Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures

Dongqiang Zeng M.D.1*, Meiyi Li M.D.1*, Rui Zhou, M.D.1, Jingwen Zhang, M.D.1, Huiying Sun, M.D.1, Min Shi, M.D., Ph.D.1, Jianping Bin, M.D., Ph.D.2, Yulin Liao, M.D., Ph.D.2, Jinjun Rao M.D., Ph.D.3, Wangjun Liao, M.D., Ph.D.1#

1Department of Oncology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, P. R. China
2Department of Cardiology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, P. R. China
3Key laboratory of new drug screening of Guangdong province.
School of Pharmaceutical Sciences, Southern Medical University, Guangzhou, Guangdong, P. R. China

* These authors contributed equally to this work

Running title
Cellular landscape of GC TME and relevant signatures

Abbreviations


Manuscript category: original article
#Corresponding author

Wangjun Liao, Department of Oncology, Nanfang Hospital, Southern Medical University, 1838 North Guangzhou Avenue, Guangzhou 510515, China. Tel: +86-20-62787731; Fax: +86-20-61360657; E-mail: nfyyliaowj@163.com.

Conflicts of Interest

No conflicts of interest relevant to this article were disclosed.
ABSTRACT

Tumor microenvironment (TME) cells constitute a vital element of tumor tissue. Increasing evidence has elucidated their clinicopathological significance in predicting outcomes and therapeutic efficacy. Nonetheless, no studies have reported a systematic analysis of cellular interactions in the tumor microenvironment. In this study, we comprehensively estimated the TME infiltration patterns of 1,524 gastric cancer patients and systematically correlated the TME phenotypes with genomic characteristics and clinicopathological features of gastric cancer using two proposed computational algorithms. Three TME phenotypes were defined, and the TMEscore was constructed using principal component analysis algorithms. The high TMEscore subtype was characterized by immune activation and response to virus and interferon-gamma. Additionally, activation of transforming growth factor β, epithelial mesenchymal transition, and angiogenesis pathways were observed in the low TMEscore subtype, which are considered T-cell suppressive and may be responsible for significantly worse prognosis in gastric cancer (hazard ratio [HR], 0.42; 95% confidence interval [CI], 0.33–0.54; P < 0.001). Furthermore, multivariate analysis revealed that TMEscore was an independent prognostic biomarker, and its value in predicting immunotherapeutic outcomes was also confirmed (IMvigor210 cohort: HR, 0.63; 95% CI, 0.46–0.89; P = 0.008; GSE78220 cohort: HR, 0.25; 95% CI, 0.07–0.89; P = 0.021). Depicting a comprehensive landscape of the TME characteristics of gastric cancer may therefore help to interpret the responses of gastric tumors to immunotherapies and provide new strategies for the treatment of cancers.
**Introduction**

Genomic analysis has been the primary methodology used in international efforts to discover novel biological targets in gastric cancer (1,2), although this method has not led to the successful discovery of distinct mechanisms. However, some studies have revealed the significance of tumor-related structures as well as upregulated signaling pathways in both cancer cells and the tumor microenvironment (TME) (3,4), suggesting that intercellular relationships are more important than genomic factors at the single-cell level (5,6). In addition, an increasing body of literature suggests a crucial role for the TME in cancer progression and therapeutic responses (7,8). For example, differences in the compositions of resident cell types within the TME, including cytotoxic T cells, helper T cells, dendritic cells (DCs), tumor-associated macrophages (TAMs), mesenchymal stem cells (MSCs), and associated inflammatory pathways have been reported in patients with cancer (5,6,9,10). The TME context determined at diagnosis reflects the immune response (11) and chemotherapy benefit (8), and changes in the numbers of CD8+ T cells, CD4+ T cells, macrophages, and cancer-associated fibroblasts infiltrating in the TME correlate with clinical outcomes in various malignancies, including gastric cancer, melanoma, urothelial cancer, lung cancer, and breast cancer (10,12-14).

Because gastric cancers are significantly associated with infectious agents, most notably *Helicobacter pylori* and Epstein-Barr virus (EBV), biomarkers that can predict responsiveness to immune-checkpoint blockade are being extensively investigated to further improve precision immunotherapy (15). Moreover, the abundance of immune cells and other cells in the TME can be estimated using computational methods (16-18). Although several studies using these methodologies have explored the clinical utility of TME infiltrates (7,19), and although several mechanisms associated with the role of TME in immunotherapy response and resistance have been experimentally identified for some tumor types (4,13,14,20,21), to date, the comprehensive landscape of cells infiltrating the TME has not yet been elucidated.
In the present study, two proposed computational algorithms (16,17) were employed to estimate the fractions of 22 immune cell types and cancer-associated fibroblasts based on clinically annotated gastric cancer gene expression profiles (1,22). We estimated the TME infiltration patterns of 1524 tumors from patients with gastric cancer and systematically correlated the TME phenotypes with genomic characteristics and clinical and pathological features of gastric cancer. As a result, we established a methodology to quantify the TME infiltration pattern (TMEscore). TMEscore was found to be a robust prognostic biomarker and predictive factor for response to immune-checkpoint inhibitors.

Materials and Methods

Gastric cancer datasets and preprocessing

We systematically searched for gastric cancer gene expression datasets that were publicly available and reported full clinical annotations. Patients without survival information were removed from further evaluation. In total, we gathered six treatment-naive cohorts of samples from patients with gastric cancer for this study: ACRG/GSE62254, GSE57303, GSE84437, GSE15459, GSE26253, GSE29272, and TCGA-STAD. The raw data from the microarray datasets generated by Affymetrix and Illumina were downloaded from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). The raw data for the dataset from Affymetrix were processed using the RMA algorithm for background adjustment in the Affy software package (23). RMA was used to perform background adjustment, quantile normalization, and final summarization of oligonucleotides per transcript using the median polish algorithm. The raw data from Illumina were processed using the lumi software package. Data from The Cancer Genome Atlas (TCGA) were downloaded from the UCSC Xena browser (GDC hub), as detailed in the supplementary methods. For TCGA dataset, RNA-sequencing data (FPKM values) were transformed into transcripts per kilobase million (TPM) values, which are more similar to those resulting from microarrays and more comparable between samples (24). The criteria
used for dataset selection, platform and source of each dataset, numbers of samples, and clinical end points are summarized in the Supplementary Methods and Supplementary Table S1. Data were analyzed with the R (version 3.4.0) and R Bioconductor packages.

Collection of clinical data

The corresponding clinical data from these datasets were retrieved and manually organized when available. For some series, clinical data not attached to gene expression profiles were obtained through one of the following three methods: i) directly downloaded from the corresponding item page in the GEO dataset website, ii) from the supplementary materials in the relative literature, and iii) using the GEOquery package in R. Corresponding authors were contacted for further information where necessary. Updated clinical data and sample information for TCGA-STAD samples were obtained from the Genomic Data Commons (https://portal.gdc.cancer.gov/) using the R package TCGAbiolinks (25). Overall survival information of all TCGA datasets was obtained from the supplementary data of recently published research (26).

Inference of infiltrating cells in TME

To quantify the proportions of immune cells in the gastric cancer samples, we used the CIBERSORT algorithm (16) and the LM22 gene signature, which allows for highly sensitive and specific discrimination of 22 human immune cell phenotypes, including B cells, T cells, natural killer cells, macrophages, DCs, and myeloid subsets. CIBERSORT is a deconvolution algorithm that uses a set of reference gene expression values (a signature with 547 genes) considered a minimal representation for each cell type and, based on those values, infers cell type proportions in data from bulk tumor samples with mixed cell types using support vector regression. Gene expression profiles were prepared using standard annotation files, and data were uploaded to the CIBERSORT web portal (http://cibersort.stanford.edu/), with the
algorithm run using the LM22 signature and 1,000 permutations. Proportions of stromal cells were estimated by applying the Microenvironment Cell Populations-counter method, which allows for robust quantification of the absolute abundance of eight immune and two stromal cell populations in heterogeneous tissues from transcriptomic data (17).

**Consensus clustering for TME-infiltrating cells**

Tumors with qualitatively different TME cell infiltration patterns were grouped using hierarchical agglomerative clustering (based on Euclidean distance and Ward's linkage). Unsupervised clustering methods (K-means) (27) for dataset analysis were used to identify TME patterns and classify patients for further analysis. A consensus clustering algorithm was applied to determine the number of clusters in the meta-dataset and Asian Cancer Research Group (ACRG) cohort to assess the stability of the discovered clusters. This procedure was performed using the ConsensusClusterPlus R package (28) and was repeated 1000 times to ensure the stability of classification.

**Differentially expressed genes (DEGs) associated with the TME phenotype**

To identify genes associated with TME cell infiltrating patterns, we grouped patients into TMEcluster-A, TMEcluster-B, and TMEcluster-C. DEGs among these three groups were determined using the R package limma (29), which implements an empirical Bayesian approach to estimate gene expression changes using moderated t-tests. DEGs among TME subtypes were determined by significance criteria (adjusted $P$ value $< 0.05$) as implemented in the R package limma. The adjusted $P$ value for multiple testing was calculated using the Benjamini-Hochberg correction (30).

**Dimension reduction and generation of TME gene signatures**
The construction of TME metagenes was performed as follows. First, each DEG among TMEcluster-ABC was standardized across all samples in the ACRG cohort. An unsupervised clustering method (K-means) (27) for analysis of DEGs was used to classify patients into three groups for further analysis. Then, the random forest classification algorithm was used to perform dimension reduction in order to reduce noise or redundant genes(31). Next, the clusterProfiler R package (32) was adopted to annotate gene patterns. A consensus clustering algorithm (28) was applied to define the cluster of genes, and principal component analysis (PCA) was conducted. Principal component 1 was extracted to serve as the gene signature score. After obtaining the prognostic value of each gene signature score, we applied a method similar to GGI (33) to define the TMEscore of each patient:

\[
\text{TMEscore} = \sum PC1_i - \sum PC1_j
\]

where \(i\) is the signature score of clusters whose Cox coefficient is positive, and \(j\) is the expression level of genes whose Cox coefficient is negative. Detailed data preprocessing steps are described in the Supplementary Methods.

Functional and pathway enrichment analysis

Gene-annotation enrichment analysis using the clusterProfiler R package (32) was performed on TME signature genes. Gene Ontology (GO) terms were identified with a strict cutoff of \(P < 0.01\) and false discovery rate (FDR) of less than 0.05. We also identified pathways that were up- and downregulated among TME gene clusters A and C for a certain TME-phenotype by running a gene set enrichment analysis (GSEA) (34) of the adjusted expression data for all transcripts. Gene sets were downloaded from the MSigDB database of Broad Institute (34). We included broad hallmarks and specific pathways of interest from the curated gene sets/canonical pathways collection. Enrichment \(P\) values were based on 10,000 permutations and subsequently adjusted for multiple testing using the Benjamini-Hochberg procedure to control the FDR (30).

Genomic and clinical data sets with immune-checkpoint blockade
Five genomic and transcriptomic datasets from patients with metastatic urothelial cancer (13) treated with anti-programmed death ligand 1 (PD-L1) agent (atezolizumab), patients with advanced melanoma treated with programmed death 1 (PD-1) blocker (35), patients with advanced melanoma treated with various types of immunotherapy from TCGA-SKCM cohort (36), patients with advanced melanoma treated with MAGE-A3 antigen-based immunotherapy (37) and mouse model treated with anti-CTLA4 antibody (38) were obtained, then we analyzed the predictive value of the TME signature score. Data sources and preprocessing methods are detailed in the Supplementary Methods.

**Statistical analysis**

The normality of the variables was tested by the Shapiro-Wilk normality test (39). For comparisons of two groups, statistical significance for normally distributed variables was estimated by unpaired Student t tests, and non-normally distributed variables were analyzed by Mann-Whitney U tests (also called the Wilcoxon rank-sum test). For comparisons of more than two groups, Kruskal-Wallis tests and one-way analysis of variance were used as non-parametric and parametric methods, respectively (40). Correlation coefficients were computed by Spearman and distance correlation analyses. Two-sided Fisher’s exact tests were used to analyze contingency tables. The cut-off values of each dataset were evaluated based on the association between patient overall survival and TMEscore in each separate dataset using the survminer package. The R package MaxStat (41), which iteratively tests all possible cut points to find the one achieving the maximum rank statistic, was used to dichotomize TMEscore and patients were then grouped into low and high TMEscore subtype. To identify significant genes in the differential gene analysis, we applied the Benjamini-Hochberg method to convert the P values to FDRs (30). The Kaplan-Meier method was used to generate survival curves for the subgroups in each data set, and the Log-rank (Mantel-Cox) test was used to determine the statistical significance of differences. The hazard ratios for univariate analyses were calculated using a univariate Cox
proportional hazards regression model. A multivariate Cox regression model was used to determine independent prognostic factors. The R package \textit{pROC} (42) was used to plot and visualize receiver operating characteristic (ROC) curves to calculate the area under the curve (AUC) and confidence intervals to evaluate the diagnostic accuracy of TMB, TMEscore and combination of them. For comparison of AUCs, likelihood ratio test for two correlated ROC curves was used. All statistical analyses were conducted using R (https://www.r-project.org/) or SPSS software (version 25.0), and the \( P \) values were two-sided. \( P \) values of less than 0.05 were considered statistically significant.

**Results**

**Landscape of TME in gastric cancer and clinicopathological characteristics of TME subtypes**

The construction scheme of TME cell infiltrating patterns and TME signatures was systematically evaluated (Fig. S1A). To select the optimal cluster number, we assessed clustering stability using the \textit{ConsensusClusterPlus} package (Fig. S1B) (28), which supported the existence of three robust subtypes of gastric cancer in meta-cohort (GSE57303, GSE34942, ACRG/GSE62254, GSE15459, GSE29272, and TCGA-STAD). Unsupervised hierarchical clustering of the 1524 tumors with matched TME cell expression profiles from above independent gastric cancer cohorts was performed, and the results are shown in Fig. S1C and Supplementary Table S2. Moreover, the TME cell network depicted a comprehensive landscape of tumor-immune cell interactions, cell lineages, and their effects on the overall survival of patients with gastric cancer (Figure 1A; Supplementary Table S3 and Supplementary Table S4). Moreover, three main TME cell infiltration subtypes revealed by the data showed significant differences in survival (Log-rank test, \( P < 0.001 \); Fig. 1B).

To further characterize and understand the biological and clinical differences
among these intrinsic phenotypes, we focused on the ACRG cohort (containing 299 patients with gastric cancer), not merely because it contained the most patients and provided the most comprehensive patient information (Supplementary Table S5), but also because the CIBERSORT algorithm was more suitable to deconvolve microarray data from the Affymetrix platform. Cluster analysis revealed three distinct patterns of TME cell infiltration as all GC datasets exhibited (Fig. S2A–D): TMEcluster-A was characterized by increases in the infiltration of cancer-associated fibroblasts, M2 macrophages, resting DCs, and resting MCs (43–46) and exhibited variable decreases in other TME cell types; TMEcluster-B exhibited high infiltration of M0 macrophages, neutrophils, activated DCs, and activated MCs; and TMEcluster-C showed significant increases in the infiltration of CD8+ T cells, M1 macrophages, and activated memory CD4+ T cells (43,44,47) (Fig. 1C). The significant differences in TME cell infiltration in the three main TME phenotypes were confirmed with Kruskal-Wallis tests (Fig. S2E; results of pairwise comparison were summarized in Supplementary Table S6).

In terms of clinical characteristics, TMEcluster-A was associated with a higher “Immunoscore” (Kruskal-Wallis, P < 2.2e-16; Fig. S2E; Supplementary Table S6), which we established based on a lasso immune signature score model in a previous study (7) to predict survival outcomes in patients with gastric cancer. In addition, we observed that samples in TMEcluster-A exhibited poorer tumor differentiation and enriched in EMT molecular subtype. The opposite patterns were observed in TMEcluster-C (Fig. 1C). Notably, survival analysis based on the TME phenotype showed TMEcluster-A (83 patients) to be significantly associated with poorer prognosis and TMEcluster-C (119 patients) to be associated with better prognosis (Log-rank test, P < 0.001). Of the 299 patients with gastric cancer, 97 belonged to TMEcluster-B, which was characterized by an intermediate prognosis (Log-rank test, P < 0.001; Fig. 1D).

Construction of the TME signature and functional annotation

To identify the underlying biological characteristics of each TME phenotype,
unsupervised analysis of 1033 DEGs acquired by *limma* package (48) was used to classify patients into genomic subtypes, consistent with the clustering results of the TME phenotype groups (Fig. S2F, Fig. S2G, and Supplementary Table S7). Next, we sought to employ random forest algorithms to perform dimension reduction to extract the phenotype signatures. The unsupervised hierarchical cluster analysis was based on the expression of the 238 most representative DEGs (Supplementary Table S8) and separated the ACRG cohort population into three distant patient clusters, termed Gene-clusters A–C (Fig. 2A). We visualized changes in clusters using an alluvial diagram (Fig. S3A). Analysis also revealed two significant expression gene sets (Fig. S3B and Supplementary Table S7).

GO enrichment analysis of the signature genes was conducted using the R package *clusterProfiler*. Significantly enriched biological processes are summarized in Supplementary Table S9. Gene-clusters A and C showed enrichment of distinct biological processes. Overexpression of genes involved in immune activation, which were enriched in Gene-cluster C, correlated with good prognosis in gastric cancer, and upregulated stroma-related genes, which were enriched in Gene-cluster A, were associated with poorer prognosis (Log-rank test, \( P < 0.001 \)) (10,20,21) (Fig. 2B–D and Supplementary Table S9). The *clusterProfiler* R package was employed to discover the potential regulatory relationships among TME signature mRNAs in gastric cancer, and these results suggested that pathways involved in the epithelial-mesenchymal transition (EMT) and immune activation exhibited a significant amount of overlap with other pathways (Fig. S3C). Fig. 2E indicated that the significant differences in TME cell infiltration and “Immunoscore” in the three TME gene clusters were consistent with the outcomes of TME cell infiltrating patterns (Fig. S2E), as expected. Robust correlations between TME signature scores and tumor-microenvironment cell infiltrating patterns were also validated in GSE15459 and TCGA-STAD datasets (Fig. S4A–H).

**Transcriptome traits and clinical characteristics of TME phenotypes**
in the ACRG cohort

Next, we defined two aggregate scores using the PCA algorithm: TMEscore A from TME signature genes A and TMEscore B from TME signature genes B (Fig. 2A and Supplementary Table S9). We computed TMEscore A and TMEscore B for each sample in the study as the sum of the relevant individual scores. To this end, we obtained the prognostic signature score, which we termed the TMEscore. In order to analyze the cytokine and chemokine milieu characterizing each gene cluster (Fig. 2A), we analyzed the expression of selected cytokine and chemokine mRNAs in the 299 gastric cancer samples. We considered CXCL10, CXCL9, GZMA, GZMB, PRF1, CD8A, IFNG, TBX2, and TNF to be immune-activated related transcripts; IDO1, CD274, HAVCR2, PDCD1, CTLA4, LAG3, and PDCD1LG2 to be immune-checkpoint-relevant transcripts; and VIM, ACTA2, COL4A1, TGFBR2, ZEB1, CLDN3, SMAD9, TWIST1, and TGRB1 to be transforming growth factor (TGF)-β/EMT pathway-relevant transcripts. Gene cluster A was associated with high expression of TGF-β/EMT pathway-relevant mRNAs, whereas expression of Th1/cytotoxic T lymphocyte (CTL)-related mRNAs (49) (particularly CXCL10, CXCL9, IFNG, and GZMB) was relatively low (Fig. S5A–C, results of pairwise comparison are summarized in Supplementary Table S6); this suggested that this cluster may be classified as the immune-suppressive group. In contrast, Gene cluster C, exhibiting the opposite mRNA expression profile, was classified as the immune-activated group. Moreover, we also tested known signatures within the gastric cancer dataset to better describe the functionality of the TME signature genes (Fig. 3A, Fig. S5D, and Supplementary Table S10). These analyses confirmed that TMEscore A was significantly associated with immune-relevant signatures, whereas TMEscore B was associated with stromal-relevant signatures (Fig. 3A).

Consistent with these findings, Gene cluster A with the EMT subtype (ACRG molecular subtypes) (22) was notably linked to low TMEscore (Fig. 3B, Kruskal-Wallis, $P < 2.2e$-16; Fig. 3C) and was associated with a poorer outcome (Fig. 3B, Fig. 3D). Furthermore, using GSEA with all transcripts ranked by the log2 (Fold
Change) between cluster A and cluster C, we found enriched expression of gene sets that are considered T-cell suppressive and exclusion (50-52) in TME gene cluster A (Fig. 3E, Fig. 3F, and Supplementary Table S11), including gene sets related to the EMT, TGF-β, and hypoxia.

After having identified TMEscore as an intrinsic gene expression signature closely linked to the stromal activation program and immune activation procedure, we sought to determine whether TMEscore could accurately predict outcomes. The 299 patients in the ACRG cohort were therefore assigned to groups based on high or low TMEscores using the cut-off value obtained with the survminer package. Five-year survival rates were 63% and 41% percent for the high and low TMEscore groups, respectively (HR, 0.32; 95% CI, 0.20–0.54; P < 0.001; Fig. 3D). When the TMEscore signature was evaluated as a continuous variable with the Cox regression model, the TMEscore model was determined to be an independent and robust prognostic factor (HR, 0.64; 95% CI, 0.50–0.82; P < 0.001; Fig. S5E). In addition, TMEscore was also investigated specifically in patients with stage II–III disease in the ACRG series to explore whether the application of adjuvant chemotherapy affected the ability of the TMEscore to predict survival outcomes. Patients were assigned to high and low TMEscore groups, and the survival advantage of the high TMEscore group was obvious, both in patients who received chemotherapy and in those who did not (Fig. 3G; Supplementary Table S6).

TME characteristics of TCGA subtype and cancer somatic genome

TCGA recently completed a comprehensive molecular characterization of gastric adenocarcinomas and has proposed subdividing tumors into four subtypes (1). Differences in the molecular subtypes were assessed in TCGA-STAD series, and a higher TMEscore was significantly associated with EBV infection, microsatellite instability (MSI), and good prognosis in gastric cancer, whereas the genome stable (GS) subtype had a lower TMEscore (22) and was associated with poorer prognosis (HR, 0.49; 95% CI, 0.31–0.76; P = 0.002; Fig. 4A, Fig. 4B, and Supplementary...
Table S12). The MSI-high subtype with the best prognosis had significantly higher TMEscores than the other two subtypes (Kruskal-Wallis, $P = 9.2e-12$, Fig. 4C). Correlation analyses between the known signatures and TMEscore were also validated in TCGA-STAD cohort (Fig. S5F), and the results were consistent with those of the ACRG cohort. The TMEscore model was again determined to be an independent and robust prognostic biomarker (HR, 0.74; 95% CI, 0.62–0.88; $P < 0.001$; Fig. S5G).

Notably, there was a significant positive correlation between the TMEscore and mutation load (Fig. 4D; Spearman coefficient: $R = 0.514$, $P = 2.2e-16$). Similar to MSI subtype, patients with EBV infection had significantly higher TMEscores and CTL infiltration than those with the GS and CIN molecular subtypes (Kruskal-Wallis $P < 2.2e-16$; Fig. 4B). Interestingly, several studies have indicated that EBV-positive gastric cancer does not exhibit higher tumor mutational burden (TMB) or MSI, but can respond to immune checkpoint therapy robustly (15,53), suggesting that TMEscore may be more useful for predicting clinical benefits in patients with gastric cancer treated with immunotherapy than TMB or MSI. We next investigated the distributions of somatic alterations and observed different patterns among gastric cancer clusters in terms of gene mutations. By analyzing the mutation annotation files of the TCGA-STAD cohort, we identified 33 highly variant mutated genes, which were associated with Tumor microenvironment score, using random forest algorithm with 1000 iterations (31) (Fig. 4E). Pre-clinical (54) and clinical (55) reports have described associations between individual altered genes and response or resistance to immune checkpoint blockade. Relatively few of these genes were exclusively correlated with sensitivity or resistance in TCGA-STAD series, such as PIK3CA and PCDH10. These data may provide a new perspective to study the mechanism of tumor microenvironment formation as well as explore individual mutations and their role in cancer immunity and immunotherapy.

**TMEscore was a prognostic biomarker that predicted**
immunotherapeutic benefits

Upon stratification of the samples according to specific datasets (Fig. 5A), large differences in overall survival were estimated between the low and high TMEscore groups for all gastric cancer datasets except GSE57303 (HR, 0.41; 95% CI, 0.13–1.34), as detailed in Supplementary Table S12. In addition, except for TNM stage I (HR, 0.58; 95% CI, 0.23–1.48), significant differences were observed in TMEscore among all other stages. Concurrently, the robust prognostic value of the TME signature was also validated in three other independent datasets (GSE15459: HR, 0.48; 95% CI, 0.29–0.77; GSE57303: HR, 0.41; 95% CI, 0.13–1.34; GSE84437: HR, 0.24; 95% CI, 0.13–0.45; Fig. S6A-C), as well as in a combined set of the five datasets (ACRG, TCGA-STAD, GSE15459, GSE57303, GSE84437) (Fig. S6D; HR, 0.42; 95% CI, 0.33–0.54). TMEscore was also predictive for relapse-free survival in the GSE26253 cohort (Fig. S6E; HR, 0.63; 95% CI, 0.46–0.87). Finally, we evaluated the prognostic value of TMEscore in 14 independent TCGA cancer cohorts including 7241 tumors (Supplementary Table S13). Although the results of subgroup analysis were heterogeneous, TMEscore was supported as a favorable prognostic biomarker in seven independent TCGA cohorts (Fig. 5B), which were acknowledged as hot tumors with diverse T-cell infiltration, including breast cancer, colon cancer, melanoma, lung squamous cell carcinoma, ovarian cancer, and cervical tumor.

Inhibition of immunological checkpoints with monoclonal antibodies that block the T-cell inhibitory molecules PD-L1 and PD-1 has emerged as an anticancer treatment with unprecedented and synergistic survival benefits (56). We next explored the prognostic value of TMEscore for immune-checkpoint therapy by assigning patients in the IMvigor210 and GSE78220 cohorts to high or low TMEscore groups. Patients with high TMEscores had significantly longer progression-free survival than those with lower TMEscores in both the IMvigor210 cohort (HR, 0.63; 95% CI, 0.46–0.89; Fig. 5C) and GSE78220 cohort (HR, 0.25; 95% CI, 0.07–0.89). The predictive value of TMEscore to checkpoint immunotherapy was also confirmed in IMvigor210 (Fig. 5C–F, Fig. S7A–E) and GSE78220 (Fig. 5G–J, Fig. S7F–G). TMEscore was
not well associated with overall survival and response to treatment with immunotherapy in TCGA-SKCM cohort (HR, 0.48; 95% CI, 0.17–1.41; **Fig. S7H** and **Fig. S7I–J**). However, it is worthwhile to point out that these patients in TCGA-SKCM cohort were collected from different medical centers with different study designs and received various types of immunotherapy, including cytokines, vaccine and checkpoint blockers. If bias is excluded, these results suggest potential limitation of TMEscore identifying responders to different immunotherapies. Interestingly, in good agreement with predicted outcomes of anti-PD-1 (GSE78220) and anti-PD-L1 (IMvigor210) antibody treatment, we validated the predictive value of TMEscore both in anti-MAGE-A3 (GSE35640, **Fig. S7K–L**) and anti-CTLA-4 (GSE63557, **Fig. S7M–N**) immunotherapy cohort. Additionally, patients with higher TMEscore (TMEscore of patients treated with immunotherapy were summarized in Supplementary Tables S14) were more likely to benefit from immune checkpoint therapy (IMvigor210 cohort: two-sided Fisher’s exact test, \( P < 0.001 \); **Fig. 5D**; Kruskal-Wallis test, \( P = 0.0041 \); **Fig. 5E**; GSE78220 cohort: two-sided Fisher’s exact test, \( P = 0.006 \); **Fig. 5H**; Wilcoxon test, \( P = 0.031 \); **Fig. 5I**). To investigate the biological characteristics of the TMEscore as it pertained to anti-PD-L1/PD-1 treatment, we observed that TMEscore A was highly positively correlated with a signature of CD8+ effector T cells, also known as CTLs (Spearman coefficient: \( R = 0.96, P < 2.2e^{-16} \); **Fig. S7D**), whereas TMEscore B was associated with the TGF-β response signal signature (F-TBRS; Spearman coefficient: \( R = 0.91, P < 2.2e^{-16} \); **Fig. S7D**), consistent with the results of gastric cancer datasets.

TMB (non-synonymous variants), which is significantly associated with efficacy of immunotherapy, was also evaluated with receiver operating characteristic (ROC) analysis (42) in IMvigor210 cohort (51). However, we did not observe a predictive advantage of TMB when compared with TMEscore (likelihood ratio test, \( P = 0.974 \), **Fig. 5F**). Moreover, combining TMB and TMEscore dramatically improved the predictive value compared with that of TMB or TMEscore alone using the pROC package (14) (likelihood ratio test, combination versus TMB, \( P = 0.004 \); combination
versus TMEscore, \(P = 0.019\); **Fig. 5F**). The survival advantage of patients in the high TMEscore group, for both high and low TMB groups, was evidently higher than that in the low TMEscore group (Log-rank test, \(P = 0.003\); **Fig. S7E**). In addition, the ROC analyses of GSE78220 and GSE35640 cohorts also demonstrated that TMEscore was a predictive biomarker to immunotherapeutic benefits (GSE78220: AUC = 0.731, **Fig. 5J**; GSE35640: AUC = 0.689, **Fig. S7L**). We next sought to validate the predictive value of TMEscore in mouse model treated with CTLA-4 antibody (accession number GSE63557, N=20). We got 82% conversion rate of TME signature genes A but only 6% conversion rate of TME signature genes B (human symbol) from mouse probes, thus, only the predictive value of TMEscoreA was estimated (Wilcoxon test, \(P < 0.001\), **Fig. S7M**; AUC = 1.000, **Fig. S7N**). Taken together, our data strongly suggest that TME evaluation is associated with response to different immunotherapy approaches, including anti-PD-1/PD-L1/CTLA-4 immune checkpoint inhibitors and MAGE-A3 antigen-based immunotherapy.

**Discussion**

The TME signature, a novel tool designed to evaluate the comprehensive TME, is a robust biomarker for predicting survival in gastric cancer and guiding more effective immunotherapy strategies. Our findings indicated that assessment of the immune and stromal statuses via the TME signature provided a potent predictor of survival in patients with gastric cancer and several other cancers, samples which were obtained from TCGA. Based on functional analysis of TME-relevant genes, our observations suggested that TME signature genes B were enriched for genes involved in extracellular matrix remodeling (\(DCN\), \(TIMP2\), \(FOXF2\), and \(MYH11\)), mesenchymal transition (\(ACTA2\), \(TGFB1L1\), and \(SFRP1\)), and cell adhesion and angiogenesis (\(PDGFRA\), \(GREM1\), and \(TMEM100\)), which are considered T-cell suppressive (13,35,51,57,58). Moreover, we observed high enrichment for genes involved in response to virus (\(IFNG\), \(TRIM22\), \(CXCL10\), \(CXCL9\), and \(CD8A\)), response to interferon-gamma (\(HLA-DPB1\), \(CCL4\), \(CCL5\), and \(IFNG\)), and T-cell activation.
(TRBC1, IDO1, CD2, NLRP3, and CD8A) among TME signature genes A.

Therapeutic antibodies that block the PD-1/PD-L1 pathway can induce robust and durable responses in patients with various cancers (11,12,35), including advanced gastric cancer (59). However, these responses only occur in a minority of patients, and several studies have found that PD-1 expression, PD-L1 expression, MSI status, and mutation load are not robust biomarkers for predicting the benefits of immune checkpoint blockade (15,59,60). The establishment of predictive biomarkers for checkpoint immunotherapy is therefore of the utmost importance in maximizing the therapeutic benefit (12,15,35). Emerging data support the idea that the TME plays a crucial role in checkpoint inhibitor immunotherapy (12-14,61). Here, we have elucidated the comprehensive landscape of interactions between the clinical characteristics of gastric cancer and infiltrating TME cells. With the help of several computational algorithms, a methodology was established to quantify the TME infiltration pattern-TME gene signature.

Integrated analysis revealed that TMEscore is a prognostic biomarker for gastric cancer and is significantly elevated in patients with MSI and EBV molecular subtypes (1,22,53), which have been confirmed to be more sensitive to immune-checkpoint blockade (15,62). In line with previous research, EBV-positive tumors had low mutation burden, but stronger evidence of immune infiltration (15,53), suggesting that our methodology to evaluate tumor microenvironment is a more predictive biomarker to further advance precision immunotherapy of gastric cancer. We also observed that TMEscore showed a strong positive correlation with mutation burden and predicted neo-antigen load in TCGA gastric cancer cohort. Furthermore, our data indicated that patients with EMT and GS subtypes exhibited the lowest TMEscores, consistent with studies (13,51,63) emphasizing that stromal activation is the core mechanism of resistance to checkpoint blockade. This resource may also help to facilitate the development of precision immunotherapy and the combined approach of both immunotherapy and inhibition of the EMT signaling pathway.

Notably, by applying ROC curve analysis (42), we also demonstrated the
predictive value of the TMEscore for checkpoint blockade in four separate cohorts of patients with metastatic urothelial cancer (13) treated with an anti-PD-L1 agent (atezolizumab), metastatic melanoma treated with anti-PD-1 (pembrolizumab), advanced melanoma treated with MAGE-A3 blocker (37), and mouse model treated with anti-CTLA-4 immunotherapy (38). Consistent with previous study about immune signature score (64), we observed a significantly higher TMEscore in responders than in non-responders undergoing checkpoint blockade therapy. However, the IS of previous research was trained and obtained from the transcriptome profile directly and only enriched in immune relevant pathways. We focused on the tumor microenvironment infiltrating patterns and accessed the subtype relevant gene signatures including immune signature (TMEscoreA) and stromal activation signature (TMEscoreB). These data offer mechanistic insights into the responses to immune checkpoint blockade, suggesting that response to PD-L1 and PD-1 blockade is not only related to enhanced cytolytic activity, antigen processing, and interferon-γ pathway components (13,65) but is also associated with inhibition of fibroblast activation, angiogenesis, the EMT, and TGF-β pathway components (13,51,57,58). This suggests that estimation of the immune TME combined with the stromal TME could potentially influence therapeutic resistance. Consistent with these findings, previous studies involving preclinical models of advanced cancer with activation of TGF-β- and EMT-relevant pathways, as well as fibroblast proliferation, demonstrated the inhibition of T cell-mediated tumor killing and a decrease in T-cell trafficking into tumors (13,51). Interestingly, in line with our findings, some preclinical studies have indicated that antibody-ligand traps (a-CTLA4-TGFβRII and a-PDL1-TGFβRII) exhibit a superior therapeutic index compared with those of their parent immune checkpoint inhibitors, which are currently in clinical use (63,66).

The results of our study should be further validated in a prospective cohort of patients receiving immunotherapy using NanoString nCounter gene expression platform (NanoString Technologies) to conquer the deficiency about uncertain cut-off values. Second, given the major clinical importance of distinct tumor regions, it is
appropriate to evaluate immune infiltration systematically in the core of the tumor and at the invasive margin. As not all patients with high TMEscore have greater benefit of immunotherapy, more clinical factors should be incorporated to prediction models for improvement of accuracy. In the current study, this comprehensive evaluation of the cellular, molecular, and genetic factors associated with TME infiltration patterns has yielded several important insights that shed light on how tumors respond to immunotherapies and may guide the development of novel drug combination strategies.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 81772580 to Wangjun Liao).

References


https://www.nature.com/articles/ng.3225#supplementary-information.

36. Network CGA. Genomic Classification of Cutaneous Melanoma. CELL 2015(1097-4172 (Electronic)) doi 10.1016/j.cell.2015.05.044.


44. Hwang ML, Lukens JR, Bullock TNJ. Cognate Memory CD4+ T Cells Generated with Dendritic Cell Priming Influence the Expansion, Trafficking, and Differentiation of Secondary CD8+ T Cells and Enhance Tumor Control. The Journal of Immunology 2007;179(9):5829-38 doi 10.4049/jimmunol.179.9.5829.


50. Stephanie C. Casey, Virginie Baylot IG, 1 Martin Eilers,2,3, Ling Tong YL, 1 Rachel Do,1 Susanne Walz,2 Dean W. Felsher1, Kelly N. Fitzgerald. MYC regulates the antitumor immune response through CD47 and PD-L1. Science 2015.


https://www.nature.com/articles/nature23643#supplementary-information.


Figure Legends

Figure 1. Landscape of the TME in gastric cancer and characteristics of TME subtypes

A, Cellular interaction of tumor micro environment cell types. Cell cluster-A, blue; Cell cluster-B, red; Cell cluster-C, brown; Cell cluster-D, orange. The size of each cell represents survival impact of each TME cell type, calculated used the formula log10(Log-rank test P value). Favor for overall survival is indicated in green, and risk for overall survival is indicated in black. The lines connecting TME cells represent cellular interactions. The thickness of the line represents the strength of correlation estimated by Spearman correlation analysis. Positive correlation is indicated in red and negative correlation in blue. B, Kaplan-Meier curves for overall survival (OS) of 1524 gastric cancer patients from seven gastric cancer cohorts (GSE15459, GSE29272, GSE34942, GSE57303, ACRG/GSE62254, GSE84437, and TCGA-STAD) with the TME infiltration classes. The number of patients in TMEcluster A, B and C phenotypes are n = 458, n = 625, and n = 441, respectively. Log-rank test showed an overall P < 0.001. C, Unsupervised clustering of TME cells for 299 patients in the ACRG cohort. Immunogroup (Immunophenotype from previous study), survival status, ACRG subtype, MSI status, histological subtype, gastric cancer grade and TMEcluster group are shown as patient annotations. D, Kaplan-Meier curves for overall survival (OS) of 299 patients in the ACRG cohort show that TME infiltration patterns are significantly associated with overall survival (Log-rank test, P < 0.001).

Figure 2. Construction of TME signatures and functional annotation

A, Unsupervised analysis and hierarchical clustering of common DEGs based on expression data derived from the ACRG cohort to classify patients into three groups, termed Gene clusters A–C. Two key intrinsic gene subtypes were found to correlate with immune or stroma activation (in Fig. 2 C–D), termed TME signature gene A and B, respectively. Immunogroup (Immunophenotype from previous study), survival status, ACRG subtype, MSI status, histological subtype, gastric cancer grade and TMEcluster group are shown as patient annotations. B, Kaplan-Meier curves for the three groups of patients (Gene clusters A–C). Log-rank test showed an overall P < 0.001. C–D, Gene Ontology (GO) enrichment analysis of the two TME relevant signature genes. The x axis indicates the number of genes within each GO term. E, The fraction of TME cells in three gene clusters. Within each group, the scattered dots represent TME cells expression values.
We also plotted the Immunoscore of three gene clusters. The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The statistical difference of three gene clusters were compared through the Kruskal-Wallis test. The range of P values are labeled above each boxplot with asterisks. (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001).

Figure 3. Transcriptome traits and clinical characteristics of TME phenotypes in the ACRG cohort

A, Gene clusters were distinguished by different signatures (immune-relevant signature, mismatch-relevant signature and stromal-relevant signature) and TMEscore. Within each group, the scattered dots represent mean value of signature genes. The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The range of P values are labeled above each boxplot with asterisks. (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001).
B, Alluvial diagram of TME gene-clusters in groups with different ACRG subtypes, TMEscores, and survival outcomes.
C, Differences of TMEscore in the ACRG subtype. The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The differences between every two groups were compared through the Kruskal-Wallis test.
D, Kaplan-Meier curves for high and low TMEscore patient groups in the ACRG subtype. The high TMEscore class is associated with better outcomes than the low TMEscore class (Log-rank test, P < 0.001).

E, F, Gene set enrichment analysis of hallmark gene sets downloaded from MSigDB database. All transcripts were ranked by log2(Fold change) between TME gene clusters A and C (in Fig. 2A). Each run was performed with 1,000 permutations. Enrichment results with significant associations between TME gene clusters A and C are shown (see detailed information in Supplementary Table S11).
G, Kaplan-Meier curves for patients with stage II-III gastric cancer in the ACRG cohort stratified by both receipt of adjuvant chemotherapy (CT) and TMEscore (Log-rank test, P < 0.001). ACRG, Asian Cancer Research Group.

Figure 4. TME characteristics of TCGA-STAD subtype and cancer somatic genome

A, Kaplan-Meier curves for high and low TMEscore groups of TCGA-STAD cohorts (Log-rank test, P = 0.002).
B, TMEscore differences in the TCGA-STAD molecular subtypes. The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The statistical difference of four groups were compared through the Kruskal-Wallis test.
C, Violin plot showing TMEscores in groups with...
different microsatellite instability (MSI) status. The differences between every two groups were compared through the Kruskal-Wallis test. D, Scatter plots depicting the positive correlation between TMEscore and mutation load in TCGA-STAT cohort. The Spearman correlation between TMEscore and mutation load was shown ($P = 2.2e-16$). The dot color indicates the TCGA molecular subtypes (CIN, red; EBV, green; GS, blue; MSI, purple). E, Distribution of highly mutated genes correlated with TMEscore. Single nucleotide variants: green; InDel (insertion or deletion): orange; frameshift: blue. The upper barplot indicates TMB, TMEscore, and overall survival (OS) per patient, whereas the left barplot shows the mutation frequency of each gene in separate TMEscore groups. TMEscores, TCGA molecular subtypes, TMEscore, gender, and overall survival status are shown as patient annotations.

**Figure 5. TMEscore is a prognostic biomarker and predicted immunotherapeutic benefits**

A, Subgroup analyses estimating clinical prognostic value between low/high TMEscore groups in independent gastric cancer datasets and cancer stage. Hazard ratios (HR) $< 1.0$ indicate that high TMEscore is a favorable prognostic biomarker. B, Subgroup analyses estimating prognostic value of TMEscore in different cancer types from TCGA datasets. HR $< 1.0$ indicates that high TMEscore is a favorable prognostic biomarker. C, Kaplan-Meier curves for patients with high and low TMEscores in the IMvigor210 cohort (Log-rank test, $P = 0.008$). D, Rate of clinical response to anti-PD-L1 immunotherapy in high or low TMEscore groups in the IMvigor210 cohort (two-sided Fisher’s exact test, $P < 0.001$). E, TMEscores in groups with different anti-PD-L1 clinical response statuses (complete response [CR], partial response [PR], stable disease [SD], progressive disease [PD]). The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The differences among groups were compared through the Kruskal-Wallis test. (Kruskal-Walls, $P = 0.004$). F, Receiver-operating characteristic (ROC) curves measuring the predictive value of the TMEscore, tumor mutation burden (TMB), and combination of TMEscore and TMB in the IMvigor210 cohort. The area under the ROC curve was 0.624, 0.623, and 0.700 for the TMEscore, TMB, and TMEscore combined with TMB, respectively; combining the two provided statistically significant improvement over single terms (likelihood ratio test, $P = 0.019$, and 0.004, respectively). G, Kaplan-Meier curves for patients with high and low TMEscores in the GSE78220 cohort (Log-rank test, $P = 0.021$). H, Rate of clinical response to anti PD-1 immunotherapy in high or low TMEscore groups in the GSE78220 cohort (two-sided Fisher’s exact test, $P = 0.006$). I, TMEscores in groups with different anti PD-1 clinical response status (CR/ PR, SD/PD). The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The differences between groups were compared through the Wilcoxon test (Wilcoxon, $P = 0.031$). J,
The predictive value of TMEscore is measured by ROC curves in the GSE78220 cohort. The area under the curve is 0.731.
Cancer Immunology Research

Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures

Dongqiang Zeng, Meiyi Li, Rui Zhou, et al.

Cancer Immunol Res  Published OnlineFirst March 6, 2019.

Updated version
Access the most recent version of this article at:
doi:10.1158/2326-6066.CIR-18-0436

Supplementary Material
Access the most recent supplemental material at:
http://cancerimmunolres.aacrjournals.org/content/suppl/2019/03/06/2326-6066.CIR-18-0436.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cancerimmunolres.aacrjournals.org/content/early/2019/03/06/2326-6066.CIR-18-0436. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.