clinical outcome [6]. Stromal cells (epithelial cells, and fibroblasts) as well as immune cells (macrophages, neutrophils, dendritic cells, T cells and B cells) are recurrently reported to contribute to tumor progression and metastasis when present in TME [7]. Extensive research on TME has revealed a crucial role of the immune response genes and tumor-infiltrating immune cells in patient survival outcome, tumor dissemination, relapse, metastasis, and therapeutic response to immunotherapy [8–11]. Immune checkpoint molecules PD-1, PD-L1, CTLA-4 and even their combination showed excellent predictive performance with immune checkpoint blockade response. For instance, Sun et al. investigated the relevance between the RNA expression of current biomarkers with the response of immunological therapy, and assessed predictive performance in different cancer types and therapeutic strategies [12].

Numerous studies have confirmed that immune and stromal cell infiltration and their crosstalk in the TME modulate cancer progression and are attractive therapeutic targets [13, 14]. Moreover, the effects of infiltrating immune and stromal cells on prognosis and have been extensively reported [15, 16]. Therefore, designing the therapeutics simultaneously target multiple components of the TME will benefit for increasing the likelihood of favorable patient outcomes. For predicting patient prognosis or immunotherapy response in esophageal cancer, gastric cancer and colorectal cancer, large studies have revealed many signatures or score systems based on gene expression [17–19], cell infiltration level [13, 20] or their combinations [21, 22], which providing potential biomarkers and therapeutic targets. In spite of these, no studies have been analyzed intensively in pan-gastrointestinal cancers. Meanwhile, the vast majority were not validated in additional datasets, the reliability was questionable.

Thus, we aimed to uncover a robust rating system predicting patient prognosis and immunotherapy response in GIAC based on the RNA expression of the signature genes reflecting TME. In this study, we analyzed the gene-expression profiles of GIAC patients

and acquired a comprehensive landscape about TME. Based on immune infiltration and immune pathway expression patterns, we classified the GIAC into three subtypes with distinct clinical and immune characteristics. Further, we determined 6 stromal or immune genes representative for the TME of GIAC patients, and established a TME score, which could precisely predict patient survival outcome and response to immunotherapy in multiple immunotherapeutic datasets.

MATERIALS AND METHODS

Data source

We obtained RNASeq expression data and clinical data of 1,199 GIAC patients in PanCanAtlas [23] website from The Cancer Genome Atlas (TCGA) project. GIAC included colorectal adenocarcinoma (CRC, namely, colon adenocarcinoma (COAD), rectum adenocarcinoma (READ), gastric adenocarcinoma (GAD) and esophageal adenocarcinoma (ESAD). For validation of the prognostic value of our marker, we additionally got two datasets (GSE17536 and GSE39582) from Gene Expression Omnibus (GEO).

For validation of the immunotherapeutic predictive value of our marker, we collected six datasets, i.e., VanAll [24] (42 melanoma patients treated with CTLA4 inhibitor (CR/PR = 14, PD/SD = 23), Riaz [25] (51 melanoma patients treated with PD1 inhibitor (CR/PR = 10, PD/SD = 39), Mariathasan [26] (also named IMvigor210, 348 urothelial carcinoma patients treated with PD1 inhibitor (CR/PR = 68, PD/SD = 230), Auslander [27] (14 melanoma patients treated with PD1-CTLA4 inhibitor (CR/PR = 2, PD/SD = 12), Gide [28] (41 melanoma patients treated with PD1-CTLA4 inhibitor (CR/PR = 19, PD/SD = 22) and 32 melanoma patients treated with PD1 inhibitor(CR/PR = 21, PD/SD = 11) and Kim [29] (45 gastric cancer patients treated with PD1 inhibitor (CR/PR = 12, PD/SD = 33). The gene expression was detected using RNA transcriptional sequencing on patients prior to immunotherapy.

Immunophenotyping for GIAC patients

We analyzed the gene-expression profiles of GIAC tumor and adjacent samples and utilized TIMER2.0 [30] and CIBERSORT [31] method to quantify infiltration level of immune cells for GIAC patients. To further investigate TME with GIAC patient classification, we firstly obtained 160 immune related signatures of 9,131 patients of multiple cancer types from panImmune [32] resource (https://gdc.cancer.gov/about-data/publications/panimmune) and only retained 1,021 GIAC (including CRC, GAD and ESAD) patients. The 160 signatures

were from 7 resources, namely, Attractors (9 signatures), Bindea (25), c7atoms (32), CIBERSORT (20), ICR (3), Senbabaoglu (3) and Wolf (68) (Supplementary Tables 1, 2).

Based on the tumors with the 160 signatures, three unsupervised clustering algorithms (Lee, brunet and nsNMF) were used to identify TME patterns and classify patients. Kappa value was used to assess classification consistency of the three algorithms. Multiple classification indexes (silhouette, dispersion and cophenetic coefficient) were applied to determine the best number of clusters. These procedures were performed using the NMF R package, which adopted non-negative matrix factorization (NMF) method [33] and was repeated 1,000 times to ensure the stability of classification. Finally, the nsNMF algorithm and optimal cluster number 3 was selected, and patients were clustered into 3 TME subtypes, i.e., 403 patients from TME-C1, 276 from TME-C2, 324 from TME-C3.

Clinical and immune cell were extracted from PanCanAtlas resource (https://gdc.cancer.gov/about-data/publications/pancanatlas) (including survival information, cancer type, tumor stage, cancer cell fraction, and Shannon score) and the immune cell infiltration fractions (leukocyte, CD8 T cell, regulatory T cell, resting NK cell, and activated NK cell) estimated by CIBERSORT algorithm were obtained from panImmune resource. Fisher exact test or Pearson's chi-square test for discrete variables and Wilcoxon rank-sum test for continuous variables were used to assess the correlation of these features and three TME subtypes.

Differential expression analysis and functional enrichment analysis

To identify genes differentiating TME clusters, differentially expressed genes (DEGs) across these clusters were firstly identified using the R package Limma [34], which implements an empirical Bayesian approach to estimate gene-expression changes using moderated t tests. The 54 genes in signature A and 85 genes in signature B were used for gene enrichment analysis. Gene Ontology (GO) terms were identified with a strict cutoff of $P < 0.01$ and false discovery rate (FDR) of less than 0.05. Enriched pathways and cancer hallmarks were identified by running Gene Set Enrichment Analysis [35] (GSEA) of the adjusted expression data for all genes. Enrichment P values were based on 10,000 permutations and subsequently adjusted for multiple testing using the Benjamini-Hochberg (BH) procedure to control the FDR. Gene sets were downloaded from MSigDB [36] database of the Broad Institute. R function "clusterProfiler" and

"enrichplot" was adopted using the clusterProfiler [37] R package was performed on TME signature A and B genes.

Establishment of TMEscore in GIAC

For identify prognostic and immunotherapeutic gene signature, we firstly screened some representative genes distinguishing the three immune subtypes. And then the random forest classification algorithm packaged in Boruta [38] was used to perform dimension reduction in order to reduce noise or redundant genes. Univariate Cox regression analysis was utilized to determine candidate prognostic genes for PFS and OS. GIAC patients were stratified into two subgroups based on the expression of each gene above or below the median. The survival curves were plotted using Kaplan-Meier (KM) curve and the survival difference of two patient groups were estimated using log-rank test (p value < 0.05).

The TMEScore was established with a formula: TMEScore = $\exp(G_i) + \exp(G_j)$, where the G_i and G_j represent the average expression of genes in signature A and signature B. TMEscore = $\alpha\sum(\exp\_I_A)/\text{len}(A) + \beta\sum(\exp\_J_B)/\text{len}(B)$, where $\exp\_I_A$, $\exp\_J_B$ indicates the expression of gene I in signature A, the expression of gene J in signature B; α, β were set as 1 if gene I or J were protective factors and -1 otherwise. Here, TMEscore = −(C6orf223 expression + EPHX4 expression + HES6 expression + NKD2 expression)/4 + (OLR1 expression + ONECUT2 expression)/2.

Tissue specimens

Our in-house GIAC cohort included 96 pairs of fresh GIAC tumor and adjacent normal tissue specimens without radiotherapy or chemotherapy, which were immediately stored in liquid nitrogen after surgery (Supplementary Table 3), all enrolled subjects were pathologically diagnosed as adenocarcinoma. All specimens were collected from Zhengzhou Central Hospital Affiliated to Zhengzhou University between 2019 and 2020 and this study was approved by the Zhengzhou Central Hospital Affiliated to Zhengzhou University. All subjects have undergone rigorous screening and underwent informed consent.

Quantitative RT-PCR (qRT-PCR)

QRT-PCR was employed to detect the RNA level of 6 genes (C6orf223, EPHX4, HES6, NKD2, OLR1 and ONECUT2). In brief, total RNAs of 96 pairs of fresh GIAC tissue specimens were extracted by Trizol method. After testing for concentration, purity, and integrity, an equal amount of RNAs was used to

synthesize cDNA. Finally, SYBR Green Quantitative Kit (DBI, Germany) and 7500 Fast Quantitative PCR System (AB, USA) were used for detection. The housekeeping gene GAPDH was used as internal reference, and the relative gene expression was expressed as $2^{-\Delta\Delta Ct}$. Primer sequences were shown in Supplementary Table 4. The relative gene expression of the 6 genes in 96 GIAC patients (matched adjacent normal and tumor samples) were showed in Supplementary Table 5.

Western blot

Western blot was used to detect protein levels of corresponding genes in 27 pairs of tumor and matched adjacent normal sample from our center. In brief, total protein was extracted using a lysate mixture containing RIPA and protease inhibitors. After concentration determination, an equal amount of total proteins from each sample was used for polyacrylamide gel electro-phoresis. After transfer to PVDF membranes, the membranes were blocked and incubated with the primary antibody overnight at 4°C. After incubated with the secondary antibody at room temperature for 1 hour, ECL was added for exposure and development.

Primary antibodies used in this study include rabbit anti-β-tubulin (Abcam, ab6046, 1:1000 dilution for WB), rabbit anti-EPHX4 (Abcam, ab183739, 1:1000 dilution for WB), mouse anti-HES6 (Abcam, ab172800, 1:1000 dilution for WB), rabbit anti-NKD2 (CST, 2073T, 1:1000 dilution for WB), rabbit anti-OLR1 (Bioss, bs-2044R, 1:1000 dilution for WB) and anti-ONECUT2 (Bioss, bs-19643R, 1:1000 dilution for WB). A goat-anti-mouse-HRP (Bioss, bs-40296G-HRP, 1:10000 dilution for WB) antibody (Bioss, bs-19643R, 1:1000 dilution for WB) and a goat-anti-rabbit-HRP antibody (Bioss, bs-80295G-HRP, 1:10000 dilution for WB) were used in WB. The relative protein expression of the 5 genes in 27 GIAC patients (matched adjacent normal and tumor samples) were showed in Supplementary Table 6.

Prognostic evaluation using TMEscore

For overall survival (OS), disease specified survival (DSS) and progression free survival (PFS), GIAC patients were stratified into two subgroups based on TMEscore above or below the median. We analyzed and validated the prognostic value of TMEscore in TCGA and two GEO datasets (GSE17536 and GSE39582). Receiver operating characteristic (ROC) curve analysis and Area Under Curve of ROC (AUC) was utilized to show prediction power according TMEscore and other factors. And multi-variate Cox regression analysis was employed to determine the independent prognostic factors for OS and PFS with adjustment for other potential clinicopathological factors, i.e., age, gender, and tumor stage. We adopted nomogram and calibration plot to display the predictive ability and power of multiple features using R package rms.

Prediction of immunotherapeutic response using TMEscore

To explore the correlation between TMEscore and immunotherapeutic response, the expression profiles of six immunotherapeutic datasets were normalized into FPKM (Fragments Per Kilobase of exon model per Million mapped fragments). TMEscore was constructed using the RNA expression of the six genes. The immunotherapeutic response contained four status, complete response (CR), partial response (PR), stable disease (SD) and progression disease (PD), where the former two indicate response to immunotherapy, while the latter two indicate none-response to immunotherapy. Fisher exact test or Pearson's chi-square test was used to measure the relevance of TMEscore (high/low) and immunotherapeutic response (responder/non-responder). Wilcoxon rank-sum test was adopted to detect the statistical difference of TMEscore between the responders and non-responders. Based on immuno-therapeutic response status and patient TMEscore, ROC curve analysis and AUC were utilized to assess the immunotherapeutic predictive value of TMEscore. To predict patients' likelihood of responding to ICBs, we utilized "EaSIeR" R package [39] to estimate the immune response score based on hallmarks of immune response (CYT, TLS, IFNy, Ayers_expIS, Tcell_inflamed, Roh_IS, Davoli_IS and chemokines) in TCGA GIAC cohort and compute the integrated immune response score. Pearson correlation test was used to measure TMEscore and these immune response scores.

Availability of data and materials

The datasets supporting our results are available in the public database GEO, TCGA and data source in method. The data of our in-house cohort is provided in Supplementary Tables.

RESULTS

TME landscape of GIAC

To quantify infiltration level of immune cells for GIAC patients, we analyzed the gene-expression profiles of GIAC tumor and adjacent samples by TIMER2.0. Compared with 93 adjacent normal samples, we observed that the infiltration fraction of Dendritic cell and Neutrophil in 1,068 GIAC tumors were significantly elevated, while that of B cell, CD8 T cell and

macrophage were relatively decreased (Figure 1A, Supplementary Tables 7–9), suggesting the formation of a complex and differentiated tumor immune micro-environment during the progression process from normal tissue to tumor. We wondered whether there was potential for cell communication between different cell types and found that CD4 T cell and CD8 T cell (r = −0.69), Dendritic cell and B cell/macrophage (r = −0.55, −0.47), Neutrophil and B cell (r = −0.45) had significantly negative relationship (Figure 1B, Supplementary Table 10). For OS, DSS and PFS, GIAC patients with high infiltration of CD8 T cell had favor survival rate, while those with high infiltration of CD4 T cell had poor survival rate (Figure 1C–1E, Supplementary Table 11). The TME landscape depicted the infiltration fraction and interactions of GIAC immune cells, as well as their effects on patient survival outcome.

Three GIAC subtypes were determined based on TME pattern

Immunotyping can mirror the immune status in tumors and their TME, and thus benefit for identifying suitable patients for immunotherapy. Based on the TME pattern of 1,100 GIAC patients with matched 160 immune features from panImmune project, unsupervised clustering was performed using three distinct algorithms (Lee, brunet and nsNMF), and their classification consistency (Kappa value) was close to and above 0.7 (Supplementary Table 12), which indicated the stability of classification based on TME pattern of immune infiltration and expression. The nsNMF algorithm was selected due to the highest average classification consistency than the others, and the optimal cluster number was set to three according to multiple clustering indicators (Figure 2A, 2B, Supplementary Figures 1, 2, Supplementary Table 13). The GIAC patients from three main groups (termed as TME C1 (403 patients), C2 (276 patients) and C3 (324 patients) were determined (Figure 2C). They showed significant differences in OS, DSS and PFS, where TME C3 and C2 had the best and worst survival, respectively (Figure 2D–2F). When considering the cancer type, COAD and READ were almost evenly distributed in the three groups, while TME C2 and C3 were enriched in GAD and ESAD, respectively (Figure 2G, Supplementary Table 14). Meanwhile, TME C2 had higher fraction of patients with tumor stage III and IV (Figure 2H, Supplementary Table 14), which was consistent with the above survival outcome, reflecting its high degree of malignancy.

TME-C2 subtype was associated with tumor immunity

By association with genomic characteristics, we discovered that TME C2 was closely associated with lower cancer DNA fraction and larger Shannon score (Figure 3A, 3B, Supplementary Table 14), suggesting relatively its high tumor heterogeneity. Considering the crucial role of immune checkpoints (ICPs) and immunogenic cell death (ICD) modulators in cancer immunity, we next investigated their expression level in the three subtypes. Most of the ICPs and ICDs related genes were differentially expressed between the immune subtypes (Supplementary Figures 3, 4). For instance, the current approved and potential immunotherapy targets, PDCD1 (also known as PD-1), CD274 (also known as PD-L1), CTLA4, HAVCR2 (also known as TIM3), LAG3 and BTLA's expression were significantly elevated in TME-C2 (Figure 3C, Supplementary Tables 15, 16). In the context of cancer, PD-L1 is usually highly expressed on tumor cells, thereby evading immune surveillance, but it has also been reported that high PD-L1 expression can make tumor cells more sensitive to PD-1/PD-L1 inhibitors. TME C2 was also related to the higher fraction of leukocyte, CD8 T cell, regulatory T cell and activated NK cell, and the lower fraction of resting NK cell (Figure 3D–3H, Supplementary Table 17), indicating their potential benefit for immunotherapy, such as PD1 inhibitors or other combination therapies. Therefore, immunotyping could reflect the level of immune modulators and guild the selection of population suitable for immunotherapy.

Two TME signatures were involved in stromal and immune functions

To uncover the underlying biological mechanism of the TME phenotypes, high-confidence differential expressed genes (DEGs) were acquired across TME phenotypes, where TBC1D3G was specifically highly expressed in TME C1, 54 genes (termed as signature A) was expressed in TME C2, and 85 genes (termed as signature B) was expressed in TME C3 (Figure 4A, Supplementary Table 18). TBC1D3G was reported to regulate the payload of macrophage-released extra-cellular vesicles to mediate inflammation [40]. We focused our attention on signature A and B, which showed enrichment in distinct molecular functions by literature annotation and overlapping with panImmune gene sets (Figure 4B, Supplementary Table 19). Signature B was involved in tumor immunity in GIAC, while signature A was involved in stromal-related function (Figure 4C, Supplementary Tables 20, 21). GO functional enrichment analysis also proved that signature B, instead of signature A, was enriched in chemokine activity and chemokine receptor binding (Figure 4D). Chemokines and their receptors play a key role in tumor growth, invasion and metastasis, as well as differentiation and development of immune cells and the regulation of immune response [41]. Gene-set
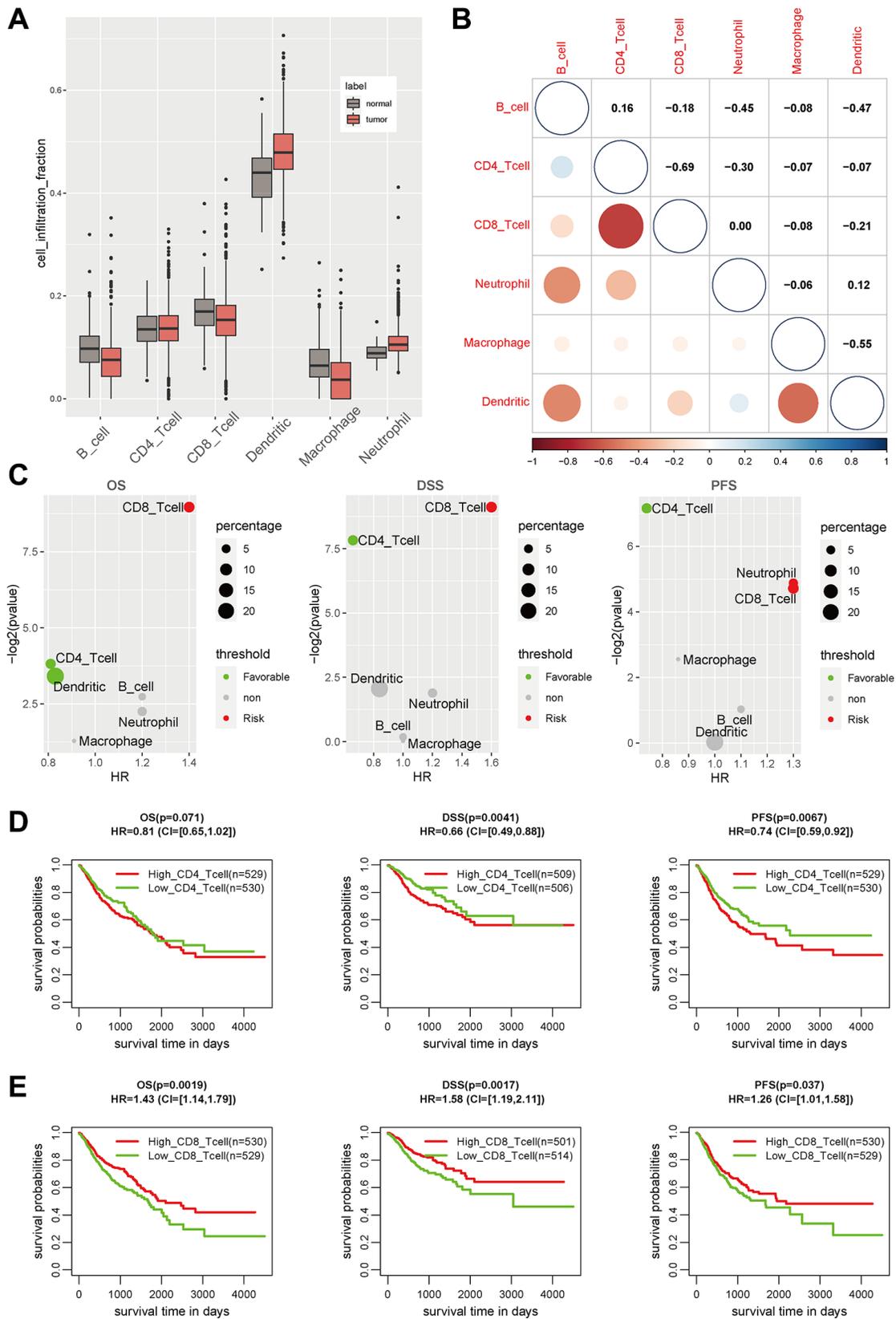
Figure 1. The immune landscape in GIAC tumors. (**A**) The cell infiltration level of six main immune cell types in GIAC tumor and adjacent normal samples. (**B**) The cell infiltration correlation of six immune cell types. (**C**) The P-value and HR of cell infiltration level of six immune cell types in survival analysis. (**D**) The KM plot of CD8 T cell infiltration in GIAC patients for OS, DSS and PFS. (**E**) The KM plot of CD4 T cell infiltration in GIAC patients for OS, DSS and PFS.

enrichment analysis (GSEA) also demonstrated that signature B was indeed involved in inflammatory and immune processes (Figure 4E).

TMEscore was a prognostic biomarker for predicting GIAC patient outcome

To refine the gene signature A and B, the random forest classification algorithm was used to perform dimension reduction in order to reduce noise or redundant genes. Overlapping with the common up-regulated genes in three cancer types (CRC, GAD and ESAD), 11 genes in signature A and 34 genes in signature B were kept. To investigate the relationship between these genes with patient survival, univariate Cox regression analysis was performed and 6 prognostic genes (C6orf223, EPHX4, HES6 and NKD2 in signature A, OLR1 and ONECUT2 in signature B) were identified to correlate to both
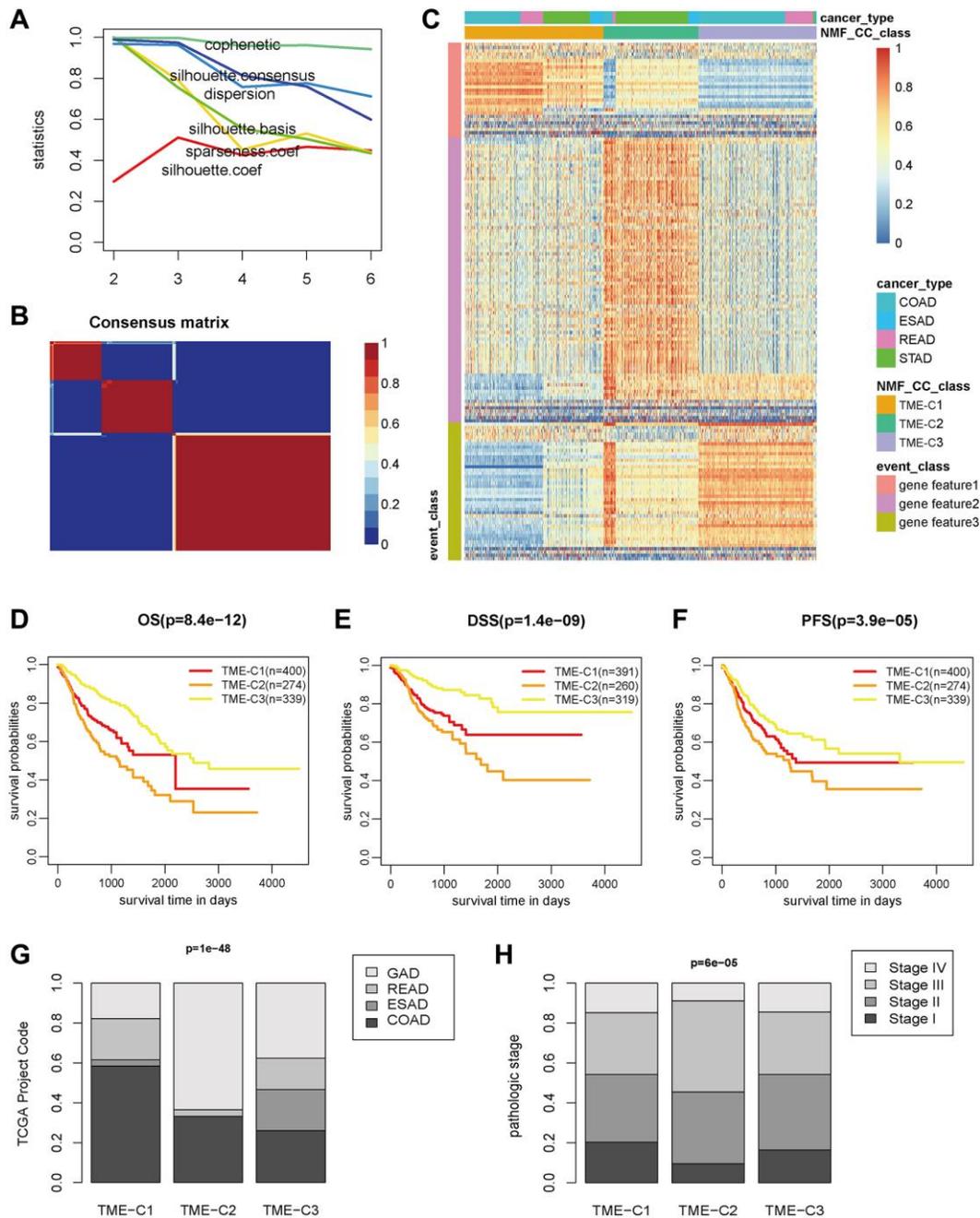
Figure 2. The three TME subtypes in GIAC tumors. (**A**) The classification indexes of different clusters in NMF results. (**B**) The consensus matrix when the cluster set as 3. (**C**) The heatmap of immune features for three TME subtypes determined by NMF classification analysis. (**D–F**) The KM plot of three TME subtypes in GIAC patients for OS, DSS and PFS. (**G**, **H**) The proportion of cancer types and tumor grade for the three TME subtypes.

patient OS and PFS (Figure 5A, Supplementary Table 22). The patients with high expression of C6orf223, EPHX4, HES6 and NKD2 had significantly longer survival time than other patients, while patients with high expression of OLR1 and ONECUT2 had significantly shorter survival time (Supplementary Figures 5, 6), which suggested C6orf223, EPHX4, HES6 and NKD2 were protective factors and OLR1 and ONECUT2 were risk factors.

We then focused on their expression status at the RNA and protein levels. Their RNA expression in GIAC tumors were significantly elevated compared with adjacent normal samples in TCGA cohort (Figure 5B). We subsequently explored their RNA expression status in 96 GIAC patients (55 CRC, 32 GAD and 9 ESAD) from our in-house cohort by qRT-PCR assay. Their RNA expression levels were significantly up-regulated in GIAC compared with matched adjacent normal

samples (Figure 5C, Supplementary Figures 7, 8). Since C6orf223 was annotated as a long noncoding RNA in GeneCards, thus we investigated the protein expression status of other five protein-coding genes in 27 GIAC patients (10 CRC, 10 GAD and 7 ESAD) from our in-house cohort by Western blot assay. Their protein expression levels were also significantly elevated in GIAC compared with matched adjacent normal samples (Figure 5D, Supplementary Figures 9, 10; Figure 5E and Supplementary Figures 11, 12). Similar results in GIAC tumors were also confirmed in multiple previous studies (Supplementary Table 23). Interestingly, though EPHX4 has hardly been studied in GIAC, the immunohistochemical (IHC) data from the HPA database revealed that its staining showed moderate to strong cytoplasmic immunoreactivity in CRC and GAD, and weak immunoreactivity in normal colon/rectum and stomach tissues (Supplementary Figures 13, 14).
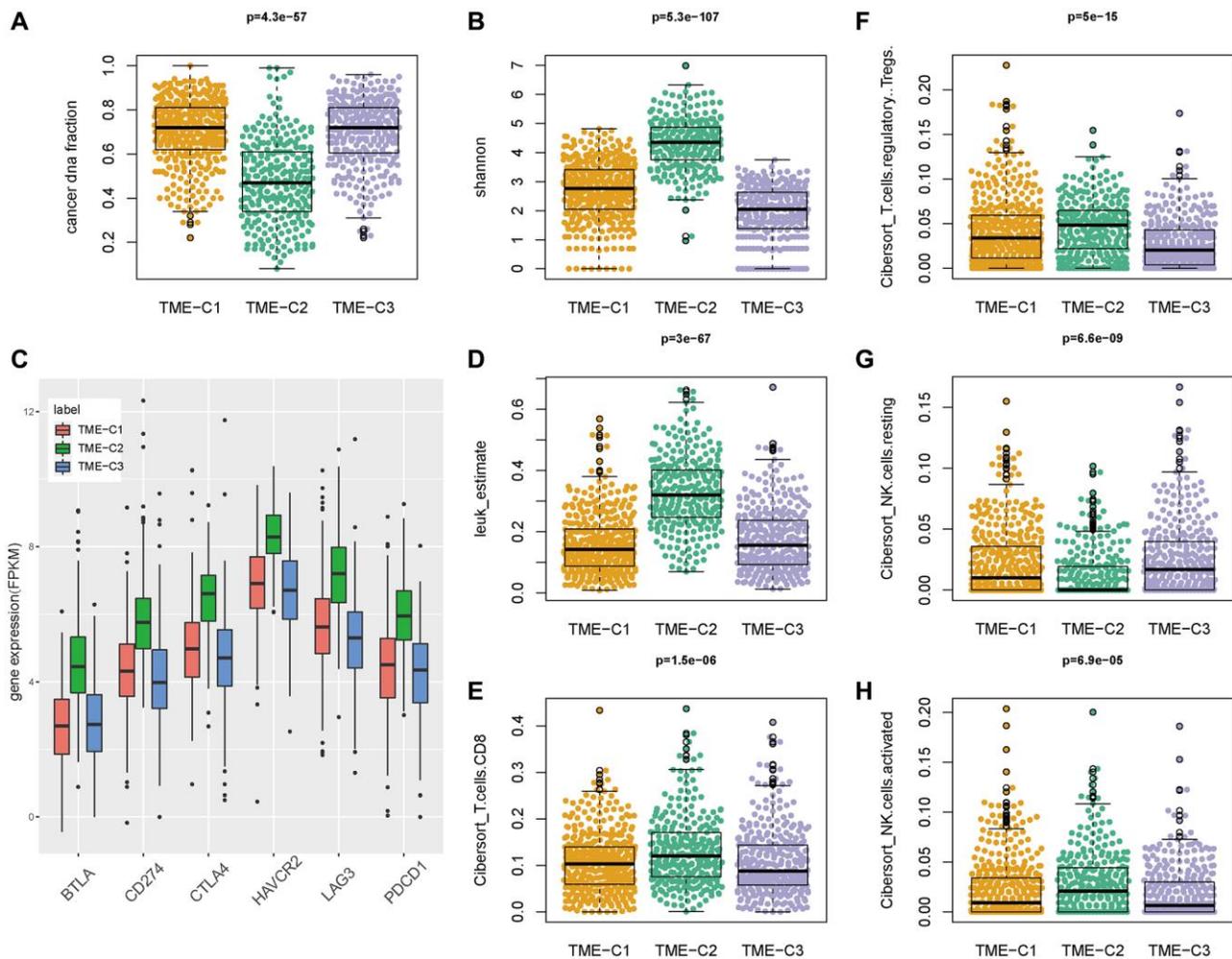

Figure 3. TME-C2 subtype was associated with tumor immunity. (**A**, **B**) The cancer DNA fraction and Shannon score across three TME subtypes. (**C**) The expression of current approved and potential immunotherapy targets (PDCD1, CD274, CTLA4, HAVCR2, LAG3 and BTLA) across three TME subtypes. (**D–H**) The immune cell infiltration fraction (leukocyte, CD8 T cell, regulatory T cell, resting NK cell, and activated NK cell) across three TME subtypes.

We used the weighted average expression of these six genes to construct a TMEscore. Univariate and multivariate Cox regression analysis showed TMEscore was as an independent prognostic factor in GIAC even considering the confounding factors such as age, gender and tumor stage (Supplementary Tables 24, 25). In TCGA, we observed that the patients with high TMEscore had significantly better prognosis than others, suggesting the predictive value of TMEscore for patient survival in GIAC (HR = 2.5, P = 3.3e-16 for OS and HR = 1.81, P = 2.6e-8 for PFS, Figure 5F, 5G). The nomogram analysis for OS also showed the good predictive ability of TMEscore as well as gender and tumor stage (Supplementary Figure 15), thus ROC curve analysis based on the independent prognostic factors (TMEscore, age, gender, and tumor stage) was performed. The AUC of TMEscore (0.81, 0.8, and 0.8 for 1-, 3- and 5-year OS) were obviously higher that of
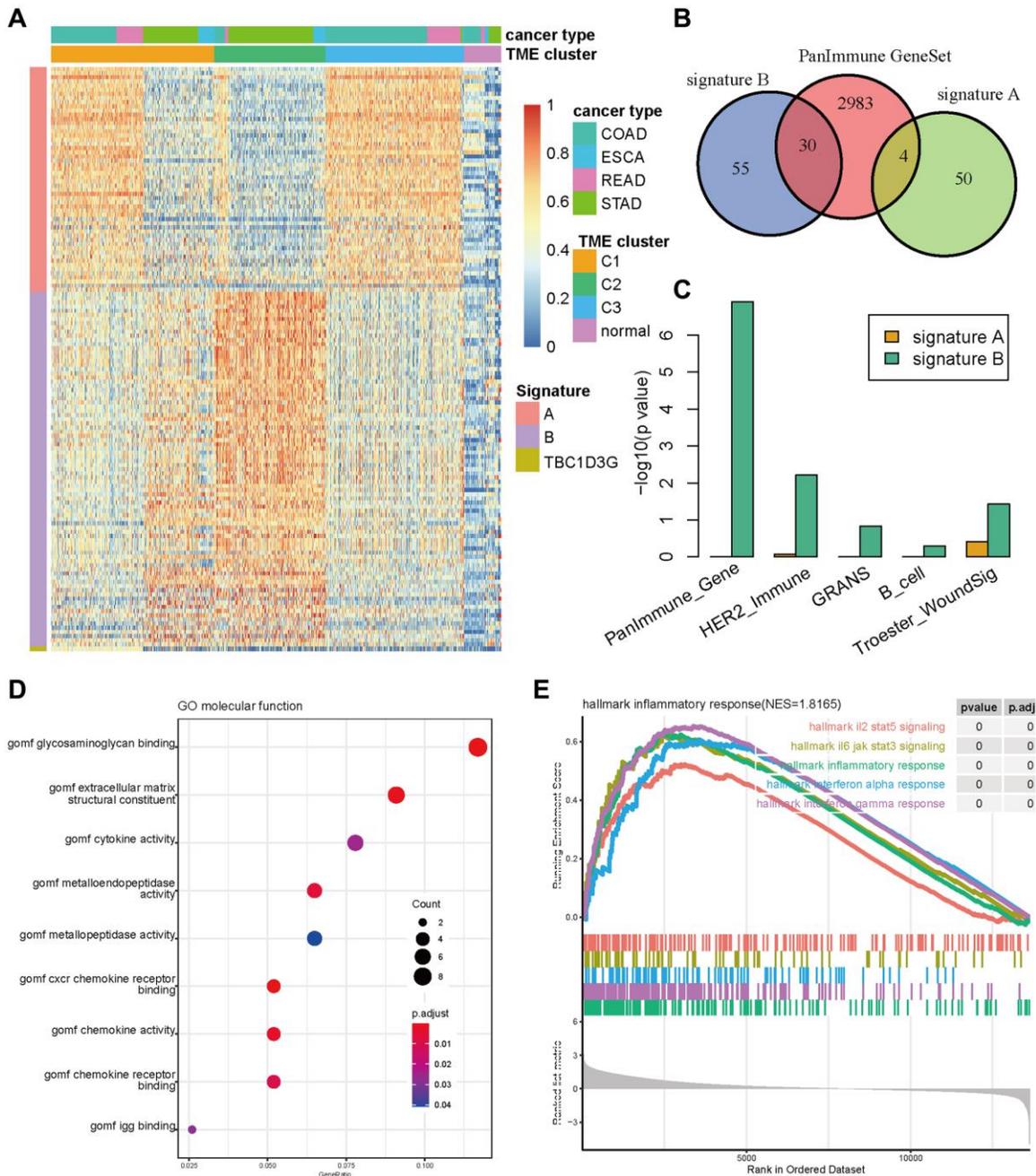
Figure 4. The signature genes differentiating TME subtypes. (**A**) Gene expression of signature A, B and TBC1D3G in three TME subtypes and normal samples. (**B**) The Venn diagram of signature A, B and PanImmune gene sets. (**C**) The enrichment of signature A and B in immune related gene sets. (**D**) The bubble diagram of signature A enriched in GO terms. (**E**) The GSEA plot of signature A enriched in cancer hallmark gene sets.
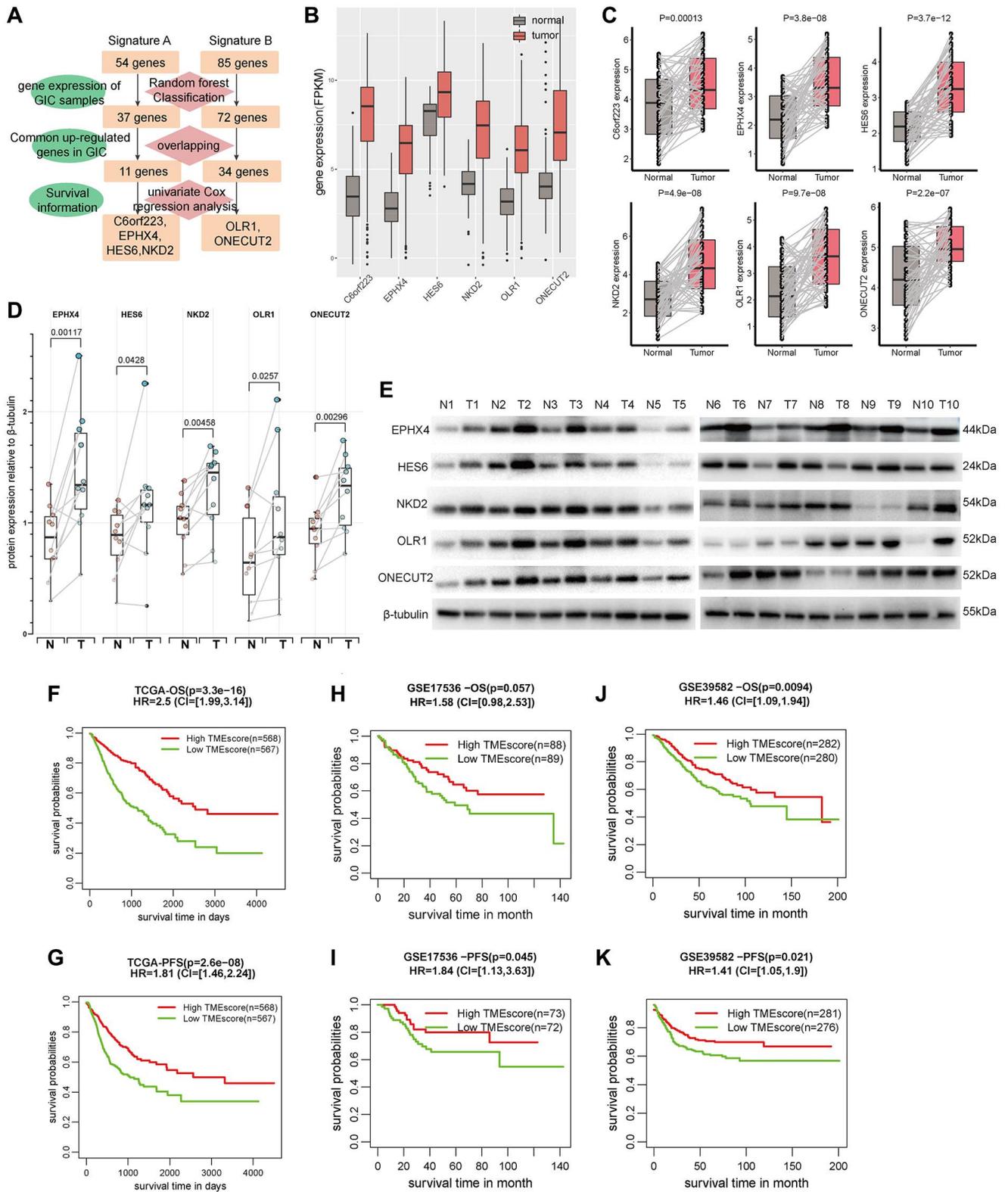
Figure 5. TMEscore was an independent factor for predicting patient survival. (**A**) The identification flowchart of six prognostic genes (C6orf223, EPHX4, HES6, NKD2 OLR1 and ONECUT2). (**B**) The expression of the six genes in GIAC tumor and normal samples from TCGA cohort. (**C**) The RNA expression of the six genes in 55 pairs of CRC tumor and normal samples from our center (GAD and ESAD see in Supplementary Figures 7, 8). (**D**) The expression of the five proteins in 10 pairs of CRC tumor and normal samples from our center (GAD and ESAD see in Supplementary Figures 9, 10). (**E**) The WB image of protein expression for CRC (GAD and ESAD see in Supplementary Figures 11, 12). (**F–K**) The survival time (OS and PFS) of patients with high/low TMEscore in TCGA cohort (**F**, **G**), GSE17536 (**H**, **I**) and GSE39582 (**J**, **K**).

age, gender, and tumor stage (Supplementary Figure 16). The calibration plot also showed good consistency between observation and predictive values for 1-, 3- and 5-year OS (Supplementary Figure 17). The nomogram analysis and ROC curve analysis for PFS had similar results (Supplementary Figures 18–20). In addition, we analyzed the relevance of TMEscore and survival outcome in GSE17536 and GSE39582 from GEO and validated the prognostic value of TMEscore (Figure 5H–5K). These results declared that TMEscore was a potential biomarker for predicting GIAC patient survival.

TMEscore predicted patient survival and immunotherapeutic benefits

In order to illustrate the role of TMEscore in tumor immunotherapy, we collected six immunotherapy datasets (Gide, Mariathasan, Auslander, Riaz, VanAll and Kim cohort) in melanoma, urothelial carcinoma and gastric cancer. Then we computed TMEscore based on the RNA expression of the six genes (C6orf223, EPHX4, HES6, NKD2, OLR1 and ONECUT2) and assessed TMEscore's relevance of patient survival time and immunotherapy response.

In Gide cohort, the 41 melanoma patients treated with PD1-CTLA4 combination inhibitor with high TMEscore had significantly better OS and PFS than other patients ($P = 0.0012$, HR = 2.387, CI = [1.22, 6.77] for OS; $P = 0.0063$, HR = 1.31, CI = [1.31, 6.15] for PFS, respectively, Figure 6A, 6B). TMEscore of the responders ($n = 19$) was apparently higher than that of non-responders ($n = 22$) ($P = 0.0097$, Figure 6C), and the patients with high TMEscore had significantly higher response rate than those with low TMEscore ($P = 0.0048$, Figure 6D). The TMEscore had a good immunotherapeutic predictive value (AUC = 0.73, Figure 6E). Similarly, the 32 melanoma patients treated with PD1 inhibitor with high TMEscore had significantly better PFS than other patients ($P = 0.045$, HR = 2.89, CI = [1.11, 9.41], Figure 6G), while OS had no statistical difference (Figure 6F). TMEscore of the responders ($n = 21$) was higher than that of non-responders ($n = 11$) ($P = 0.0048$, Figure 6H), and the patients with high TMEscore had increased response rate compared with the others ($P = 0.0023$, Figure 6I). TMEscore also had an outstanding immunotherapeutic predictive value (AUC = 0.8, Figure 6J). We observed a similar phenomenon in the other four data sets (Mariathasan, Auslander, Riaz, and VanAll cohort) (Supplementary Figure 21).

Besides, in Kim cohort, 45 gastric cancer patients were treated with PD1 inhibitor and 12 patients have complete response or partial response. The TMEscore

of the responders ($n = 33$) was higher than that of non-responders ($n = 12$) ($P = 0.0015$, Figure 6K), and the patients with high TMEscore had increased response rate compared with the others ($P = 0.0074$, Figure 6L). The immunotherapeutic predictive value was AUC = 0.8 (Figure 6M). Furthermore, TMEscore was highly correlated with eight response scores based on hallmarks of immune response estimated by "EaSIeR" and the integrated score in TCGA GIAC dataset (Supplementary Figure 22, Figure 6N). In summary, these results demonstrated that TMEscore could predict tumor survival and immunotherapy response status.

DISCUSSION

Despite numerous clinical trials about gastrointestinal cancer treatment in the last two decades, especially for metastatic patients, the clinical outcome is still not optimistic, and the survival usually is less than 30 months [42]. Innovative ideas, such as the gut microbiota imbalance and tumor immune microenvironment, have been introduced into the basic research of GIAC and tentative clinical treatment [43, 44]. In addition to traditional radiotherapy and chemotherapy, immunotherapy, cellular therapy, molecular targeted therapy and microbial therapy have developed rapidly [45], thus forming some experience and consensus guiding clinical practice. Hence, the development of novel and effective strategies to control GIAC is an urgent need in GIAC prognosis and treatment.

Immune checkpoint blocking (ICB) therapy has caused a great change of the therapeutic landscape, making some advanced-stage cancer patients achieve clinical benefits [46]. It is quite attracting and necessary to develop biomarkers for ICB response for clinical trials and applications [47]. Sun et al. systematically assessed the predictive power of 22 current transcriptomic biomarkers for ICB responses involving immune checkpoints and lymphocyte infiltration in multiple ICB treatment baseline datasets [12]. They revealed that these biomarkers exhibited distinct predictive value for ICB response, where some performed superior overwhelmingly or slightly just in certain circum-stances. Unfortunately, some accepted biomarkers still did not have any predictive value of ICB response in benchmark datasets, indicating the possibility of the combination of multiple biomarkers for predicting ICB response in future.

The underlying mechanisms of immunotherapy have been widespread explored and some consensus has emerged. When T cell is continuously stimulated, T cell will become exhausted and continuously express high PD-1. In the tumor microenvironment, tumor cells can

express PD-L1 or PD-L2. As a result, T cell function is continuously inhibited, which makes it unable to kill tumor cells [48]. One-way cancer cells evade destruction by the immune system is through ligands attached to the PD-1 protein of T cells. When the ligand binds to PD-1, T cells are unable to detect tumors and deliver signal the immune system to attack them [49]. The mechanism of PD-L1/PD-1 antibody is that tumor
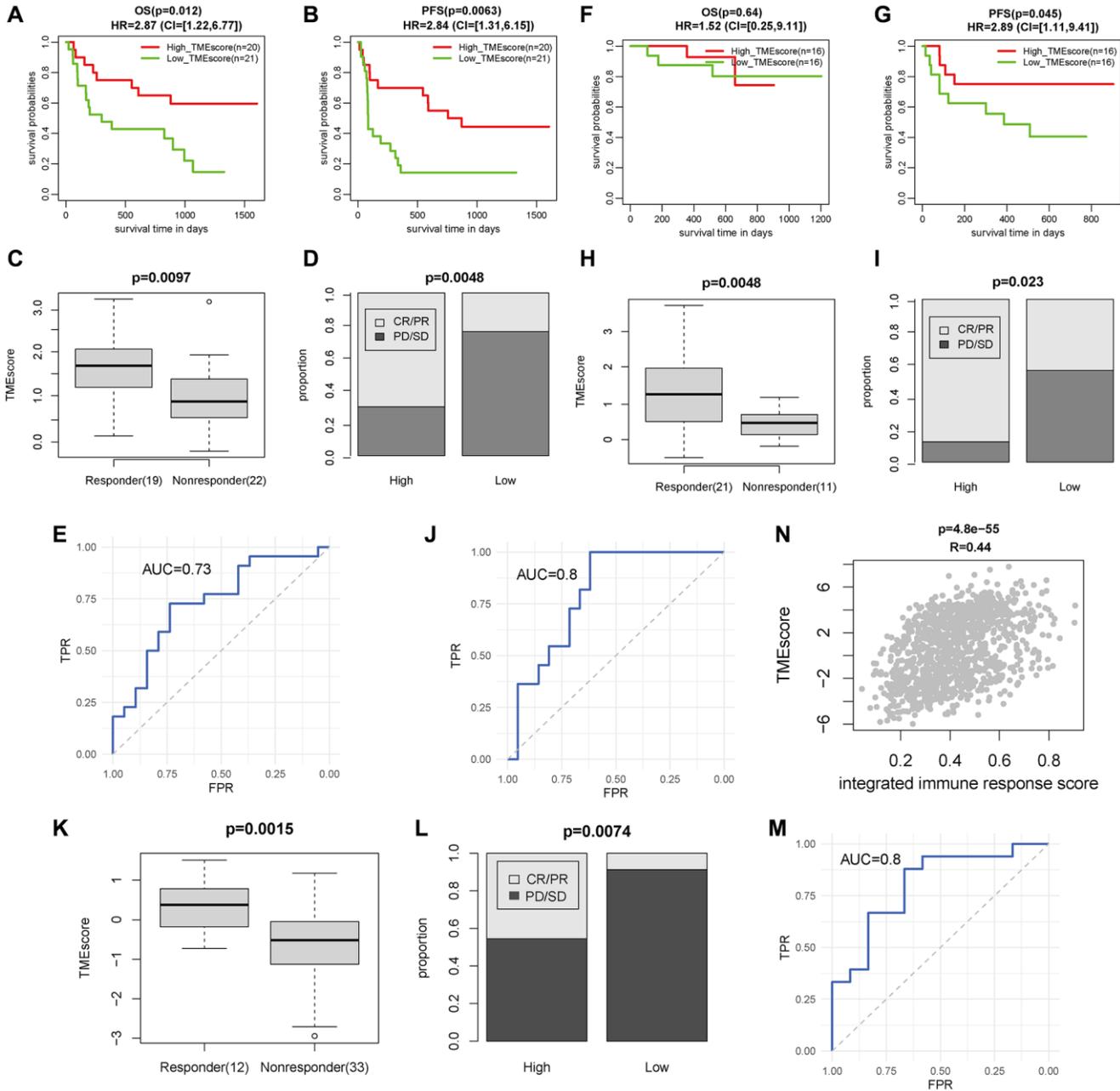

Figure 6. TMEscore was a prognostic and immunotherapeutic biomarker in Gide, Kim and TCGA GIAC cohort. (**A–C**) The survival time (OS and PFS) and response rate of patients with high/low TMEscore for 41 patients treated with PD1+CTLA4 inhibitor in Gide cohort. (**D**) The TMEscore of patients with different response status for patients treated with PD1+CTLA4 inhibitor. (**E**) The ROC curve of TMEscore predicting immunotherapeutic response for patients treated with PD1+CTLA4 inhibitor. (**F–H**) The survival time (OS and PFS) and response rate of patients with high/low TMEscore for 32 patients treated with PD1 inhibitor in Gide cohort. (**I**) The TMEscore of patients with different response status for patients treated with PD1 inhibitor. (**J**) The ROC curve of TMEscore predicting immunotherapeutic response for patients treated with PD1 inhibitor. (**K**) The response rate of patients with high/low TMEscore for 45 gastric cancer patients treated with PD1 inhibitor in Kim cohort. (**L**) The TMEscore of patients with different response status for patients treated with PD1 inhibitor. (**M**) The ROC curve of TMEscore predicting immunotherapeutic response for patients treated with PD1 inhibitor. (**N**) The correlation plot of TMEscore with integrated immune response score in TCGA GIAC cohort.

cells use PD-L1 to bind to the PD-1 of T cells, "trick" T cells, evade the recognition of T cells, and continue to roam in the body. However, PD-L1/PD-1 inhibitors can help T cells uncover the hypocrisy of tumor cells and restore their recognition and killing of tumor cells [50]. In recent years, immunotherapy has shown great success in improving response to some solid cancers, especially melanoma [51]. Thus, numerous clinical and preclinical studies have motivated researchers to focus their attention on gastrointestinal cancers, expecting that patients will benefit from immunotherapy, as represented by PD1 and CTLA4 blockers. In 2021, PD-1 antibody combined with chemotherapy achieved a substantial breakthrough in the first-line treatment of esophageal cancer. In Keynote-590 phase 3 trial, first-line PD-L1 inhibitor pembrolizumab plus 5-fluorouracil or cisplatin significantly improved median OS and PFS in esophageal cancer patients [52]. The first-line treatment regimen of PD-1 monoclonal antibody combined with chemotherapy or HER2 monoclonal antibody were found to improve the survival and effective rate of HER2-negative gastric cancer patients. In CheckMate 649 phase 3 trial, PD-L1 inhibitor nivolumab combined with FOLFOX or CAPE-OX as first-line therapy had better OS and PFS than first-line chemotherapy [53]. For MSI-H/dMMR CRC, immune monotherapy showed consistent advantages and moved forward to become the new standard of first-line treatment. In Keynote-177 phase 3 trial, pembrolizumab as first-line therapy in MSI-H/dMMR metastatic CRC tended to reduce the risk of death and increase PFS compared with standard therapy [54]. The success of these studies heralded a shift in the treatment strategy for certain types of gastrointestinal cancer, and prompting the National Comprehensive Cancer Network (NCCN) or Chinese Society of Clinical Oncology (CSCO) guidelines to incorporate new protocols.

In CRC, the level of T cell infiltration to tumor is directly related to the therapeutic effect of tumor patients [55], suggesting that immune cell infiltration in tumor microenvironment plays a key role in inhibiting tumor growth. The MSI-H/dMMR CRC showed a higher accumulation of tumor mutations, accompanied by a higher level of immune cell infiltration. Pembrolizumab, nivolumab, and ipilimumab were approved to treat MSI-H/dMMR tumors mainly due to their severe infiltration with CD8+/CD4+ T cells T cells [56]. Metastatic CRC generally has a higher level of mutation accumulation and thus has a response effect to immunotherapy. Although advanced (Stage IV) MSI-H/dMMR tumors account for 2–4% of all metastatic CRC, they have high expression of PD1, PDL1 and CTLA4, making this subtype more sensitive to immune checkpoint inhibitors. The low mutation load and immune cell infiltration are considered to be the main causes of immune resistance and non-response to immunotherapy for MSI-L/pMMR CRC [57]. Currently, immune checkpoint inhibitors in combination with other therapies are being explored in various preclinical and clinical trials for treatment of CRC with MSI-L/pMMR [58].

Although we have revealed TMEscore to be a biomarker for predicting prognosis and immuno-therapeutic response, there are still some limitations and deficiencies in this study. First, the expression and biological function of the six key genes need further experimental investigation, including expression assays in more patients, as well as cell and animal function assays. Second, at present, immunotherapy datasets mainly focus on melanoma and urothelial carcinoma, there is indeed few immunotherapy datasets in GIAC patients except for Kim cohort (gastric cancer). We consider that many antitumor drugs have a broad spectrum and sometimes can be applied in multiple tumors, which is no exception for the immune-checkpoint inhibitors targeting PD1 and CTLA4. Thus, we extra used the transcriptional sequencing data of GIAC patients and immunotherapy response status of melanoma and urothelial carcinoma patients to assess the value of TMEscore. Though this method often used in previous studies [59–64] can be serve as a tentative and preliminary expansion application, it is indeed biased and not rigorous. When relevant data is publicly available in the future, we can precisely assess the predicted value of TMEscore in matched datasets. Finally, whether TMEscore can be serve as a prognostic and immunotherapeutic biomarker of pan-cancer is worth exploring in multiple human cancers.

In conclusion, we comprehensively analyzed the pattern of immune infiltration and immune pathway expression in over 1,000 GIAC patients and revealed three TME associated patient subgroups accompanied with quite distinct immune and clinicopathologic features. Based on the RNA expression of six key genes from two immune/stromal signatures, TMEscore was established and validated to be predictive to patient survival outcome and response to immune-checkpoint inhibitors in multiple immunotherapeutic datasets. In summary, depicting a comprehensive TME landscape will benefit for understanding the underlying mechanisms in GIAC, such as cell communication and immunosuppression or activation. TMEscore will be useful to account for the responses of GIAC to immunotherapies and provide new strategies for the treatment of cancers.

AUTHOR CONTRIBUTIONS

Kunkun Li, Huili Wu and Yong Zhang offered scientific idea and designed the study. Yong Zhang, Lu

Li and Feifei Chu conducted molecular experiments and data analysis, designed the figures and drafted the manuscript. Lei Zhang, Li Zhang, Kunkun Li, and Huili Wu inspected and modified the manuscript. All authors read and approved the final version.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

ETHICS STATEMENT AND CONSENT

All methods were carried out in accordance with relevant guidelines and regulations. The study was reviewed and approved by Ethics Committee of Zhengzhou Central Hospital Affiliated to Zhengzhou University. Informed consent was obtained from all subjects.

FUNDING

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021; 71:209–49.
https://doi.org/10.3322/caac.21660
PMID:33538338

2. Schepers A, Clevers H. Wnt signaling, stem cells, and cancer of the gastrointestinal tract. Cold Spring Harb Perspect Biol. 2012; 4:a007989.
https://doi.org/10.1101/cshperspect.a007989
PMID:22474007

3. Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F, Seoane JA, Farshidfar F, Bowlby R, Islam M, Kim J, Chatila W, Akbani R, et al, and Cancer Genome Atlas Research Network. Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. Cancer Cell. 2018; 33:721–35.e8.
https://doi.org/10.1016/j.ccell.2018.03.010
PMID:29622466

4. Bijlsma MF, Sadanandam A, Tan P, Vermeulen L. Molecular subtypes in cancers of the gastrointestinal tract. Nat Rev Gastroenterol Hepatol. 2017; 14:333–42.
https://doi.org/10.1038/nrgastro.2017.33
PMID:28400627

5. DeBerardinis RJ. Tumor Microenvironment, Metabolism, and Immunotherapy. N Engl J Med. 2020; 382:869–71.
https://doi.org/10.1056/NEJMcibr1914890
PMID:32101671

6. Wu T, Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett. 2017; 387:61–8.
https://doi.org/10.1016/j.canlet.2016.01.043
PMID:26845449

7. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. Cancer Res. 2019; 79:4557–66.
https://doi.org/10.1158/0008-5472.CAN-18-3962
PMID:31350295

8. Zeng D, Zhou R, Yu Y, Luo Y, Zhang J, Sun H, Bin J, Liao Y, Rao J, Zhang Y, Liao W. Gene expression profiles for a prognostic immunoscore in gastric cancer. Br J Surg. 2018; 105:1338–48.
https://doi.org/10.1002/bjs.10871
PMID:29691839

9. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L, Ye G, Deng H, Mou T, Cai S, Zhou Z, Liu H, Chen G, et al. ImmunoScore Signature: A Prognostic and Predictive Tool in Gastric Cancer. Ann Surg. 2018; 267:504–13.
https://doi.org/10.1097/SLA.0000000000002116
PMID:28002059

10. Zhang C, Li D, Yu R, Li C, Song Y, Chen X, Fan Y, Liu Y, Qu X. Immune Landscape of Gastric Carcinoma Tumor Microenvironment Identifies a Peritoneal Relapse Relevant Immune Signature. Front Immunol. 2021; 12:651033.
https://doi.org/10.3389/fimmu.2021.651033
PMID:34054812

11. Zhuge L, Huang B, Xie J, Gao Z, Zheng D, Zheng S, Xiang J, Zhang J. Immunoscore Signature Predicts Postoperative Survival and Adjuvant Chemotherapeutic Benefits in Esophageal Squamous Cell Carcinoma. Cancer Manag Res. 2020; 12:12885–94.
https://doi.org/10.2147/CMAR.S279684
PMID:33364836

12. Sun S, Xu L, Zhang X, Pang L, Long Z, Deng C, Zhu J, Zhou S, Wan L, Pang B, Xiao Y. Systematic Assessment of Transcriptomic Biomarkers for Immune Checkpoint Blockade Response in Cancer Immunotherapy. Cancers (Basel). 2021; 13:1639.
https://doi.org/10.3390/cancers13071639
PMID:33915876

13. Xiong D, Wang Y, You M. A gene expression signature of TREM2[hi] macrophages and γδ T cells predicts immunotherapy response. Nat Commun. 2020; 11:5084.

https://doi.org/10.1038/s41467-020-18546-x
PMID:33033253

14. Zhang X, Shi M, Chen T, Zhang B. Characterization of the Immune Cell Infiltration Landscape in Head and Neck Squamous Cell Carcinoma to Aid Immunotherapy. Mol Ther Nucleic Acids. 2020; 22:298–309.
https://doi.org/10.1016/j.omtn.2020.08.030
PMID:33230435

15. Xu WH, Xu Y, Wang J, Wan FN, Wang HK, Cao DL, Shi GH, Qu YY, Zhang HL, Ye DW. Prognostic value and immune infiltration of novel signatures in clear cell renal cell carcinoma microenvironment. Aging (Albany NY). 2019; 11:6999–7020.
https://doi.org/10.18632/aging.102233
PMID:31493764

16. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m^{6}A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. Mol Cancer. 2020; 19:53.
https://doi.org/10.1186/s12943-020-01170-0
PMID:32164750

17. Carino A, Graziosi L, Marchianò S, Biagioli M, Marino E, Sepe V, Zampella A, Distrutti E, Donini A, Fiorucci S. Analysis of Gastric Cancer Transcriptome Allows the Identification of Histotype Specific Molecular Signatures With Prognostic Potential. Front Oncol. 2021; 11:663771.
https://doi.org/10.3389/fonc.2021.663771
PMID:34012923

18. Li X, Xie M, Yin S, Xiong Z, Mao C, Zhang F, Chen H, Jin L, Lan P, Lian L. Identification and Validation of a Six Immune-Related Genes Signature for Predicting Prognosis in Patients With Stage II Colorectal Cancer. Front Genet. 2021; 12:666003.
https://doi.org/10.3389/fgene.2021.666003
PMID:34017356

19. Zhu X, Tian X, Ji L, Zhang X, Cao Y, Shen C, Hu Y, Wong JWH, Fang JY, Hong J, Chen H. A tumor microenvironment-specific gene expression signature predicts chemotherapy resistance in colorectal cancer patients. NPJ Precis Oncol. 2021; 5:7.
https://doi.org/10.1038/s41698-021-00142-x
PMID:33580207

20. Mori T, Kumagai K, Nasu K, Yoshizawa T, Kuwano K, Hamada Y, Kanazawa H, Suzuki R. Clonal Expansion of Tumor-Infiltrating T Cells and Analysis of the Tumor Microenvironment within Esophageal Squamous Cell Carcinoma Relapsed after Definitive Chemoradiation Therapy. Int J Mol Sci. 2021; 22:1098.
https://doi.org/10.3390/ijms22031098
PMID:33499345

21. Guo W, Tan F, Huai Q, Wang Z, Shao F, Zhang G, Yang Z, Li R, Xue Q, Gao S, He J. Comprehensive Analysis of PD-L1 Expression, Immune Infiltrates, and m6A RNA Methylation Regulators in Esophageal Squamous Cell Carcinoma. Front Immunol. 2021; 12:669750.
https://doi.org/10.3389/fimmu.2021.669750
PMID:34054840

22. Fakih M, Ouyang C, Wang C, Tu TY, Gozo MC, Cho M, Sy M, Longmate JA, Lee PP. Immune overdrive signature in colorectal tumor subset predicts poor clinical outcome. J Clin Invest. 2019; 129:4464–76.
https://doi.org/10.1172/JCI127046
PMID:31524634

23. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM, and Cancer Genome Atlas Research Network The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013; 45:1113–20.
https://doi.org/10.1038/ng.2764
PMID:24071849

24. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, Utikal J, Hassel JC, Weide B, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science. 2015; 350:207–11.
https://doi.org/10.1126/science.aad0095
PMID:26359337

25. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, Hodi FS, Martín-Algarra S, Mandal R, Sharfman WH, Bhatia S, Hwu WJ, Gajewski TF, et al. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. Cell. 2017; 171:934–49.e16.
https://doi.org/10.1016/j.cell.2017.09.028
PMID:29033130

26. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE III, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018; 554:544–8.
https://doi.org/10.1038/nature25501
PMID:29443960

27. Auslander N, Zhang G, Lee JS, Frederick DT, Miao B, Moll T, Tian T, Wei Z, Madan S, Sullivan RJ, Boland G, Flaherty K, Herlyn M, Ruppin E. Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. Nat Med. 2018; 24:1545–9.
https://doi.org/10.1038/s41591-018-0157-9
PMID:30127394

28. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J, Madore J, Lim SY, Velickovic R, Wongchenko M, Yan Y, Lo S, Carlino MS, et al. Distinct Immune Cell

Populations Define Response to Anti-PD-1 Monotherapy and Anti-PD-1/Anti-CTLA-4 Combined Therapy. Cancer Cell. 2019; 35:238–55.e6.
https://doi.org/10.1016/j.ccell.2019.01.003
PMID:30753825

29. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JI, Kim K, Liu XQ, Sher X, Jung H, Lee M, Lee S, Park SH, Park JO, et al. Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. Nat Med. 2018; 24:1449–58.
https://doi.org/10.1038/s41591-018-0101-z
PMID:30013197

30. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. 2020; 48:W509–14.
https://doi.org/10.1093/nar/gkaa407
PMID:32442275

31. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12:453–7.
https://doi.org/10.1038/nmeth.3337
PMID:25822800

32. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, et al, and Cancer Genome Atlas Research Network. The Immune Landscape of Cancer. Immunity. 2018; 48:812–30.e14.
https://doi.org/10.1016/j.immuni.2018.03.023
PMID:29628290

33. Hofree M, Shen JP, Carter H, Gross A, Ideker T. Network-based stratification of tumor mutations. Nat Methods. 2013; 10:1108–15.
https://doi.org/10.1038/nmeth.2651
PMID:24037242

34. Diboun I, Wernisch L, Orengo CA, Koltzenburg M. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC Genomics. 2006; 7:252.
https://doi.org/10.1186/1471-2164-7-252
PMID:17029630

35. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–50.
https://doi.org/10.1073/pnas.0506580102
PMID:16199517

36. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015; 1:417–25.

https://doi.org/10.1016/j.cels.2015.12.004
PMID:26771021

37. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb). 2021; 2:100141.
https://doi.org/10.1016/j.xinn.2021.100141
PMID:34557778

38. Speiser JL, Miller ME, Tooze J, Ip E. A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling. Expert Syst Appl. 2019; 134:93–101.
https://doi.org/10.1016/j.eswa.2019.05.028
PMID:32968335

39. Lapuente-Santana Ó, van Genderen M, Hilbers PAJ, Finotello F, Eduati F. Interpretable systems biomarkers predict response to immune-checkpoint inhibitors. Patterns (N Y). 2021; 2:100293.
https://doi.org/10.1016/j.patter.2021.100293
PMID:34430923

40. Leung CCT, Wong CKC. Characterization of stanniocalcin-1 expression in macrophage differentiation. Transl Oncol. 2021; 14:100881.
https://doi.org/10.1016/j.tranon.2020.100881
PMID:33074126

41. Mollica Poeta V, Massara M, Capucetti A, Bonecchi R. Chemokines and Chemokine Receptors: New Targets for Cancer Immunotherapy. Front Immunol. 2019; 10:379.
https://doi.org/10.3389/fimmu.2019.00379
PMID:30894861

42. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, Aranda Aguilar E, Bardelli A, Benson A, Bodoky G, Ciardiello F, D'Hoore A, Diaz-Rubio E, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. Ann Oncol. 2016; 27:1386–422.
https://doi.org/10.1093/annonc/mdw235
PMID:27380959

43. Lin A, Zhang J, Luo P. Crosstalk Between the MSI Status and Tumor Microenvironment in Colorectal Cancer. Front Immunol. 2020; 11:2039.
https://doi.org/10.3389/fimmu.2020.02039
PMID:32903444

44. Gagnière J, Raisch J, Veziant J, Barnich N, Bonnet R, Buc E, Bringer MA, Pezet D, Bonnet M. Gut microbiota imbalance and colorectal cancer. World J Gastroenterol. 2016; 22:501–18.
https://doi.org/10.3748/wjg.v22.i2.501
PMID:26811603

45. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, Diaz LA Jr. Immunotherapy in colorectal

cancer: rationale, challenges and potential. Nat Rev Gastroenterol Hepatol. 2019; 16:361–75.
https://doi.org/10.1038/s41575-019-0126-x
PMID:30886395

46. Kishore C, Bhadra P. Current advancements and future perspectives of immunotherapy in colorectal cancer research. Eur J Pharmacol. 2021; 893:173819.
https://doi.org/10.1016/j.ejphar.2020.173819
PMID:33347822

47. Lee HO, Hong Y, Etlioglu HE, Cho YB, Pomella V, Van den Bosch B, Vanhecke J, Verbandt S, Hong H, Min JW, Kim N, Eum HH, Qian J, et al. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. Nat Genet. 2020; 52:594–603.
https://doi.org/10.1038/s41588-020-0636-z
PMID:32451460

48. Pauken KE, Torchia JA, Chaudhri A, Sharpe AH, Freeman GJ. Emerging concepts in PD-1 checkpoint biology. Semin Immunol. 2021; 52:101480.
https://doi.org/10.1016/j.smim.2021.101480
PMID:34006473

49. Ren D, Hua Y, Yu B, Ye X, He Z, Li C, Wang J, Mo Y, Wei X, Chen Y, Zhou Y, Liao Q, Wang H, et al. Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy. Mol Cancer. 2020; 19:19.
https://doi.org/10.1186/s12943-020-1144-6
PMID:32000802

50. Luoma AM, Suo S, Williams HL, Sharova T, Sullivan K, Manos M, Bowling P, Hodi FS, Rahma O, Sullivan RJ, Boland GM, Nowak JA, Dougan SK, et al. Molecular Pathways of Colon Inflammation Induced by Cancer Immunotherapy. Cell. 2020; 182:655–71.e22.
https://doi.org/10.1016/j.cell.2020.06.001
PMID:32603654

51. Dhatchinamoorthy K, Colbert JD, Rock KL. Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation. Front Immunol. 2021; 12:636568.
https://doi.org/10.3389/fimmu.2021.636568
PMID:33767702

52. Sun JM, Shen L, Shah MA, Enzinger P, Adenis A, Doi T, Kojima T, Metges JP, Li Z, Kim SB, Cho BC, Mansoor W, Li SH, et al, and KEYNOTE-590 Investigators. Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. Lancet. 2021; 398:759–71.
https://doi.org/10.1016/S0140-6736(21)01234-4
PMID:34454674

53. Janjigian YY, Shitara K, Moehler M, Garrido M, Salman P, Shen L, Wyrwicz L, Yamaguchi K, Skoczylas T, Campos Bragagnoli A, Liu T, Schenker M, Yanez P, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. Lancet. 2021; 398:27–40.
https://doi.org/10.1016/S0140-6736(21)00797-2
PMID:34102137

54. André T, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, Smith D, Garcia-Carbonero R, Benavides M, Gibbs P, de la Fouchardiere C, Rivera F, Elez E, et al, and KEYNOTE-177 Investigators. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. N Engl J Med. 2020; 383:2207–18.
https://doi.org/10.1056/NEJMoa2017699
PMID:33264544

55. IJsselsteijn ME, Sanz-Pamplona R, Hermitte F, de Miranda NFC. Colorectal cancer: A paradigmatic model for cancer immunology and immunotherapy. Mol Aspects Med. 2019; 69:123–9.
https://doi.org/10.1016/j.mam.2019.05.003
PMID:31136750

56. Oliveira AF, Bretes L, Furtado I. Review of PD-1/PD-L1 Inhibitors in Metastatic dMMR/MSI-H Colorectal Cancer. Front Oncol. 2019; 9:396.
https://doi.org/10.3389/fonc.2019.00396
PMID:31139574

57. Overman MJ, Lonardi S, Wong KYM, Lenz HJ, Gelsomino F, Aglietta M, Morse MA, Van Cutsem E, McDermott R, Hill A, Sawyer MB, Hendlisz A, Neyns B, et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. J Clin Oncol. 2018; 36:773–9.
https://doi.org/10.1200/JCO.2017.76.9901
PMID:29355075

58. Zhang L, Zhao Y, Dai Y, Cheng JN, Gong Z, Feng Y, Sun C, Jia Q, Zhu B. Immune Landscape of Colorectal Cancer Tumor Microenvironment from Different Primary Tumor Location. Front Immunol. 2018; 9:1578.
https://doi.org/10.3389/fimmu.2018.01578
PMID:30042763

59. Ma J, Zhang M, Yu J. Identification and Validation of Immune-Related Long Non-Coding RNA Signature for Predicting Immunotherapeutic Response and Prognosis in NSCLC Patients Treated With Immunotherapy. Front Oncol. 2022; 12:899925.
https://doi.org/10.3389/fonc.2022.899925
PMID:35860577

60. Wang L, Ge X, Zhang Z, Ye Y, Zhou Z, Li M, Yan H, Wu L, Bai Q, Li J, Zhu J, Wang Y. Identification of a Ferroptosis-Related Long Noncoding RNA Prognostic

Signature and Its Predictive Ability to Immunotherapy in Hepatocellular Carcinoma. Front Genet. 2021; 12:682082.
https://doi.org/10.3389/fgene.2021.682082
PMID:34745200

61. Song P, Li W, Guo L, Ying J, Gao S, He J. Identification and Validation of a Novel Signature Based on NK Cell Marker Genes to Predict Prognosis and Immunotherapy Response in Lung Adenocarcinoma by Integrated Analysis of Single-Cell and Bulk RNA-Sequencing. Front Immunol. 2022; 13:850745.
https://doi.org/10.3389/fimmu.2022.850745
PMID:35757748

62. Jiang H, Ning G, Wang Y, Lv W. Identification of an m6A-Related Signature as Biomarker for Hepatocellular Carcinoma Prognosis and Correlates with Sorafenib and Anti-PD-1 Immunotherapy Treatment Response. Dis Markers. 2021; 2021:5576683.
https://doi.org/10.1155/2021/5576683
PMID:34221187

63. Feng C, Li T, Xiao J, Wang J, Meng X, Niu H, Jiang B, Huang L, Deng X, Yan X, Wu D, Fang Y, Lin Y, et al. Tumor Microenvironment Profiling Identifies Prognostic Signatures and Suggests Immunotherapeutic Benefits in Neuroblastoma. Front Cell Dev Biol. 2022; 10:814836.
https://doi.org/10.3389/fcell.2022.814836
PMID:35493068

64. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, Bin J, Liao Y, Rao J, Liao W. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. Cancer Immunol Res. 2019; 7:737–50.
https://doi.org/10.1158/2326-6066.CIR-18-0436
PMID:30842092

Supplementary Figures
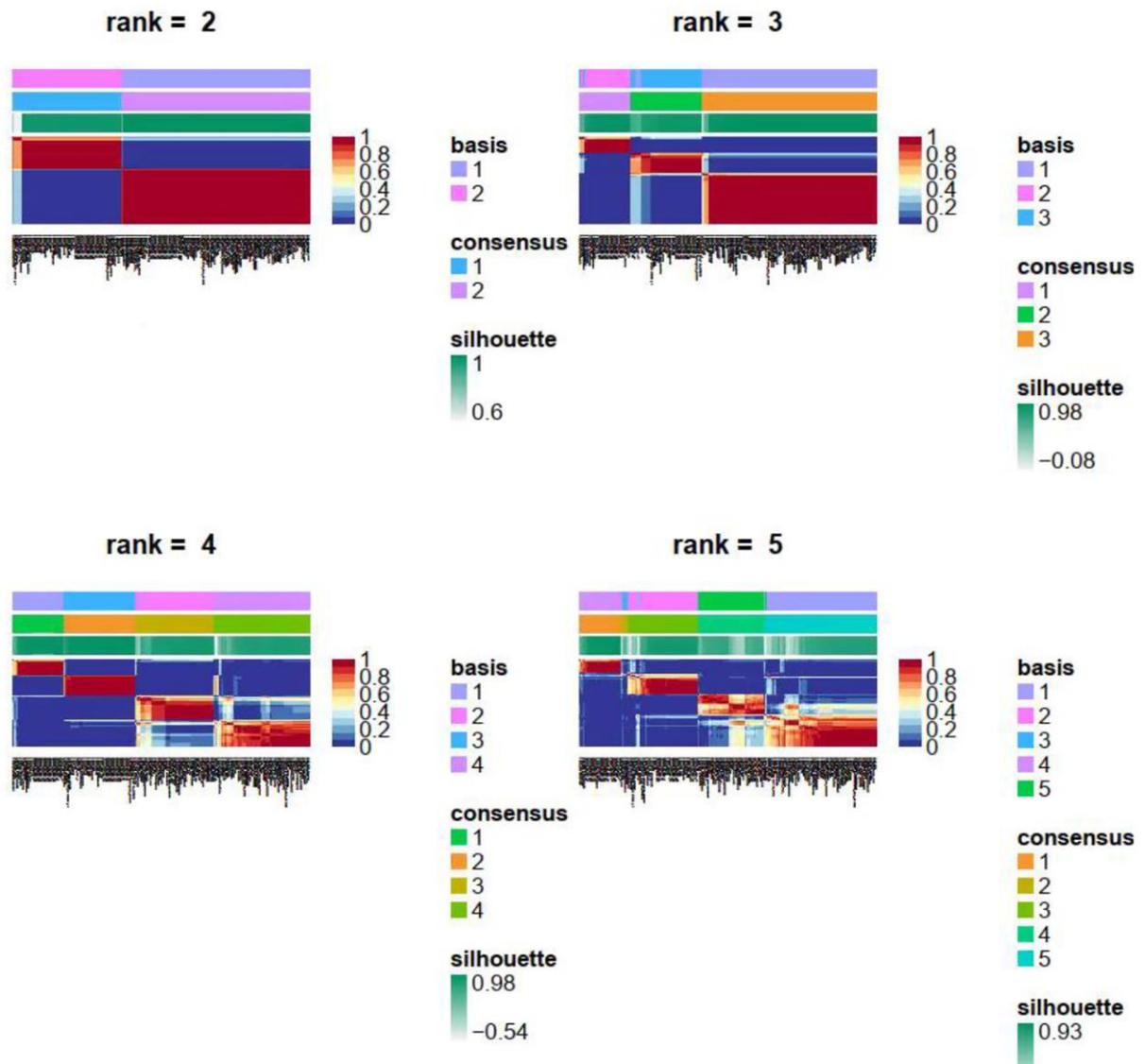
**Supplementary Figure 1. The consensus map of nsNMF clustering for cluster number = 2,3,4,5.**
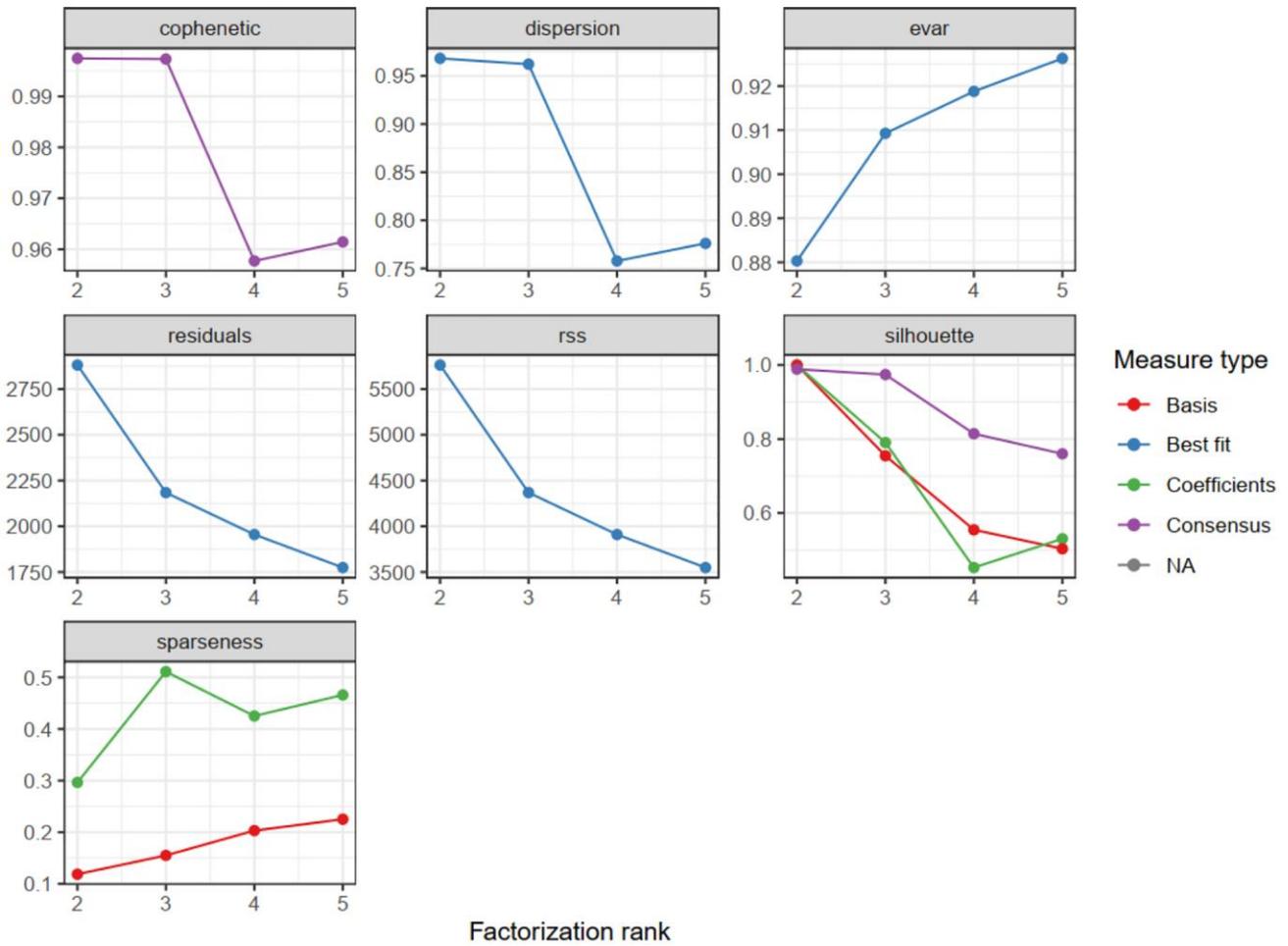
Supplementary Figure 2. The cluster indicators of nsNMF clustering for cluster number = 2,3,4,5.
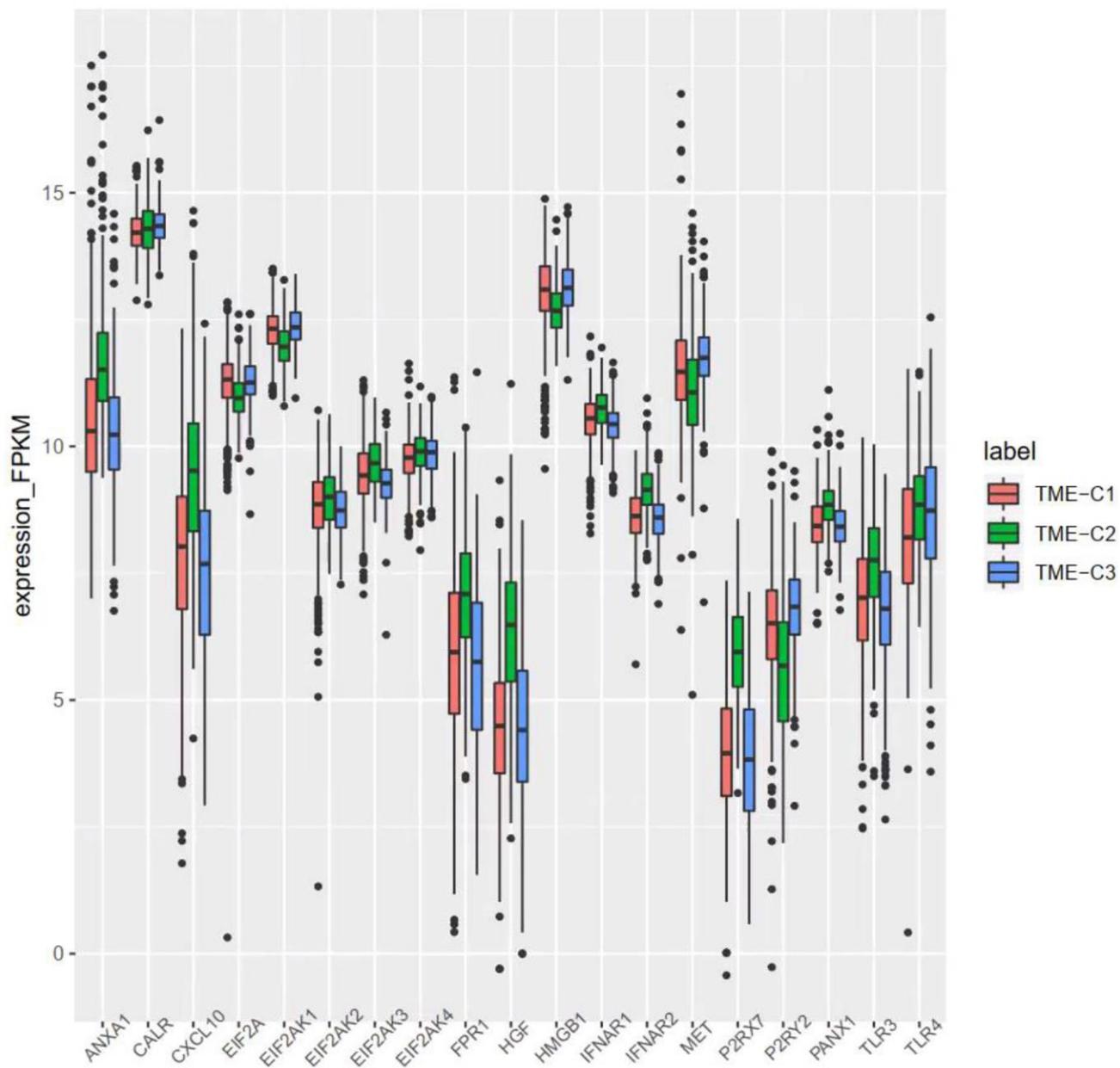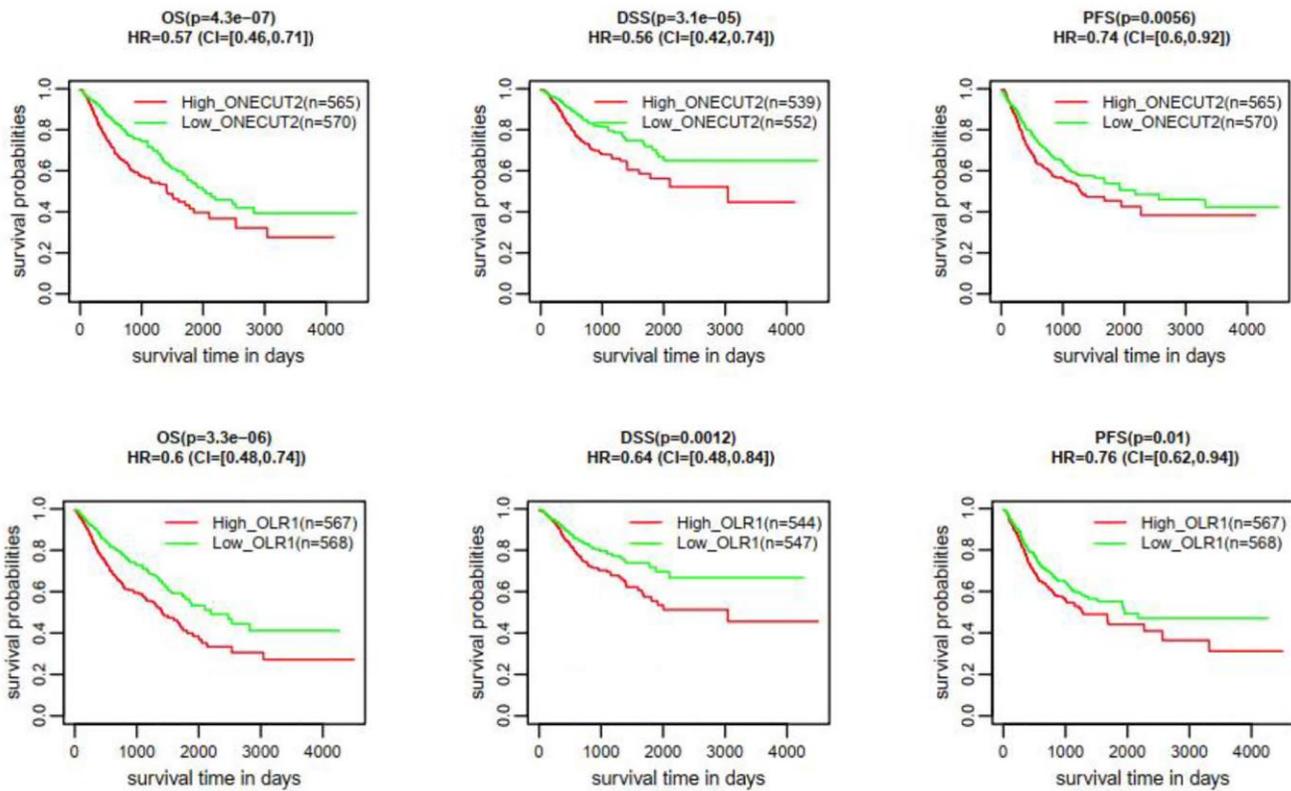
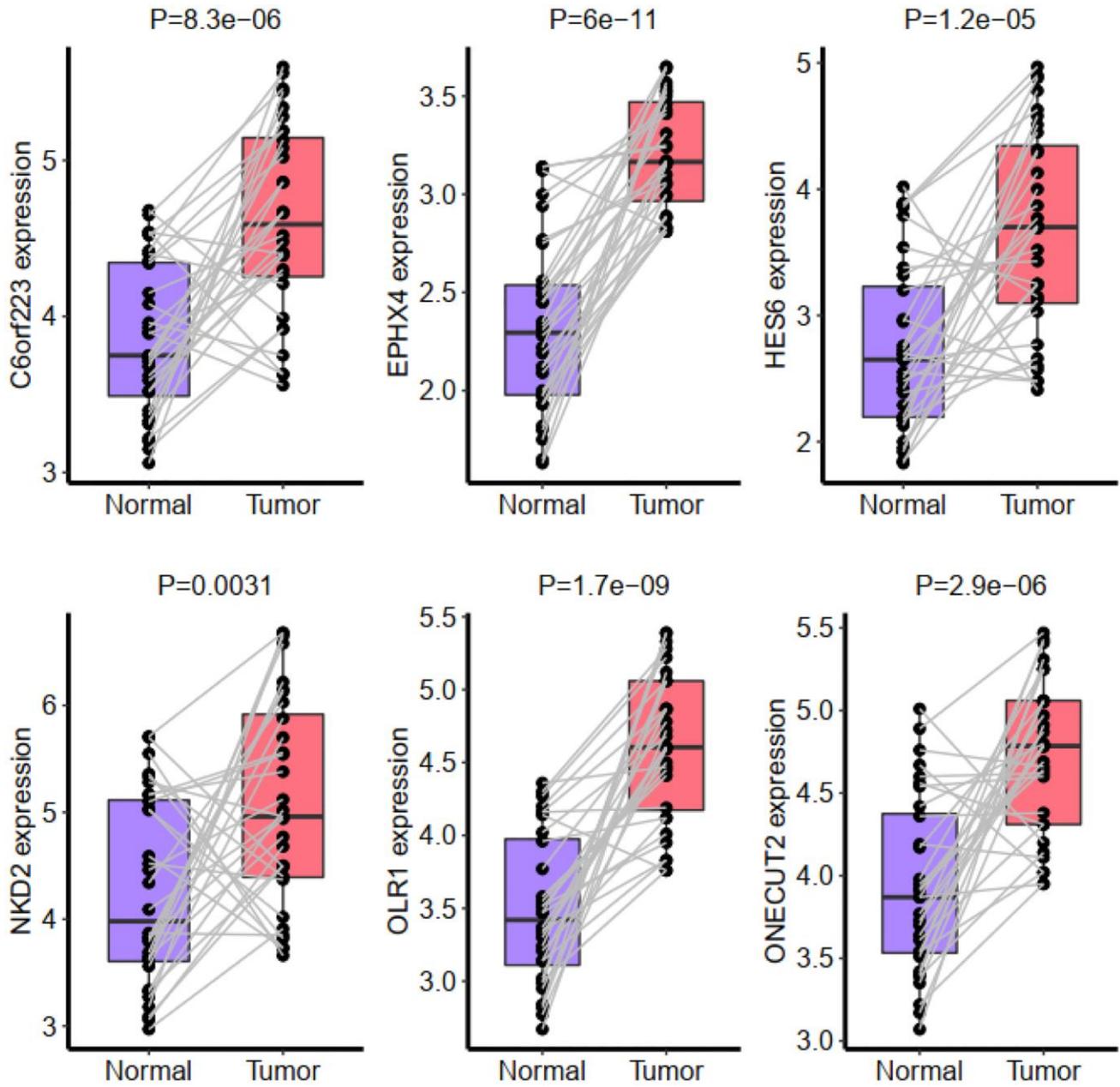**Supplementary Figure 3. The expression of 21 immunogenic cell death (ICD) modulators across three immune subtypes.**
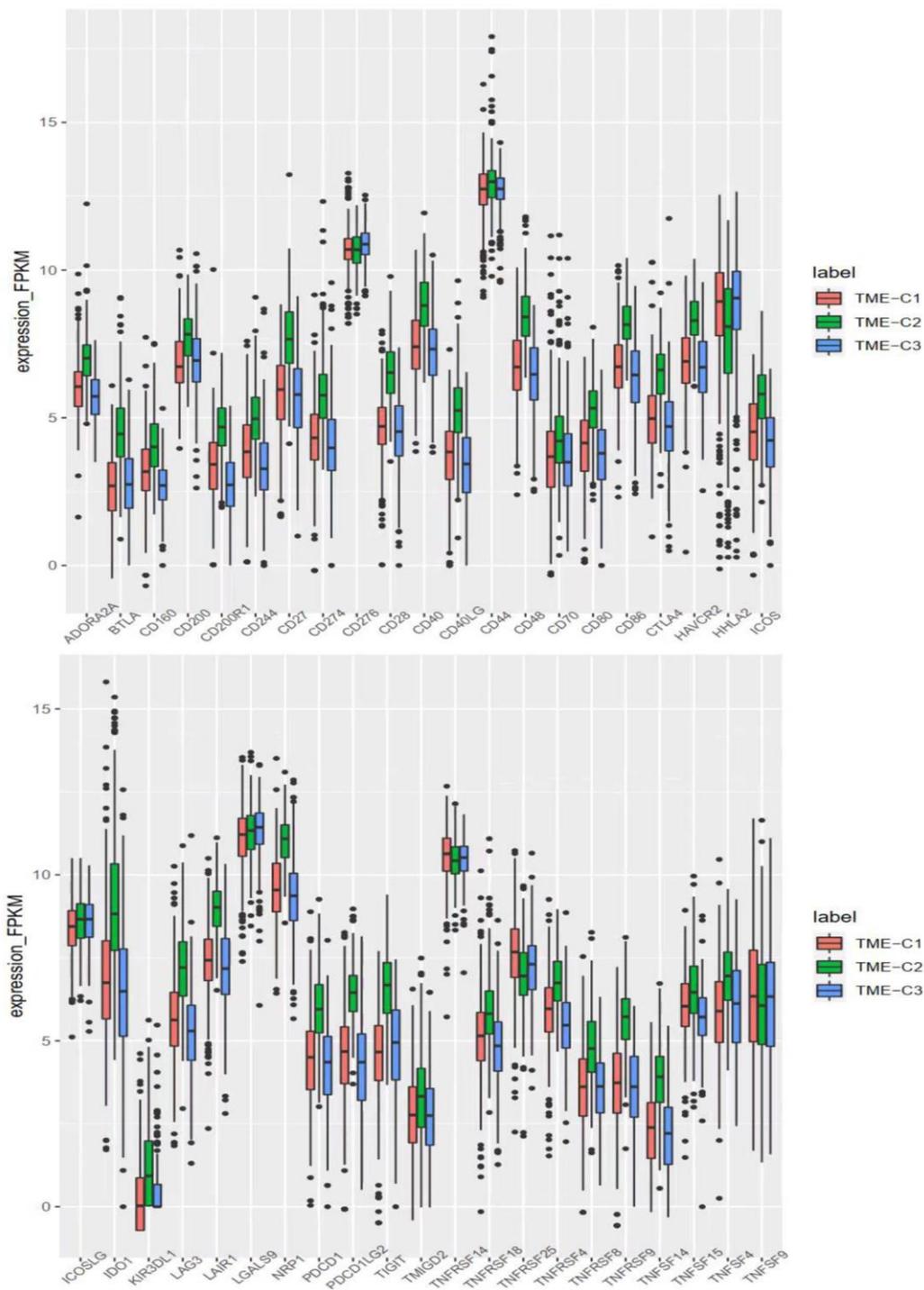
Supplementary Figure 4. The expression of 40 immune checkpoints (ICPs) across three immune subtypes.
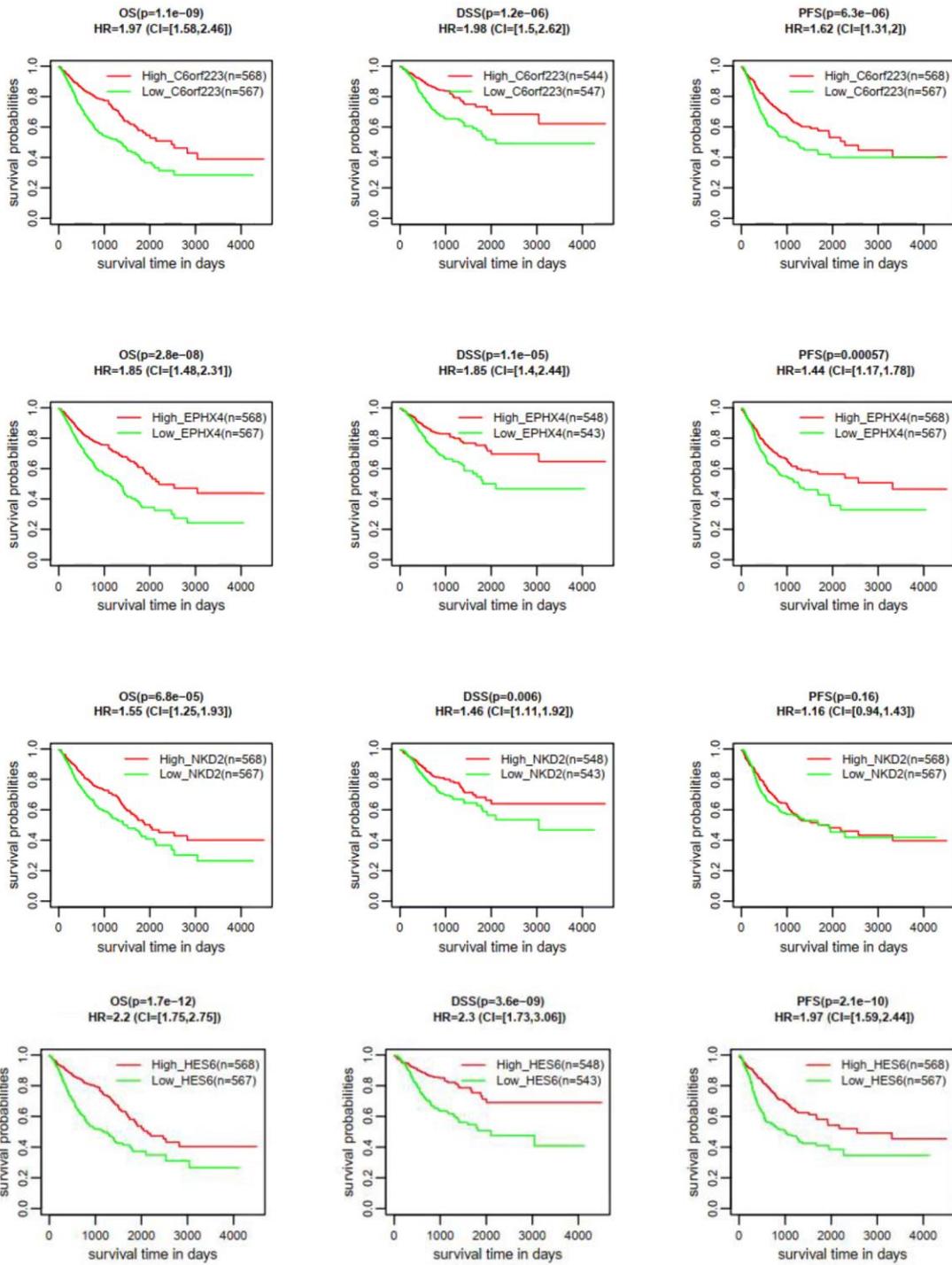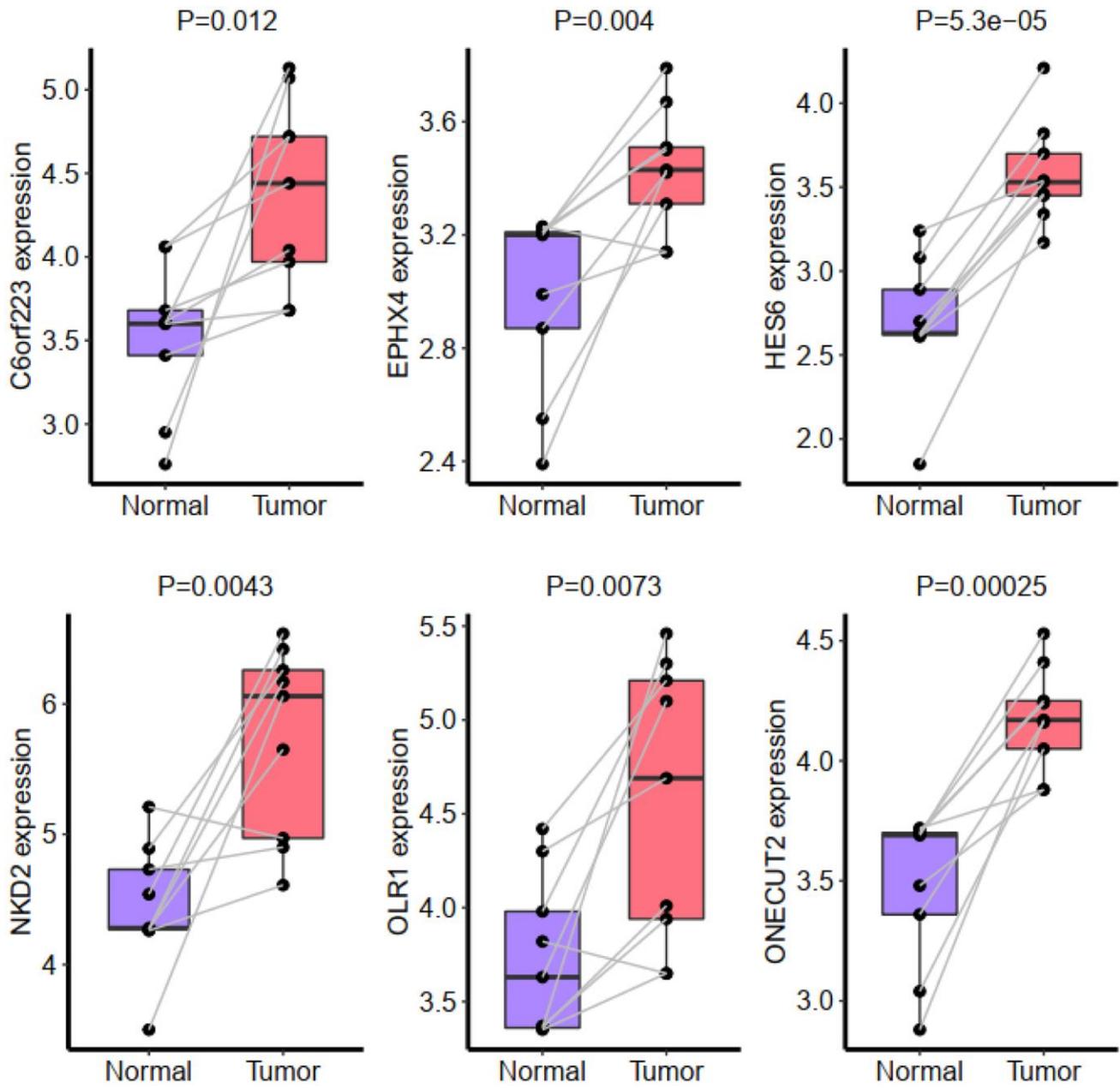
Supplementary Figure 5. The survival plot of patient groups with different C6orf223, EPHX4, HES6 and NKD2 expression.

Supplementary Figure 6. The survival plot of patient groups with different OLR1 and ONECUT2 expression in TCGA GIAC cohort.

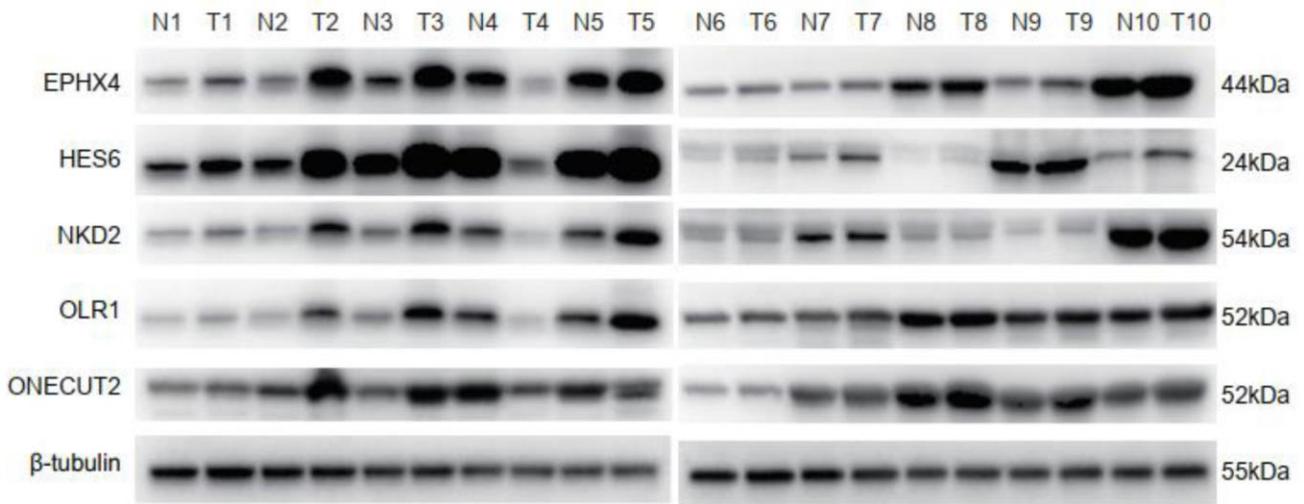**Supplementary Figure 7. The RNA expression of the six genes in 32 pairs of GAD tumor and normal samples from our center.**

Supplementary Figure 8. The RNA expression of the six genes in 9 pairs of ESAD tumor and normal samples from our center.
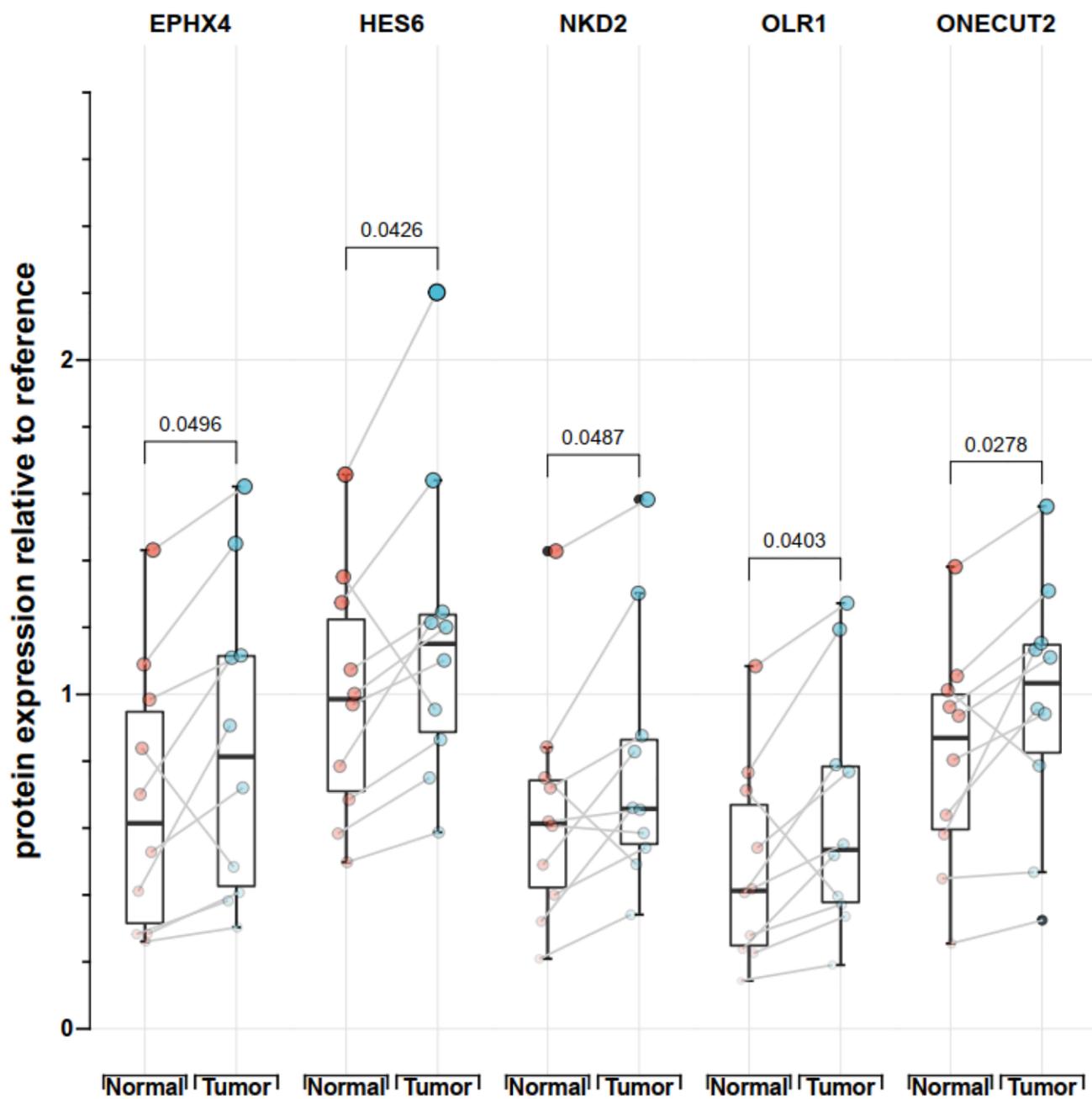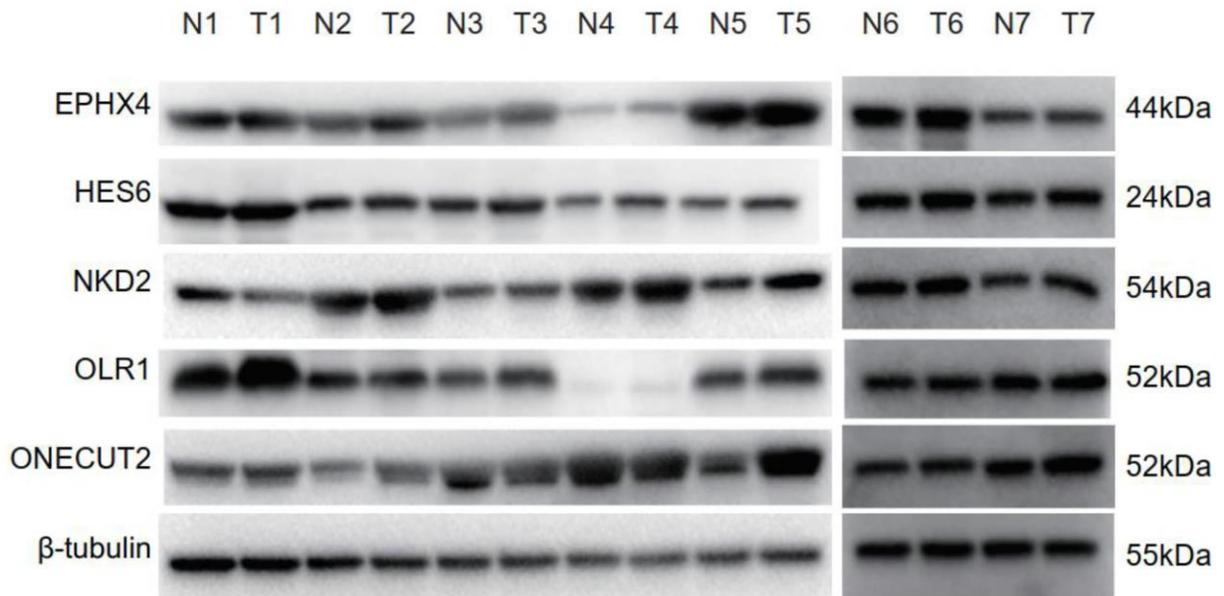
Supplementary Figure 9. The expression of the five proteins in 10 pairs of GAD tumor and normal samples from our center.
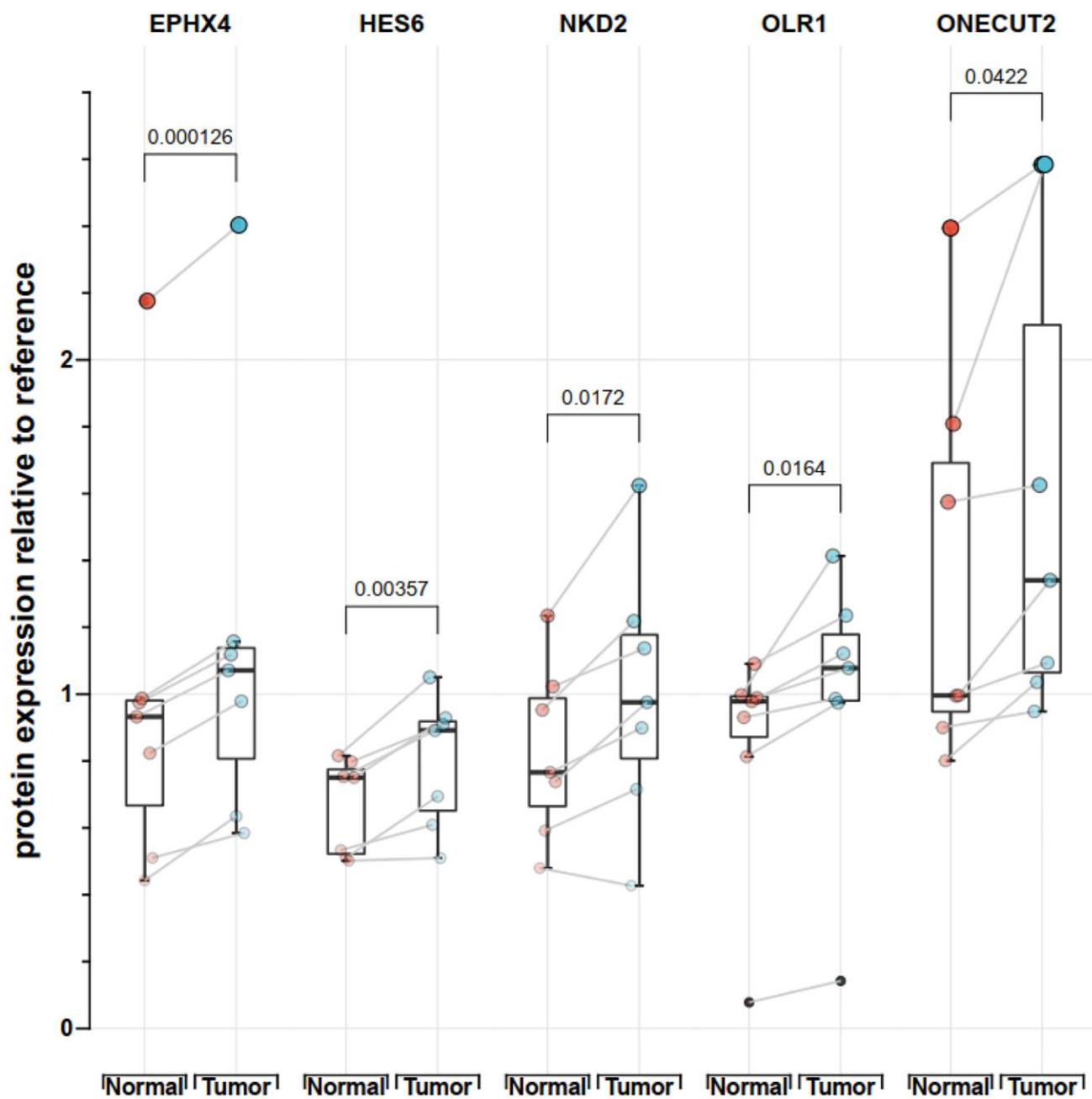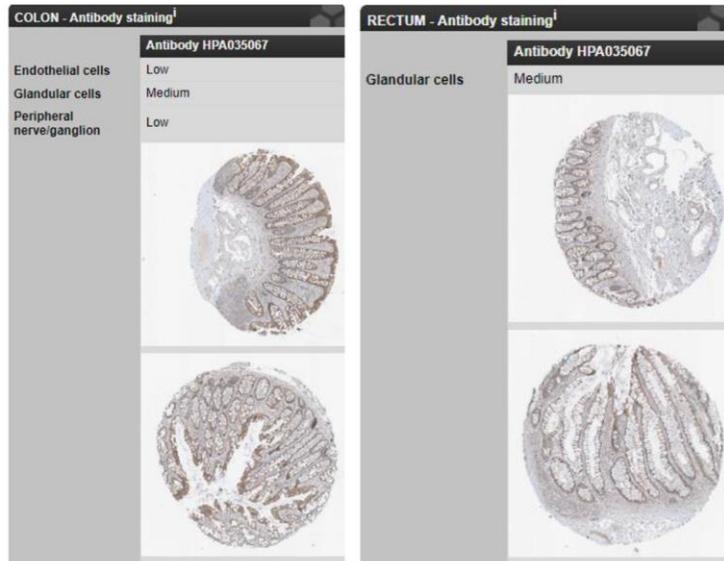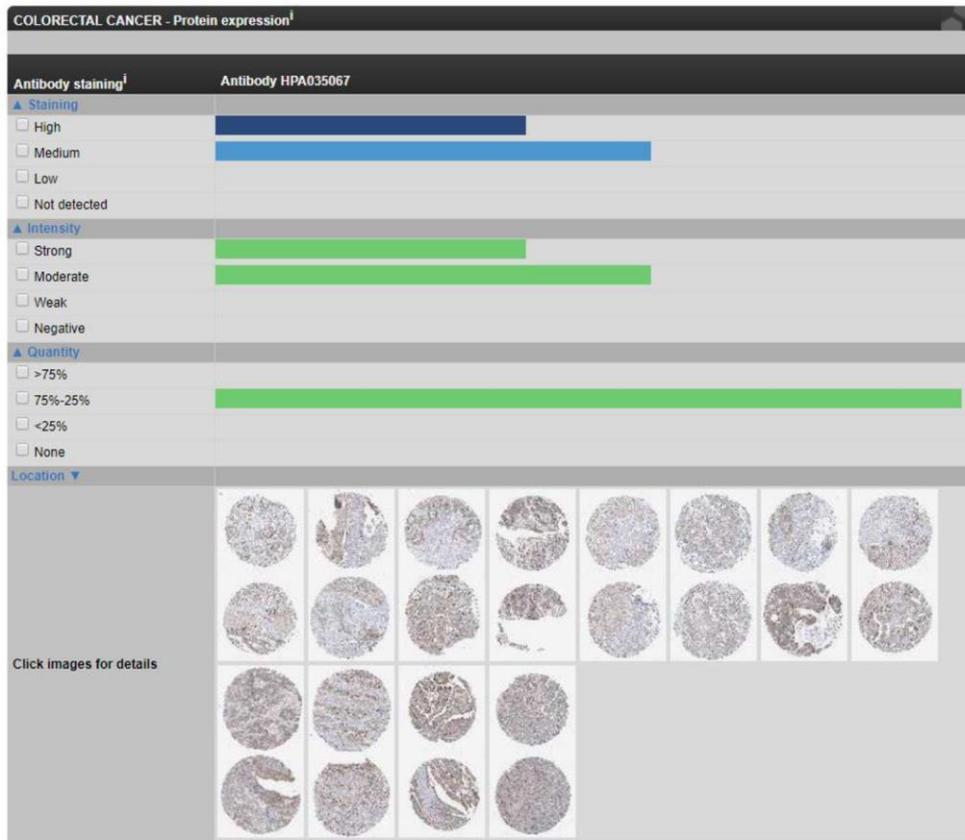
Supplementary Figure 10. The expression of the five proteins in 7 pairs of ESAD tumor and normal samples from our center.
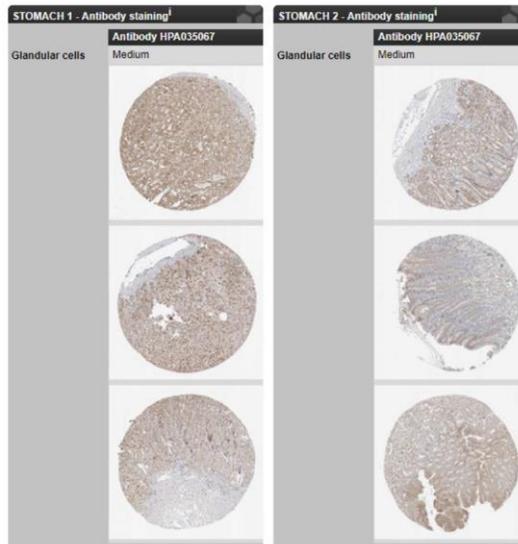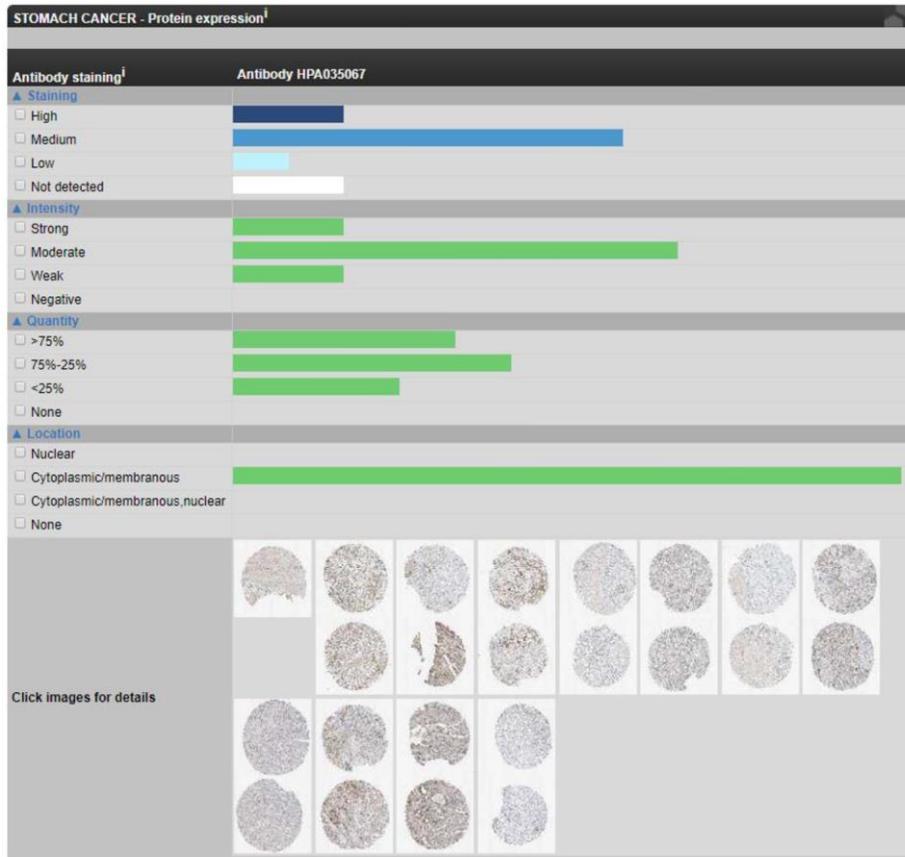
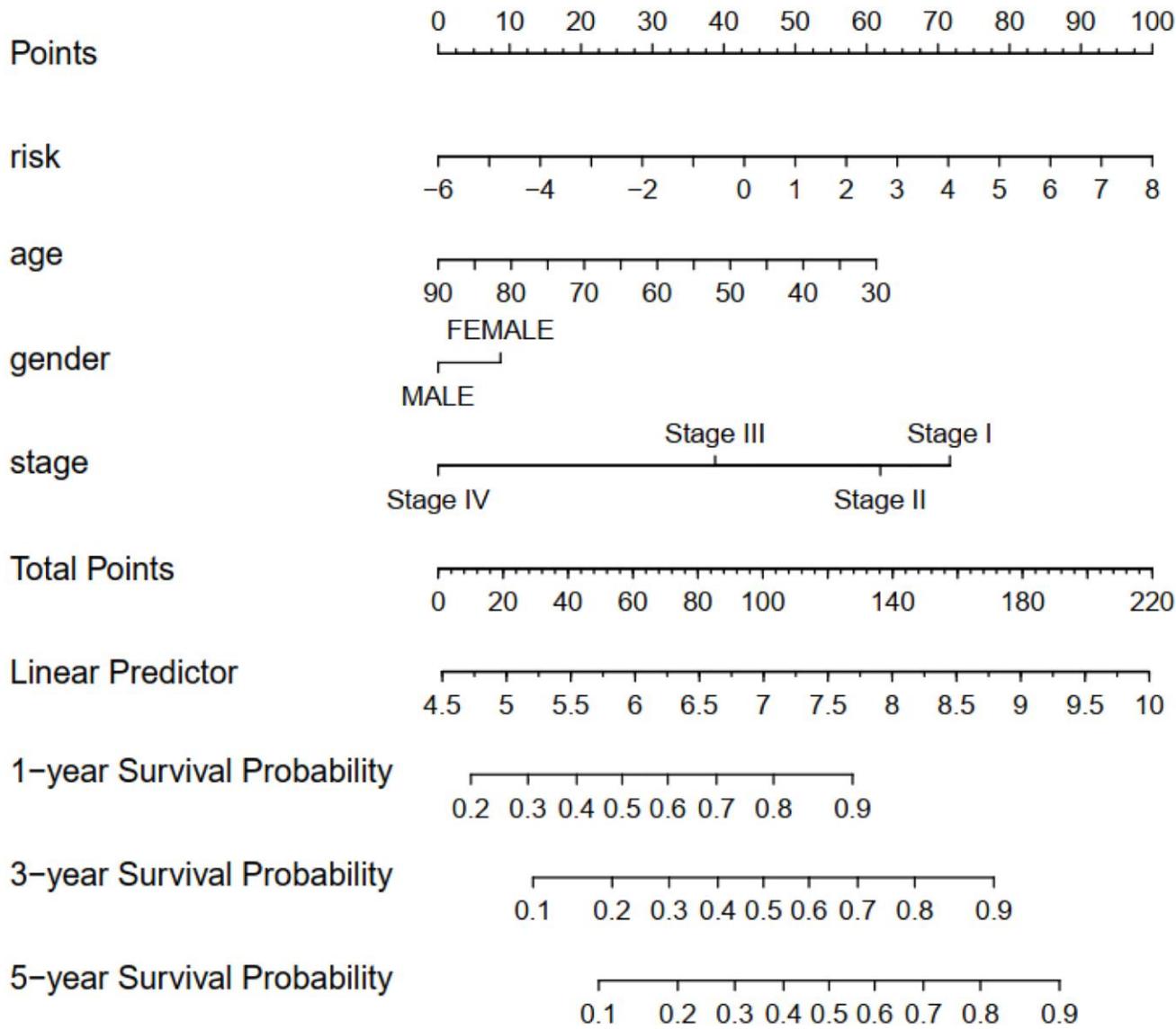**Supplementary Figure 11. The WB image of protein expression for GAD.**

Supplementary Figure 12. The WB image of protein expression for ESAD.

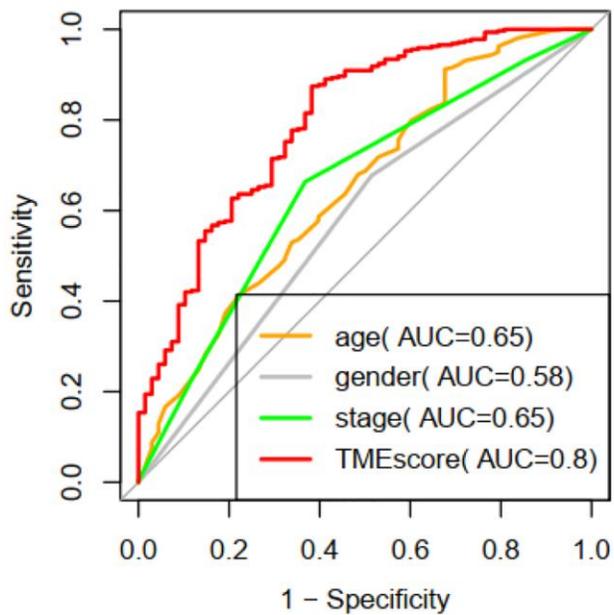**Supplementary Figure 13. EPHX4 staining of colorectal cancer and normal colon/rectum tissues in HPA database.**

Supplementary Figure 14. EPHX4 staining of stomach cancer and normal stomach tissues in HPA database.

Supplementary Figure 15. The nomogram plot of TMEscore for predicting OS.
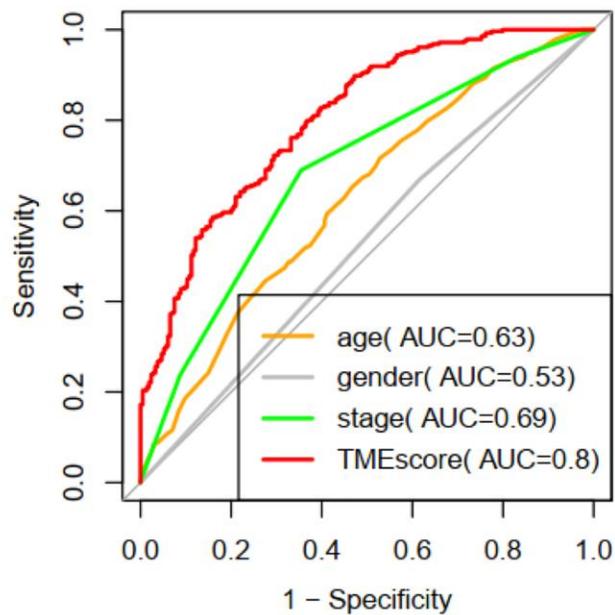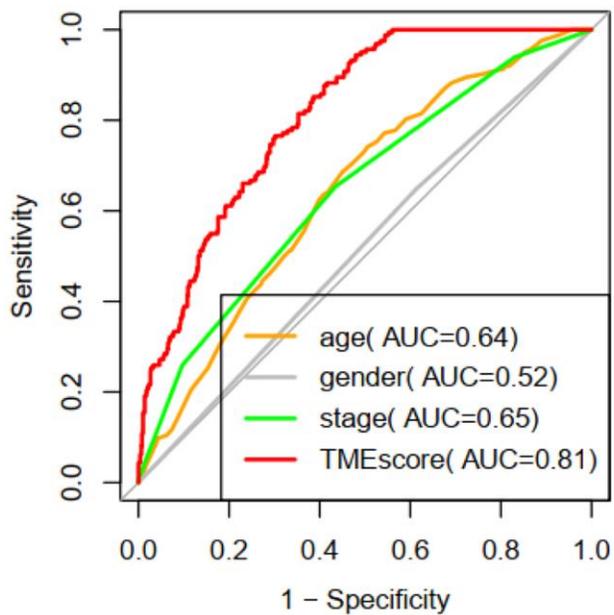
Supplementary Figure 16. The ROC curve plot of TMEscore for predicting OS compared to tumor stage.
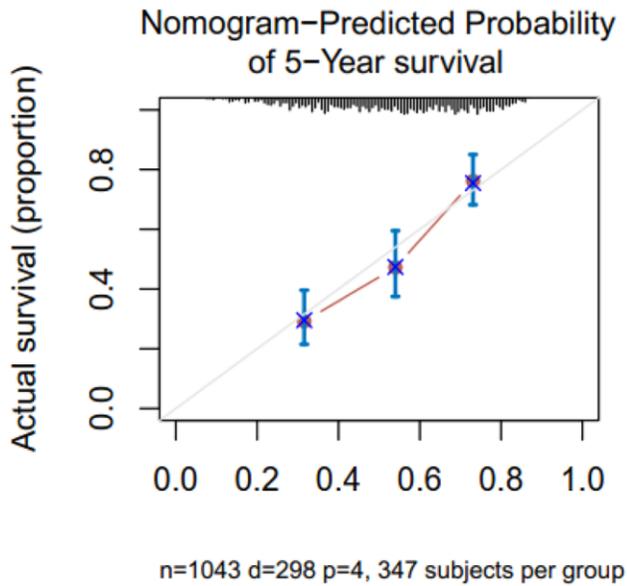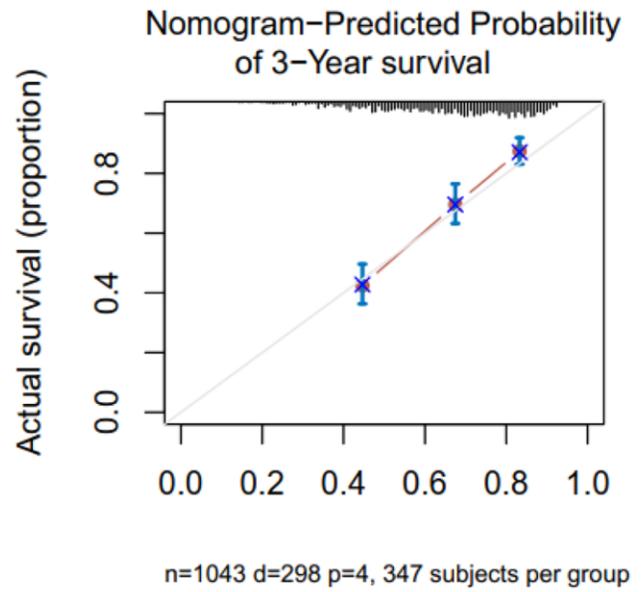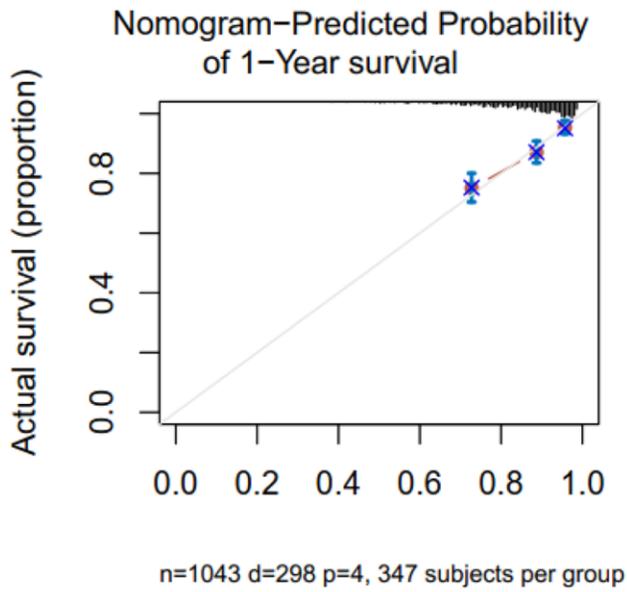
Supplementary Figure 17. The calibration plot of the model integrating TMEscore and tumor grade for predicting OS.
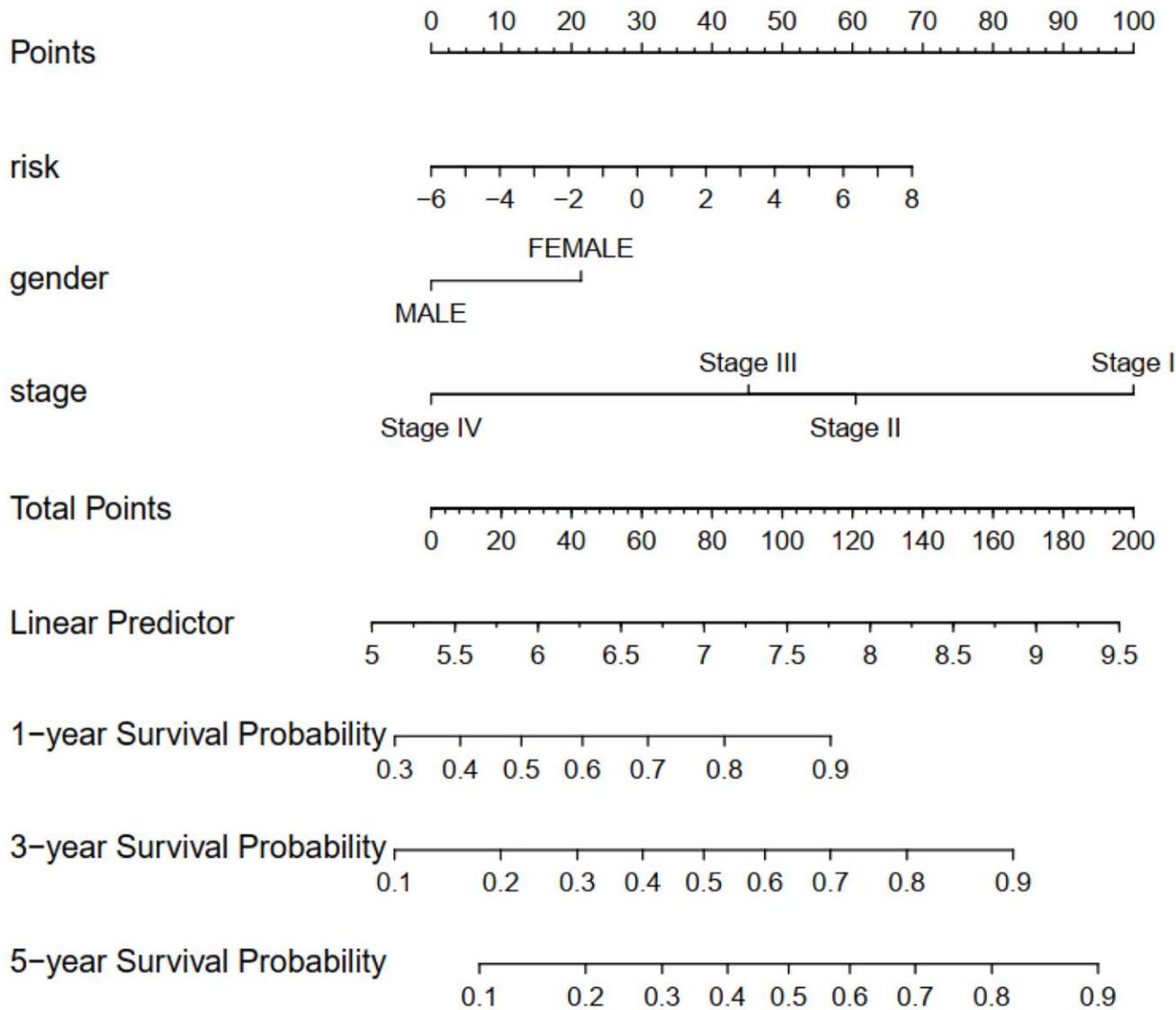
Supplementary Figure 18. The nomogram plot of TMEscore for predicting PFS.
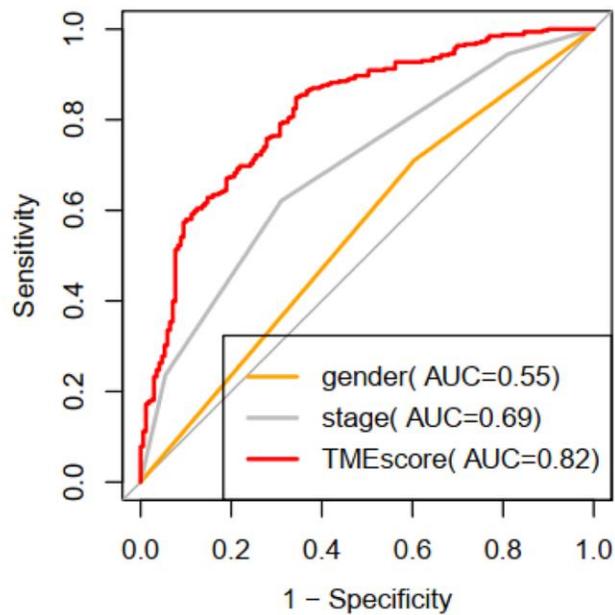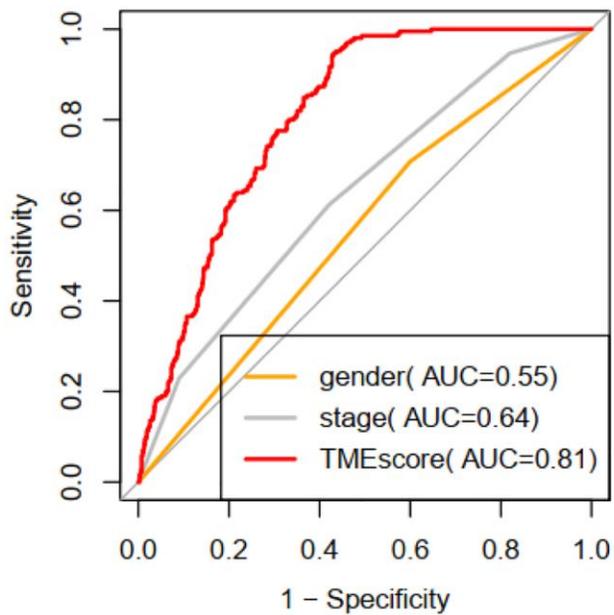
Supplementary Figure 19. The ROC curve plot of TMEscore for predicting PFS compared to tumor stage.
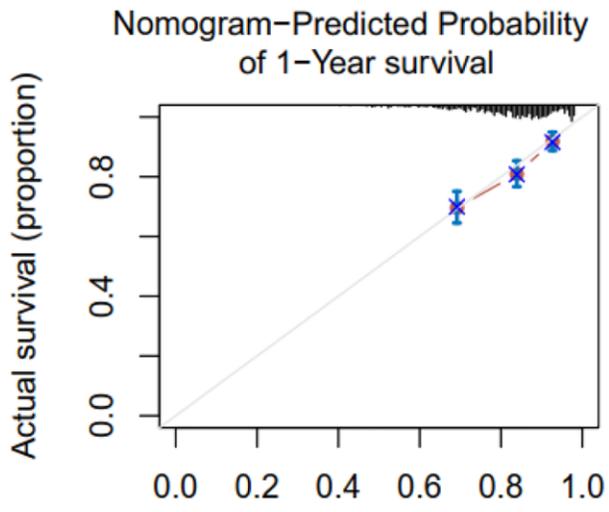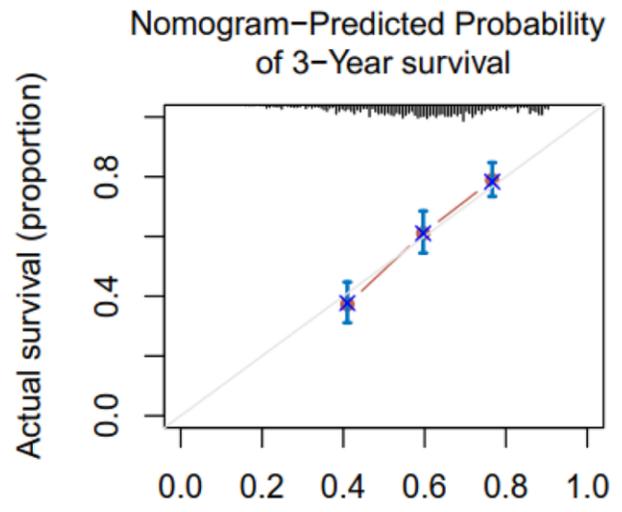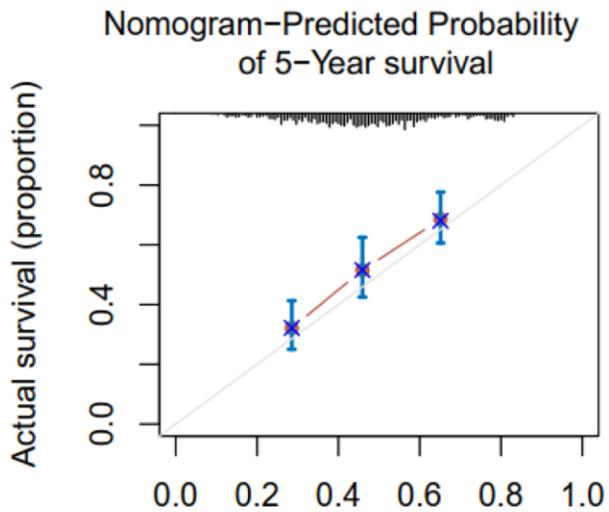
Supplementary Figure 20. The calibration plot of the model integrating TMEscore and tumor grade for predicting PFS.
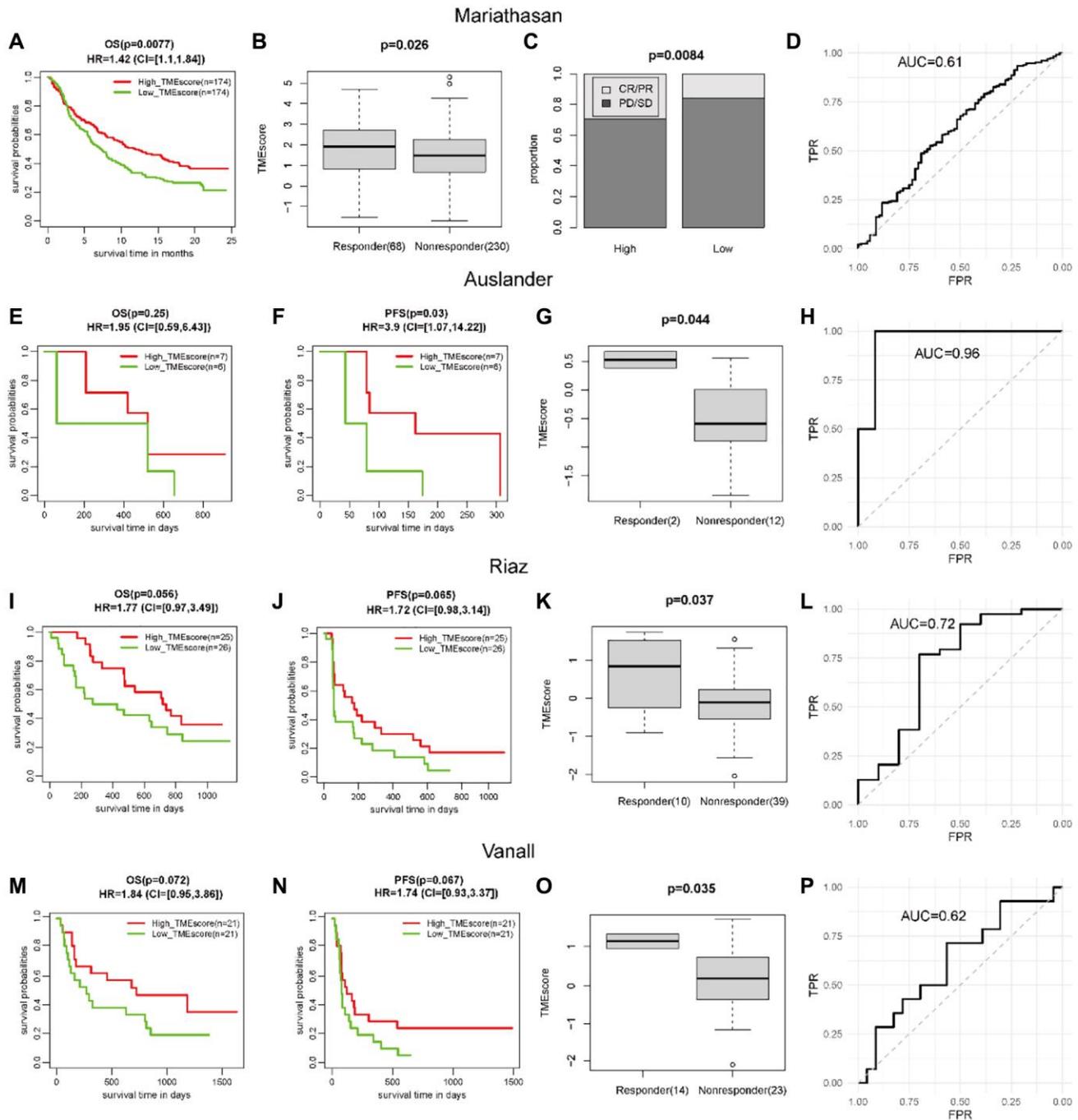
Supplementary Figure 21. TMEscore was a prognostic and immunotherapeutic biomarker in other four immunotherapeutic data sets. (**A**, **B**) The survival time and response rate of patients with high/low TMEscore in Mariathasan cohort. (**C**) The TMEscore of patients with different response status. (**D**) The ROC curve of TMEscore predicting immunotherapeutic response. (**E**, **F**) The survival time (OS and PFS) of patients with high/low TMEscore in Auslander cohort. (**G**) The TMEscore of patients with different response status (**H**) The ROC curve of TMEscore predictin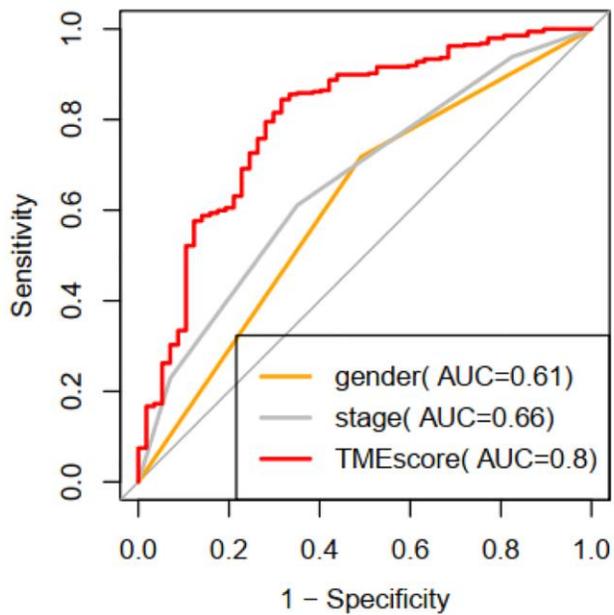g immunotherapeutic response. (**I**, **J**) The survival time (OS and PFS) of patients with high/low TMEscore in Riaz cohort. (**K**) The TMEscore of patients with different response status (**L**) The ROC curve of TMEscore predicting immunotherapeutic response. (**M**, **N**) The survival time (OS and PFS) of patients with high/low TMEscore in Vanall cohort. (**O**) The TMEscore of patients with different response status. (**P**) The ROC curve of TMEscore predicting immunotherapeutic response.
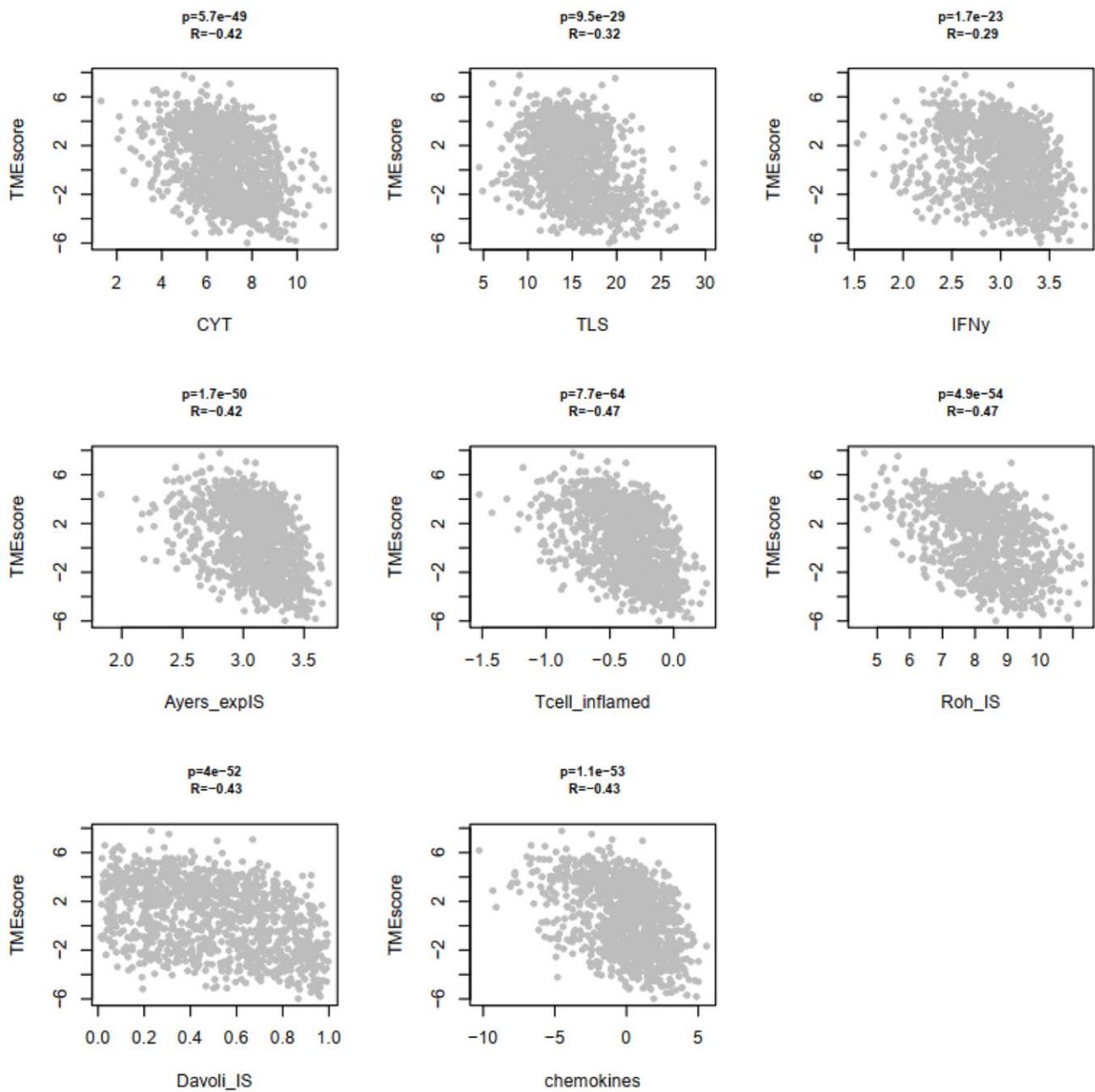
Supplementary Figure 22. The correlation plot of TMEscore with eight immune response scores estimated by "EaSIeR" in TCGA GIAC cohort.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–3, 5–8, 12–15, 17, 18, 20 and 21.

Supplementary Table 1. The 160 signatures from 7 sources download from PanImmune_GeneSet.

Supplementary Table 2. The genes from 160 signatures download from PanImmune_GeneSet.

Supplementary Table 3. The basic clinical information of 96 GIAC patients from our center.

Supplementary Table 4. The primers used in this study.

Name	Direction	Sequence(5′–3′)
C6orf223	Forward	GCTGTGAGAGTTTTCCCCGA
	Reverse	CCTTTTGACCCCAAAGCACG
EPHX4	Forward	CCTGGTCTACTGCTACTGCG
	Reverse	TTTCAACAGCTGAGCAGGGT
HES6	Forward	CCGAGCTCCTGAACCATCTG
	Reverse	GCTGCAGGGTCCCTAAAAGT
NKD2	Forward	AAGTCCGGGAAAGCCTTCAG
	Reverse	AGATGTGTTCACCGCCATGT
OLR1	Forward	CATTATGGTGCTGGGCATGC
	Reverse	TGGGGCATCAAAGGAGAACC
ONECUT2	Forward	GGAATCCAAAACCGTGGAGTAA
	Reverse	CTCTTTGCGTTTGCACGCTG
GAPDH	Forward	CCGGGAAACTGTGGCGTGATGG
	Reverse	AGGTGGAGGAGTGGGTGTCGCTGTT

Supplementary Table 5. The RNA expression of 6 genes in 96 pairs of tumor and matched adjacent normal samples in GIAC patients detected by qRT-PCR.

Supplementary Table 6. The protein expression of 5 genes in 27 pairs of tumor and matched adjacent normal samples in GIAC patients detected by Western blot.

Supplementary Table 7. The infiltration fraction of six immune cell types in GIAC tumor samples.

Supplementary Table 8. The infiltration fraction of six immune cell types in GIAC adjacent normal samples.

Supplementary Table 9. The statistic information about the infiltration level of 6 immune cell types in GIAC tumor and normal samples.

	B_cell	CD4_Tcell	CD8_Tcell	Neutrophil	Macrophage	Dendritic
Tumor (average)	0.07	0.14	0.15	0.11	0.04	0.48
Normal (average)	0.1	0.14	0.17	0.09	0.07	0.44
log2 (Fold change)	−0.46	\	−0.11	0.31	−0.73	0.15
p value	1.60E-06	1	0.033	4.70E-23	4.20E-06	4.80E-11

Supplementary Table 10. The infiltration correlation of 6 immune cell types in in GIAC tumor samples.

	B_cell	CD4_Tcell	CD8_Tcell	Neutrophil	Macrophage	Dendritic
B_cell	1	0.16	−0.18	−0.45	−0.08	−0.47
CD4_Tcell	0.16	1	−0.69	−0.3	−0.07	−0.07
CD8_Tcell	−0.18	−0.69	1	0	−0.08	−0.21
Neutrophil	−0.45	−0.3	0	1	−0.06	0.12
Macrophage	−0.08	−0.07	−0.08	−0.06	1	−0.55
Dendritic	−0.47	−0.07	−0.21	0.12	−0.55	1

Supplementary Table 11. The univariate Cox regression result of 6 immune cell types for OS, DSS and PFS in TCGA.

	Survival_index	HR	log10 (*p*-value)	Threshold	Percentage
B_cell	OS	1.2	2.7	non	3.64
CD4_Tcell	OS	0.81	3.8	Favorable	6.85
CD8_Tcell	OS	1.4	9	Risk	7.72
Neutrophil	OS	1.2	2.3	non	5.53
Macrophage	OS	0.91	1.3	non	2.18
Dendritic	OS	0.83	3.4	Favorable	24.09
B_cell	DSS	1	0.18	non	3.64
CD4_Tcell	DSS	0.66	7.8	Favorable	6.85
CD8_Tcell	DSS	1.6	9.1	Risk	7.72
Neutrophil	DSS	1.2	1.9	non	5.53
Macrophage	DSS	1	0.029	non	2.18
Dendritic	DSS	0.84	2.1	non	24.09
B_cell	PFS	1.1	1	non	3.64
CD4_Tcell	PFS	0.74	7.2	Favorable	6.85
CD8_Tcell	PFS	1.3	4.7	Risk	7.72
Neutrophil	PFS	1.3	4.9	Risk	5.53
Macrophage	PFS	0.86	2.6	non	2.18
Dendritic	PFS	1	0.029	non	24.09

Supplementary Table 12. The patient classification consistency of the three NMF clustering methods.

Supplementary Table 13. The signatures that differentiate three TME clusters.

Supplementary Table 14. The clinical and genomic characteristics of patients from three subgroups.

Supplementary Table 15. The expression of 6 immunotherapy targets of patients from three subgroups.

Supplementary Table 16. The statistics of 6 immunotherapy targets' expression from three subgroups.

	TMEC1	**TMEC2**	**TMEC3**	*p*_**value**
BTLA	2.61	4.56	2.8	5.00E-55
CD274	4.31	5.83	4.1	1.00E-55
CTLA4	5.01	6.5	4.68	7.40E-67
HAVCR2	6.92	8.34	6.64	1.80E-66
LAG3	5.66	7.19	5.23	6.80E-63
PDCD1	4.46	5.92	4.24	6.40E-66

Supplementary Table 17. The immune cell infiltration of patients from three subgroups.

Supplementary Table 18. The genes in signature A, B and C that differentiate three TME clusters.

Supplementary Table 19. The *p*-value in functional enrichment of signature A and B on pan-immune genesets.

	gene_sigA	gene_sigB
PanImmune_GeneSet	1	1.30E-07
HER2_Immune_PCA_18006808	0.85	0.0061
CSR_Activated_15701700	1	1
GRANS_PCA_16704732	1	0.15
B_cell_PCA_16704732	1	0.51
CHANG_CORE_SERUM_RESPONSE_UP	1	1
T_cell_PCA_16704732	1	1
Troester_WoundSig_19887484	0.38	0.037
Module11_Prolif_score	1	1
CSF1_response	1	0.028
STAT1_score	1	1.60E-08
KEGG_HEMATOPOIETIC_CELL_LINEAGE	1	0.41
Module4_TcellBcell_score	1	0.39
TGFB_score_21050467	0.27	0.0011
Module5_TcellBcell_score	1	1.80E-07
Immune_cell_Cluster_21214954	1	7.20E-05
Immune_NSCLC_score	1	0.068
IFNG_score_21050467	1	1
LCK_19272155	1	1
Minterferon_Cluster_21214954	1	0.032
MCD3_CD8_21214954	1	1
LYMPHS_PCA_16704732	1	1
Angiogenesis	1	1
Tcm cells	1	1
CD8 T cells	1	1
Interferon_Cluster_21214954	1	0.19
NK cells	1	1
Macrophages	1	0.015
B cells	1	1
Tfh cells	1	1
Neutrophils	1	0.17
MDACC.FNA.1_20805453	1	1
iDC	1	0.17
Eosinophils	1	1
TAMsurr_TcClassII_ratio	1	0.00055
Mast cells	1	1
Th1 cells	1	0.15
Th2 cells	1	1
TcClassII_score	1	1
Bcell_21978456	1	1
Tcell_21978456	1	1

	1	1
T helper cells	1	1
Module3_IFN_score	1	0.0078
ZHANG_INTERFERON_RESPONSE	1	1
IL12_score_21050467	1	1
IL2_score_21050467	1	1
IgG_19272155	1	1
ICR_SCORE	1	1.80E-08
T cells	1	1
CD8_PCA_16704732	1	1
LIexpression_score	1	1

Supplementary Table 20. The functional enrichment analysis of signature A on GO, KEGG and cancer hallmark.

Supplementary Table 21. The functional enrichment analysis of signature B on GO, KEGG and cancer hallmark.

Supplementary Table 22. The univariate Cox regression result of 6 genes for OS and PFS in TCGA.

Gene	Survival Index	HR	coef	*P* value
C6orf223	OS	2	0.68	2.30E-09
EPHX4	OS	1.9	0.62	4.50E-08
HES6	OS	2.2	0.79	6.00E-12
NKD2	OS	1.6	0.44	7.80E-05
ONECUT2	OS	0.57	-0.56	5.90E-07
OLR1	OS	0.6	-0.51	4.20E-06
C6orf223	PFS	1.6	0.48	7.80E-06
EPHX4	PFS	1.4	0.37	0.00062
HES6	PFS	2	0.68	4.50E-10
NKD2	PFS	1.2	0.15	0.016
ONECUT2	PFS	0.74	-0.29	0.0058
OLR1	PFS	0.76	-0.27	0.01

Supplementary Table 23. The high expression of the five genes in literatures.

Gene	PMID	Cancer Type
OLR1(LOX-1)	34921134	Colon Cancer
OLR1(LOX-1)	30483757	Gastric Cancer
OLR1(LOX-1)	28345638	Gastric Cancer
OLR1(LOX-1)	22641216	Colorectal Cancer
HES6	30015909	Colorectal Cancer
HES6	33681178	Colorectal Cancer
ONECUT2(OC2)	32801861	Gastric Cancer
ONECUT2(OC2)	34253241	Colorectal Cancer
ONECUT2(OC2)	35874630	Colorectal Cancer
ONECUT2(OC2)	32129880	Gastric Cancer
ONECUT2(OC2)	33015779	Gastric Cancer
ONECUT2(OC2)	34365839	Gastric Cancer
NKD2	33455110	Colon Cancer
NKD2	27246976	Gastric Cancer
NKD2	26396173	Gastric Cancer
NKD2	26985708	Colon Cancer
EPHX4	31404997	Rectal Cancer

Supplementary Table 24. The Cox regression result of TMEscore for OS in TCGA.

Univariate Cox regression for OS				
	HR	lower .95	upper .95	*P* value
TMEscore	0.85	0.82	0.88	1.50E-16
age	1	1	1	0.0024
Gender = MALE (ref = FEMALE)	1.4	1.1	1.8	0.0024
Stage = II (ref=I)	1.6	1	2.7	0.04
Stage = III (ref=I)	3	1.9	4.8	2.80E-06
Stage = IV (ref=I)	5.5	3.4	9	9.20E-12
Multi-variate Cox regression for OS				
	HR	lower .95	upper .95	*P* value
TMEscore	0.84	0.8	0.88	4.60E-14
age	1	1	1	6.90E-09
Gender = MALE (ref = FEMALE)	1.1	0.84	1.4	0.51
Stage = II (ref = I)	2	1.1	3.8	0.023
Stage = III (ref = I)	3.7	2	6.8	1.70E-05
Stage = IV (ref = I)	10	5.5	19	1.80E-13

Supplementary Table 25. The Cox regression result of TMEscore for PFS in TCGA.

Univariate Cox regression for PFS

	HR	lower .95	upper .95	*P* value
TMEscore	0.9	0.87	0.93	2.80E-08
age	0.99	0.98	1	0.14
Gender = MALE (ref = FEMALE)	1.7	1.4	2.2	5.50E-06
Stage = II (ref = I)	2.4	1.5	3.8	0.00039
Stage = III (ref = I)	3.3	2.1	5.3	5.70E-07
Stage = IV (ref = I)	7.6	4.6	12	9.50E-16

Multi-variate Cox regression for PFS

	HR	lower .95	upper .95	*P* value
TMEscore	0.91	0.87	0.95	3.70E-05
Gender = MALE (ref = FEMALE)	1.4	1.1	1.8	0.0086
Stage = II (ref = I)	2.3	1.3	4	0.0026
Stage = III (ref = I)	3.2	1.8	5.5	3.70E-05
Stage = IV (ref = I)	8.3	4.8	15	1.20E-13