Prognostic Values of Radiosensitivity Genes and CD19 Status in Gastric Cancer: A Retrospective Study Using TCGA Database

Li-Bo Liang,1,* Xin-Yan Huang,2,* He He3 Ji-Yan Liu4,5

1Department of International Medical Centre, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, People's Republic of China; 2West China School of Stomatology, Sichuan University, Chengdu 610041, Sichuan Province, People's Republic of China; 3Department of Laboratory Medicine, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, People's Republic of China; 4Department of Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, People's Republic of China; 5Department of Oncology, The First People's Hospital of Ziyang, Ziyang 641300, Sichuan Province, People's Republic of China

*These authors contributed equally to this work

Background: The correlation between the radiosensitivity genes combined with CD19 status and clinical outcome was investigated to identify gastric cancer (GC) patients who would benefit from radiotherapy combined with CAR-T cell therapy.

Methods: The gene expression and clinical features were downloaded from The Cancer Genome Atlas (TCGA) Stomach Cancer (STAD). To identify the hub radiosensitivity genes and CD19 status, 407 patients were categorized into two groups: radiosensitivity (RS) and radioresistance (RR) based on the prognosis. The chi-square test, Mann–Whitney test, and Kaplan–Meier survival analysis were applied to compare the differential expression in these groups and analyze the correlation between the gene expression and clinical outcome and features. Finally, the influencing factors for the prognosis of GC were investigated by multiple Cox regression, especially in RS patients.

Results: A total of 15 differential expression genes, containing two communities with 8 hub radiosensitivity genes, were identified. We also identified a 2-gene signature model with a negative coefficient and calculated the risk score for the prognosis of GC. Also, Helicobacter pylori infection was validated, and the high-risk score of radiosensitivity genes was the risk factor, and high CD19 expression was the protective factor for the prognosis.

Conclusion: The radiosensitivity gene signature and CD19 expression predicted the clinical outcome of GC patients.

Keywords: gastric cancer, radiosensitivity genes, CAR-T cell therapy, TCGA database

Introduction

Gastric cancer (GC) is associated with high morbidity and mortality rates worldwide, ranking as the fourth most common cancer worldwide and the second leading cause of cancer-related deaths.1 The dismal prognosis results from rare individualized precision diagnosis and treatment.2 In recent years, an increasing number of studies have demonstrated that the genetic background was associated with the occurrence and prognosis of GC.3 For unresectable GC, radiotherapy is the primary locoregional treatment. Nonetheless, several patients failed radiotherapy due to intrinsic or acquired resistance of tumor cells to radiotherapy and nonspecific toxicity to the mucosa and surrounding normal gastric tissues.4 Several studies have shown that the effect of radiotherapy on the survival rate of patients with GC is controversial. Therefore, it is essential to find the potential innovative therapeutic strategies, such as specific immunotherapy combined with traditional therapy for beneficial progress.5,6
Although there are no immunotherapies, including the immune checkpoint inhibitors (PDL-1), chimeric antigen receptor (CAR)-T cells, and tumor vaccines, for GC that has been put into clinical use, cancer immunotherapies have received considerable attention as a viable treatment option for GC. CD19 is the primary target for CAR-T, and an innovative treatment protocol to improve the overall survival (OS), and the quality of life of patients. In this study, we focused on the prediction of the prognosis of GC with respect to the genes of radiosensitivity and CD19 status and aimed to evaluate the influencing factors for the prognosis of GC.

Materials and Methods

Data Resources

We downloaded GC-related gene expression array profiles, phenotype, and survival time from the Cancer Genome Atlas (TCGA) Stomach Cancer (STAD) of the UCSC Xena website (https://xenabrowser.net/datapages/) to assess the survival analysis and construct the prognosis model. We enrolled 407 samples, and after deleting invalid data, we analyzed 351 samples. The flowchart of this study is shown in (Supplementary material Fig. S1). Furthermore, we performed the external validation based on Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). GSE84437 datasets consisted of GC and paracancerous non-tumor tissues, and 433 samples were included for further analysis. All the data analyses were conducted in the R software (version 3.6.0).

Radiosensitivity Gene Signatures

We selected the gene signatures implicated in clinical, cellular, and tumor cell radiosensitivity. The “ConsensusClusterPlus” package of R was applied to Partitioning Around Medoid (PAM) to cluster the patients into two groups with k=2 based on the gene expression profile. The two clusters were generated without any labels because of PAM, an unsupervised machine learning algorithm. Next, the clinical outcomes were compared between clusters to define the RS as better DFI than RR. The “survfit” package and “survdiff” package in R created the Kaplan–Meier survival curves and assessed the difference between the curves of the two groups, respectively; the significant difference was set at P-value <0.05.

The screening cutoff for differential expression gene (DEGs) of TCGA was |log fold change| > 1.5 and adjusted P-value < 0.05 (corresponding to -log10(adjusted P-value)>1.30), and analyzed by the “limma” package of R. The Volcano plot and heatmap showed the DEGs of the tumor tissues and adjacent tumor tissues. The Venn diagrams showed the intersection of DEGs and gene signatures of radiosensitivity (http://bioinfogp.cnb.csic.es/tools/venny/index.html).

We used the R package “clusterProfiler” to explore the biological functions, and “Goplot” facilitated the visualization of the results in Gene Ontology (GO), including biological processes (BP), cellular components (CC), and molecular features (MF). Also, the top 10 enrichment pathways were displayed.

CD19 Status

The chi-square test and Mann–Whitney test evaluated the differences in the clinicopathological characteristics and the expression levels of CD19 between the RR and RS groups, respectively. The heatmap with annotations verified the radiosensitivity and CD19 status.

Prognosis Risk Model Establishment

To further explore the association between prognosis and radiosensitivity genes expressions and clinicopathological features. We construct the risk score models through univariate and multivariate stepwise Cox analysis. According to the median score of the risk score model, samples divided into a high-risk group and low-risk group. The “survfit” package created the Kaplan–Meier survival curves for the disease-free interval (DFI), meaning for patients having new tumor events, whether it is a local recurrence, distant metastasis, new primary tumor of the cancer. The difference between curves of the two groups compared by “survdiff” package in R, and the statistical difference was P-value <0.05.

External Validation

According to the median score of the risk score model, samples were divided into high-risk and low-risk groups. The “survfit” package created the Kaplan–Meier survival curves for OS and compared the survival probability of the two groups.

Results

Validation of Hub Genes for Radiosensitivity

We clustered the datasets into two groups, including radio-resistant (RR, n=201) and radiosensitive (RS, n=150) (Figure 1A). The tracking plot (Supplementary material Fig. S2a) showed that the clustering analysis based on k=2
was stable, and the cluster-consensus plot (Supplementary material Fig. S2b) showed that the clustering effect based on the mean of pairwise consensus values is balanced in the two clusters. The survival rate of patients with radiotherapy was significantly higher than those without radiotherapy ($P=0.006$, Figure 1B). In the patients with radiotherapy, compared to RR, the survival rate of DFI in RS was significantly higher ($P=0.0255$, Figure 1C), but no significant difference was detected in non-RT ($P=0.8468$, Figure 1D). These results showed that radiosensitivity and the therapeutic effect of radiotherapy might be affected by the related genes, and hence, we explored the hub genes for radiotherapy.

Compared to the adjacent tumor tissues, there were 135 upregulated and 264 downregulated genes in gastric tumor tissues (Figure 2A). Also, 15 radiosensitivity genes included CDK1 (cyclin-dependent kinases 1), MMRN1 (recombinant multimerin 1), GDF15 (growth differentiation factor 15), CLDN1 (recombinant claudin 1), PLK1 (polo-like kinase 1), AURKA (aurora kinase A), KLF4 (Krüppel-like factor 4),

**Figure 1** The survival analysis of DFI based on radiotherapy and radiosensitivity. (A) The datasets were clustered into two groups via consensus clustering. Moreover, cluster 1 presents the RR, and cluster 2 presents the RS via comparison of the DFI. (B) The survival rate of DFI was compared between the patients undergoing or not radiotherapy. The red line represents the survival rate of patients undergoing radiotherapy, and the blue line represents the survival time of patients without radiotherapy. (C) The survival rate of DFI in patients with radiotherapy and cluster RR are represented by red line as compared to RS represented by the blue line. (D) The survival rate of DFI in the patients without radiotherapy, also cluster RR, is represented by red line as compared to RS represented by the blue line.

**Abbreviations:** DFI, disease-free interval; RT, radiotherapy; RR, radioresistant; RS, radiosensitivity.
*SPARC* (secreted protein acidic and rich in cysteine), *SPP1* (secreted phosphoprotein 1), *SERPINE1* (serpin family E member 1), *CXCL8* (CXC motif chemokine ligand 8), *FBLN1* (Fibulin 1), *CLU* (clusterin gene), *GSTA1* (glutathione S-transferase alpha 1), and *LTF* (lactotransferrin) with differential expression in GC (Figure 2B). The heatmap also verified the status of these DEGs with respect to radiosensitivity in adjacent tumor tissues and gastric tumor tissues (Figure 2C).

Furthermore, we analyzed the biological functions of the target genes. The biological processes mainly included enriched regeneration, response to lipopolysaccharide, response to molecule of bacterial origin and antibiotic, associated cell components present in the extracellular matrix, and molecular functions in extracellular matrix structural constituent (Figure 3A). The GO chord plot and cluster plot displayed the correlations between the DEGs of radiosensitivity, ranked by the logFC and GO terms (Figure 3B and C). The protein–protein interaction (PPI) and the molecular complex detection (MOCDE) displayed two central communities in these differentially expressed radiosensitivity genes in GC: five and three genes, respectively (Figure 3D).

**Identification of the CD19 Status**

The status of CD19 differed markedly in the RR and RS groups. In high CD19 group, the proportion of RR was 41.4% (n=65) and that for RS (n=92) was 58.6%, whereas in the low CD19 group, the proportion of RR was 68.24% (n=101) and that for RS (n=47) was 31.76% (P<0.001). Similarly, the expression of CD19 was significantly higher in RS than in RR (Supplementary material Fig. S3a). The heatmap shows the expression of the radiosensitivity gene signatures and clinicopathological variables in the RR and...
RS groups. Also, a significantly different expression of CD19 was observed in the RR and RS groups (Fig. S3b).

Prognostic Role of Radiosensitivity Genes and CD19 Status in GC
To investigate the effect of radiosensitivity genes and CD19 status on gastric prognosis, we performed Cox univariate analysis (Figure 4A). The risk signature was established by selecting SPARC, AURKA, SERPIN E1, and CDK1, and the risk scores of GC patients were calculated using the LASSO Cox regression model; finally, the 2-gene model was obtained: \((-0.3376) \times AURKA + (-0.1868) \times CDK1\). According to the median risk score, patients were divided into low- and high-risk groups. The OS rate of the high-risk group was
significantly lower than that of the low-risk group (Figure 4B, $P<0.001$), and the 5-year OS rate was 0.56\% in the high-risk group and 7.3\% ($P=0.0015$) in the low-risk group.

The heatmap shows the expression of the two selected radiosensitivity genes combined with CD19 status and clinicopathological variables in the high- and low-risk groups. However, no significant difference was detected in the variables between the two groups (Figure 5A).

Furthermore, we used stepwise Cox multivariate analyses to determine whether the risk signature was an independent predictor. Multivariate analyses showed that the patients’ age, female gender, M stage, *Helicobacter pylori* infection, CD19 expression, and the risk score of hub radiosensitivity genes were significantly related to OS (Figure 5B). While in RS, multivariate analyses showed that only *Helicobacter pylori* infection, family history, and 

![Figure 4](image.png) Effect of radiosensitivity genes in GC contributing to the prognosis. (A) Cox univariate analysis of radiosensitivity genes showed that SPARC, SERPINE1, and MMRN1 were the risk factors, while AURKA and CDK1 were the protective factors of the prognosis. (B) Correlation between the risk score and the OS of GC patients showed that the survival probability in the low-risk group was higher than that in the high-risk group.

![Figure 5](image.png) Effects of radiosensitivity genes and clinicopathological variables on the prognosis of GC patients. (A) The heatmap shows the expression of two hub radiosensitivity genes and the distribution of clinicopathological variables between the high- and low-risk groups. Both genes were highly expressed in the low-risk group and lowly expressed in the high-risk group. (B and C) Cox multivariate analyses of clinicopathological variables and OS in all patients (B) and in RS cluster (C) shows that *Helicobacter pylori* infection, family history, and risk score of hub radiosensitivity genes in GC are the risk factors and the CD19 expression was the protective factor of prognosis.
CD19 expression, and the risk score of hub radiosensitivity genes were significantly related to OS (Figure 5C). These findings suggested that the risk signature and CD19 are influencing factors for GC patients and can independently predict the prognosis of the disease.

External Validation Based on GEO Datasets
In order to ensure the comparability of the results, we downloaded the dataset (GSE84437) from GEO database for external validation. The survival analysis based on the risk score of the 2-gene signature showed that the survival probability in low-risk score was significantly higher than that in the high-risk score (Fig. S4a). The expression of CKD1 and AURKA in GSE84437 dataset was similar to that in TCGA, indicating radiosensitivity in the low-risk group. Also, a significant difference was noted in the T stage of GC, indicating that T stage was high in the high-risk group (Fig. S4b).

Discussion
The TCGA dataset GC (STAD) described a set of gene signatures that reflected radiosensitivity, and the result suggested that a 2-gene model could be a predictive marker for the response of these patients to RT. We also investigated the correlation between CD19 status, radiosensitivity genes, clinicopathological features, and prognosis of GC. These studies, for the first time, demonstrated the correlation between radiosensitivity and CD19 status of GC patients and identified the potential factors that affect survival.

Firstly, we applied unsupervised learning to divide the dataset into two groups and used DFI to determine RR and RS. The results showed that the patients in RS treated with RT provided a better DFI rate than those in RR, indicating that the radiosensitivity genes might be related to the therapeutic effect of GC. This result was similar in breast cancer and glioblastoma multiforme.17,18 Next, we investigated the patients and tumor features to identify the hub of radiosensitivity genes. Although the trend of gene expression differed in the high- and low-risk groups, no statistical correlation was established between the expression of single-gene and clinical characteristics.

To further explore the prediction and influencing factors of the prognosis of GC patients, we combined the clinical variables, CD19 status, and the expression of hub radiosensitivity genes expressions. The therapeutic effects of radiation arise from triggering the immunological response in the tumor microenvironment by inducing interferon-gamma and the activation of T cells.19,20 Reportedly, in preclinical studies, this treatment ignites an adaptive anti-tumor immunity, known as the “abscopal effect”21 CAR-T cell therapy is one of the most promising approaches for malignant tumors, and the FDA (food and drug administration) has approved it for relapsed/refractory acute lymphoblastic leukemia.22 Recent studies have been directed at examining the efficacy of CAR-T therapy in gastrointestinal tumors.23 CD19 is the target of CAR-T therapy, and hence, we focused on the prognosis of GC patients, especially in RS. Hitherto, no studies have been reported regarding the radiosensitivity and CD19 status of GC. After adjusting the parameters, the multiple Cox regression model showed that CD19 was a protective factor, especially in RS.

Similarly, the risk score calculated by the 2-radiosensitivity genes model was the risk factor, but considering they had a negative coefficient, we ascribed it to be positive. Several studies showed that CDK1 inhibits cell proliferation and enhances the radiosensitivity in cancers.24–26 CDK1-encoding proteins drive the cell cycle by phosphorylation of Ser/Thr proteins.27 The potential mechanism enhances the radiosensitivity of cancer cells related to homologous recombination repair28 and genomic stability.29 In different cancer types, AURKA drives radiosensitization from chromosomal missegregation on tripolar mitotic spindles.30 Both the genes showed a favorable function on radiotherapy treated as RS genes. However, the mechanism of radiosensitivity in GC needs to be explored further.

We speculated that the patients with high expression of CD19 consist of several immunogenic tumors before radiotherapy, inducing a robust adaptive immune response, which is associated with enhanced GC survival.7 Especially in RS, RT promotes immunological response in the tumor microenvironment, enhancing the eradication of the tumor cells to extend the rate of OS.31 However, the crosstalk between radiosensitivity genes and CD19 status needs to be confirmed by further experiments and clinical trials.32

Conclusion
In conclusion, we first validated the predictive value of radiosensitivity genes in GC patients undergoing RT in the TCGA dataset and also deduced the correlation between these gene models and CD19 status. Furthermore, we found that the radiosensitivity gene signature and CD19 expression were significant influencing factors for the
clinical outcome of GC patients and may be used for selecting patients who will benefit from RT combined with CAR-T cell therapy. We also showed that CD19-high-RS patients had improved outcome, suggesting that these patients may potentially benefit from RT combined with CAR-T.

Ethics approval
This article does not contain any studies with human participants performed by any of the authors.

Author Contributions
All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Funding
The current study was supported by the National Natural Science Foundation of China (81902142). The datasets generated and analyzed during the current study are available in The Cancer Genome Atlas (TCGA) repository (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga).

Disclosure
The authors report no conflicts of interest for this work.

References


