# Evaluating prognostic value and relevant gene signatures of tumor microenvironment characterization in esophageal carcinoma

**Wang Sheng[1#], Xiaoyu Li[2#], Jiayi Li[1], Yanjun Mi[1], Fan Li[3^]**

[1]Department of Medical Oncology, Xiamen Key Laboratory of Antitumor Drug Transformation Research, The First Affiliated Hospital of Xiamen University, Xiamen, China; [2]Department of Oncology, Tongji Hospital Affiliated to Tongji Medical College of Huazhong University of Science and Technology, Wuhan, China; [3]Department of Thoracic Surgery, Tongji Hospital Affiliated to Tongji Medical College of Huazhong University of Science and Technology, Wuhan, China

*Contributions:* (I) Conception and design: Y Mi, F Li; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: W Sheng; (V) Data analysis and interpretation: Y Mi, F Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Fan Li. Tongji Hospital affiliated to Tongji Medical College of Huazhong University of Science and Technology, No. 1095 Jiefang Avenue, Wuhan 430030, China. Email: tjhtsdrli@163.com.

**Background:** Tumor microenvironment (TME) cells are an important part of tumor tissues. There is increasing evidence that the TME plays a vital role in tumor prognosis, and is associated with patient survival in various kinds of malignances. To date, very little research has been conducted on how to effectively use TME to better evaluate the prognosis of patients with esophageal carcinoma (EC). The concept of a "TME score" was introduced to better distinguish the prognosis of patients.

**Methods:** We employed bioinformatic methods to investigate the TME infiltration patterns of 160 patients with EC from the Cancer Genome Atlas (TCGA) cohort. TME clusters were identified using k-means clustering methods with 1,000 resampling times. The significance of the survival difference among patients belonging to different TME clusters was assessed by the log-rank test and Kaplan-Meier survival curves. Correlations between immune cell types and survival were calculated by a Cox regression, and the Pearson correlation coefficient (PCC) was used to measure the relationship among different immune cell types. We classified patient into 2 subtypes based on the optimal breakpoint of TME score determined by R package maxstat.

**Results:** Two TME phenotypes were defined based on the immune cell type fractions, and patients with a high TME score phenotype had a better prognosis than those with a low TME score phenotype. Kaplan-Meier analysis for differentially expressed micro ribonucleic acids (RNAs) and messenger RNAs also showed that different TME score subtypes were significantly associated with the prognosis of EC. Just as tumor mutational burden can predict the efficacy of immunotherapy, the TME score can predict the efficacy of immune checkpoint inhibitors (ICIs). The genomic alterations of 2 TME score subtypes of EC further revealed that genomic instability is prevalent in TMEs, and patients with a low TME score subtype have a more unstable chromosome status than those with a high subtype.

**Conclusions:** Thus, TME score is an emerging prognostic biomarker for predicting the efficacy of ICIs.

**Keywords:** Esophageal carcinoma (EC); tumor microenvironment (TME); TME score; prognosis

^ ORCID: 0000-0002-6591-8264.

## Introduction

Esophageal carcinoma (EC) is one of the most aggressive cancers. It was ranked 9th and 6th for incidence and mortality, respectively (1). The standard therapies for EC are surgery, chemotherapy, and radiotherapy. At present, 3 therapeutic agents that target human epidermal growth factor receptor 2 (HER2), vascular endothelial growth factor 2 (VEGF-2), and programmed cell death protein 1 (PD-1) have been approved by the Food and Drug Administration for use in the treatment of EC (2). Despite the great improvement in therapeutics, the 5-year survival rate remains poor at about 20% (3). Previous genomic studies have explored the molecular mechanism underlying EC, but have mainly focused on the genetic landscape (4,5). The tumor microenvironment (TME), which contains diverse immune cell populations, is thought to play a crucial role in the development of different cancers, including breast cancer (6), gastric cancer, (7) and esophageal squamous cell carcinoma (ESCC) (8), which suggest that immune cell infiltration may be associated with the clinical outcome of cancers. During the development of tumors, oncogenic mutations allow cells to proliferate independently, and cancer cells interact with the TME to favor its expansion and evade immune control (9).

A greater understanding of the tumor-immune microenvironment may provide more advanced prognostic biomarkers for the immunotherapy and chemotherapy of EC patients and reveal additional novel targets. Immunotherapy is a well-known treatment that works by modulating the TME. In addition, research suggests that the success of chemotherapy and radiotherapy relies on the induction of a durable anti-cancer immune response, and that targeted therapies also have immuno-stimulatory effects (10-12). It may be that the TME could change as a result of therapies, such that there is a shift from a pre-existing immune response to a therapy-induced immune response (10). Thus, information about the prediction of a treatment response and prognosis could be gained from the TME. A number of studies have highlighted the role of the TME in prognosis and treatment. For example, a high density of T lymphocytes in TME is associated with a favorable prognosis, while a high density of M2 macrophages is associated with a poor prognosis (10,13,14). The TME context determined at the time of diagnosis could also reflect the immune response and chemotherapy benefits (15,16). PD-L1 blockade is used in the microsatellite instability (MSI)/deficient mismatch repair (dMMR) treatment of colorectal cancers (17). Thus, it is necessary to integrate existing medical and

targetable biomarkers, including the MSI biomarker, for the comprehensive evaluation of EC.

Recently, several bioinformatic approaches have been developed to analyze the TME and have been successfully applied to some tumor types (7,18-22). However, little research has been conducted on the use of TME in evaluating the therapeutic efficacy of treatments for patients with EC. In relation to estimations of the TME, immunohistochemistry (IHC) is the only technique to yield quantitative information on the TME. In this study, we analyzed the immune infiltrate of EC by using a proposed computational algorithm, and developed an approach to quantify the TME infiltration pattern (i.e., the TME score). Our results showed that the TME score has potential clinical application and value in evaluating the prognosis of patients with EC. We further showed significant correlations between some micro ribonucleic acids (miRNAs) and differentially expressed genes (DEGs) in TME patterns with survival.

### *Purpose*

This study designed a TME signature to evaluate the comprehensive TME of EC, and a TME score classifier to effectively predict the prognosis of EC patients. The results of this study may have important implications for identifying subgroups of EC patients with a low or high risk of poor survival. We present the following article in accordance with the REMARK reporting checklist (available at https://dx.doi.org/10.21037/jgo-21-371).

## Methods

### *Data collection*

Molecular data from The Cancer Genome Atlas (TCGA) for EC (ESCA) type, including RNA-sequencing data, single nucleotide variation (SNV) date, copy number variations (CNV) date, miRNA microarray data, methylation microarray data, and corresponding clinical data with survival information, were downloaded (https://portal.gdc.cancer.gov/projects/TCGA-ESCA). Samples without survival information were excluded from further analysis. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

### *Data pre-processing*

Genes with a low expression level were removed by

supplying RNA-sequence raw read count data to the filterByExpr function of edgeR. The read counts of the remaining genes were then normalized by Voom in the Limma package to facilitate the deconvolution of immune cell types.

### Estimation of immune cell type fractions in the TME

CIBERSORT is an algorithm for enumerating the immune cell composition of complex tissues from bulk RNA-sequencing data. To infer the fraction of 22 immune cell types in tumor samples, a leukocyte gene signature matrix containing 547 genes (LM22) was used. We uploaded the pre-treated RNA-sequencing data (n=160) to the CIBERSORT web portal (http://cibersort.stanford.edu/), and ran the algorithm using the LM22 gene signature and 1,000 iterations.

### Consensus clustering for TME-immune cells

TME clusters were identified using k-means clustering methods in the ConsensuClusterPlus R package with 1,000 resampling times to ensure the stability of classification. The optimal number of clusters was determined using the elbow method. The significance of the survival difference among patients belonging to different TME clusters was assessed by the log-rank test and Kaplan-Meier survival curves. Correlations between immune cell types and survival were calculated by a Cox regression, and the Pearson correlation coefficient (PCC) was used to measure the relationship among different immune cell types.

### DEG analysis

DEGs between 2 TME clusters were identified by the Limma package in R with the significance criteria of a P value <0.001 and a $|\log_2 FC| > 1$. To explore the biological implications of these DEGs, functional enrichment was carried out using the R package cluster Profiler.

### Generation of TME gene signatures

The signature genes most closely related to the TME clusters were selected by reducing redundant DEGs using a random-forest algorithm. These selected signature genes were then separated into 2 categories according to the correlation between the Cox coefficients and survival. After calculating the coefficient, we defined the TME score,

which is similar to the previously reported Gene expression Grade Index (GGI), to summarize the total expression level of survival-related genes. The TME score is calculated according to the following formula:

$$\text{TME score} = \Sigma \text{ voom}(X) - \Sigma \text{ voom}(Y) \qquad [1]$$

where X is the expression level of genes whose Cox efficient is positive, and Y is the expression level of genes whose Cox efficient is negative.

### Exploration of the predictive value of the TME score for prognosis

To explore the relationship between the TME score and the survival of patients, we classified patient into 2 groups based on the optimal breakpoint of TME score determined by R package maxstat. The survival difference between the high and low TME score groups of patients was assessed using the log-rank test and Kaplan-Meier survival curve.

### Identification of somatic mutations

SNV data, comprising 160 samples, were used to obtain the mutation spectrum. The maftools R package was used.

### Identification of mutational signatures

SNV data, including 160 samples, were used to obtain the mutational spectra and identify mutational signatures. The maftools and SomaticSignatures R packages were used.

### Inference of the copy number variation (CNV)

The CNV regions among the 160 samples were detected by the Genomic Identification of Significant Targets in Cancer (GISTIC) module of GenePattern. The confidence interval was set as 0.95, and the other parameters were default values.

### Identification of prognosis-related differentially expressed mRNA and miRNA

As stated above, the patients were divided into high and low TME score groups. The differentially expressed messenger RNA (mRNA) and miRNA of these 2 groups were identified using the Limma R package. Next, a gene annotation enrichment analysis was performed. For the analysis of mRNA, the significance criteria were a P<5e-2 and a $|\log_2 FC| > 0.5$. For the analysis of miRNA,

the significance criteria were a P<0.05 and a |log$_2$FC| >0.25. For the analysis of methylated sites, the significance criteria were a P<0.05 and a |log$_2$FC| >0.5. Finally, the differentially expressed mRNA and miRNA were further analyzed to explore their correlation with survival.

### Statistical analysis

The data analyses were performed with R software (version 3.5.1) and R Bioconductor packages. The normality of the variables was examined using the Shapiro-Wilk normality test. Comparisons between the 2 groups were performed using unpaired Student *t*-tests if the variables were normally distributed, or otherwise, Wilcoxon rank-sum tests. The survival curves for the subgroups were generated using the Kaplan-Meier method, and the statistical significance of the differences was determined using the log-rank test. A P<0.05 was considered statistically significant.

### Data availability

All the data used in this study are available for download from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/projects/TCGA-ESCA).

## Results

### TME score subtypes as biomarkers for predicting the prognosis of EC

All 160 EC samples with overall survival (OS) information were used for the prognostic model construction. A univariate Cox regression analysis was performed to estimate the prognostic relationship between the TME score and OS in EC patients. Through the LASSO Cox regression model with minimized lambda, the TME score subtypes of 22 immune cell signatures were selected to build the TME score model in the training cohort. The associations between the TME score subtypes of 22 immune cell signatures and OS are presented in *Figure 1A*. Additionally, the TME cell network depicted tumor-immune cell interactions and their effects on the OS of patients with EC using Cytoscape (see *Figure 1B* and Table S1). The TME cell network revealed that eosinophils (23), M0 macrophages, M2 macrophages, natural killer cells, and memory cluster of differentiation (CD)4+ T cells (24) were significantly negatively correlated with prognosis in EC, while activated dendritic cells (DCs), CD8+ T cells, and naive CD4+ T cells were significant

positively correlated with prognosis in EC. Next, to achieve optimal clustering stability, the unsupervised clustering method (K-means) and elbow method were applied to identify the optimal K value. As a result, 2 clusters were determined (K=2; see Figure S1A, S1B). 2 distinct patterns of TME cell infiltration were determined (K=2) by unsupervised clustering (see Figure S1C), and the proportions of immune cells were slightly different between TME cluster 1 and TME cluster 2 (see Figure S1D).

To further understand the biological and clinical differences of the TME phenotypes, the DEGs between TME cluster 1 and TME cluster 2 were determined using the Limma R package. As a result, a total of 687 DEGs were included in the TME classification. An unsupervised clustering analysis was also performed to group patients into genomic subtypes (see Figure S2A, S2B), and the clustering results were significantly consistent with those of the TME phenotype groups (Pearson's Chi-squared test, P<0.001). We then applied a random-forest algorithm to reduce redundant DEGs, and obtained 119 signature genes (see Tables S2, S3). The biological processes of gene ontology (GO) terms were mainly enriched in immune-related processes, such as the humoral immune response mediated by circulating immunoglobulin, the immunoglobulin-mediated immune response, and B cell-mediated immunity (see *Figure 1C*). Based on the Cox regression model, the Cox coefficient for each cluster signature gene was obtained and used to calculate TME scores. As a result, the samples were classified into high or low TME scores. A further survival analysis stratified by TME score demonstrated that patients with a high TME score phenotype had a better prognosis than those with a low TME score phenotype (see *Figure 1D*).

In terms of clinical characteristics, we performed unsupervised hierarchical clustering on 202 survival-related genes, and the significant prognostic effect of the immune subtypes in 160 patients with EC were classified as either high or low TME score subtypes (see *Figure 2A* and Table S4). The survival difference was independent of grade, gender, purity, ploidy, tumor mutational burden (TMB), and OS status. The DEG analysis revealed 15 differentially expressed miRNAs and 202 differentially expressed mRNAs between the high and low TME score subtypes with a threshold adjusted P<0.05 and a |log2FC| >0.25 for miRNA, and an adjusted P<0.05 and a |logFC| >0.5 for mRNA, respectively.

The biological processes of GO terms revealed that differentially expressed mRNAs were mainly enriched

1232

Sheng et al. Evaluating prognosis with TME score
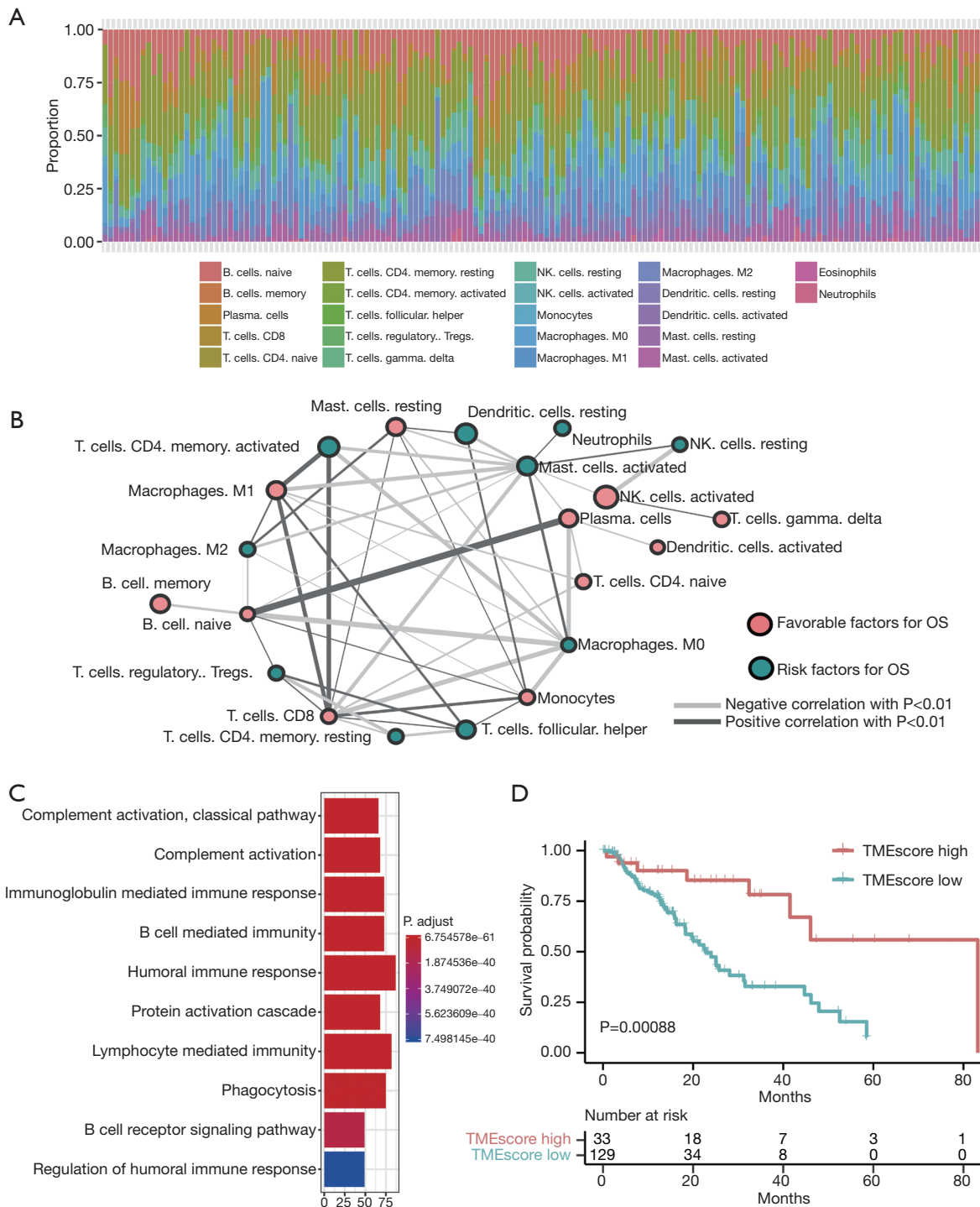


**Figure 1** Landscape of the TME in EC and the characteristics of TME subtypes. (A) A hierarchical heatmap of different immune cells. (B) Tumor-immune cell interaction network. The lines connecting the cells represent cellular interactions. The thickness of the line represents the strength of the correlation. The positive correlation is indicated in dark gray, and the negative in light gray. The size of each cell represents their survival effect. Favorable factors for overall survival are indicated in red, and risk factors are indicated in green. (C) The GO enrichment analysis of the 119 signature genes. (D) Kaplan-Meier curves for high (n=33) and low (n=129) TME score patient groups. Log-rank test, P<0.001. TME, tumor microenvironment; EC, esophageal carcinoma.
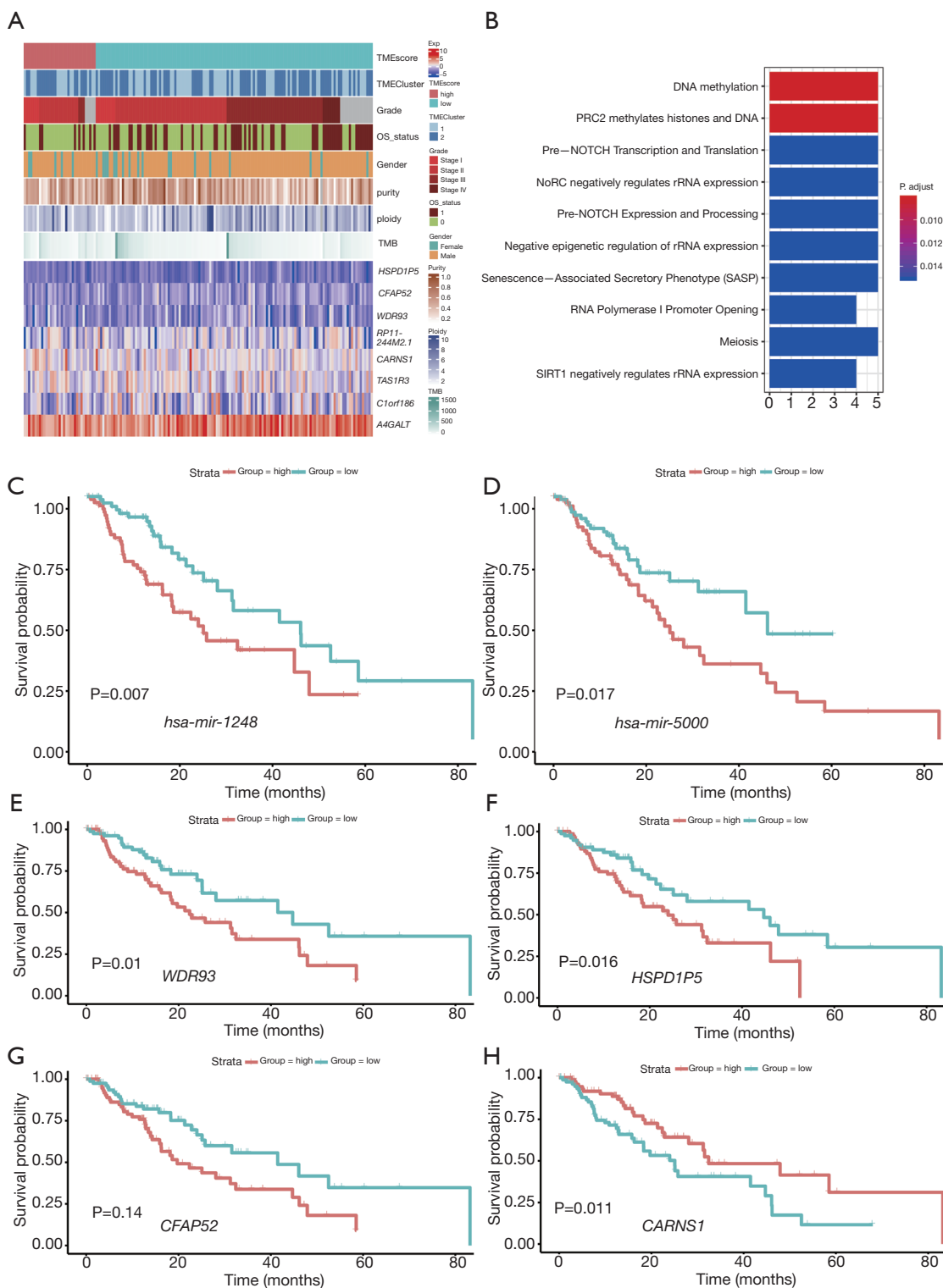
**Figure 2** Identification of novel genomic prognostic biomarkers. (A) Comprehensive genome landscape of ESCA. (B) The expression values of differentially expressed mRNA in TME clusters. Higher expression levels of *has-mir-1248* (C), *has-mir-5000* (D), *WDR93* (E), *CFAP52* (F), and *HSPD1P5* (G) were associated with shorter OS, while a higher expression level of CARNS1 (H) was associated with longer OS. TME, tumor microenvironment; OS, overall survival.

in deoxyribonucleic acid (DNA) methylation, polycomb repressive complex 2 (PRC2) methylates histones, and DNA (see *Figure 2B*). However, we did not find any differentially expressed methylated site. Based on the Cox regression model, 2 differentially expressed miRNAs and 8 differentially expressed mRNAs were identified that had a statistically significant effect on the prognosis of EC. In relation to these, the 2 miRNAs were hsa-mir-1248 and hsa-mir-5000, and the top 4 genes were *WDR93*, *CARNS1*, *CFAP52*, and *HSPD1P5*. The OS of patients with a high expression level of has-mir-1248 or has-mir-5000 was significantly shorter than that of patients with a lower expression level (see *Figure 2C,2D*). Similarly, a high expression level of *WDR93*, *CFAP52*, and *HSPD1P5* was a negative prognostic factor of EC (see *Figure 2E-2G*), while a high expression level of *CARNS1* was a good prognostic factor in EC (see *Figure 2H*).

### Genomic alterations of high and low TME score subtypes of EC

The landscape of the somatic mutations of EC are presented in *Figure 3A*. Missense mutation mainly caused by SNV was the major type of mutation, and C>T was the major type of base substitution (5). The most frequently mutated genes were TP53 in the high and low TME score subtypes. There were 28 genes with significant differences in mutation frequencies between the high and low TME score subtypes, among which the top 10 differential genes were *TP53*, *TTN*, *CSMD3*, *SYNE1*, *FLG*, *RIMS2*, *KMT2D*, *MUC16*, *MUC4*, *RYR2*, *LRP1B*, *DNAH5*, and *PCLO* (see *Figure 3A, B, C*). *TTN-AS1*, which is a potential diagnostic and prognostic biomarker (25), is highly expressed in ESCC tissues, and promotes ESCC cell proliferation and metastasis (26). *PCLO*, which is frequently mutated and amplified in ESCC, represents a novel prognostic biomarker and therapeutic target for patients with ESCC (27). These results further confirmed that these high-frequency mutated genes play a key role in the prognosis of EC tumors.

The mutational signature analysis showed that the high TME score subtype was associated with Signature 1, Signature 13, Signature 17, and Signature 29 (see *Figure 4A*). Conversely, the a TME score subtype was associated with Signature 1, Signature 3, and Signature 17 (see *Figure 4B*). Signature 1 was associated with the spontaneous deamination of 5-methylcytosine, Signature 3 was associated with DNA-double-strand breaks repair, and Signature 29 was associated with exposure to tobacco mutagens.

We also investigated genomic instability. The CNVs analysis by GISTIC showed that the significantly amplified regions included 7p, 8q, and 20q. Further, deletions, including at 3p, 5q, and 18q, frequently occurred in the low TME score subtype (see Figure S3A), whereas deletions, including at 9p, 4p, 5q, 18q, and 21q, frequently occurred in the high TME score subtype (see Figure S3B). The minimal common region analysis showed that amplifications of 8q24.21 and 11q13.3, and deletions of 2q22.1, 9p21.3, and 16q23.1 frequently occurred in the high TME score subtype, while amplifications of 11q13.3, and deletions of 7q31.1, 9p21.3, 16q23.1, and 4q22.1 frequently occurred in the low TME score subtype (see Figure S3C and Table S5). Thus, the results indicated a more instable status of chromosomes in patients with the low TME score subtype. Based on the CNV results, the tumor purity and ploidy analysis by ABSOLUTE showed that the estimated tumor purity ranged from 0.19 to 1.00, and tumor ploidy ranged from 1.75 to 10.12 (see *Figure 4C,4D* and Table S6). Thus, genomic disorder appears to be a common phenomenon in the process of tumorigenesis. There were no significant differences between high and low TME score subtypes in terms of tumor purity and ploidy (Wilcoxon-test, $P=0.62$ and $P=0.55$, respectively).

### The therapeutic benefits of the TME score

The TME score is a significant biomarker that can add prognostic information beyond that provided by standard clinicopathological properties (6,28). The predictive value of the TME score in EC patients who received adjuvant chemotherapy was examined in a TCGA-ESCA data set. The EC cohort included the mRNA expression levels of patients (n=42) before and after treatment with adjuvant chemotherapy. However, there was no statistically significant differences between patients before and after treatment with adjuvant chemotherapy (P=1.0 for TCGA-ESCA; see Figure S4A). In addition, MSI/dMMR was more frequently detected in well-differentiated tumors and younger patients (≤60 years) (29). We also tried to determine whether the TME score and MSI were correlated. The relationship between the TME score and MSI was identified in EC patients who were examined in the Sensor score and MANTIS (n=157) cohorts. The TME score was not correlated with MSI/dMMR (see Figure S4B,S4C).

The inhibition of the immunological checkpoints that block T-cell inhibitory molecules programmed cell death protein-1 (PD-1) and PD-L1, including pembrolizumab
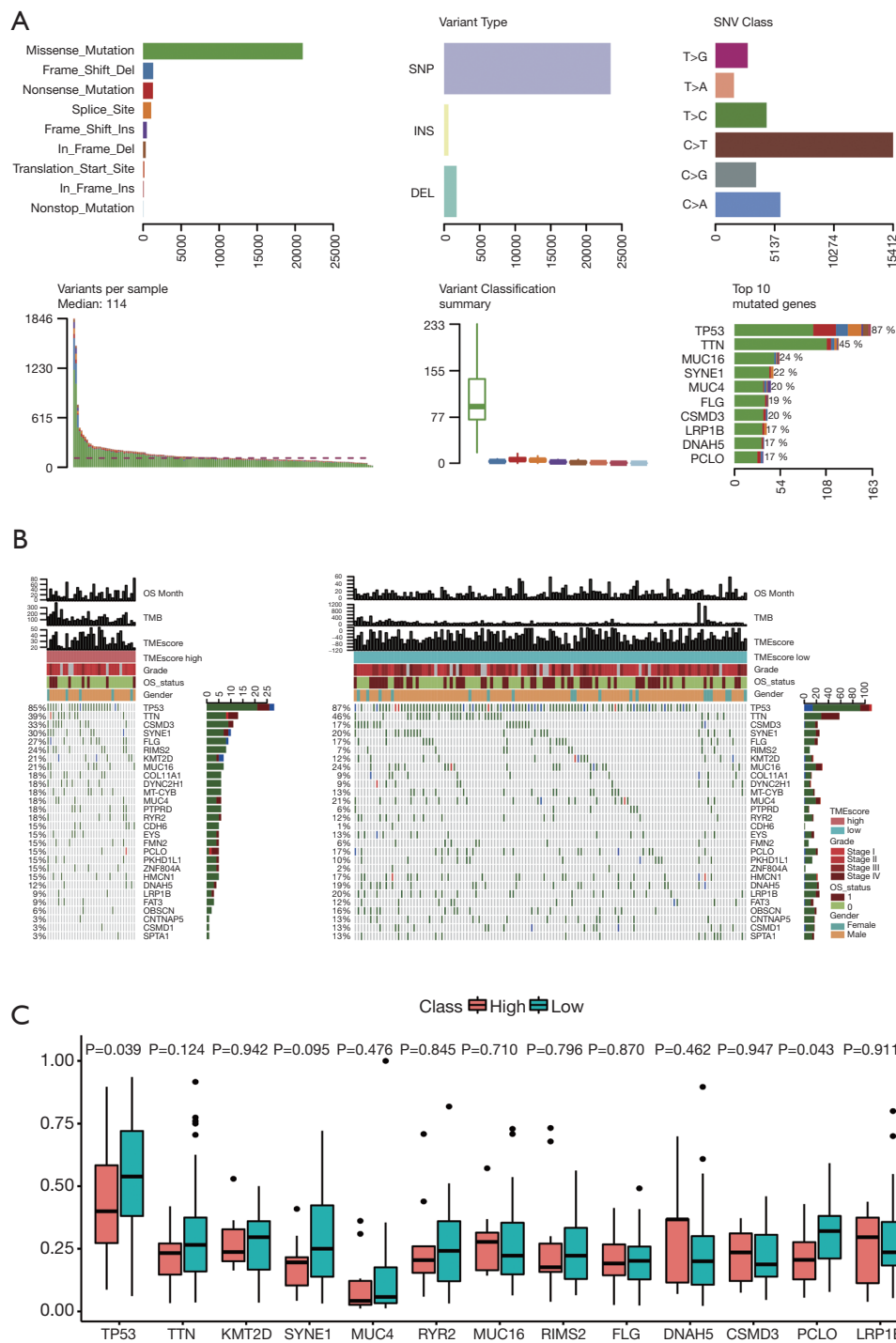
**Figure 3** TME characteristics and the cancer somatic genome in EC. (A) The landscape of somatic mutations of EC. (B) The oncoPrint was constructed by those with low TME scores on the right (blue) and those with high TME scores on the left (red). Individual patients are represented in each column. The top bar plot indicates OS, TMB, and TME score per patient. The right bar plot shows the mutation frequency of each gene in the separate TME score groups. Clinical stage, gender, and OS status are shown as patient annotations. (B) Mutation frequency of the genes in the 2 TME score groups. (C) Mutation frequency of the genes in the 2 TME score groups. TME, tumor microenvironment; EC, esophageal carcinoma; OS, overall survival.

*J Gastrointest Oncol* 2021;12(4):1228-1240 | https://dx.doi.org/10.21037/jgo-21-371
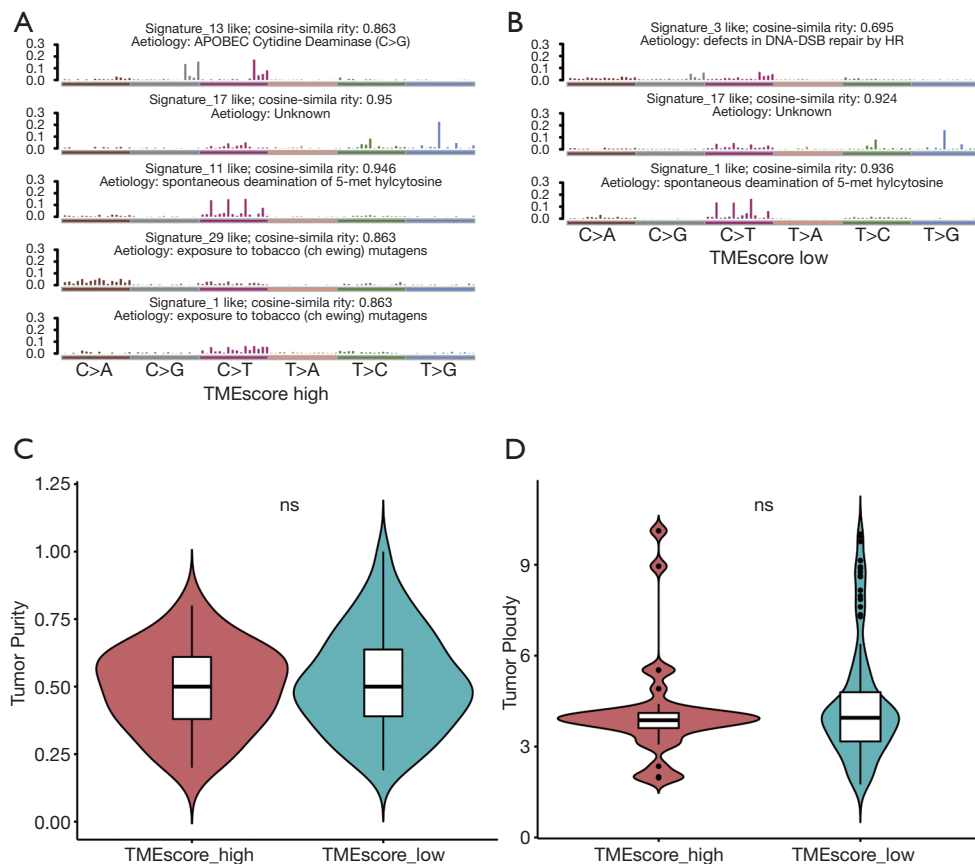
**Figure 4** Mutational signatures associated with TME score subtypes. (A) Mutational signatures associated with the high TME score subtype on the left and the low TME score subtype on the right. (B) The tumor purity and ploidy in the 2 TME score subtypes. (C) Tumor purity of TMEscore-high and TMEscore-low groups. (D) Ploidy of tumor samples in the TMEscore-high and TMEscore-low groups. The 'ns' mean 'no significant difference'. TME, tumor microenvironment.

and nivolumab, has emerged as an anti-cancer treatment with unprecedented and synergistic survival benefits (30). Previous research has implied that the Tumor Immune Dysfunction and Exclusion (TIDE) (http://tide.dfci.harvard. edu/) score model was significantly associated with the efficacy of immunotherapy (31). Further, the TIDE model was used to evaluate the clinical efficacy of immunotherapy for high and low TME score subtypes. By Kaplan-Meier analysis, the high TIDE score subtype had much shorter progression free survival then low TIDE score subtype in EC (Figure S4D). However, there was not significant difference in survival prediction efficiency between the low TIDE score subtype and the high TIDE score subtype (Wilcoxon-test, P=0.054) (*Figure 5A*).

The TMB, which may predict clinical responses to immune checkpoint inhibitors (ICIs) (32), was evaluated

by a receiver operating characteristic (ROC) analysis. The TMB is an emerging biomarker for predicting the efficacy of ICIs. The predictive efficiency of the TME score was not significantly different to that of the TMB based on ROC curve analysis (in which the AUC of the TME score was 0.6089 and the AUC of the TMB was 0.6181; P>0.05l; see *Figure 5B*). Thus, both TME score and TMB can be used to predict the survival of patients with EC, and the results were consistent.

## Discussion

Accurate prognostic assessment is critical in selecting the appropriate treatment. Oncologists routinely use the tumor, node, metastasis (TNM) staging system in prognosis. However, as clinical outcomes vary among patients in the
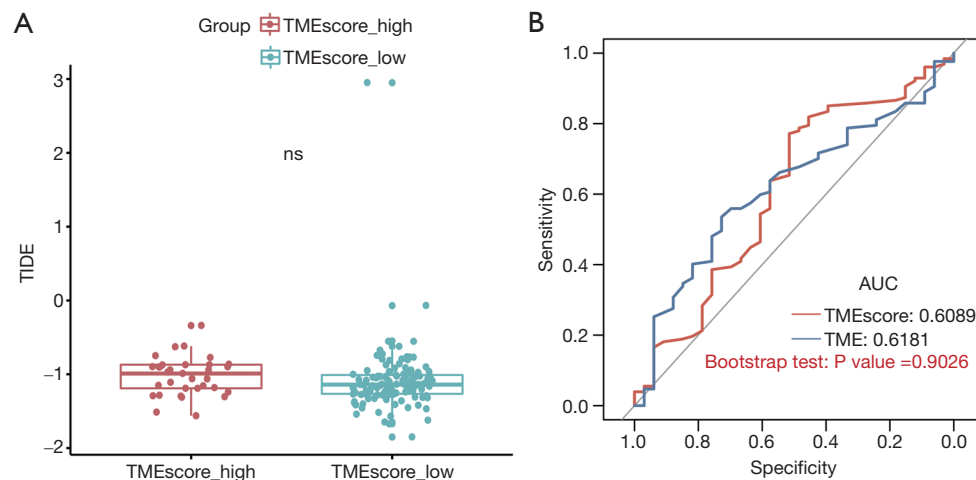
**Figure 5** TME score is a prognostic biomarker that predicts the effect of immunotherapy. (A) The distribution of TIDE in the 2 TME score subtypes. (B) The ROC curves of TME score and the TMB. TME, tumor microenvironment; ROC, receiver operating characteristic; TMB, tumor mutational burden. The 'ns' mean 'no significant difference'.

same TNM stages, a more accurate prognostic biomarker is needed (15,33,34). Recently, research has revealed that the TME is a promising predictive and prognostic biomarker. IHC has been widely used in various types of cancers, including EC, to characterize the tumor-immune infiltrates within the TME. However, IHC has a number of limitations; for example, it cannot assay cell types simultaneously. In the present study, we applied a bioinformatics method called CIBERSORT that integrates the expression signatures of immune cells to rapidly profile the immune infiltration patterns in EC. Notably, we found a significantly negative correlation between eosinophils and EC survival. Tumor infiltrating eosinophils have been reported in a variety of solid tumors, including EC; however, the function of eosinophils may depend on the cellular composition of the local TME and the presence of a costimulatory factor (35). Eosinophils are regulated by numerous molecules in which interleukin-5 (IL-5) is the most specific and critical cytokine that is responsible for eosinophils biology (36-38). IL-5 has been reported to be positively regulated by miR-1248 (39). MiR-1248 expression in the low TME score group was higher than that of the high TME score group, and high miR-1248 expression was associated with poor survival. These results further support the findings of this study.

We also constructed a classifier, referred to as the TME score, and examined its use as a prognostic tool. We found a significant correlation between the TME score and patients' survival, which suggests that the TME score

is a prognostic biomarker for EC. However, due to the scarcity of RNA-sequencing data sets containing clinical information, validations using other datasets could not be conducted in the current study. Thus, we explored the biological underpinnings of the TME signature to support our findings. Previous studies have suggested that *TP53* mutations lead to poor clinical outcomes in a number of cancers (40), and *PCLO*, which is a frequently mutated and amplified gene in ESCC, contributes to ESCC aggressiveness by stabilizing *EGFR* and promoting *EGFR* dependent signaling (27). Our in-depth analysis also revealed a more instable chromosome status in the low TME score group, which may also contribute to the association between low TME scores and poor clinical prognosis.

There are increasing evidences that the TME has substantial promise as a predictive biomarker for the ICI response (41-43). Jiang *et al.* developed a computational method, called the TIDE tool, to predict the immune checkpoint blockade response (44). However, the lack of treatment information in the study suggests that TIDE tool may not be applicable. Therefore, it is necessary to accept immunotherapy-related RNA data sets to evaluate the value of TME score in predicting the benefit of immunotherapy.

## Conclusions

We designed a TME signature to evaluate the comprehensive TME. Our TME score classifier effectively predicted

the prognosis of EC patients, which may have important implications for identifying subgroups of EC patients who are at low or high risk of poor survival.

## Footnote

*Reporting Checklist:* The authors have completed the REMARK reporting checklist (available at https://dx.doi.org/10.21037/jgo-21-371).

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/jgo-21-371). All authors reported that this work had received funding from the National Natural Science Foundation of China (No. 81871877), The Foundation for Young Scholars of The First Affiliated Hospital of Xiamen University (No. ZLYYA201705) and The National Major Scientific and Technological Special Project for "Significant New Drugs Development" (No. 2020ZX09201005), and got technical support from Shanghai Tongshu Biotechnology Co., Ltd. The authors have no other conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Institutional ethical approval and informed consent were waived.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

## References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Ajani JA, D'Amico TA, Bentrem DJ, et al. Esophageal and Esophagogastric Junction Cancers, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2019;17:855-83.
3. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. Lancet 2017;390:2383-96.
4. Gao YB, Chen ZL, Li JG, et al. Genetic landscape of esophageal squamous cell carcinoma. Nat Genet 2014;46:1097-102.
5. Wang F, Liu DB, Zhao Q, et al. The genomic landscape of small cell carcinoma of the esophagus. Cell Res 2018;28:771-4.
6. Wang S, Xiong Y, Zhang Q, et al. Clinical significance and immunogenomic landscape analyses of the immune cell signature based prognostic model for patients with breast cancer. Brief Bioinform 2020;bbaa311.
7. Zeng D, Li M, Zhou R, et al. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. Cancer Immunol Res 2019;7:737-50.
8. Zheng Y, Chen Z, Han Y, et al. Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. Nat Commun 2020;11:6268.
9. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646-74.
10. Fridman WH, Zitvogel L, Sautès-Fridman C, et al. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol 2017;14:717-34.
11. Galluzzi L, Buqué A, Kepp O, et al. Immunological Effects of Conventional Chemotherapy and Targeted Anticancer Agents. Cancer Cell 2015;28:690-714.
12. Galluzzi L, Senovilla L, Zitvogel L, et al. The secret ally: immunostimulation by anticancer drugs. Nat Rev Drug Discov 2012;11:215-33.
13. Fridman WH, Galon J, Pagès F, et al. Prognostic and predictive impact of intra- and peritumoral immune

infiltrates. Cancer Res 2011;71:5601-5.

14. Fridman WH, Pagès F, Sautès-Fridman C, et al. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 2012;12:298-306.

15. Jiang Y, Zhang Q, Hu Y, et al. ImmunoScore Signature: A Prognostic and Predictive Tool in Gastric Cancer. Ann Surg 2018;267:504-13.

16. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet 2016;387:1909-20.

17. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 2017;357:409-13.

18. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016;17:218.

19. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013;39:782-95.

20. Gentles AJ, Newman AM, Liu CL, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med 2015;21:938-45.

21. Stoll G, Bindea G, Mlecnik B, et al. Meta-analysis of organ-specific differences in the structure of the immune infiltrate in major malignancies. Oncotarget 2015;6:11894-909.

22. Stoll G, Enot D, Mlecnik B, et al. Immune-related gene signatures predict the outcome of neoadjuvant chemotherapy. Oncoimmunology 2014;3:e27884.

23. O'Sullivan JA, Bochner BS. Eosinophils and eosinophil-associated diseases: An update. J Allergy Clin Immunol 2018;141:505-17.

24. Biswas SK, Mantovani A. Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. Nat Immunol 2010;11:889-96.

25. Zheng QX, Wang J, Gu XY, et al. TTN-AS1 as a potential diagnostic and prognostic biomarker for multiple cancers. Biomed Pharmacother 2021;135:111169.

26. Lin C, Zhang S, Wang Y, et al. Functional Role of a Novel Long Noncoding RNA TTN-AS1 in Esophageal Squamous Cell Carcinoma Progression and Metastasis. Clin Cancer Res 2018;24:486-98.

27. Zhang W, Hong R, Xue L, et al. Piccolo mediates EGFR signaling and acts as a prognostic biomarker in esophageal squamous cell carcinoma. Oncogene 2017;36:3890-902.

28. Zhang H, Song J, Dong J, et al. Tumor Microenvironment Analysis Identified Subtypes Associated With the Prognosis and the Tumor Response to Immunotherapy in Bladder Cancer. Front Genet 2021;12:551605.

29. Lee HK, Kwon MJ, Ra YJ, et al. Significance of druggable targets (PD-L1, KRAS, BRAF, PIK3CA, MSI, and HPV) on curatively resected esophageal squamous cell carcinoma. Diagn Pathol 2020;15:126.

30. Curran MA, Montalvo W, Yagita H, et al. PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. Proc Natl Acad Sci U S A 2010;107:4275-80.

31. TIDE: Tumor Immune Dysfunction and Exclusion. Available online: http://tide.dfci.harvard.edu/

32. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019;51:202-6.

33. Galon J, Pagès F, Marincola FM, et al. Cancer classification using the Immunoscore: a worldwide task force. J Transl Med 2012;10:205.

34. Galon J, Pagès F, Marincola FM, et al. The immune score as a new possible approach for the classification of cancer. J Transl Med 2012;10:1.

35. Reichman H, Karo-Atar D, Munitz A. Emerging Roles for Eosinophils in the Tumor Microenvironment. Trends Cancer 2016;2:664-75.

36. Collins PD, Marleau S, Griffiths-Johnson DA, et al. Cooperation between interleukin-5 and the chemokine eotaxin to induce eosinophil accumulation in vivo. J Exp Med 1995;182:1169-74.

37. Rothenberg ME. Eosinophilia. N Engl J Med 1998;338:1592-600.

38. Sanderson CJ. Interleukin-5, eosinophils, and disease. Blood 1992;79:3101-9.

39. Panganiban RP, Pinkerton MH, Maru SY, et al. Differential microRNA epression in asthma and the role of miR-1248 in regulation of IL-5. Am J Clin Exp Immunol 2012;1:154-65.

40. Wang X, Sun Q. TP53 mutations, expression and interaction networks in human cancers. Oncotarget 2017;8:624-43.

41. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat Rev Cancer 2019;19:133-50.

42. Riaz N, Havel JJ, Makarov V, et al. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. Cell 2017;171:934-949.e16.

43. Tumeh PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature 2014;515:568-71.

44. Jiang P, Gu S, Pan D, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med 2018;24:1550-8.

(English Language Editor: L. Huleatt)
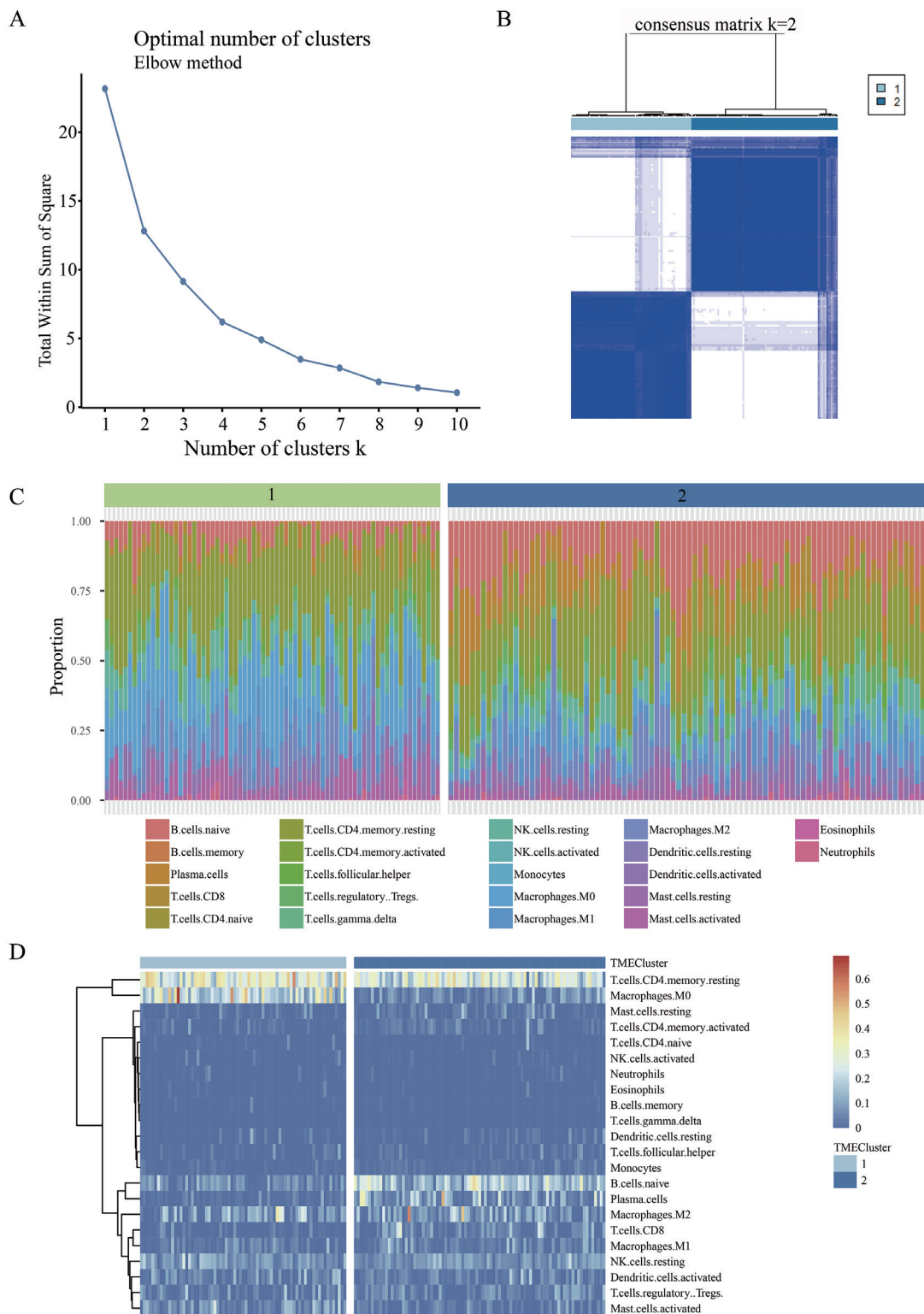
**Figure S1** TME clusters of tumor-immune cells. (A) Estimation of the optimal k value using the elbow method. (B) Consensus matrixes for K=2, displaying the clustering stability using 1,000 iterations of hierarchical clustering. (C) The proportion of tumor-immune cells in the TME cluster. (D) The distribution of tumor-immune cells in the TME cluster. TME, tumor microenvironment.
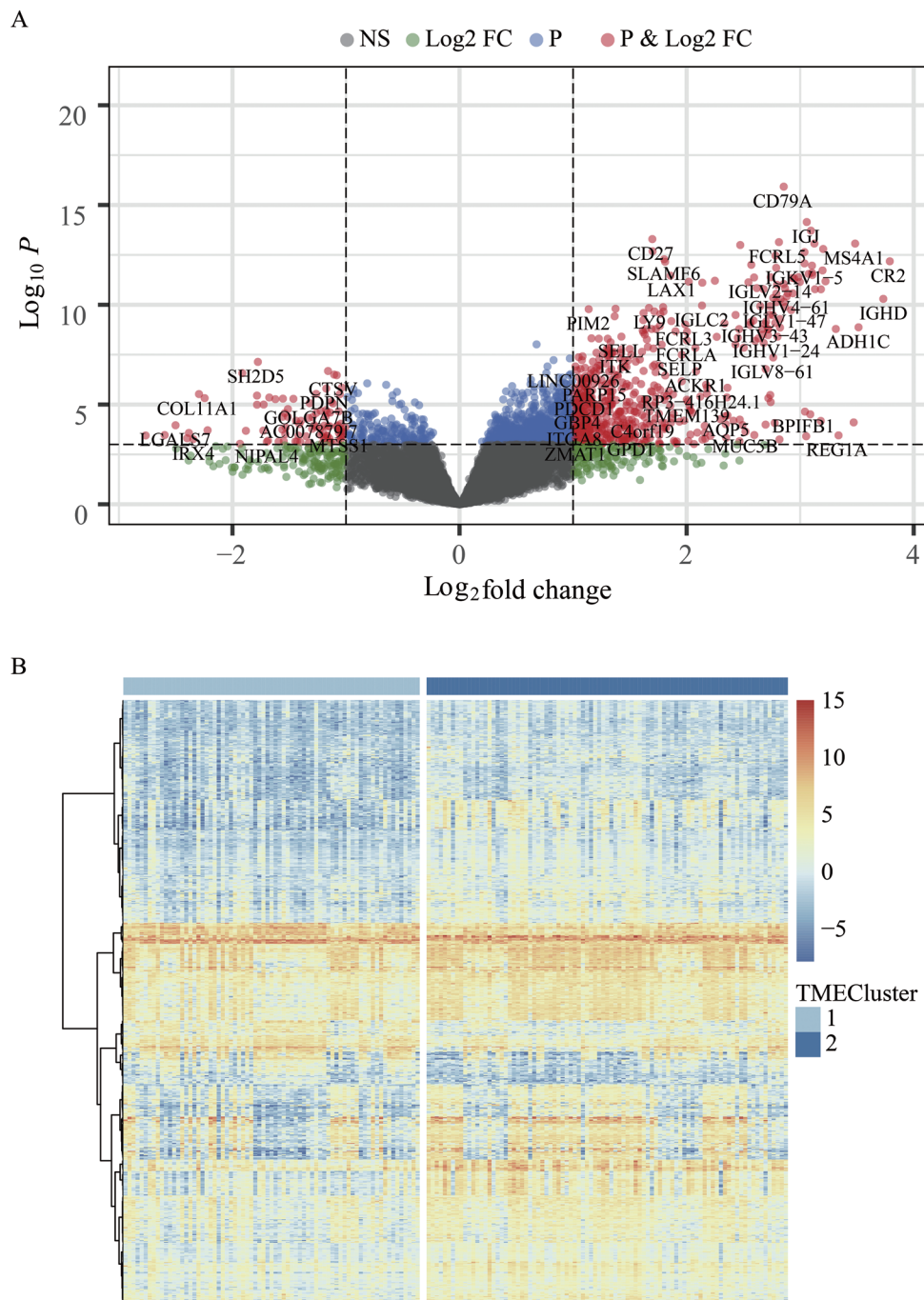
**Figure S2** DEGs in the TME cluster. (A) Volcano plot of DEGs in the TME cluster. (B) The expression values of DEGs in the TME cluster. DEG, differentially expressed gene; TME, tumor microenvironment.
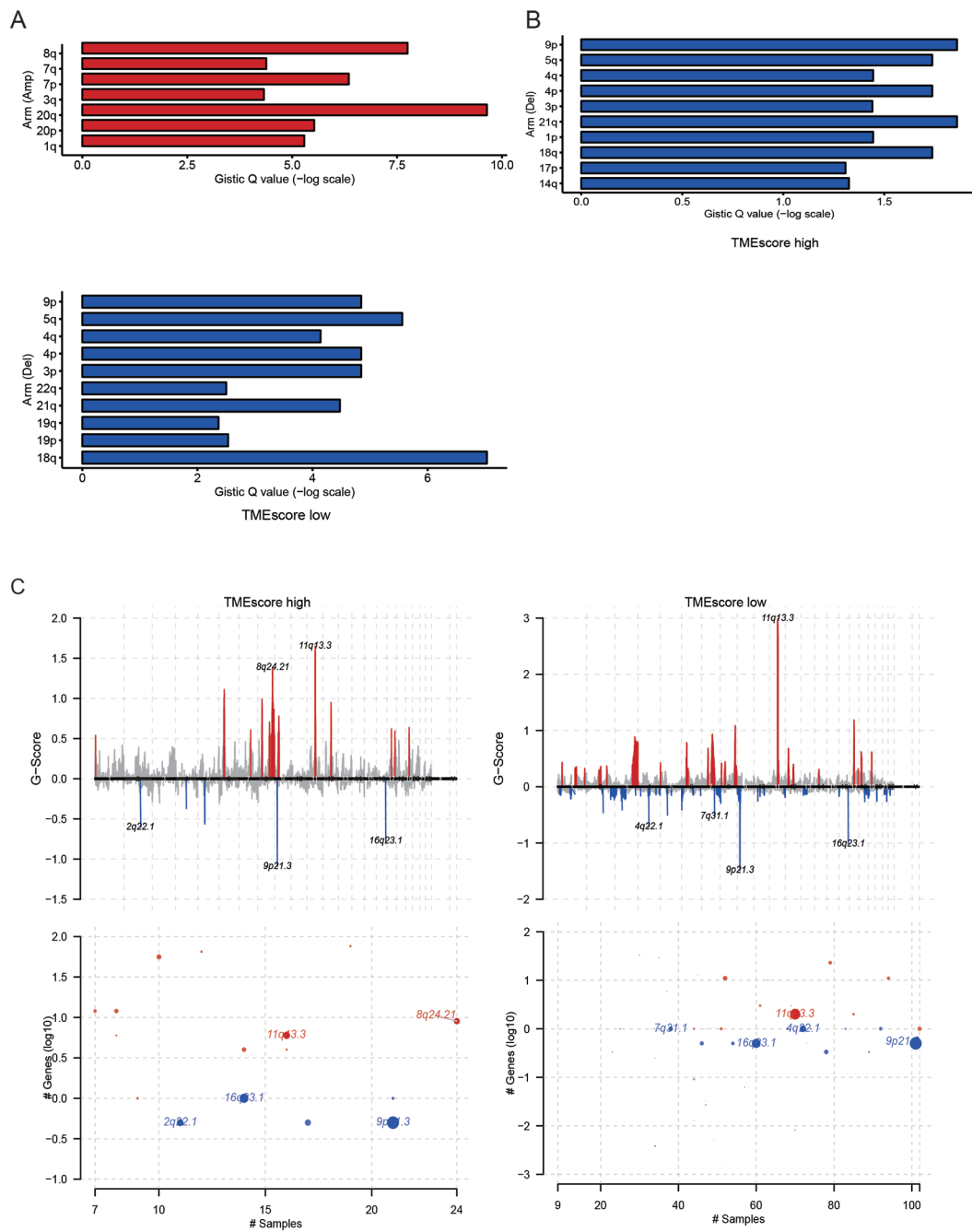
**Figure S3** CNV analysis and Kaplan-Meier analysis. (A) Amplification and deletion of chromosome arm in both TME score-low. (B) Deletion of chromosome arm in both TME score-high. (C) Distribution of localized copy number amplification and deletion regions (red and blue represent amplification and deletion, respectively). CNV, copy number variation; TME, tumor microenvironment.
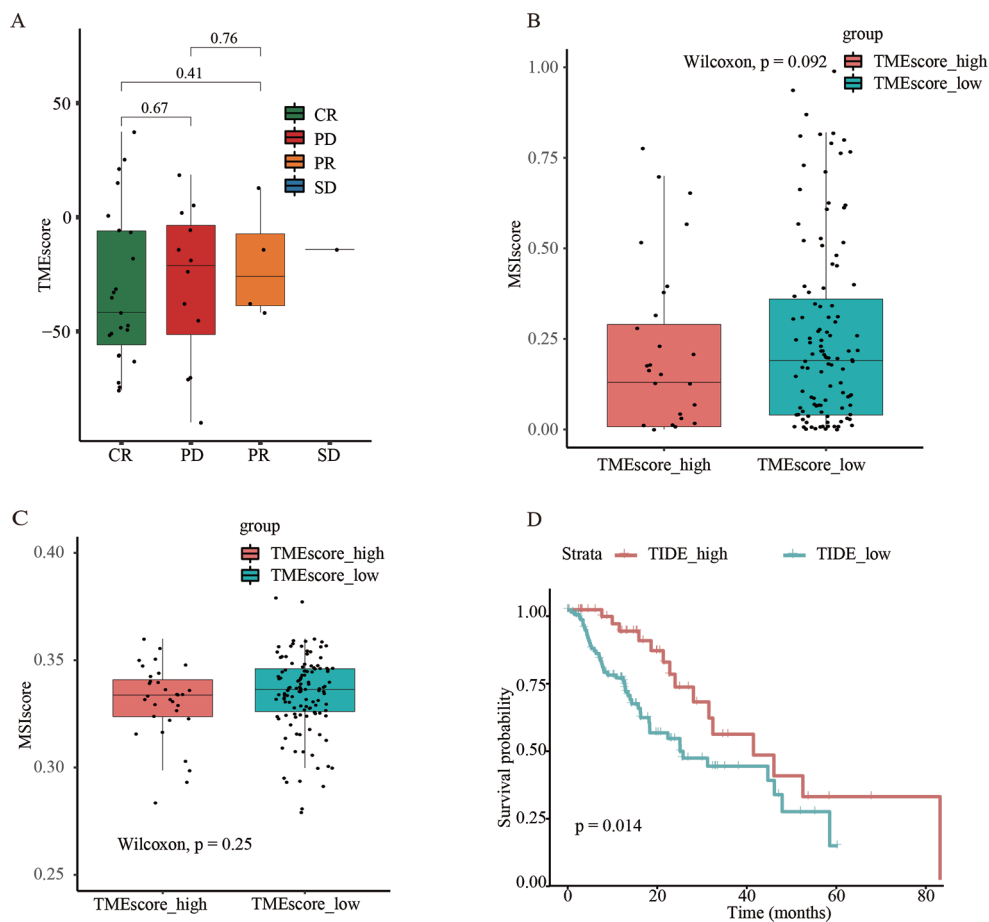
**Figure S4** The therapeutic benefits of the TME score. (A) The predictive value of the TME score in EC patients who received adjuvant chemotherapy. (B and C) The relationship between the TME score and MSI (Sensor score and MANTIS cohorts). (D) Kaplan-Meier curves of TIDE. TME, tumor microenvironment; EC, esophageal carcinoma; MSI, microsatellite instability; TIDE, the Tumor Immune Dysfunction and Exclusion; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

**Table S1** Association between immune cells and OS

| Cell type | Risk | Log rank P value |
|---|---|---|
| Eosinophils | Negative | 0.033 |
| NK.cells.activated | Positive | 0.106 |
| Dendritic.cells.resting | Negative | 0.179 |
| T.cells.CD4.memory.activated | Negative | 0.202 |
| Mast.cells.activated | Negative | 0.254 |
| B.cells.memory | Positive | 0.292 |
| T.cells.follicular.helper | Negative | 0.297 |
| Macrophages.M1 | Positive | 0.306 |
| Mast.cells.resting | Positive | 0.315 |
| Plasma.cells | Positive | 0.320 |
| Neutrophils | Negative | 0.565 |
| T.cells.gamma.delta | Positive | 0.579 |
| Monocytes | Positive | 0.608 |
| T.cells.CD4.naive | Positive | 0.619 |
| T.cells.regulatory.Tregs. | Negative | 0.623 |
| T.cells.CD8 | Positive | 0.629 |
| Macrophages.M2 | Negative | 0.648 |
| T.cells.CD4.memory.resting | Negative | 0.655 |
| NK.cells.resting | Negative | 0.690 |
| Macrophages.M0 | Negative | 0.719 |
| B.cells.naive | Positive | 0.775 |
| Dendritic.cells.activated | Positive | 0.833 |

OS, overall survival.


**Table S2** Signature genes positively associated with the TME

| Genes |
|---|
| *IGHD, SSTR1, PKDCC, ACSL5, HOXB-AS3, PPP1R16B, DPCR1, SLC44A4, GNA14, DSC3, MECOM, PRR26, C16orf74, CHDH, DUOX2, HOMER3, SLC1A1, CTSE, ALDOB, TRHDE, CD5, SNAI2, MYRF, SHROOM3, GSDMB, BCL2L14, CLDN2, USH1C, ERN2, HABP2, CA13, WDR66, DMBT1, TMPRSS2, GPD1, AC007386.4, TJP3, SULT1B1, ZG16B, PPP1R1B, FAM221A, RP11-324O2.3, AKR7L, ADH6, GATA6, KCNK5, FUT4, TESC, VILL, S1PR5, IRF8, PIWIL4, DNAJC22, PRR15L, FA2H, C9orf152* |

TME, tumor microenvironment.


**Table S3** Signature genes negatively associated with the TME

| Genes |
|---|
| *IGJ, CAPN5, CCNI2, HID1, ATP10B, VNN2, PLEKHA6, ARHGEF38, AGT, MYO1A, KLRB1, RP11-1220K2.2, HOXB6, CES3, LGALS4, HSH2D, C4orf19, MUC5B, ARTN, BTNL8, ICA1, TMC5, IQGAP2, VSIG2, SULT1C2, SLC37A1, CDX2, TRIM31, PLS1, METTL7B, SH3BGRL2, SAMD5, SMIM24, MYO15B, ADAM28, GTF2IRD2P1, KIAA1244, NCMAP, PPARG, RP3-395M20.8, SLC4A4, DDAH1, HNF4G, TOX3, AC006042.6, CTSS, PIGR, IL17RB, RAB17, OLFM4, ONECUT2, XK, SHH, SLC3A1, GATM, SPNS3, NOSTRIN, DLX2, HNF1A-AS1, GPR35, CYP3A5, PIP5K1B, RP11-739N20.2* |

TME, tumor microenvironment.

**Table S4** Landscape of the TME in EC

| | Age | Sex | Grade | OS (months) | OS status | Score | Group | TME cluster | Purity | Ploidy | TMB |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TCGA-JY-A93E-01 | 61 | Male | Stage III | 25.2 | 0 | 54.21742431 | High | 2 | 0.46 | 3.75 | 147 |
| TCGA-L5-A8NT-01 | 69 | Male | Stage II | 27.1 | 0 | 53.3347741 | High | 2 | 0.39 | 3.45 | 87 |
| TCGA-L5-A4OU-01 | 81 | Male | Stage II | 28.98 | 0 | 48.7687703 | High | 2 | 0.42 | 4.14 | 146 |
| TCGA-2H-A9GR-01 | 80 | Male | Stage II | 32.42 | 1 | 45.99521387 | High | 2 | 0.2 | 3.87 | 380 |
| TCGA-2H-A9GK-01 | 43 | Male | Stage III | 7.62 | 1 | 45.73813682 | High | 2 | 0.29 | 3.62 | 153 |
| TCGA-JY-A938-01 | 75 | Male | Stage II | 34.82 | 0 | 43.65594148 | High | 2 | 0.7 | 3.85 | 115 |
| TCGA-JY-A6FD-01 | 51 | Female | Stage II | 67.97 | 0 | 43.03214985 | High | 2 | 0.65 | 3.93 | 114 |
| TCGA-IG-A4P3-01 | 48 | Male | Stage II | 18.63 | 1 | 40.02176844 | High | 2 | 0.34 | 4.04 | 227 |
| TCGA-Z6-A8JD-01 | 53 | Male | Stage II | 3.42 | 0 | 36.12807077 | High | 2 | 0.25 | 4.11 | 182 |
| TCGA-L5-A4OH-01 | 71 | Male | Stage I | 32.59 | 0 | 29.05047095 | High | 2 | 0.62 | 3.08 | 174 |
| TCGA-JY-A939-01 | 77 | Male | Stage II | 21.68 | 0 | 25.80574654 | High | 2 | 0.34 | 4.05 | 52 |
| TCGA-KH-A6WC-01 | 82 | Male | Stage I | 6.27 | 0 | 25.41678036 | High | 2 | 0.61 | 2 | 26 |
| TCGA-L5-A8NW-01 | 55 | Male | | 46.06 | 1 | 23.96660129 | High | 2 | 0.4 | 4.4 | 130 |
| TCGA-L5-A8NE-01 | 77 | Male | Stage II | 55.45 | 0 | 22.37891614 | High | 2 | 0.28 | 4.91 | 214 |
| TCGA-LN-A49S-01 | 59 | Male | Stage II | 13.14 | 0 | 21.97945873 | High | 2 | 0.48 | 8.95 | 126 |
| TCGA-V5-AASX-01 | 74 | Male | | 8.97 | 0 | 20.61834862 | High | 2 | 0.61 | 3.84 | 267 |
| TCGA-L5-A8NU-01 | 84 | Male | Stage II | 83.18 | 1 | 18.42952882 | High | 2 | 0.38 | 3.76 | 52 |
| TCGA-L5-A4OG-01 | 79 | Female | Stage I | 4.7 | 0 | 17.35606076 | High | 2 | 0.5 | 3.85 | 148 |
| TCGA-L5-A43C-01 | 81 | Male | | 3.15 | 0 | 47.81625028 | High | 1 | 0.62 | 3.61 | 121 |
| TCGA-VR-A8EY-01 | 44 | Female | Stage II | 33.67 | 0 | 45.20710463 | High | 1 | 0.74 | 2.35 | 91 |
| TCGA-JY-A6FB-01 | 77 | Male | Stage I | 60.35 | 0 | 42.01339107 | High | 1 | 0.54 | 3.94 | 75 |
| TCGA-IG-A5B8-01 | 72 | Male | Stage I | 0.79 | 1 | 41.79390784 | High | 1 | 0.38 | 3.91 | 231 |
| TCGA-LN-A49U-01 | 62 | Male | Stage II | 15.34 | 0 | 37.46944181 | High | 1 | 0.57 | 4.17 | 115 |
| TCGA-S8-A6BW-01 | 51 | Male | Stage I | 20.37 | 0 | 36.18138671 | High | 1 | 0.56 | 3.29 | 85 |
| TCGA-L5-A88Z-01 | 70 | Female | Stage II | 7.39 | 0 | 35.72041456 | High | 1 | 0.8 | 1.97 | 59 |
| TCGA-JY-A6FG-01 | 50 | Male | Stage III | 41.49 | 1 | 32.61960224 | High | 1 | 0.57 | 4.11 | 186 |
| TCGA-L5-A891-01 | 51 | Male | | 3.75 | 0 | 31.09107672 | High | 1 | 0.45 | 5.52 | 144 |
| TCGA-JY-A6FH-01 | 53 | Male | Stage II | 47.34 | 0 | 30.47473299 | High | 1 | 0.22 | 5.53 | 108 |
| TCGA-JY-A93F-01 | 58 | Female | Stage I | 24.01 | 0 | 27.08134281 | High | 1 | 0.55 | 4.09 | 84 |
| TCGA-IG-A3QL-01 | 54 | Male | Stage II | 35.18 | 0 | 25.3023902 | High | 1 | 0.64 | 3.07 | 77 |
| TCGA-LN-A8HZ-01 | 56 | Male | Stage II | 12.32 | 0 | 21.31145906 | High | 1 | 0.5 | 10.12 | 91 |
| TCGA-V5-A7RC-06 | 55 | Male | | 3.42 | 1 | 18.62300766 | High | 1 | 0.47 | 3.67 | 75 |
| TCGA-IG-A97I-01 | 58 | Male | Stage II | 12.16 | 0 | 17.30201828 | High | 1 | 0.65 | 2 | 99 |
| TCGA-IG-A625-01 | 60 | Male | Stage III | 12.81 | 1 | 16.51455926 | Low | 2 | 0.66 | 4.39 | 45 |

**Table S4** (*continued*)

**Table S4** (*continued*)

| | Age | Sex | Grade | OS (months) | OS status | Score | Group | TME cluster | Purity | Ploidy | TMB |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TCGA-JY-A93D-01 | 51 | Male | Stage III | 31.54 | 1 | 15.18941367 | Low | 2 | 0.25 | 2.25 | 95 |
| TCGA-Z6-AAPN-01 | 57 | Male | Stage II | 2.66 | 0 | 14.52304483 | Low | 2 | 0.41 | 3.81 | 330 |
| TCGA-V5-AASV-01 | 67 | Male | Stage II | 15.34 | 0 | 12.97941016 | Low | 2 | 0.75 | 2.26 | 89 |
| TCGA-Z6-A9VB-01 | 53 | Male | Stage III | 1.31 | 0 | 12.9078017 | Low | 2 | 0.34 | 7.61 | 81 |
| TCGA-IG-A3YB-01 | 61 | Male | Stage III | 2.63 | 0 | 12.72295909 | Low | 2 | 0.61 | 4.02 | 65 |
| TCGA-V5-A7RB-01 | 59 | Male | | 5.29 | 1 | 12.17730994 | Low | 2 | 0.44 | 4.17 | 243 |
| TCGA-L5-A8NV-01 | 75 | Male | Stage II | 52.53 | 1 | 11.44863349 | Low | 2 | 0.38 | 3.14 | 100 |
| TCGA-JY-A93C-01 | 47 | Male | Stage III | 23.16 | 0 | 11.09493906 | Low | 2 | 0.55 | 3.16 | 80 |
| TCGA-L5-A4OX-01 | 60 | Male | Stage II | 7.42 | 1 | 7.248113325 | Low | 2 | 0.65 | 3.87 | 115 |
| TCGA-2H-A9GJ-01 | 57 | Male | Stage I | 58.51 | 1 | 6.561082013 | Low | 2 | 0.7 | 2.1 | 97 |
| TCGA-L5-A88T-01 | 86 | Male | Stage II | 22.8 | 0 | 4.715614504 | Low | 2 | 0.45 | 4.01 | 77 |
| TCGA-L5-A43E-01 | 74 | Male | Stage I | 30.22 | 0 | 4.53116118 | Low | 2 | 0.51 | 3.36 | 135 |
| TCGA-L5-A893-01 | 71 | Female | Stage I | 3.02 | 0 | 3.793991204 | Low | 2 | 0.37 | 4.2 | 120 |
| TCGA-2H-A9GL-01 | 74 | Male | Stage III | 5.91 | 1 | 3.728631352 | Low | 2 | 0.4 | 4.8 | 186 |
| TCGA-L7-A6VZ-01 | 62 | Male | Stage III | 10.35 | 0 | 0.837002711 | Low | 2 | 0.53 | 8.16 | 98 |
| TCGA-XP-A8T6-01 | 54 | Male | Stage II | 25.07 | 1 | 0.530405828 | Low | 2 | 0.58 | 3.8 | 71 |
| TCGA-R6-A8W5-01 | 60 | Male | Stage IV | 15.77 | 1 | –5.454195281 | Low | 2 | 0.2 | 4.15 | 93 |
| TCGA-IC-A6RF-01 | 69 | Female | Stage I | 15.67 | 0 | –5.605140944 | Low | 2 | 0.55 | 2.14 | 95 |
| TCGA-2H-A9GO-01 | 58 | Male | Stage IV | 16.23 | 1 | –6.066961882 | Low | 2 | 0.43 | 8.63 | 118 |
| TCGA-IG-A8O2-01 | 62 | Male | Stage III | 4.66 | 1 | –8.354187901 | Low | 2 | 0.61 | 3.2 | 109 |
| TCGA-VR-AA7I-01 | 70 | Male | Stage III | 15.9 | 1 | –10.56812262 | Low | 2 | 0.41 | 3.21 | 90 |
| TCGA-IC-A6RE-01 | 59 | Male | Stage II | 7.69 | 0 | –11.77930855 | Low | 2 | 0.2 | 4.17 | 462 |
| TCGA-JY-A6FE-01 | 49 | Male | Stage III | 3.68 | 1 | –12.74473462 | Low | 2 | 0.38 | 5.28 | 51 |
| TCGA-L5-A4OE-01 | 81 | Male | Stage III | 23.98 | 1 | –13.04247581 | Low | 2 | 0.49 | 4.6 | 261 |
| TCGA-L7-A56G-01 | 65 | Male | | 10.84 | 1 | –14.06502577 | Low | 2 | 0.55 | 3.38 | 63 |
| TCGA-L5-A4OI-01 | 79 | Male | Stage III | 19.97 | 0 | –16.44914795 | Low | 2 | 0.66 | 4.33 | 1263 |
| TCGA-L5-A8NG-01 | 77 | Male | Stage III | 35.94 | 0 | –17.24047186 | Low | 2 | 0.3 | 7.86 | 118 |
| TCGA-LN-A7HY-01 | 50 | Male | Stage III | 12.02 | 0 | –17.91795058 | Low | 2 | 0.35 | 7.35 | 142 |
| TCGA-L5-A8NH-01 | 54 | Male | Stage IV | 12.91 | 1 | –23.41078436 | Low | 2 | 0.27 | 3.96 | 112 |
| TCGA-VR-AA4D-01 | 53 | Male | Stage II | 46.16 | 1 | –23.69132545 | Low | 2 | 0.62 | 3.41 | 77 |
| TCGA-LN-A9FP-01 | 60 | Female | Stage II | 12.02 | 0 | –31.30560784 | Low | 2 | 1 | 4.02 | 415 |
| TCGA-LN-A49W-01 | 73 | Male | Stage III | 13.24 | 0 | –32.73325872 | Low | 2 | 0.43 | 3.88 | 73 |
| TCGA-Q9-A6FW-01 | 61 | Male | Stage III | 7.82 | 0 | –32.85359156 | Low | 2 | 0.5 | 4.21 | 185 |
| TCGA-LN-A8I1-01 | 67 | Female | Stage II | 13.17 | 0 | –33.6048471 | Low | 2 | 0.63 | 3.46 | 87 |

**Table S4** (*continued*)

| | Age | Sex | Grade | OS (months) | OS status | Score | Group | TME cluster | Purity | Ploidy | TMB |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TCGA-M9-A5M8-01 | 58 | Male | Stage II | 33.08 | 0 | −34.38105001 | Low | 2 | 0.93 | 2.79 | 65 |
| TCGA-VR-A8EP-01 | 51 | Male | Stage III | 27.07 | 0 | −37.48722031 | Low | 2 | 0.42 | 3.08 | 86 |
| TCGA-V5-AASW-01 | 72 | Male | | 9.26 | 0 | −38.88859241 | Low | 2 | 0.31 | 6.12 | 117 |
| TCGA-2H-A9GI-01 | 68 | Male | Stage III | 14.29 | 1 | −39.75879743 | Low | 2 | 0.55 | 2.66 | 164 |
| TCGA-L5-A4OJ-01 | 70 | Female | Stage I | 20.99 | 0 | −39.98917851 | Low | 2 | 0.86 | 1.84 | 218 |
| TCGA-JY-A6FA-01 | 51 | Male | Stage II | 44.71 | 1 | −41.72588264 | Low | 2 | 0.72 | 2.09 | 106 |
| TCGA-L5-A4OS-01 | 86 | Female | Stage II | 58.54 | 0 | −44.40856809 | Low | 2 | 0.19 | 2.7 | 82 |
| TCGA-LN-A4A8-01 | 52 | Male | Stage II | 15.51 | 0 | −49.16551343 | Low | 2 | 0.71 | 2.03 | 99 |
| TCGA-IG-A7DP-01 | 50 | Female | Stage III | 14.85 | 0 | −50.76160756 | Low | 2 | NA | NA | 18 |
| TCGA-2H-A9GH-01 | 44 | Male | Stage II | 31.24 | 1 | −51.05706433 | Low | 2 | 1 | 8.76 | 106 |
| TCGA-L5-A4OT-01 | 77 | Male | Stage IV | 4.89 | 1 | −53.32740443 | Low | 2 | 0.49 | 3.92 | 205 |
| TCGA-VR-A8Q7-01 | 60 | Male | Stage III | 52.23 | 0 | −54.60812908 | Low | 2 | 0.29 | 2.89 | 72 |
| TCGA-L5-A88V-01 | 60 | Male | Stage III | 2.6 | 0 | −55.41109354 | Low | 2 | 0.72 | 3.28 | 144 |
| TCGA-L5-A4OP-01 | 67 | Female | Stage I | 7.16 | 0 | −55.62208884 | Low | 2 | 0.69 | 1.81 | 100 |
| TCGA-L5-A8NS-01 | 76 | Male | Stage II | 13.4 | 0 | −55.69961448 | Low | 2 | 0.39 | 3.52 | 339 |
| TCGA-L5-A4OO-01 | 75 | Male | Stage III | 3.32 | 0 | −56.64870948 | Low | 2 | 0.33 | 4.28 | 88 |
| TCGA-R6-A6KZ-01 | 42 | Male | | 5.06 | 1 | −56.70320347 | Low | 2 | 0.35 | 5.13 | 135 |
| TCGA-VR-A8EZ-01 | 47 | Male | Stage III | 18.17 | 1 | −59.50710675 | Low | 2 | 0.41 | 6.18 | 201 |
| TCGA-R6-A6XQ-01 | 58 | Male | | 6.34 | 1 | −60.02680856 | Low | 2 | 0.8 | 9.95 | 65 |
| TCGA-LN-A7HZ-01 | 49 | Male | Stage II | 13.17 | 0 | −60.29923058 | Low | 2 | 0.5 | 10.01 | 60 |
| TCGA-IG-A51D-01 | 63 | Male | Stage II | 17.02 | 0 | −60.4165586 | Low | 2 | 0.5 | 4.27 | 134 |
| TCGA-L5-A88S-01 | 84 | Male | Stage I | 15.47 | 0 | −60.78792761 | Low | 2 | 0.82 | 3.9 | 75 |
| TCGA-L5-A8NM-01 | 84 | Female | Stage II | 7.75 | 1 | −61.38506461 | Low | 2 | 0.39 | 2.05 | 1085 |
| TCGA-LN-A7HV-01 | 58 | Male | Stage II | 10.51 | 0 | −63.02904193 | Low | 2 | 0.47 | 3.47 | 93 |
| TCGA-R6-A8W8-01 | 72 | Male | | 2.89 | 1 | −64.20673312 | Low | 2 | 0.63 | 2.75 | 110 |
| TCGA-L5-A8NN-01 | 81 | Male | Stage III | 5.49 | 0 | −69.31018669 | Low | 2 | 0.31 | 5.21 | 98 |
| TCGA-LN-A4A3-01 | 61 | Male | Stage III | 12.75 | 0 | −70.93140118 | Low | 2 | 0.31 | 3.97 | 64 |
| TCGA-LN-A4MQ-01 | 46 | Male | Stage III | 12.32 | 0 | −74.28300264 | Low | 2 | 0.65 | 8.93 | 73 |
| TCGA-L5-A8NI-01 | 79 | Male | Stage III | 13.47 | 1 | −76.40098706 | Low | 2 | 0.34 | 5.4 | 167 |
| TCGA-L5-A8NJ-01 | 77 | Male | Stage III | 16.46 | 0 | −81.51012839 | Low | 2 | 0.75 | 2.83 | 194 |
| TCGA-LN-A5U5-01 | 57 | Male | Stage IV | 4.47 | 1 | −84.47230082 | Low | 2 | 0.53 | 4.18 | 45 |
| TCGA-IG-A5S3-01 | 69 | Female | Stage II | 23.39 | 0 | −87.4784893 | Low | 2 | 0.73 | 9.77 | 62 |
| TCGA-VR-A8EX-01 | 63 | Male | Stage IV | 28.09 | 1 | −87.52962306 | Low | 2 | 0.4 | 3.52 | 115 |
| TCGA-LN-A49Y-01 | 77 | Male | Stage II | 12.45 | 0 | −89.85395773 | Low | 2 | 0.88 | 3.84 | 199 |

Table S4 (*continued*)

| | Age | Sex | Grade | OS (months) | OS status | Score | Group | TME cluster | Purity | Ploidy | TMB |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TCGA-R6-A6DQ-01 | 74 | Female | | 7.59 | 1 | –114.0412025 | Low | 2 | 0.46 | 3.25 | 69 |
| TCGA-2H-A9GQ-01 | 80 | Male | Stage III | 4.2 | 1 | 13.06857923 | Low | 1 | 0.75 | 2.84 | 231 |
| TCGA-L5-A8NF-01 | 57 | Male | Stage IV | 2.66 | 1 | 12.69563845 | Low | 1 | 0.46 | 2.77 | 110 |
| TCGA-LN-A49P-01 | 71 | Male | Stage II | 12.32 | 0 | 12.53733958 | Low | 1 | 0.49 | 4.39 | 63 |
| TCGA-2H-A9GM-01 | 53 | Male | Stage II | 13.93 | 1 | 9.873298261 | Low | 1 | 0.51 | 3.96 | 91 |
| TCGA-VR-AA4G-01 | 51 | Female | Stage III | 18.04 | 0 | 8.866454168 | Low | 1 | 0.33 | 3.94 | 62 |
| TCGA-LN-A49O-01 | 47 | Male | Stage II | 13.4 | 0 | 5.367156069 | Low | 1 | 0.51 | 3.16 | 54 |
| TCGA-L5-A43J-01 | 90 | Male | Stage II | 4.3 | 1 | 3.731963763 | Low | 1 | 0.57 | 3.52 | 512 |
| TCGA-S8-A6BV-01 | 76 | Male | Stage III | 20.01 | 0 | 2.090819664 | Low | 1 | 0.45 | 2.99 | 112 |
| TCGA-IG-A4QS-01 | 71 | Male | Stage III | 3.88 | 1 | –1.83201075 | Low | 1 | 0.53 | 5.8 | 198 |
| TCGA-L5-A4OW-01 | 56 | Female | Stage II | 7.13 | 1 | –2.787619378 | Low | 1 | 0.52 | 4.12 | 240 |
| TCGA-R6-A8WG-01 | 60 | Male | | 12.68 | 1 | –3.098706595 | Low | 1 | 0.82 | 3.59 | 93 |
| TCGA-LN-A49M-01 | 62 | Male | Stage II | 12.65 | 0 | –6.422834234 | Low | 1 | 0.76 | 1.75 | 192 |
| TCGA-LN-A4A4-01 | 36 | Male | Stage III | 12.58 | 0 | –6.813949388 | Low | 1 | 0.62 | 2.53 | 82 |
| TCGA-L5-A88Y-01 | 76 | Male | | 0.36 | 0 | –8.955868983 | Low | 1 | 0.33 | 4.64 | 197 |
| TCGA-Z6-A8JE-01 | 57 | Male | Stage III | 2.1 | 0 | –10.84671593 | Low | 1 | 0.62 | 9.14 | 180 |
| TCGA-RE-A7BO-01 | 72 | Female | Stage II | 7 | 1 | –12.49851705 | Low | 1 | 0.5 | 3 | 185 |
| TCGA-R6-A8WC-01 | 56 | Male | | 2.3 | 0 | –13.70848969 | Low | 1 | 0.84 | 2.59 | 109 |
| TCGA-L5-A88W-01 | 67 | Male | Stage II | 25.1 | 1 | –14.06053101 | Low | 1 | 0.69 | 2 | 124 |
| TCGA-V5-A7RC-01 | 55 | Male | | 3.42 | 1 | –18.72974578 | Low | 1 | 0.57 | 3.74 | 75 |
| TCGA-VR-A8EO-01 | 49 | Male | Stage II | 25.79 | 0 | –23.7775511 | Low | 1 | 0.78 | 3.87 | 44 |
| TCGA-L5-A4ON-01 | 65 | Male | Stage II | 18.33 | 1 | –26.34589442 | Low | 1 | 0.3 | 5.53 | 135 |
| TCGA-LN-A8I0-01 | 52 | Male | Stage II | 13.37 | 0 | –29.53674198 | Low | 1 | 0.35 | 4.59 | 43 |
| TCGA-R6-A6Y0-01 | 54 | Male | | 53.91 | 0 | –32.32474332 | Low | 1 | 0.2 | 4.78 | 151 |
| TCGA-LN-A5U7-01 | 46 | Male | Stage II | 25.23 | 0 | –35.06070703 | Low | 1 | 0.73 | 3.15 | 102 |
| TCGA-V5-A7RE-01 | 45 | Male | Stage I | 16.43 | 0 | –37.76349412 | Low | 1 | 0.28 | 3.59 | 139 |
| TCGA-VR-A8ET-01 | 64 | Male | Stage II | 1.54 | 1 | –39.55431646 | Low | 1 | 0.52 | 4.1 | 22 |
| TCGA-IG-A3YA-01 | 53 | Male | Stage III | 20.76 | 0 | –42.55182807 | Low | 1 | 0.27 | 4.11 | 80 |
| TCGA-LN-A9FO-01 | 42 | Male | Stage II | 0.13 | 0 | –44.22696249 | Low | 1 | 0.57 | 3.37 | 92 |
| TCGA-IG-A3YC-01 | 62 | Male | Stage III | 20.11 | 0 | –45.12874276 | Low | 1 | 0.22 | 3.16 | 48 |
| TCGA-IG-A97H-01 | 36 | Male | Stage II | 14.49 | 0 | –46.98362277 | Low | 1 | 0.25 | 6.39 | 87 |
| TCGA-LN-A9FQ-01 | 62 | Male | Stage II | 12.84 | 0 | –47.01459067 | Low | 1 | 0.53 | 7.97 | 86 |
| TCGA-LN-A4A9-01 | 58 | Male | Stage II | 11.53 | 1 | –47.05091624 | Low | 1 | 0.68 | 2.39 | 182 |
| TCGA-LN-A5U6-01 | 54 | Male | Stage II | 12.32 | 0 | –47.24856243 | Low | 1 | 0.63 | 8.61 | 92 |

Table S4 (*continued*)

| | Age | Sex | Grade | OS (months) | OS status | Score | Group | TME cluster | Purity | Ploidy | TMB |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TCGA-XP-A8T8-01 | 49 | Male | Stage II | 14.36 | 0 | −48.16359601 | Low | 1 | 0.8 | 3.87 | 87 |
| TCGA-IG-A3I8-01 | 51 | Female | Stage II | 33.25 | 0 | −48.2079175 | Low | 1 | 0.6 | 3.79 | 95 |
| TCGA-VR-A8EU-01 | 51 | Male | Stage IV | 18.3 | 1 | −48.50146976 | Low | 1 | 0.4 | 7.3 | 82 |
| TCGA-R6-A6L4-01 | 27 | Male | | 16.29 | 1 | −49.71571801 | Low | 1 | 0.44 | 5.03 | 118 |
| TCGA-VR-A8EQ-01 | 73 | Male | Stage III | 22.8 | 1 | −51.52771905 | Low | 1 | 0.47 | 5.08 | 146 |
| TCGA-L5-A8NR-01 | 81 | Female | Stage III | 8.71 | 0 | −52.01169279 | Low | 1 | 0.45 | 5.74 | 270 |
| TCGA-2H-A9GF-01 | 67 | Male | Stage III | 25.76 | 1 | −53.23604225 | Low | 1 | 0.32 | 7.29 | 228 |
| TCGA-LN-A49X-01 | 44 | Male | Stage II | 12.61 | 0 | −54.12930235 | Low | 1 | 0.4 | 4.3 | 60 |
| TCGA-L5-A4OM-01 | 54 | Female | Stage I | 47.9 | 1 | −55.03814574 | Low | 1 | 0.48 | 4.35 | 61 |
| TCGA-LN-A7HW-01 | 59 | Male | Stage II | 11.99 | 0 | −57.65183778 | Low | 1 | 0.56 | 3.2 | 86 |
| TCGA-R6-A6XG-01 | 64 | Male | | 38.37 | 0 | −58.51737922 | Low | 1 | 0.47 | 5.12 | 161 |
| TCGA-IG-A6QS-01 | 54 | Male | Stage II | 9.95 | 1 | −59.31469209 | Low | 1 | 0.7 | 1.95 | 60 |
| TCGA-L5-A8NK-01 | 84 | Female | Stage II | 13.53 | 0 | −59.54210181 | Low | 1 | 0.27 | 3.7 | 172 |
| TCGA-LN-A9FR-01 | 70 | Male | Stage II | 12.25 | 0 | −60.50119561 | Low | 1 | 0.65 | 8.88 | 64 |
| TCGA-L5-A8NQ-01 | 71 | Male | Stage II | 21.35 | 1 | −61.21314301 | Low | 1 | 0.51 | 4.25 | 216 |
| TCGA-LN-A4A5-01 | 49 | Male | Stage II | 22.37 | 1 | −67.35119522 | Low | 1 | 0.44 | 3.83 | 63 |
| TCGA-IG-A50L-01 | 58 | Male | Stage III | 0.53 | 0 | −67.3520416 | Low | 1 | 0.5 | 4.27 | 102 |
| TCGA-R6-A6DN-01 | 58 | Male | | 7.98 | 1 | −68.73193408 | Low | 1 | 0.45 | 5.6 | 80 |
| TCGA-LN-A4A1-01 | 60 | Male | Stage II | 12.58 | 0 | −70.14206226 | Low | 1 | 0.64 | 4.15 | 104 |
| TCGA-LN-A7HX-01 | 72 | Male | Stage II | 12.22 | 0 | −72.31835444 | Low | 1 | 0.58 | 3.88 | 88 |
| TCGA-VR-A8ER-01 | 54 | Male | Stage III | 12.42 | 1 | −75.7553627 | Low | 1 | 0.88 | 3.05 | 52 |
| TCGA-L5-A8NL-01 | 56 | Male | Stage III | 13.21 | 0 | −78.04101838 | Low | 1 | 0.22 | 4.33 | 84 |
| TCGA-2H-A9GN-01 | 70 | Male | Stage III | 8.94 | 1 | −80.40115402 | Low | 1 | 0.23 | 4.63 | 97 |
| TCGA-VR-A8EW-01 | 57 | Male | Stage III | 8.11 | 1 | −81.96905515 | Low | 1 | 0.39 | 5.24 | 101 |

TME, tumor microenvironment; EC, esophageal carcinoma.

**Table S5** The recurrent occurring CNV regions

| Group | Cytoband | q value |
|---|---|---|
| TMEscore high | 11q13.3 (amp) | 3.65E-08 |
| | 8q24.21 (amp) | 1.16E-06 |
| | 6p21.1 (amp) | 8.24E-05 |
| | 9p21.3 (del) | 1.54E-19 |
| | 16q23.1 (del) | 1.49E-11 |
| | 2q22.1 (del) | 1.75E-07 |
| | 5q11.2 (del) | 1.50E-06 |
| TMEscore low | 11q13.3 (amp) | 9.69E-74 |
| | 17q12 (amp) | 4.45E-19 |
| | 8q24.21 (amp) | 1.54E-16 |
| | 7q21.2 (amp) | 5.56E-13 |
| | 3q26.2 (amp) | 5.17E-12 |
| | 6p21.1 (amp) | 6.06E-10 |
| | 7p11.2 (amp) | 3.46E-08 |
| | 12p12.1 (amp) | 4.30E-08 |
| | 18q11.2 (amp) | 5.82E-07 |
| | 19q12 (amp) | 6.65E-07 |
| | 9p21.3 (del) | 6.86E-95 |
| | 16q23.1 (del) | 3.06E-54 |
| | 4q22.1 (del) | 1.09E-31 |
| | 5q12.1 (del) | 4.10E-18 |
| | 7q31.1 (del) | 1.55E-19 |
| | 2q22.1 (del) | 1.06E-15 |
| | 6p25.3 (del) | 6.02E-13 |
| | 3p14.2 (del) | 6.02E-13 |
| | 8p23.2 (del) | 4.74E-07 |
| | 20p12.1 (del) | 1.47E-05 |
| | 21q22.12 (del) | 1.18E-05 |
| | 18q21.2 (del) | 1.01E-05 |
| | 10p11.21 (del) | 1.98E-05 |
| | 3p14.3 (del) | 2.23E-05 |
| | 7q36.3 (del) | 1.98E-05 |
| | 7q36.3 (del) | 1.98E-05 |
| | 9p23 (del) | 2.80E-05 |

CNV, copy number variation.

**Table S6** Purity and ploidy of patients with EC

| Sample | Purity | Ploidy |
|---|---|---|
| TCGA-LN-A49M-01 | 0.76 | 1.75 |
| TCGA-L5-A4OP-01 | 0.69 | 1.81 |
| TCGA-L5-A4OJ-01 | 0.86 | 1.84 |
| TCGA-IG-A6QS-01 | 0.7 | 1.95 |
| TCGA-L5-A88Z-01 | 0.8 | 1.97 |
| TCGA-KH-A6WC-01 | 0.61 | 2 |
| TCGA-IG-A97I-01 | 0.65 | 2 |
| TCGA-L5-A88W-01 | 0.69 | 2 |
| TCGA-LN-A4A8-01 | 0.71 | 2.03 |
| TCGA-L5-A8NM-01 | 0.39 | 2.05 |
| TCGA-JY-A6FA-01 | 0.72 | 2.09 |
| TCGA-2H-A9GJ-01 | 0.7 | 2.1 |
| TCGA-IC-A6RF-01 | 0.55 | 2.14 |
| TCGA-JY-A93D-01 | 0.25 | 2.25 |
| TCGA-V5-AASV-01 | 0.75 | 2.26 |
| TCGA-VR-A8EY-01 | 0.74 | 2.35 |
| TCGA-LN-A4A9-01 | 0.68 | 2.39 |
| TCGA-LN-A4A4-01 | 0.62 | 2.53 |
| TCGA-R6-A8WC-01 | 0.84 | 2.59 |
| TCGA-2H-A9GI-01 | 0.55 | 2.66 |
| TCGA-L5-A4OS-01 | 0.19 | 2.7 |
| TCGA-R6-A8W8-01 | 0.63 | 2.75 |
| TCGA-L5-A8NF-01 | 0.46 | 2.77 |
| TCGA-M9-A5M8-01 | 0.93 | 2.79 |
| TCGA-L5-A8NJ-01 | 0.75 | 2.83 |
| TCGA-2H-A9GQ-01 | 0.75 | 2.84 |
| TCGA-VR-A8Q7-01 | 0.29 | 2.89 |
| TCGA-S8-A6BV-01 | 0.45 | 2.99 |
| TCGA-RE-A7BO-01 | 0.5 | 3 |
| TCGA-VR-A8ER-01 | 0.88 | 3.05 |
| TCGA-IG-A3QL-01 | 0.64 | 3.07 |
| TCGA-VR-A8EP-01 | 0.42 | 3.08 |
| TCGA-L5-A4OH-01 | 0.62 | 3.08 |
| TCGA-L5-A8NV-01 | 0.38 | 3.14 |

**Table S6** (*continued*)

**Table S6** (*continued*)

| Sample | Purity | Ploidy |
|---|---|---|
| TCGA-LN-A5U7-01 | 0.73 | 3.15 |
| TCGA-IG-A3YC-01 | 0.22 | 3.16 |
| TCGA-LN-A49O-01 | 0.51 | 3.16 |
| TCGA-JY-A93C-01 | 0.55 | 3.16 |
| TCGA-LN-A7HW-01 | 0.56 | 3.2 |
| TCGA-IG-A8O2-01 | 0.61 | 3.2 |
| TCGA-VR-AA7I-01 | 0.41 | 3.21 |
| TCGA-R6-A6DQ-01 | 0.46 | 3.25 |
| TCGA-L5-A88V-01 | 0.72 | 3.28 |
| TCGA-S8-A6BW-01 | 0.56 | 3.29 |
| TCGA-L5-A43E-01 | 0.51 | 3.36 |
| TCGA-LN-A9FO-01 | 0.57 | 3.37 |
| TCGA-L7-A56G-01 | 0.55 | 3.38 |
| TCGA-VR-AA4D-01 | 0.62 | 3.41 |
| TCGA-L5-A8NT-01 | 0.39 | 3.45 |
| TCGA-LN-A8I1-01 | 0.63 | 3.46 |
| TCGA-LN-A7HV-01 | 0.47 | 3.47 |
| TCGA-L5-A8NS-01 | 0.39 | 3.52 |
| TCGA-VR-A8EX-01 | 0.4 | 3.52 |
| TCGA-L5-A43J-01 | 0.57 | 3.52 |
| TCGA-V5-A7RE-01 | 0.28 | 3.59 |
| TCGA-R6-A8WG-01 | 0.82 | 3.59 |
| TCGA-L5-A43C-01 | 0.62 | 3.61 |
| TCGA-2H-A9GK-01 | 0.29 | 3.62 |
| TCGA-V5-A7RC-06 | 0.47 | 3.67 |
| TCGA-L5-A8NK-01 | 0.27 | 3.7 |
| TCGA-V5-A7RC-01 | 0.57 | 3.74 |
| TCGA-JY-A93E-01 | 0.46 | 3.75 |
| TCGA-L5-A8NU-01 | 0.38 | 3.76 |
| TCGA-IG-A3I8-01 | 0.6 | 3.79 |
| TCGA-XP-A8T6-01 | 0.58 | 3.8 |
| TCGA-Z6-AAPN-01 | 0.41 | 3.81 |
| TCGA-LN-A4A5-01 | 0.44 | 3.83 |
| TCGA-V5-AASX-01 | 0.61 | 3.84 |
| TCGA-LN-A49Y-01 | 0.88 | 3.84 |

**Table S6** (*continued*)

| Sample | Purity | Ploidy |
|---|---|---|
| TCGA-L5-A4OG-01 | 0.5 | 3.85 |
| TCGA-JY-A938-01 | 0.7 | 3.85 |
| TCGA-2H-A9GR-01 | 0.2 | 3.87 |
| TCGA-L5-A4OX-01 | 0.65 | 3.87 |
| TCGA-VR-A8EO-01 | 0.78 | 3.87 |
| TCGA-XP-A8T8-01 | 0.8 | 3.87 |
| TCGA-LN-A49W-01 | 0.43 | 3.88 |
| TCGA-LN-A7HX-01 | 0.58 | 3.88 |
| TCGA-L5-A88S-01 | 0.82 | 3.9 |
| TCGA-IG-A5B8-01 | 0.38 | 3.91 |
| TCGA-L5-A4OT-01 | 0.49 | 3.92 |
| TCGA-JY-A6FD-01 | 0.65 | 3.93 |
| TCGA-VR-AA4G-01 | 0.33 | 3.94 |
| TCGA-JY-A6FB-01 | 0.54 | 3.94 |
| TCGA-L5-A8NH-01 | 0.27 | 3.96 |
| TCGA-2H-A9GM-01 | 0.51 | 3.96 |
| TCGA-LN-A4A3-01 | 0.31 | 3.97 |
| TCGA-L5-A88T-01 | 0.45 | 4.01 |
| TCGA-IG-A3YB-01 | 0.61 | 4.02 |
| TCGA-LN-A9FP-01 | 1 | 4.02 |
| TCGA-IG-A4P3-01 | 0.34 | 4.04 |
| TCGA-JY-A939-01 | 0.34 | 4.05 |
| TCGA-JY-A93F-01 | 0.55 | 4.09 |
| TCGA-VR-A8ET-01 | 0.52 | 4.1 |
| TCGA-Z6-A8JD-01 | 0.25 | 4.11 |
| TCGA-IG-A3YA-01 | 0.27 | 4.11 |
| TCGA-JY-A6FG-01 | 0.57 | 4.11 |
| TCGA-L5-A4OW-01 | 0.52 | 4.12 |
| TCGA-L5-A4OU-01 | 0.42 | 4.14 |
| TCGA-R6-A8W5-01 | 0.2 | 4.15 |
| TCGA-LN-A4A1-01 | 0.64 | 4.15 |
| TCGA-IC-A6RE-01 | 0.2 | 4.17 |
| TCGA-V5-A7RB-01 | 0.44 | 4.17 |
| TCGA-LN-A49U-01 | 0.57 | 4.17 |
| TCGA-LN-A5U5-01 | 0.53 | 4.18 |

**Table S6** (*continued*)

| Sample | Purity | Ploidy |
|---|---|---|
| TCGA-L5-A893-01 | 0.37 | 4.2 |
| TCGA-Q9-A6FW-01 | 0.5 | 4.21 |
| TCGA-L5-A8NQ-01 | 0.51 | 4.25 |
| TCGA-IG-A51D-01 | 0.5 | 4.27 |
| TCGA-IG-A50L-01 | 0.5 | 4.27 |
| TCGA-L5-A4OO-01 | 0.33 | 4.28 |
| TCGA-LN-A49X-01 | 0.4 | 4.3 |
| TCGA-L5-A8NL-01 | 0.22 | 4.33 |
| TCGA-L5-A4OI-01 | 0.66 | 4.33 |
| TCGA-L5-A4OM-01 | 0.48 | 4.35 |
| TCGA-LN-A49P-01 | 0.49 | 4.39 |
| TCGA-IG-A625-01 | 0.66 | 4.39 |
| TCGA-L5-A8NW-01 | 0.4 | 4.4 |
| TCGA-LN-A8I0-01 | 0.35 | 4.59 |
| TCGA-L5-A4OE-01 | 0.49 | 4.6 |
| TCGA-2H-A9GN-01 | 0.23 | 4.63 |
| TCGA-L5-A88Y-01 | 0.33 | 4.64 |
| TCGA-R6-A6Y0-01 | 0.2 | 4.78 |
| TCGA-2H-A9GL-01 | 0.4 | 4.8 |
| TCGA-L5-A8NE-01 | 0.28 | 4.91 |
| TCGA-R6-A6L4-01 | 0.44 | 5.03 |
| TCGA-VR-A8EQ-01 | 0.47 | 5.08 |
| TCGA-R6-A6XG-01 | 0.47 | 5.12 |
| TCGA-R6-A6KZ-01 | 0.35 | 5.13 |
| TCGA-L5-A8NN-01 | 0.31 | 5.21 |
| TCGA-VR-A8EW-01 | 0.39 | 5.24 |
| TCGA-JY-A6FE-01 | 0.38 | 5.28 |
| TCGA-L5-A8NI-01 | 0.34 | 5.4 |
| TCGA-L5-A891-01 | 0.45 | 5.52 |
| TCGA-JY-A6FH-01 | 0.22 | 5.53 |
| TCGA-L5-A4ON-01 | 0.3 | 5.53 |
| TCGA-R6-A6DN-01 | 0.45 | 5.6 |
| TCGA-L5-A8NR-01 | 0.45 | 5.74 |
| TCGA-IG-A4QS-01 | 0.53 | 5.8 |
| TCGA-V5-AASW-01 | 0.31 | 6.12 |

**Table S6** (*continued*)

| Sample | Purity | Ploidy |
|---|---|---|
| TCGA-VR-A8EZ-01 | 0.41 | 6.18 |
| TCGA-IG-A97H-01 | 0.25 | 6.39 |
| TCGA-2H-A9GF-01 | 0.32 | 7.29 |
| TCGA-VR-A8EU-01 | 0.4 | 7.3 |
| TCGA-LN-A7HY-01 | 0.35 | 7.35 |
| TCGA-Z6-A9VB-01 | 0.34 | 7.61 |
| TCGA-L5-A8NG-01 | 0.3 | 7.86 |
| TCGA-LN-A9FQ-01 | 0.53 | 7.97 |
| TCGA-L7-A6VZ-01 | 0.53 | 8.16 |
| TCGA-LN-A5U6-01 | 0.63 | 8.61 |
| TCGA-2H-A9GO-01 | 0.43 | 8.63 |
| TCGA-2H-A9GH-01 | 1 | 8.76 |
| TCGA-LN-A9FR-01 | 0.65 | 8.88 |
| TCGA-LN-A4MQ-01 | 0.65 | 8.93 |
| TCGA-LN-A49S-01 | 0.48 | 8.95 |
| TCGA-Z6-A8JE-01 | 0.62 | 9.14 |
| TCGA-IG-A5S3-01 | 0.73 | 9.77 |
| TCGA-R6-A6XQ-01 | 0.8 | 9.95 |
| TCGA-LN-A7HZ-01 | 0.5 | 10.01 |
| TCGA-LN-A8HZ-01 | 0.5 | 10.12 |
| TCGA-IG-A7DP-01 | NA | NA |

EC, esophageal carcinoma.