Identification of ferroptosis-related genes for the prediction of prognosis and chemotherapy benefit of gastric cancer

J.-J. ZHENG, J.-P. WANG

Department of Colorectal and Anal Surgery, Lishui Municipal Central Hospital, Lishui, China

Abstract. – OBJECTIVE: As a newly reported programmed cell death, ferroptosis plays a crucial role in tumor progression. We aimed to comprehensively analyze the prognostic value of ferroptosis-related genes and identify the genes for the prediction of prognosis and chemotherapy benefit of gastric cancer (GC) patients.

MATERIALS AND METHODS: Public microarray data and corresponding clinical information of GC were downloaded from GEO databases. Patients in GSE66229 were randomly separated into discovery and internal validation at a ratio of 2:1. GSE15459 was set as the external validation set. LASSO Cox regression model was performed to identify the most significant prognostic ferroptosis-related genes.

RESULTS: Twenty ferroptosis-related genes were finally identified to establish the predictive signature. In the discovery data set, the signature could divide patients into low- and high-risk groups with significantly different overall survival (HR: 2.11, 95% CI: 1.40-3.17, \( p < 0.001 \)). These results were successfully validated in the internal validation cohort (HR 2.04, 95% CI 1.18-3.52; \( p = 0.008 \)) and external validation cohort (HR 4.87, 95% CI 2.99-7.93; \( p = 0.008 \)). Survival ROC at 5-year revealed a remarkably higher predictive ability of the ferroptosis classifier (AUC=0.835) compared with other clinical features in predicting prognosis. Further GSEA analysis showed that samples of high risk were related to several established tumor invasion and metastatic signaling pathways. Further experiments revealed that VLDLR and GCH1 were two newly identified genes associated with chemotherapy sensitivity in GC.

CONCLUSIONS: The developed ferroptosis gene-set based prognostic signature indicated superior prognostic and predictive value, suggesting new possibilities for individualized treatment of GC patients.

Key Words: Gastric cancer, Ferroptosis, GEO, Signature.

Introduction

Gastric cancer (GC) is one of the leading causes of cancer-associated death worldwide.

Though the current treatment strategies including surgery and neoadjuvant/adjuvant chemotherapy have improved the prognosis of GC significantly, long-term survival remains undesirable. Nowadays, the American Joint Committee on Cancer (AJCC) TNM stage system is the only established prognosis evaluation tool. However, increasing evidence have revealed its insufficiency for prognosis prediction. Consequently, novel molecular biomarkers and prognostic models are warranted to be identified to advance the prediction of high-risk GC patients.

In the past decades, great efforts have been made to seek sensitive biomarkers for identifying high-risk patients with poor survival. Currently, with the development of transcriptome profiling, several multi-gene signatures have been constructed to predict the prognosis for GC patients. However, only a few signatures with high prognostic accuracy have been developed in GC.

Programmed cell death (PCD) has been confirmed to be related to tumorigenesis and progression. Ferroptosis is a newly discovered iron-dependent form of PCD that is driven by the lethal accumulation of lipid peroxidation. Since it was first reported, it has received widespread attention for its potential as a therapeutic target for cancer treatment. Numerous studies have revealed the critical role of ferroptosis in killing tumor cells and various ferroptosis promoting and inhibiting genes have been identified in cancer. The ferroptosis-related genes can be classified into three categories including driver, suppressor, and marker. In GC, several previous studies have confirmed the important significance of ferroptosis regulatory genes in cancer treatment and prognosis evaluation. However, no previous studies analyzed the prognostic value of ferroptosis-related genes comprehensively in GC.
In this study, a total of 491 GC patients with microarray data from the Gene Expression Omnibus (GEO) database were identified to develop and validate a novel ferroptosis-related gene-set based prognostic signature. The signature we developed may help to improve the risk stratification in GC.

**Materials and Methods**

**Data Source**

We downloaded the raw data of two microarray datasets (GSE66229 and GSE15459) produced by the same Affymetrix Human Genome U133 Plus 2.0 platform from the GEO database and the method of Robust Multichip Average was used to make background adjustments. All the probes were mapped according to their original EntrezGeneID. ComBat method was used to remove the potential batch effects when combining the two datasets.

**Source of Ferroptosis-Related Gene**

Gene list of Ferroptosis related genes was downloaded from world’s first database of ferroptosis regulators (FerrDb, http://www.zhounan.org/ferrdb) comprising a total of 259 ferroptosis-related genes.

**Establishment of a Risk Score Model**

First, patients from GSE66229 set were randomized 2:1 into the discovery and internal validation set. Univariate Cox regression analysis was then used in the discovery set to identify factors significantly associated with OS ($p < 0.05$). The least absolute shrinkage and selection operator method (LASSO) Cox regression model at 10-fold cross-validation was then performed to screen the most significant prognostic ferroptosis-related genes. Lastly, based on the specific risk score formula, the risk score of each patient was calculated. The formula was as follows:

$$\text{Risk score} = \sum_{i=1}^{n} \exp_i \times \beta_i$$

**GSEA**

JAVA program (http://software.broadinstitute.org/gsea/index.jsp) was used to perform gene set enrichment analysis (GSEA). To determine the relationship between risk score and the given gene sets, GSEA was then performed on all genes ranked by enrichment score. It was conducted at 1000 random permutations and $p$-value less than 0.05 was set as significance. The positive enrichment score for an enriched gene set means that most of the genes in this set have higher expression levels with a higher risk score.

**Western Blotting**

Western blotting was performed using whole-cell protein lysates of GC cells using primary antibodies against VLDLR (ab203271, 1:1000; Abcam, Cambridge, MA, USA) and GCH1 (28501-1-AP, 1:1000; Proteintech, Rosemont, IL, USA) and a secondary antibody (anti-rabbit IgG, 1:7500; Cell Signaling Technology, Danvers, MA, USA). Equal loading of protein samples was monitored using an anti-β-actin antibody (66009-1-lg, 1:2500; Proteintech).

**Cell Viability Assay and Cell Apoptosis Measurement**

Cell viability was measured by CCK-8 assay. According to the manufacturer’s instructions, FITC Annexin V Apoptosis Detection Kit (BD, La Jolla, CA, USA) was used to detect the apoptotic rate of cells.

**Statistical Analysis**

Kaplan-Meier analysis with the log-rank test was used to compare the survival differences between low and high-risk patients. Multivariate Cox regression analysis was performed to evaluate the independent prognostic role of the signature in predicting overall survival (OS). Survival ROC was used to assess the predictive ability of signature. $p<0.05$ suggested that the difference was statistically significant. All statistical analyses were performed with R (version 3.5.0, https://www.r-project.org/).

**Results**

**GC Datasets**

Two GC data sets (GSE66229 and GSE15459) and their corresponding clinical information were downloaded from the GEO database. After excluding samples without follow-up information, 300 GC patients in GSE66229 (200 patients from the test series and 100 from the validation series) and 191 patients in GSE15459 were finally identified. The basic information about these patients was shown in Table I.
Establishment of Ferroptosis-Related Gene Set-Based Predictive Signature

First of all, univariate Cox survival analysis was performed in the discovery set and a total of 206 ferroptosis-related genes were found as prognosis-associated genes (Supplementary Table 1). Further LASSO Cox analysis screened 20 prognostic genes (Figure 1). Then, risk score was calculated based on the expression value of the 20 genes and risk regression coefficients for each patient: Risk score=(−0.071×expression level of ANGPTL7) + (0.358×expression level of CDKN2A) + (0.172×expression level of FADS2) + (−0.398×expression level of GCH1) + (0.183×expression level of GDF15) + (0.100×expression level of IL6) + (−0.020×expression level of LINC00472) + (−0.251×expression level of MAPK3) + (0.049×expression level of NNMT) + (0.385×expression level of NOX4) + (−0.241×expression level of PTGS2) + (−0.016×expression level of RGS4) + (0.128×expression level of SCD) + (−0.606×expression level of SLC1A4) + (0.310×expression level of SLC2A3) + (−0.304×expression level of SOCS1) + (0.108×expression level of TAZ) + (−0.152×expression level of TF) + (−0.008×expression level of TP63) + (0.366×expression level of VLDLR).

Then, in the discovery set, GC patients were categorized into low- and high-risk groups based on the median risk score. As illustrated in Figure 2A (left panel), the risk score was positively associated with the emergence of death events, indicating that a higher risk score was correlated with poorer OS. Survival ROC

Table 1. Basic information of patients from GEO datasets.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Discovery set (N = 200)</th>
<th>Internal validation set (N = 100)</th>
<th>External validation set (N = 191)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 60</td>
<td>67</td>
<td>39</td>
<td>56</td>
</tr>
<tr>
<td>≥ 60</td>
<td>133</td>
<td>61</td>
<td>135</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>139</td>
<td>60</td>
<td>67</td>
</tr>
<tr>
<td>Female</td>
<td>61</td>
<td>40</td>
<td>124</td>
</tr>
<tr>
<td>TNM stage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>19</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>II</td>
<td>62</td>
<td>35</td>
<td>29</td>
</tr>
<tr>
<td>III</td>
<td>67</td>
<td>29</td>
<td>72</td>
</tr>
<tr>
<td>IV</td>
<td>52</td>
<td>25</td>
<td>59</td>
</tr>
<tr>
<td>Adjuvant chemotherapy status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>150</td>
<td>69</td>
<td>NA</td>
</tr>
<tr>
<td>Yes</td>
<td>50</td>
<td>31</td>
<td>NA</td>
</tr>
<tr>
<td>Lauren type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intestinal</td>
<td>98</td>
<td>46</td>
<td>98</td>
</tr>
<tr>
<td>Diffuse</td>
<td>89</td>
<td>48</td>
<td>75</td>
</tr>
<tr>
<td>Mixed</td>
<td>13</td>
<td>6</td>
<td>18</td>
</tr>
</tbody>
</table>

NA: Not available.

Figure 1. Selection of prognostic genes based on LASSO Cox model. A, LASSO coefficient profiles of the 206 OS associated ferroptosis related genes. B, Tuning parameter (l) selection in the LASSO model used 10-fold cross-validation via minimum criteria.
Ferroptosis and gastric cancer

at 3-year, 5-year, and 7-year revealed that this signature had considerable predictive ability and prognostic accuracy (Figure 2A, middle panel). The next survival analysis showed a notable survival difference between patients in low- and high-risk groups (Figure 2A, right panel, \( p < 0.001 \)).

**Internal and External Validation of Ferroptosis-Related Gene Set-Based Predictive Signature**

To internally validate the prognostic value of this signature, we performed the same analyses in the internal validation set and similar results were attained. This signature showed high prognostic accuracy too and successfully divided patients into two groups with significantly different OS (Figure 2B, right panel, \( p < 0.001 \)). The 5-year survival rates for low- and high-risk patients were 64.9% and 40.9%, respectively.

To further confirm its performance, external validation was performed. Based on the same risk formula, similar death risk distribution was observed (Figure 2C, left panel). Consistent with the above findings, patients in the high-risk group had significantly shorter OS than those in the low-risk group (Figure 2C, right panel, \( p < 0.001 \)).
Prognostic Independence and Performance of the Signature in Predicting OS

Multivariate Cox was conducted to further test the independence of the signature. The result showed that the ferroptosis-related gene set-based predictive signature maintained its independence in predicting OS in the discovery, internal validation, and external validation sets after adjusting other clinicopathological risk features (Table II). Further subgroup analysis was then performed based on AJCC tumor stage and chemotherapy status. The stratification analysis showed that the ferroptosis signature could distinguish patients with varied outcomes despite the same tumor stage and chemotherapy status (Figure 3).

Additionally, we performed survival ROC analysis to compare the prognostic accuracy of survival prediction among the ferroptosis signature, AJCC, and postoperative chemotherapy status in the entire cohort. The result showed that the ferroptosis signature had the highest prognostic accuracy (AUC=0.835) among other clinical factors (Figure 4).

Ferroptosis Signature Associated with Biological Signaling Pathway

GSEA analysis in the entire cohort was performed to identify the ferroptosis signature associated biological signaling pathway. Significant gene sets (p<0.05) were visualized in Figure 5. The risk score was accompanied by exceptional

<table>
<thead>
<tr>
<th></th>
<th>HR</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery set</td>
<td>82.15 (0.12-74.16)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Internal validation set</td>
<td>89.73 (4.95-1627.1)</td>
<td>0.002</td>
</tr>
<tr>
<td>External validation set</td>
<td>14.04 (5.56-35.47)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

*Adjusting for age, sex, TNM stage and, lauren type and adjuvant chemotherapy status.

Figure 3. Subgroup analysis of the prognostic value of signature. Kaplan-Meier survival analysis for the entire dataset (N=491) based on the ferroptosis related gene set based signature stratified by AJCC TNM stage (A-C) and adjuvant chemotherapy status (D-E).
regulation of several important tumor initiations and metastasis networks, namely Focal adhesion signaling, ECM receptor interaction signaling, Adherens junction, and TGF-β signaling.

Identification and Experimental Validation of Chemotherapy Benefit-Associated Genes

To identify the chemotherapy benefit-associated genes incorporated into ferroptosis signature, we divided patients into the long-term survival group (patients survived more than five years without postoperative chemotherapy) and the early death group (patients survived less than one year regardless of the receipt of postoperative chemotherapy) and performed differential gene expression analysis. It was found that very low-density lipoprotein receptor (VLDLR) was significantly elevated in the early death group while GTP cyclohydrolase-1 (GCH1) was strikingly over-expressed in the long-term survival group (Figure 6A). Therefore, we speculated that these two genes from ferroptosis signature may be two critical factors influencing chemotherapy sensitivity of GC patients.

To further experimentally validate the effect of VLDLR and GCH1 on cell viability and cisplatin-induced cell apoptosis in vitro, we silenced VLDLR (Figure 6B) and GCH1 (Figure 6C) expression in GC cells by transfecting siRNAs. It was found that VLDLR attenuation inhibited cell viability (Figure 6D) and augmented the rate of apoptosis induced by cisplatin notably (Figure 6E).

Figure 4. Prognostic accuracy of the developed ferroptosis related gene set based signature. Survival ROC curves at 5-year of the ferroptosis related gene set based signature, debulking status, AJCC TNM stage and adjuvant chemotherapy status in the entire cohort.

Figure 5. GSEA analysis of high-risk score associated signaling pathways.
On the contrary, GCH1 silence promoted cell viability (Figure 6F) and decreased the rate of apoptosis induced by cisplatin strikingly (Figure 6G). Together, VLDLR and GCH1 were proved as two newly identified genes associated with chemotherapy sensitivity in GC.

Discussion

At the point of the first diagnosis, the majority of GC patients are identified as advanced tumor stage. To predict the long-term prognosis of GC patients, TNM staging system is currently the most widely accepted risk stratification factor. However, for the inherent genetic and molecular variation, it is easy to notice that the prognosis of GC patients with the same stage always varies significantly. Therefore, novel valuable predictive models are warranted to be developed to make accurate risk stratification for GC. In this study, we for the first time analyzed the prognostic value of ferroptosis-related genes comprehensively and identified 20 ferroptosis regulatory genes to construct a prognostic signature in GC. According to the risk model, patients can be...
Ferroptosis and gastric cancer

divided into low- and high-risk groups with significantly different OS. Additionally, this result was successfully validated in the internal and external validation set. Multivariate and stratified analysis showed that the predictive value of ferroptosis-related gene set-based signature was independent of AJCC tumor stage. More importantly, the prognostic accuracy of ferroptosis signature was proved to be better than other clinicopathological factors, indicating that incorporation of this ferroptosis-related gene set-based signature into the current risk evaluation system will increase the predictive performance in predicting OS for GC patients.

Increasing evidence has highlighted the critical role of ferroptosis in cancer initiation, invasion, and cancer therapeutics. However, the profiling of ferroptosis in GC has not yet been clarified. As the high risk of overfitting cannot be completely avoided in the predictive models based on high-dimension data, it will weak the significance of the estimator when the model was applied to an independent set. Consequently, LASSO Cox regression model was performed in this study to facilitate the detection of ferroptosis-related genes with robust predictive ability, strong expression variances, and weak correlation among each other. Based on this established procedure, we constructed a 20 ferroptosis-related genes-based prognostic classifier in GC.

Many of the ferroptosis-related genes integrated into this risk model have been experimentally validated to have oncogenic or tumor-inhibiting effects in GC. CDKN2A is frequently loss in GC and has been considered as a promising therapeutic target. FADS2 is a kind of fatty acid desaturate, which was proved to be associated with the risk of GC. The role of GDF15 and SCD in regulating ferroptosis has been investigated in GC and it was found that GDF15 knockdown promotes erastin-induced ferroptosis in GC cells by attenuating the expression of SLC7A11. In addition, SCD was proved to be able to promote tumor growth, migration, and anti-ferroptosis of gastric cancer. The interleukin (IL)-6 family of cytokines, especially IL-6, is highly up-regulated in GC and considered as one of the most important cytokine families during tumorigenesis and metastasis. As a long non-coding RNA, LINC00472 expression is epigenetically regulated and plays a crucial role in modulating GC cell growth and motility. MAPK3 is a member of mitogen-activated protein kinase. It has been widely tested to be related to invasion, metastasis, and drug resistance of GC. NNMT, a major metabolic regulator, has been identified as a predictor of cancer prognosis in GC and higher expression of NNMT is significantly correlated with poorer survival. NOX4, a member of the NOX family, has emerged as a significant source of reactive oxygen species, playing an important role in GC cell proliferation and apoptosis. The expression of PTGS2 was found to be elevated in GC and it plays an essential role in mediating the process of resistance against cisplatin. SOCS1 is a negative regulator of various cytokines and it was confirmed as a tumor suppressor on GC cell proliferation. TAZ plays oncogenic roles in various cancers and activation of YAP/TAZ signaling pathway contributes to the initiation and progression of GC. As for the rest ferroptosis-related genes identified in this classifier, further basic research is needed to be performed to investigate their function in GC cells. In this study, further experiments were performed to validate the effect of VLDLR and GCH1 on chemotherapy sensitivity. These two genes are potentially targeted for the reverse of chemoresistance in GC.

To our best knowledge, we analyzed the prognostic value of ferroptosis-related genes comprehensively and developed the novel ferroptosis-related gene set-based signature in GC for the first time. However, some limitations are inevitable. First, all the data used in this study were downloaded from the public GEO database and external validations based on PCR data are not available. More prospective real-world data are warranted to verify the prognostic value of our signature. In addition, the functions and mechanisms of the identified ferroptosis-related genes in the progression of GC deserve further investigation.

Conclusions

We successfully developed a powerful ferroptosis-related gene-set based signature in GC. This novel signature may facilitate the personalized prediction of GC prognosis and serve as a reliable risk stratification system for GC prognostication.

Conflict of Interest

The Authors declare that they have no conflict of interests.
Acknowledgements
We thank the GEO databases for providing these valuable data.

Data and Materials Availability
Source data and reagents are available from the corresponding author upon reasonable request.

Funding
This work was supported by Zhejiang Medical and Health Technology Program (2021ZH074).

Authors’ Contribution
Jiujian Zheng made the study conception. Jiujian Zheng wrote the main manuscript text. Jiujian Zheng and Jianping Wang performed experiments and collect the data. Jiujian Zheng and Jianping Wang prepared Figures 1-6. All listed authors reviewed and approved the manuscript.

Ethics Approval and Consent to Participate
Not applicable.

Consent for Publication
Not applicable.

References


