Introduction

Esophageal cancer is one of the most common malignant tumors in the world (1). The two main histological types of esophageal cancer are squamous cell carcinoma and adenocarcinoma, which have different clinical features and prognoses (2). Though multidisciplinary approaches are widely applied in esophageal cancer, the survival rates are dismal for advanced esophageal cancer (3,4).

The tumor microenvironment (TME), including immune cells and stromal cells, is associated with the disease progression, treatment response or prognosis of esophageal cancer (5,6). A comprehensive analysis of TME in pretherapeutic biopsy samples has shown that tumor-infiltrating M2 macrophage predicted poor pathological response to neoadjuvant chemotherapy and unfavorable overall survival (OS) (7). In another report, cancer-associated fibroblasts in clinical samples were found to promote lymph node
metastasis in esophageal cancers, which was confirmed in vitro and in vivo (8). Additionally, increases in interferon γ, activated lymphocytes and programmed death ligand-1 were shown to reshape the immune microenvironment after chemoradiation on esophageal cancers (9).

Estimation of STromal and Immune cells in MA lignant Tumors using Expression data (ESTIMATE) is a method to evaluate tumor purity by estimating stromal and immune scores using gene expression signatures (10). The ESTIMATE algorithm has been applied to analyzing gene expression profiles of bladder cancer, glioblastoma and pancreatic cancer from The Cancer Genome Atlas (TCGA) database (11-13).

In this study, we applied the ESTIMATE algorithm to the gene expression profiles of esophageal cancer from TCGA database, and found some microenvironment-associated functions or pathways. More importantly, we identified some TME-associated genes that correlated with poor esophageal cancer prognosis. We present the following article in accordance with the MDAR checklist (available at http://dx.doi.org/10.21037/tcr-20-2288).

Methods

Database

The gene expression profiles and clinical data of esophageal cancer were extracted from TCGA database (https://portal.gdc.cancer.gov/, February 16, 2020). Immune scores and stromal scores were calculated based on ESTIMATE algorithm using the “estimate” package (https://bioinformatics.mdanderson.org/estimate/rpackage.html) in R (R version 3.6.2, https://www.r-project.org/). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). For this type of study, ethical approval and formal consent are not required.

Differentially expressed genes (DEGs)

To obtain DEGs, the “limma” package (https://bioconductor.org/packages/l limma/) was used (fold change =2; adjusted P value <0.05). A heatmap of DEGs was created using the “pheatmap” package (https://CRAN.R-project.org/package=pheatmap). Shared DEGs were detected using the “VennDiagram” package (https://CRAN.R-project.org/package=VennDiagram).

Gene enrichment and protein interactions analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using the “clusterProfiler” package (https://bioconductor.org/packages /clusterProfiler/). Protein-protein interaction (PPI) analysis was performed in the online STRING database (https://string-db.org/).

Survival analysis

Survival analysis was performed using the “survival” package (https://CRAN.R-project.org/package=survival).

Statistical analysis

The patients’ characteristics were described as n (%) for categorical variables and median with range for continuous variables. Comparisons of scores and comparisons of gene expressions in different groups were performed using non-parametric test. OS was defined as the time from the diagnosis of esophageal cancer to the death or the end of the study. OS was calculated using the Kaplan-Meier method and comparison of survival was conducted using the log-rank test. A P value <0.05 was considered statistically significant.

Results

Patients’ characteristics

The gene expression profiles and clinical data of 158 patients with esophageal cancer from the TCGA database were used for analysis. The patients’ characteristics are shown in Table 1. The median age was 60 years (range, 27–90 years), and 135 (85.4%) patients were male. There were 17 (10.8%), 72 (45.6%), 55 (34.8%) and 14 (8.9%) patients in stages I, II, III and IV, respectively. Seventy-eight (49.4%) patients were diagnosed with esophageal adenocarcinoma, and 80 (50.6%) were diagnosed with esophageal squamous cell carcinoma.

Immune scores and stromal scores

There was no significant difference in immune scores among the four disease stages (Figure 1A, P=0.695). Patients in stage III had the highest stromal scores (Figure 1B, P=0.015). Between patients with adenocarcinoma and squamous cell carcinoma, no significant difference was observed in immune scores (Figure 1C, P=0.161) or in stromal scores (Figure 1D, P=0.192). All patients were divided into high and low scores group based on the
estimate scores with the median value as the cut-off. The median OS time was not significantly different between patients with high and low immune scores (Figure 2A, \(P=0.434\)), and also not different between patients with high and low stromal scores (Figure 2B, \(P=0.655\)).

**DEGs, gene enrichment and protein interactions analysis**

Compared with the low immune score group, 958 up-regulated genes and 62 down-regulated genes were observed in the high immune scores group. A heatmap of DEGs is shown in Figure 3A. Further, compared with the low stromal score group, 1,165 up-regulated genes and 55 down-regulated genes were observed in the high stromal score group. A heatmap of the DEGs is shown in Figure 3B. In the high immune and stromal score groups, 603 shared up-regulated genes were found (Figure 4A), and no shared down-regulated genes were found.

The function and pathway enrichment of the 603 shared genes were analyzed. The 10 most significant GO terms in biological process (BP), cellular component (CC) and molecular function (MF), including regulation of lymphocyte activation, external side of plasma membrane,
cytokine binding, etc. are shown in Figure 4B. The 20 most significant KEGG pathways, including viral protein interaction with cytokine and cytokine receptor, hematopoietic cell lineage, chemokine signaling pathway, etc. are shown in Figure 4C.

PPI analysis was performed to elucidate the interactions of the shared DEGs. The PPI network consisted of 162 nodes and 276 edges (Figure 4D). The 20 genes with the most connections with other genes, such as ITGAM with 16 connections, CXCL10 with 15 connections and CCR2 with 14 connections, are shown in Figure 4E.

**DEG expression and OS**

All patients were subdivided into two groups according to the expression of each of the 603 shared DEGs. Patients with higher than average of expression level of gene were regarded as the high expression group, and the other patients were regarded as the low expression group. OS comparisons were performed between the two groups for each gene. A total of 11 genes were found to be associated with OS. High expression of these genes predicted shorter OS (Figure 5). The 11 prognostic genes included MS4A7, TMIGD3, MS4A4A, EVI2A, MS4A6A, FCER1G, AIF1, GNGT2, LCP2, DNAJC5B and RNASE6.

**Discussion**

M1-like macrophages and eosinophils play an important role in Barrett esophagus and chronic esophagitis, while M2-like macrophages and regulatory T cells are correlated with esophageal cancer development (14, 15). The expression of RALDH2 and FOXP3 were reported to be associated with the aberrant immune microenvironment of Barrett esophagus (16). However, the role of TME in esophageal cancer is unclear.

In this study, we analyzed the gene expression profiles and clinical information of 158 patients with esophageal cancer in the TCGA database. In regard to the association of the immune/stromal scores and disease stage, pathological type and OS, only the stromal scores among the different stages were significantly different. A total of 603 shared up-regulated DEGs in the high immune and stromal score groups were found. The function and pathway enrichment, and PPI of the genes were analyzed. Moreover, high expression levels of 11 genes were found to be associated with the short survival time of esophageal cancer patients.

The stromal scores differed among the various stages, suggesting that TME plays a part in esophageal cancer development. Despite the different pathogenesis and prognosis in esophageal adenocarcinoma and esophageal squamous cell carcinoma, neither immune scores nor stromal scores were significantly different between the two pathological types in this study. The results might be explained by the similar TME in esophageal adenocarcinoma and esophageal squamous cell carcinoma (5). Besides, estimate scores were not found to be correlated with the survival of esophageal cancer patients in this study, though they have been reported to be prognostic in a few
Figure 3 Heatmaps of differentially expressed genes in different immune/stromal scores. (A) High immune scores vs. low immune scores; (B) high stromal scores vs. low stromal scores (fold change =2; adjust P value <0.05).
malignant tumors (12,13). Thus, further studies of immune and stromal components in TME of esophageal cancer are needed.

In the high immune and stromal scores groups, 603 shared regulated genes were identified. The shared genes underwent GO and KEGG analyses. All the GO terms in BP were associated with immune functions, involving leukocyte activation, adhesion and proliferation. Most of the other GO terms were also related to the immune system, such as chemokine activity, chemokine receptor binding, cytokine binding and cytokine receptor activity. In addition, most KEGG pathways were also associated with the immune system, such as chemokine signaling pathway, cytokine-cytokine receptor interaction, cell adhesion molecules and B cell receptor signaling pathway.

A PPI network was created to show the interactions of the shared DEGs. The genes with at least eight connections to other genes are shown in Figure 4E. ITGAM gene had the most connections, followed by CXCL10 and CCR2. ITGAM, also known as CD11B, was a risk factor of systemic lupus erythematosus (17), but its role in cancer is unclear. In patients with esophageal cancer, the lower levels of CXCL10 in the tumor tissue and serum were reported to be associated with favorable survival and better histopathological response, respectively (18). Further, CCL2-CCR2 axis was reported to increase tumor associated macrophages in esophageal carcinogenesis, and predicted adverse outcome (19).

Of the 603 shared upregulated genes, 11 were found to be correlated with the OS time of patients with esophageal cancer. We found that high expression of these genes predicted poor prognosis. In the prognostic genes, high expression of MS4A4E and MS4A4E was associated with poor prognoses in gastric cancer patients (20,21). TMIGD3 isoform 1 could suppress the malignant progression of osteosarcoma cells (22), whereas high expression of EVI2A.
Figure 5 Overall survival of esophageal cancer patients in different expression levels of each gene in differentially expressed genes. (A) MS4A7; (B) TMIGD3; (C) MS4A4A; (D) EVI2A; (E) MS4A6A; (F) FCER1G; (G) AIF1; (H) GNGT2; (I) LCP2; (J) DNAJC5B; (K) RNASE6.

indicated worse OS in osteosarcoma (23). MS4A6A was related to the Alzheimer’s disease risk and pathology (24). FCER1G was involved in the progression of lung adenocarcinoma and clear cell renal cell carcinoma (25,26). AIF1 was upregulated in M2-like macrophages and associated with the advanced stages and poor survival of hepatocellular carcinoma (27). GNGT2 was also detected as a prognostic marker in esophageal cancer in another report (28). LCP2, an inflammatory factor, was revealed as the potential prognostic factor in colorectal cancer and diffuse large B-cell lymphoma (29,30). In contrast to our results, the high expression levels of DNAJC5B and RNASE6 were reported to be associated with better survival in esophageal squamous cell carcinoma (31,32). In brief, the prognostic genes in our paper have not yet been thoroughly studied in previous literatures. More researches are required to elucidate and confirm the role of these genes in patients with esophageal cancer. In conclusion, microenvironment-associated functions and pathways were analyzed in patients with esophageal cancer from the TCGA database. Moreover, we found 11 microenvironment-associated genes that were correlated to poor outcomes of esophageal cancer patients. These genes might be used as prognostic biomarkers for esophageal cancer. Further studies on these genes may be helpful to understand the TME and provide new therapies for esophageal cancer.

Acknowledgments

Funding: None.
Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at http://dx.doi.org/10.21037/tcr-20-2288

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr-20-2288). The authors have no 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). For this type of study, ethical approval and formal consent are not required.

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


Cite this article as: Zhou MH, Wang XK. Microenvironment-related prognostic genes in esophageal cancer. Transl Cancer Res 2020;9(12):7531-7539. doi: 10.21037/tcr-20-2288