Original Research

Bioinformatics Analysis Reveals Prognostic Significance of the Macrophage Marker Gene Signature in Gastric Adenocarcinoma

Zhipeng Li¹,²,³,†, Hui Chen¹,²,³, Zhongqing Chen¹,²,³, Lihe Xie¹,²,³, Dun Pan¹,²,³,*

¹Department of Gastrointestinal Surgery, The First Affiliated Hospital, Fujian Medical University, 350000 Fuzhou, Fujian, China
²Department of Gastrointestinal Surgery, National Regional Medical Center, Binhai Campus of the First Affiliated Hospital, Fujian Medical University, 350212 Fuzhou, Fujian, China
³Fujian Research Institute of Abdominal Surgery, The First Affiliated Hospital, Fujian Medical University, 350000 Fuzhou, Fujian, China

*Correspondence: pandun0229@fjmu.edu.cn (Dun Pan)
†These authors contributed equally.

Abstract

Background: Gastric adenocarcinoma (GAC) is a malignant tumor with the highest incidence in the digestive system. Macrophages have been proven to play important roles in tumor microenvironment. Methods: Herein, single-cell RNA sequencing (scRNA-seq) profiles from the Gene Expression Omnibus (GEO) and bulk RNA-seq data from the Cancer Genome Atlas (TCGA) database were utilized to construct a macrophage marker gene signature (MMGS) to predict the prognosis of GAC patients. Subsequently, a risk score model based on the MMGS was built to predict the prognosis of GAC patients; further, this was validated in the GEO cohort. The risk score categorized patients into the high- and low-risk groups. A nomogram model based on the risk score and clinic-pathological characteristics was developed. Results: Seven genes, ABCA1, CTHRC1, GADD45B, NPC2, PLTP, PRSS23, and RNASE1, were included in the risk score model. Patients with a low-risk score showed a better prognosis. The MMGS had good sensitivity and specificity for predicting the prognosis in GAC patients. The risk score was an independent prognostic factor. The constructed nomogram exhibited favorable predictability and reliability for predicting GAC prognosis. Conclusion: In conclusion, the risk score model based on the seven MMGSs performed well in the predicting prognosis of GAC patients. Our study may provide new insights into clinical decision-making for the personalized treatment of patients with gastric cancer (GC).

Keywords: gastric adenocarcinoma; single-cell RNA sequencing; macrophage marker gene signature; survival; prognosis

1. Introduction

Gastric cancer (GC) is a malignant tumor originating from the gastric mucosal epithelium, with more than half of the cases occurring in Eastern Asia [1]. Gastric adenocarcinoma (GAC) is a common malignancy of the digestive system, accounting for more than 40% of all cancer cases worldwide [2]. Although the treatment of GC has improved, its prognosis remains undesirable, and the 5-year relative survival rate of patients diagnosed at an advanced stage is less than 10% [3]. Increasing evidence has shown that the interaction between tumor and stromal cells and the tumor microenvironment (TME) evolving during disease progression have a bearing on the survival and treatment response of patients [4]. The wide heterogeneity of the TME creates difficulties in cancer management [5]. Immunotherapy, especially immune checkpoint inhibitors, prolonged their overall survival (OS) time, which provides the possibility of new treatment options [6].

As part of the immune system, macrophages play indispensable roles in resisting diseases and recruiting other immune cells to the scene [7]. Macrophage phagocytosis leads to tumor elimination, inflammatory body activation, and antigen presentation in tumors, which can induce adaptive immunity [8,9]. In addition to their anti-tumor effects, macrophages contribute to tumor metastasis and resistance to treatment [10]. Tumor-associated macrophages (TAMs) undergo phenotypic polarization in response to various microenvironment signals [11]. The functional plasticity of TAMs, that is, macrophage polarization (M1 and M2), can affect the therapeutic effect and prognosis of patients with cancer [12]. M1 macrophages mediate the inflammatory response against invading tumor cells, whereas M2 macrophages exhibit an anti-inflammatory role, favoring tumor progression [13]. Colorectal cancer-conditioned macrophages with a mixed M1/M2 phenotype induce epithelial-mesenchymal transition process to enhance migration and invasion [14]. High M2/M1 ratios are related to poor prognosis in oral squamous cell carcinoma patients [15]. Therefore, it is necessary to explore the role of immune cells (especially macrophages) in GC and their association with prognosis and treatment prediction.

Single-cell RNA sequencing (scRNA-seq) is an emerging technique providing an unprecedented opportunity to identify specific cell types and characterize diverse immune cell populations in the TME, thus, providing a new way to develop gene signatures to predict the prognosis of cancer patients [16,17]. The Cancer Genome Atlas (TCGA)
and Gene Expression Omnibus (GEO) databases provide massive amounts of publicly available transcriptome and scRNA-seq data. In the present study, we performed systematic bioinformatics analyses using scRNA-seq and bulk RNA-seq data to construct macrophage marker gene signatures (MMGSs) for GAC, with an external validation cohort to validate its prognostic value. Our study may advance our understanding of the mechanisms underlying GAC progression and offer insights into potential individualized GAC treatments.

2. Materials and Methods

2.1 Dataset and Data Preprocessing

The scRNA-seq data of 26 gastric tumor samples from GSE183904 downloaded from the GEO database were used for screening for macrophage marker genes. RNA sequencing and clinical data were acquired from UCSC Xena (https://xenabrowser.net/datapages/). To verify the prognostic value of the model, GSE84437 dataset was downloaded from the GEO database. The average value of multiple probes corresponding to the same gene was used as the gene expression level. Samples with OS < 60 days were excluded. Subsequently, 328 GAC samples from the TCGA were adopted as the training set, and 426 GC samples from GSE84437 were adopted as the validation set. The clinical-pathological features of the included cases are shown in the Supplementary Table 1.

2.2 Identification and Functional Analysis of Macrophage Marker Genes

ScRNA-seq data was analyzed using “Seurat” and “SingleR” packages, with 119,878 cells from 26 tumor samples included in this analysis. Quality control was performed as follows: (1) genes detected in less than three cells were excluded; (2) cells expressing less than 50 genes were excluded; (3) cells with mitochondrial gene expression ≥5% were excluded; (4) cells with ribosomal gene expression ≥5% were excluded; and (5) all cells with gene expression less than 10,000 were excluded. Gene expression of the remaining 15,116 cells was standardized using the “LogNormalize” method in the NormalizeData function. The first 2000 highly variable genes were obtained via the “vst” method in the FindVariableFeatures function. The first 2000 highly variable genes were screened above with a p-value of <0.05 and log2 fold change (FC) >1. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment and Gene Ontology (GO) analysis were performed using “ClusterProfiler” (with “simplify” function to remove the redundant GO terms), “GPlot”, “org.Hs.eg.db”, and “enrichplot” packages.

2.3 Prognostic Risk Model of Macrophage Marker Genes

Patients with more than 60 days of follow-up in the TCGA database were included to establish a risk score model. Macrophage marker gene expression data were combined with the OS time and status of each patient. To screen genes with the greatest effect on OS, the p-value of less than 0.01 was set in the univariate Cox analysis. Using the “glmnet” package in R, a macrophage marker gene prognostic risk model to predict the prognosis of patients with GC was constructed using the least absolute shrinkage and selection operator (LASSO) method. Based on the median score, the patients were divided into the high- and low-risk groups. The “Survival” and “survminer” packages in R were applied to compare the OS of high- and low-risk groups through the Kaplan-Meier analysis. The receiver operating characteristic (ROC) curves were used to verify the accuracy of the risk model for predicting patient survival with the “timeROC” package. In addition, we detected the expression of genes in the risk score model in each cell cluster and tumor tissues in the TCGA dataset. Correlation between genes and macrophages were evaluated using TIMER (https://cistrome.shinyapps.io/timer/).

2.4 Construction of Nomogram

Nomogram, a convenient statistical tool, was used to evaluate the clinical characteristics and risk scores of the patients. Nomogram was constructed using the cph and nomogram functions in the “rms” package of R software. A calibration curve was used to visualize the deviation between the predicted probability and actual occurrence.

2.5 Evaluation of Drug Therapy

To assess the relationship between MMGS and anti-tumor drug sensitivity, the “pRophetic” R package was applied to calculate the half-maximal inhibitory concentration (IC50) of the samples. The IC50 values were compared between the high- and low-risk groups using Wilcoxon signed-rank test.

2.6 Expression Validation of MMGS Using Real time-Polymerase Chain Reaction (RT-PCR)

The expression of MMGSs was validated in tumor and para-carcinoma tissues from 10 GC patients using RT-PCR. This study was approved by the Ethics Committee of the First Affiliated Hospital of Fujian Medical University (2020-219). Written informed consent was obtained from each patient. RNA was extracted using the TRIzol reagent (DP424, TIANGEN, Beijing, China) for RT-PCR. FastK- ing cDNA First Strand Synthesis Kit (KR116, TIANGEN,
Table 1. The primer sequences used in RT-PCR.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequences (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>CTGGGGCTACACTGAGCACC AAGTGGTCGTTGAGGGCAATG</td>
</tr>
<tr>
<td>APCA1</td>
<td>CTACATCTCCCTTCGCCAGACA CCGCAGACAATACGAGACAC</td>
</tr>
<tr>
<td>CTHRC1</td>
<td>CACACACCTCTGGTAAACACATCCA GCCCTTTGAACACATCCA</td>
</tr>
<tr>
<td>GADD45B</td>
<td>TGCTGTGACAACGACATCAAC GTGAGGGTCGCGACCGAG</td>
</tr>
<tr>
<td>NPC2</td>
<td>CAAAGGACAGTCTTTACAGCGT GGATAGGGCAGTTAATTCCACTC</td>
</tr>
<tr>
<td>PLTP</td>
<td>TGTTGTATGGCCCTTTTGGT AATTGGGAGTGAATTCATCTGCT</td>
</tr>
<tr>
<td>PRSS23</td>
<td>GGAACGCGCAATTTATGCTG TCCATCGTGATTCAGGTGAG</td>
</tr>
<tr>
<td>RNASE1</td>
<td>GAATCCCGGGCAAGAAATTC GTGTCTCTCCTCGGTGGT</td>
</tr>
</tbody>
</table>

RT-PCR, Real-time-Polymerase Chain Reaction.

Beijing, China) was used for mRNA reverse transcription. RT-qPCR was performed in an ABI 7300 Real-time PCR Detection System with SuperReal PreMix Plus. The reaction was set at 95 °C for 15 min for pre-denaturation, then at 95 °C for 10 s, at 55 °C for 30 s, and 72 °C for 32 s repeating 40 cycles. Relative gene expression was analyzed by \(2^{-\Delta\Delta CT}\) method. GAPDH was used as an internal reference. The primer sequences used in this study are listed in Table 1.

2.7 Statistical Analysis

R software (v4.0.5, R Foundation for Statistical Computing, Vienna, Austria) was used for all statistical analysis. Univariate Cox regression analysis was performed using the “survival” and “survminer” packages. LASSO analysis and model construction were carried out using the “glmnet” software package in R. The Wilcoxon test was utilized to determine statistical differences in categorical variables. A Kaplan-Meier curve was drawn, and the log-rank test was applied to analyze significant differences in the OS between the two groups. Univariate and multivariate Cox proportional hazards regression analyses were performed to understand the relationship between the risk score and clinical indicators and OS. ROC analysis was used to evaluate the sensitivity and specificity of the risk score in predicting prognosis. The area under the ROC curve (AUC) was used to assess the accuracy of prognosis. \(p\)-value < 0.05
Fig. 2. Identification of macrophage marker genes by single-cell sequencing. (A) Quantity and expression distribution of selected genes. (B) Variance of variable (red) and non-variable (black) gene. (C) PCA results. (D) 16 principal components for subsequent analysis. (E) 20 clusters were visualized based on the tSNE algorithm. (F) The heatmap showed the relative expression of genes in 20 clusters. Yellow represents high expressed genes and purple represents low expressed genes; (G) the annotation of 20 clusters.

was considered statistically significant for all analyses. The list of package applied in the analysis, including the name, version number and website address was shown in Supplementary Table 2.

3. Results

3.1 Macrophage Marker Gene Expression Profile

The flow chart of this study is shown in Fig. 1. A total of 15,116 cells from 26 GC samples, including 25,288 corresponding genes, were selected for analysis (Fig. 2A). Analysis of variance showed 2000 highly variable genes (Fig. 2B). PCA was used to identify the available dimensions and screen for related genes. The PCA results did not demonstrate a significant separation between cells in human GC (Fig. 2C). The first 16 PCs were selected for subsequent analysis (Fig. 2D). The heat map of the first 16 PCs is shown in Supplementary Fig. 1. Human GC cells were divided into 20 independent clusters (Fig. 2E). Results showed that different cell clusters had different gene expression profiles, and some genes were differentially expressed among the 20 clusters (Fig. 2F). Based on the expression patterns of cell-type-specific marker genes, these clusters were annotated using singleR (Fig. 2G). Cluster 11, including 424 cells, was annotated as macrophages. Differential expression analysis was performed between cluster 11 and the other clusters. Based on these criteria, 347 GC-related macrophage marker genes were identified.

3.2 Functional Annotation of Macrophage Marker Genes

Totally, 347 macrophage marker genes were screened from the GSE183904 dataset. GO (Fig. 3A–C) and KEGG pathway enrichment analysis (Fig. 3D) were performed to explore biological functions of these genes. GO analysis showed that these genes were involved in immune-related processes, such as neutrophil activation involved in the immune response, humoral immune response, antigen binding, and immunoglobulin binding. Each functional classification enriched in the first three pathways are shown in Supplementary Table 3.
3.3 Construction of the MMGS Prognostic Model

Totally, 328 patients were included in the training cohort and followed up for 61 to 3720 days. Univariate Cox regression analysis was performed on 347 macrophage marker genes, and 14 genes significantly related to prognosis were screened \( (p < 0.01, \text{Fig. 4A}) \). Result indicated that all 14 genes were risk factors for OS, that is, higher expression of these genes was associated with shorter OS. Using LASSO Cox regression analysis performed on 14 characteristic genes, seven genes, including \textit{ABCA1}, \textit{CTHRC1}, \textit{GADD45B}, \textit{NPC2}, \textit{PLTP}, \textit{PRSS23}, and \textit{RNASE1}, were obtained to construct the MMGS prognostic model (Fig. 4B). The risk score of each tumor sample was calculated using the coefficients obtained by the LASSO algorithm, and the formula was as follows: MMGS risk score = \textit{ABCA1} expression \times 0.096 \textit{+ CTHRC1} expression \times 0.051 \textit{+ GADD45B} expression \times 0.005 \textit{+ NPC2} expression \times 0.109 \textit{+ PLTP} expression \times 0.038 \textit{ + PRSS23} expression \times 0.039 \textit{ + RNASE1} expression \times 0.099. As the coefficients in the formula were all positive, all seven genes contributed positively to the risk score. Based on the median risk score, patients were divided into the high- and low-risk groups. The survival status of patients with GC is shown in Fig. 4C. Kaplan-Meier survival analysis showed a longer OS in the low-risk group than in the high-risk group \( (p = 0.0015, \text{Fig. 4D}) \). To assess the predictive efficiency of the signa-
Fig. 4. Construction of MMGS prognostic model. (A) Univariate Cox regression analysis identified 14 genes related to prognosis. (B) LASSO Cox regression analysis of 14 characteristic genes. (C) survival status of GC patients with the high- and low-risk score. (D) Risk score distribution, survival overview, and hierarchical clustering for GAC patients in training set. (E) The Receiver operating characteristic (ROC) curve of 1-, 3-, and 5-year survival.

Table 2. Characteristics and clinical information of GC patients used in RT-PCR study.

<table>
<thead>
<tr>
<th>Number</th>
<th>Gender</th>
<th>Age (years)</th>
<th>BMI (kg/m²)</th>
<th>Stage</th>
<th>Clinical staging</th>
<th>Differentiated degree</th>
<th>Family history</th>
<th>Tumor metastasis</th>
<th>Tumor infiltrating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td>34</td>
<td>20.81</td>
<td>IIIA</td>
<td>T4</td>
<td>N1</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>58</td>
<td>17.93</td>
<td>IIIA</td>
<td>T3</td>
<td>N2</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>38</td>
<td>20.70</td>
<td>IIIA</td>
<td>T4</td>
<td>N1</td>
<td>M0</td>
<td>G2</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>64</td>
<td>22.41</td>
<td>IIIIC</td>
<td>T4</td>
<td>N3b</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>77</td>
<td>17.07</td>
<td>IIIB</td>
<td>T4</td>
<td>N3a</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>71</td>
<td>21.76</td>
<td>IIIA</td>
<td>T3</td>
<td>N0</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>50</td>
<td>21.47</td>
<td>IIIIC</td>
<td>T4</td>
<td>N3b</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>64</td>
<td>24.22</td>
<td>IIIIC</td>
<td>T4</td>
<td>N3b</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>76</td>
<td>27.55</td>
<td>IIIB</td>
<td>T3</td>
<td>N1</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>Male</td>
<td>54</td>
<td>22.84</td>
<td>IIIB</td>
<td>T3</td>
<td>N1</td>
<td>M0</td>
<td>G3</td>
<td>No</td>
</tr>
</tbody>
</table>

BMI, body mass index.

3.4 Validation and Prognostic Value of the MMGS Model

We further verified whether the signature had the same effect on the prognosis of other GC cases. The GEO dataset was used as the verification set, and the risk score was calculated using the abovementioned formula. Consistent with the results of TCGA analysis, the OS of patients with low-risk score was longer than that of patients with high-risk score (Fig. 6A,B, *p* = 0.027). Survival prediction of the risk score was evaluated using a time-dependent ROC analysis at 1, 3, and 5 years. The results showed that the MMGS prognostic model also demonstrated good discrimination performance in the validation set (Fig. 6C).
3.5 Correlation between the MMGS Risk Score and Classical Clinical Variables

TNM staging is a commonly used tumor staging standard. To further test whether the prognostic value of the risk score was related to clinical indicators, hierarchical analysis was performed at different stages to explore the prognostic value of the MMGS model in different subgroups. Stratified analysis showed that the MMGS model could identify the prognostic value of different patient subgroups (Fig. 7). The high-risk groups of patients aged <60 years, with N0+1, G3, or T3+4 disease, showed lower OS.

3.6 Construction of Nomogram

A nomogram (Fig. 8A) to predict the survival probability in patients by evaluating the risk score, T, N, M, age, and sex was constructed in TCGA-GAC. The nomogram was calibrated using a calibration curve. The results
showed that the survival prediction probability (solid line) at 1, 3, and 5 years in the nomogram matched well with the ideal reference line (dotted line, Fig. 8B). We also evaluated the predictive discrimination of the nomogram and used the C-index to quantify the consistency level between the probability derived from the nomogram and observed death. The C-index of the prognostic nomogram reached 0.67.

3.7 Drug Sensitivity Prediction

The predictive effect of the MMGS model on drug therapy was compared between the high- and low-risk groups. The high-risk group in the TCGA was more sensitive to axitinib, cisplatin, DMOG, imatinib, nilotinib, and rapamycin (Fig. 9A). Similarly, the low-risk group in GSE84437 exhibited higher IC50 values for DMOG, imatinib, nilotinib, and rapamycin (Fig. 9B).

3.8 Expression Validation of MMGS by RT-PCR

The expression of ABCA1, CTHRC1, GADD45B, NPC2, and RNASE1 in tumor samples was validated using RT-PCR. Clinical characteristics of the patients are shown in Table 2. The results indicated that ABCA1, CTHRC1, and NPC2 were up-regulated without statistical significance, and GADD45B and RNASE1 were significantly down-regulated in GC (Fig. 10), which was consistent with the results in training set (Fig. 5B–H).
4. Discussion

GAC is one of the tumors with the highest incidence in the digestive system. Studies have shown that the residual state and pathological stage are independent risk factors for the prognosis of postoperative patients with GAC, but the predictive value of postoperative survival is not high [18,19]. Herein, we generated a signature based on seven genes, including \textit{ABCA1}, \textit{CTHRC1}, \textit{GADD45B}, \textit{NPC2}, \textit{PLTP}, \textit{PRSS23}, and \textit{RNASE1}, and \textit{RNASE1}, which achieved good performance in predicting the prognosis of patients with GAC. In addition, the risk score based on the signature of the seven genes was an independent prognostic factor for GAC.

An imbalance in lipid metabolism is an important factor promoting tumor macrophage reprogramming and tumor progression. ATP binding cassette transporter A1 (ABCA1) mediates intracellular lipid outflow and maintains cellular lipid homeostasis. In addition, ABCA1 mediates cholesterol outflow, reduces plasma membrane-free cholesterol and lipid content in lipid rafts, inhibits the immune system activated by toll-like receptors, and weakens the chemotactic function of macrophages [20]. Overexpressed \textit{ABCA1} in colorectal cancer is related to poor prognosis, which could favor tumorigenesis [21]. In addition, hypermethylation of \textit{ABCA1} had been linked to worse prognosis of ovarian cancer patients [22]. An integration anal-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig7.png}
\caption{Correlation between MMGS risk score and classical clinical variables. (A) Age <60. (B) N0+1. (C) G3. (D) T3+4.}
\end{figure}
ysis indicated that *ABCA1* has diagnostic and prognostic value, and is associated with poorer prognosis in GAC [23]. Our study identified increased *ABCA1* in GAC, which was statistically different from that in adjacent normal tissues. We speculate that this up-regulation is related to *ABCA1* regulating the production of intracellular anti-inflammatory substances through other signaling pathways, thereby inhibiting inflammation [24].

Niemann Pick type C2 (NPC2) protein, a small molecule soluble protein, is involved in lipid metabolism in vertebrates [25]. In mice with hepatocellular carcinoma (HCC), Liao et al. [26] found decreased NPC2 expression in mice with hepatocellular carcinoma (HCC), and blocking NPC2 promoted the proliferation, migration, and xenograft tumorigenesis of HCC. Our analysis showed that NPC2 was overexpressed in GAC tissue. We speculate that inconsistency in expressions among organs may be attributed to the crucial role of the liver in lipid metabolism, and the difference in lipid metabolism function between the liver and stomach leads to differences in tumorigenesis and development. Consistent with our results, NPC2 is overexpressed in the most common cancer types, such as glioblastoma and pancreatic cancer [27].

Yao et al. [28] reported that high NPC2 expression was associated with poor prognosis of GC patients. Similar to ABCA1 and NPC2, phospholipid transfer protein (PLTP) is a lipopolysaccharide-binding protein with a relative molecular weight of 81 kD. It specifically mediates the transport and exchange of phospholipids between plasma lipoproteins through a carrier-mediated mechanism. PLTP has pro-inflammatory and anticancer properties [29,30]. It has been reportedly associated with the prognosis of lung adenocarcinoma patients [31]. Huang et al. [32] suggested that the higher expression of PLTP in GC may serve as a biomarker for predicting the prognosis of GC patients. However, in this analysis, PLTP showed no difference in expression between normal and tumor tissues. Therefore, we speculate that PLTP is essential for the lipid metabolism and immune cell recruitment of GAC.

*CTHRC1* is located at 8q22.3, with a length of 11,435 bases. This gene encodes collagen, called CTHRC1, which is a secreted protein that exists outside the cell. *CTHRC1* is implicated in the development of a variety of human malignancies [33]. Research has indicated that *CTHRC1* is involved in macrophage recruitment in endometrial carcinoma [34]. Zhou et al. [35] reported that *CTHRC1* could act as a prognostic biomarker for renal papillary cell carcinoma and clear cell carcinoma. Further, *CTHRC1* is involved in macrophage infiltration in the TME, which can predict poor prognosis in GC [36]. Consistent with the previous reports, ascended expression of *CTHRC1* with a prognostic value was detected in GAC.

Growth arrest and DNA damage-inducible 45 (GADD45) proteins function as sensors in response to various physiological and environmental stressors. A lack of GADD45 beta (*GADD45B*) impairs the macrophage response to inflammatory stimuli [37]. In colorectal cancer, high *GADD45B* expression is correlated with poor progression-free survival (PFS) compared to that with low expression in patients with colorectal cancer [38]. In addition, increased *GADD45B* is related to shorter OS and
PFS in epithelial ovarian cancer [39]. Highly expressed GADD45B was detected in prostate cancer patients, indicating better patient survival [40]. However, our study showed that GADD45B was down-regulated in GAC. This result shows that GADD45B may play different roles in different tumors, which needs to be verified in future studies. Our study indicated that the expression of serine protease 23 (PRSS23) did not differ between normal and GAC tissues. It has been reported that the overexpression of PRSS23 is associated with poor prognosis and macrophage...
infiltration in GC patients, promoting tumorigenesis and progression [41,42]. Decreased RNASE1, a member of novel gene signatures with prognostic implications in the GC microenvironment, has been reported [43,44]. Consistent with previous studies, reduced RNASE1 was observed in GAC.

Our study had some limitations. RNA-seq data of patients were downloaded from two different databases, TCGA and GEO, which cannot guarantee the complete consistency of baseline data in terms of age, clinical stage, and treatment, thus, weakening the research conclusion to a certain extent. External validation of the model in larger cohorts is necessary. Therefore, we plan to collect a large number of samples to validate our model and extend our research to include additional clinical samples and anti-tumor drugs to further improve the accuracy and applicability of our model. In addition, further functional experiments are required to clarify the underlying mechanism of the seven genes in GAC.

5. Conclusions

In conclusion, we constructed a risk score model based on seven MMGSs to predict the prognosis of GC patients and it achieved a good predictive performance. The risk score divides patients into the high- and low-risk groups with different sensitivities to commonly used anti-tumor drugs, which may provide new insights into clinical decision-making for the personalized treatment of patients with GC.

Abbreviations

ABCA1, ATP binding cassette transporter A1; AUC, area under the ROC curve; FDR, false discovery rate; FC, fold change; FPKM, fragments per kilobase of exon per million mapped reads; GAC, gastric adenocarcinoma; GADD45, growth arrest and DNA damage-inducible 45; GADD45B, GADD45 beta; GC, gastric cancer; GEO, gene expression omnibus; GO, gene ontology; HCC, hepatocellular carcinoma; KEGG, kyoto encyclopedia of genes and genomes; MMGS, macrophage marker gene signature; NPC2, niemann Pick type C2; OS, overall survival; PC, principal component; PCA, principal component analysis; PFS, progression-free survival; PLTP, phospholipid transfer protein; PRSS23, serine protease 23; RT-PCR, real time-polymerase chain reaction; scRNA-seq, single-cell RNA sequencing; TCGA, The Cancer Genome Atlas; TLRs, toll-like receptors; TPM, transcripts-per-million; TME, tumor microenvironment; TAMs, tumor-associated macrophages; tSNE, t-distributed stochastic neighbor embedding.

Availability of Data and Materials

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

Conception and design: DP, ZL, and HC; Provision of materials and samples: HC and LX; Experimental execution: ZL, HC and ZC; Data analysis and interpretation: ZL and DP. All authors read and agree the final version of the manuscript. All authors contributed to editorial changes in the manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The study received ethical approval from the Ethics Committee of the First Affiliated Hospital of Fujian Medical University (2020-219). All patients have signed the informed consent.

Acknowledgment

We would like to thank Editage (https://www.editage.com/) for English language editing.

Funding

This study was funded by Natural Science Foundation of Fujian Province (No. 2022J01210) and Young and Middle-aged Key Personnel Training Project of Fujian Provincial Health Commission (No. 2021GGA029).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.fbl2905172.
References


