Multi-omics Analysis of the Prognostic and Biological role of Cuproptosis-Related Gene in Gastric Cancer

Ruopeng Zhang  
Sun Yat-Sen University Cancer Center

Feiyang Zhang  
Sun Yat-Sen University Cancer Center

Zekun Liu  
Sun Yat-Sen University Cancer Center

Yuqian Huang  
Sun Yat-Sen University Cancer Center

Lei Liu  
Sun Yat-Sen University Cancer Center

Shatong Li  
Sun Yat-Sen University Cancer Center

Yinghe Li  
Sun Yat-Sen University Cancer Center

Baiwei Zhao  
Sun Yat-Sen University Cancer Center

Wanqi Chen  
chenwq1@sysucc.org.cn

Sun Yat-Sen University Cancer Center

Research Article

Keywords: Gastric cancer, Cuproptosis-related genes, Methylation, MiRNA-mRNA network, TFs-mRNA network, Co-expression network

Posted Date: July 28th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3199150/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Gastrointestinal Oncology on January 1st, 2023. See the published version at https://doi.org/10.21037/jgo-23-946.
Abstract

Background

Gastric cancer (GC) remains the third leading cause of cancer-related death. Cuproptosis has a high correlation with cancer development and progression, while Cuproptosis-related genes (CRGs) are rarely reported in GC. The aim of this multi-omics study was to investigate the prognostic value and biological functions of CRGs in GC, which may help guide precision medicine-based decision-making in GC patients.

Methods

RNA sequencing (RNA-seq) data, Copy number variations (CNV) data, and Single nucleotide variant (SNV) data were obtained from The cancer genome atlas (TCGA) database. Chi-squared test was adopted to screen differentially expressed CRGs (DE-CRGs) between samples from 14 kinds of carcinoma and adjacent tissue samples. Then, gastric cancer (GC) samples were divided into high- and low-expressed groups based on DE-CRGs for further overall survival (OS), progression-free survival (PFS), disease-free survival (DFS), and disease-special survival (DSS) analysis. After single-gene Receiver operating characteristic (ROC) analysis, biomarkers of GC was obtained eventually. Besides, methylation sites related with biomarkers were acquired and survival analysis was performed based on those sites. Next, the correlation between immune cells and biomarkers was verified. Finally, we established miRNA-mRNA, TFs-mRNA, and co-expression networks to detect factors that have a regulating effect on biomarkers.

Results

Four DE-CRGs including CDKN2A, DLD, GLS, LIAS, and PDHB in most of the 14 cancers were screened out. Seven CRGs including GLS, LIAS, CDKN2A, DLD, LDAT, MTF1 and PDHA1 have a significant difference in the survival of GC patients. Next, single-gene ROC proved that PDHB, CDKN2A, LIAS, and GLS significantly correlate with GC prognosis. Three CRGs including LIAS, GLS, and CDKN2A were remain as biomarkers based on the results we got previously, and were used to generate a nomogram. After, 3 methylation sites with a significant survival relationship which include cg13601799, 07562918, and 07253264 were found. Then, we found that B cells native is significantly correlated with CDKN2A, 4 immune cells such as T cells regulatory (Tregs) are significantly correlated with GLS, and 2 immune cells such as T cells CD4 memory activated are significantly correlated with LIAS. Moreover, we found 10 miRNA in the miRNA-mRNA network and 3 TFs in the TFs-mRNA network have a significant correlation with OS. Finally, 20 enrichment functions were obtained such as cardiac septum development, collagen fibril organization, and sensory organ morphogenesis on the basis of the co-expression network.
Conclusions

3 biomarkers with a prognosis prediction value of GC were found, and multi-factor regulatory networks was constructed to screen out 13 factors with regulating influences of biomarkers.

1. Introduction

Gastric cancer (GC) is the fifth most diagnosed malignancy worldwide and remains the third leading cause of cancer-related death due to its frequently advanced stage at diagnosis[1; 2]. Despite declining incidence rates in most countries, it can be expected to see more gastric cancer cases in the future due to ageing populations. The prognosis of GC patients is poor due to factors such as tumor recurrence, metastasis, tumor heterogeneity, and chemotherapy resistance[3]. While the development of immune checkpoint inhibitors plus chemotherapy has significantly enhanced treatment for GC patients, a considerable number of them do not receive benefits from this regimen[4; 5]. Therefore, to develop reliable molecular biomarkers for GC diagnosis, prognosis and therapeutics is of vital importance.

Copper (Cu) is an indispensable mineral nutrient involved in a wide range of physiological processes[6]. Recent studies have shown that Copper is a required cofactor for enzymes that mediate cellular functions[7]. Connections between copper and cancer have been noted with numerous observations pointing out that copper accumulation may promote malignant transformation and a higher level of copper is required for tumors, at the same time, dysregulation of copper stores can induce oxidative stress and cytotoxicity. Cuproptosis is the most recently identified copper-dependent regulated cell death form relies on mitochondria respiration, which may affect cancer development and progression[8; 9].

In recent years, research has revealed the significant role of copper in gastric cancer, leading to increased proliferation, invasion, and metastasis of cancer cells. Tang et al conducted the first and comprehensive Cu-binding proteins (CBP) analysis of GC patients and established a clinically feasible CBP signature for predicting survival and response to treatment[10]. Feng et al investigated a novel cuproptosis-related lncRNAs signature impacts on the prognosis and immunological features of GC[11]. These studies of Copper involvement in gastric cancer provides valuable insights for the development of targeted therapeutic strategies. However, the role of cuproptosis-related genes (CRGs) in tumorigenesis and tumor prognosis is still an under-explored topic.

Multi-Omics is a comprehensive study of the roles, relationships, and actions of various types of molecules in cells and allows the cost-effective and rapid elucidation of an entire genome[12]. This includes fields such as genomics, transcriptomics, proteomics, metabolomics, epigenomics and so on. Recent advances in omics technologies have led to unprecedented efforts characterizing the molecular changes[13; 14]. Through multi-omics analysis, the molecular mechanism of CRGs in tumor development and progression can be further revealed. However, truly integrated multi-omics analyses of CRGs in GC have rarely been reported. Therefore, the aim of this study was to investigate the prognostic
value and biological functions of CRGs in GC. This study screens diagnostic, prognostic and therapeutic molecular biomarkers of GC based on CRGs, which may help guide precision medicine-based decision-making in GC patients.

2. Materials and methods

2.1 Data collection

RNA sequencing (RNA-seq) data of 35 normal stomach and 415 gastric cancer (GC) samples, Copy number variations (CNV) data of 441 GC samples, and Single nucleotide variant (SNV) data of 431 GC samples were obtained from The Cancer Genome Atlas (TCGA) database (http://portal.gdc.cancer.gov/). We removed the samples without survival information such as survival states and survival times, and there are 409 GC samples remained. Differential expression information of CRGs in 14 different cancers was obtained from the “Expression” module in Gene Set Cancer Analysis (GSCA) database (http://bioinfo.life.hust.edu.cn/GSCA/). The cancers include COAD, ESCA, KIRC, HNSC, PRAD, BRCA, BLCA, THCA, GC, KIRP, LUAD, LUSC, LIHC, and KICH. Cuproptosis-related genes (CRGs) were acquired from reference[15].

2.2 Gene expression profiling of CRGs

Expression level of CRGs between normal and tumor samples were acquired from the Pan-Cancer database in TCGA and were visualized as box plots by the R package “ggpubr” (version 0.4.0). The Chi-squared test was adopted to detect the significant differences of CRGs’ expression between carcinoma and adjacent tissue samples according to p < 0.05. The expression level of CRGs between GC and normal stomach samples with pairing relationships were also visualized alone as box plots. Besides, the Spearman correlation analysis was used to detect correlation among CRGs via the R package “corrplot” (version 0.91).

2.3 Mutation and CNV of CRGs in GC samples

Gene mutation frequencies were obtained by using the R package “maftools” (version 2.6.05) on the basis of SNV data which were downloaded from TCGA[16]. The Oncoplot function was adopted to generate a waterfall plot of Gene mutation frequencies. After, CNV data were downloaded from TCGA for further analysis according to the frequencies of chromosomal amplification, chromosomal deletion, and normal diploid of genes. The R package “ggplot2” (version 3.3.3) was used to generate bar diagrams[17].

2.4 Survival analyses of CRGs

The expression level of CRGs was merged with the total survival time and survival state of related GC patients. And the patients were divided into high- and low-expression groups according to an optimal threshold that was determined via the R package “survminer” (version 0.4.8). After, survival analyses of Overall survival (OS), Disease-special survival (DSS), Disease-free survival (DFS), and Progression-free survival (PFS) between both two groups were performed by using the R package “survminer”.

Page 5/26
2.5 clinical analysis of biomarkers

ROC curves of CRGs were generated by adopting the R package “pROC” (version 1.17.0.1) and the Area under the curve (AUC) was calculated to assess the clinical value of CRGs[18]. Then, differentially expressed CRGs (DE-CRGs) with survival correlation and clinical value was screened out as biomarkers of GC. The influences of clinical characteristics on the expression level of biomarkers were detected via the Chi-square test according to \( p < 0.05 \). Next, biomarkers were used to generate a nomogram by R package “Rms” (version 6.2-0) for the prediction of 1-, 3-, and 5-year survival probability. The calibration curves were produced at the same time for verifying the performance of the nomogram. The closer the slope is to 1, the more accurate the prediction is.

2.6 Methylation analysis of biomarkers

Methylation data were downloaded from the University of California, Santa Cruz (UCSC) Xena website (http://xena.ucsc.edu/) for Annotating the methylation site of biomarkers by adopting the R package “ChAMP” (version 2.20.1)[19]. Then, the methylation sites were visualized via the R package “pheatmap” (version 1.0.12). After, Spearman correlation analysis among methylation sites of biomarkers were performed according to \( p < 0.05 \) and \( |\logFC| > 0.1 \). Besides, the R package “survival” (version 3.2-3) was used to conduct Kaplan-Meier (K-M) survival analysis to detect the survival correlation of methylation sites and to generate K-M survival curves[20].

2.7 Enrichment analysis of biomarkers

Differentially expressed genes (DEGs) between high- and low-expressed groups were screened out by the R package “limma” (version 3.44.3) with \( p < 0.05 \) and \( |\logFC| > 1 \)[21]. Gene set enrichment analysis (GSEA) was performed by the R package “ClusturProfiler” (version 3.18.1) and R package “org.Hs.eg.db” (version 3.12.0) on the basis of Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (PMID: 34557778). GO includes Biological process (BP), Molecular functions (MF), and Cellular components (CC). The significance criteria were \( p < 0.05 \) and count \( \geq 1 \).

2.8 Immune correlation analysis of biomarkers

Immune cell percentages in 409 GC samples were calculated via Cell type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) algorithm (version 1.03) and LM22 gene set, and the correlation between biomarkers and immune cell percentage was detected by adopting Spearman correlation analysis. The rank sum test was used to compare the 22 immune cells between high- and low-expression groups. Related bubble charts, box plots, and lollipop diagrams were generated via the R package “ggplot2” and “ggpubr”.

2.9 Regulation factors were obtained via Multi-factor regulatory networks

RNA-seq data of miRNA in normal and GC samples were obtained from the UCSC Xena website. Targeting correlation of miRNA-mRNA were identified using the miRWalk database
and differentially expressed miRNA (DE-miRNA) \((p < 0.05\) and \(|\logFC| > 0)\) between normal \((n = 41)\) and GC samples \((n = 387)\) were screened out via the R package “limma”. Then, miRNAs obtained previously were intersected with the DE-miRNAs, and a miRNA-mRNA network was generated by R package “Cytoscape” (version 3.8.2)[22].

Targeting correlation of TFs-mRNA were identified adopting the NetworkAnalyst database (https://www.networkanalyst.ca/), and DEGs \((p < 0.05\) and \(|\logFC| > 1)\) between normal \((n = 35)\) and GC samples \((n = 415)\) were sifted out[23]. Then, an intersection between DEGs and TFs was produced and it was used to construct a TFs-mRNA network of biomarkers via the R package “Cytoscape”. The Interaction among factors in the networks were identified by Spearman correlation analysis and visualized as a heatmap through R package “ggplot2”. Survival analyses of miRNAs in the network between high- and low-expressed groups were performed by the R package “survival”.

Correlation analysis between DEGs and biomarkers was performed to conduct further differential analysis according to \(p < 0.05\) and \(|\logFC| > 0.3\). The R package “Cytoscape” was used to construct a Co-expression network of biomarkers, and the enrichment analysis of co-expression genes and biomarkers in the network was implemented by the Metascape tool (https://metascape.org/gp/index.html#/main/step1)[24].

3. Results

3.1 Four differential expressed CRSs between GC and adjacent tissue samples

In order to identify whether cuproptosis has a high correlation with cancers progression, we obtained differential expression information of 10 CRGs including CDKN2A, DLAT, DLD, FDX1, GLS, LIAS, LIPT1, MTF1, PDHA1, and PDHB from GSCA database for differential expression analyses. The result shows that CDKN2A and LIAS have a significant difference in at least seven cancers (Fig. 1A). After, the expression level of 10 CRGs between the normal and tumor samples was acquired from the TCGA database. The box plots demonstrated that the expression level of CDKN2A and GLS were significantly up-regulated and LIAS and PDHB were significantly down-regulated in GC samples (Fig. 1B). Of note, expression level of DLD significantly changed in Bladder urothelial carcinoma, Breast invasive carcinoma, and Thyroid carcinoma (Fig. 1C-E). Besides, the box plots visually reveal the normalized expression level of each CRG between the normal stomach and GC samples that with pairing correlation (Fig. 1F). The heatmap shows that there is no significant correlation among CRGs (Fig. 1G). Moreover, tumor-specific CNVs and SNVs are helpful for exploring the molecular mechanisms of GC progression. We analyzed the mutation frequencies of CRGs of GC samples, and the result shows that the LIPT1 and CDKN2A have the highest mutation frequency of 6% and 4% respectively (Fig. 1H). After analysis of CNV was performed and we found that the CDKN2A, LIAS, and PHDB frequently occur the copy number deletion (DEL), and DLD frequently occurs the cope number amplification (AMP) (Fig. 1I).
3.2 Survival analyses between high- and low-expression groups

A series of survival analyses between high- and low-expression groups were performed to identify the correlation between the expression level of CRGs and GC prognosis. The OS curves show that the expression level of MTF1, DLAT, and LIAS are positively correlated, while the expression level of GLS is negatively correlated with the prognosis of GC (Additional File 2). After, the disease-special survival analysis results demonstrate that the expression level of MTF1, DLAT, PDHA1, DLD, and LIAS are positively correlated, while the expression level of CDKN2A is negatively correlated with the prognosis of GC (Additional File 3). Moreover, disease-free survival analysis results illustrate that the expression level of three CRGs including MTF1, PDHA1, and DLD are positively correlated with the prognosis of GC (Additional File 4). Finally, progressive-free survival curves show that the expression level of five CRGs including MTF1, PDHA1, DLD, LIAS, and DLAT is positively correlated with the prognosis of GC (Additional File 5). As shown in Table 1, there were 4 differentials expressed CRGs had a significant survival relationship.

<table>
<thead>
<tr>
<th>gene</th>
<th>OS</th>
<th>DSS</th>
<th>DFS</th>
<th>PFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>*GLS</td>
<td>TRUE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
</tr>
<tr>
<td>*LIAS</td>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE</td>
<td>TRUE</td>
</tr>
<tr>
<td>*CDKN2A</td>
<td>FALSE</td>
<td>TRUE</td>
<td>FALSE</td>
<td>FALSE</td>
</tr>
<tr>
<td>*DLD</td>
<td>FALSE</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>*PDHB</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
</tr>
<tr>
<td>DLAT</td>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE</td>
<td>TRUE</td>
</tr>
<tr>
<td>MTF1</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>FDX1</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
</tr>
<tr>
<td>LIPT1</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
</tr>
<tr>
<td>PDHA1</td>
<td>FALSE</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
</tbody>
</table>

* represents differentially expressed genes in GC; TRUE represents survival significance. OS: overall survival; DSS: disease-special survival; DFS: disease-free survival; PFS: progressive-free survival.

3.3 Three biomarkers of clinical prognosis prediction were obtained
We perform single-gene ROC analysis to evaluate the value of CRGs for GC prognosis, and the ROC curves demonstrate that PDHB, CDKN2A, LIAS, and GLS have great prognosis prediction abilities (Fig. 2A-D). Then, considering the differential expression between normal and GC samples, and the survival correlation with GC, three CRGs including CDKN2A, LIAS, and GLS were chosen as biomarkers of GC prognosis. After, the expression level of biomarkers between high- and low-risk groups in different clinical subgroups including age, gender, vital, state, grade, T stage, N stage, and M stage was compared. We found that the expression level of CDKN2A between G2 and G3 is significantly different. The expression level of GLS in T3 and T4 are significantly higher than in T1, while the expression level of LIAS in T2 and T3 are significantly lower than in T1 (Fig. 2E-G). The total information is provided in Additional File 1. Finally, clinical features were combined with the expression of biomarkers to generate a nomogram of GC prognosis prediction (Fig. 2H). The calibration curve demonstrates that the nomogram has a great prediction accuracy of 1-year survival (Fig. 2I).

3.4 Three methylation sites of CRGs significantly related to survival

Gene methylation significantly affects gene function. In this study, the methylation site of biomarkers was detected through the UCSC Xena website, and visualized in a heatmap (Fig. 3A). The scatter plots show that cg07562918 and cg13601799 were significantly related with CDKN2A, cg07253264 and cg09390371 were significantly associated with LIAS, and cg03962451 and cg19300307 were significantly correlated with GLS (Fig. 3B-G). Total information of the correlation of methylation sites were provided in Additional File 1. Finally, K-M survival analysis between high- and low-expressed groups was utilized, and we found that cg13601799 and cg07562918 of CDKN2A, and cg07253264 of LIAS have a strongly significant correlation with GC survival (Fig. 3H-J).

3.5 Functional enrichment of CDKN2A, GLS, and LIAS

In order to identify the biological functions of biomarkers, differential expression analysis was implemented to screen DEGs between high- and low-expressed groups based on the expression level of biomarkers, and then DEGs were used for enrichment analysis. Enrichment results of biomarker CDKN2A show that there are 82 functions (such as digestion, antimicrobial humoral response, and thyroid hormone generation) in GO-BP, 35 functions (such as apical plasma membrane, apical part of cell, and chylomicron) in GO-CC, 33 functions (such as endopeptidase activity, cholesterol transfer activity, and sterol transfer activity) in GO-MF (Fig. 4A), and 9 KEGG pathways such as Fat digestion and absorption, Cholesterol metabolism, and Proximal tubule bicarbonate reclamation (Fig. 4B). Enrichment results of biomarker GLS demonstrate that there are 90 functions (such as digestion, digestive system process, and maintenance of gastrointestinal epithelium) in GO-BP, 12 functions (such as digestion, digestive system process, and maintenance of gastrointestinal epithelium) in GO-CC, 42 functions (such as aspartic-type endopeptidase activity, aspartic-type peptidase activity, and carboxylic acid transmembrane transporter activity) in GO-MF (Fig. 4C), and 4 KEGG pathways include Fat digestion and absorption and Protein digestion, absorption, Pancreatic secretion, and Mineral absorption (Fig. 4D). Enrichment results of biomarker LIAS illustrate that there are 171 functions (such as uterus
development, skin development, and negative regulation of cardiac muscle tissue development) in GO-BP, 51 functions (such as serine-type endopeptidase inhibitor activity, co-receptor binding, and endopeptidase inhibitor activity) in GO-MF (Fig. 4E), and 3 KEGG pathways include Cell adhesion molecules, Wnt signaling pathway, and Fat digestion and absorption (Fig. 4F).

3.6 Expression of biomarkers is correlated with immune cells

Immune cells are an important part of the tumor microenvironment (TME) and profoundly influence the progression of GC. We calculated the percentage of the immune cells in GC samples, and the results were visualized as a stacked bar chart (Fig. 5A). The bubble charts demonstrate that biomarker GLS has a negative correlation with T cells regulatory (Tregs), NK cells resting, and Neutrophils, biomarker LIAS has a positive correlation with T cells CD4 memory activated, and has a negative correlation with Macrophages M0 and Macrophages M1, and biomarker CDKN2A has a negative correlation with B cells naive and Plasma cells (Fig. 5B). The results of differential analysis demonstrated that the percentage of B cells native in high-expressed group of CDKN2A was significant different with low-expression group (Fig. 5C). And the proportion of B cells naive, T cells regulatory (Tregs), Mast cells resting, Mast cells activated, and Dendritic cells resting in high-expressed group of GLS were significant different with the proportion in low-expressed group (Fig. 5D), and the percentage of B cells naive, T cells regulatory (Tregs), Mast cells resting, Mast cells activated, and Dendritic cells resting in high-expressed group of LIAs is significant different with low-expressed group (Fig. 5E). The correlations between expression level of biomarkers and immune cells were visualized in lollipop charts (Fig. 5F-H). Besides, we detected the correlation between biomarkers and Immune checkpoints, and the result shows that CDKN2A has a negative correlation with IDO1, TIGIT, CD274, PDCD1LG2, ICOS, and HAVCR2, LIAS is negatively related with CD27, and GLS is positively related with ICOS (Fig. 5I).

3.7 MiRNA-mRNA network of biomarkers

Mi-RNAs can affect the function of biomarkers by degrading mRNA. The volcano plot indicates that there are 724 DE-mRNAs between normal and GC samples (Fig. 6A). Then, a miRNA-mRNA network of biomarkers including 22 miRNAs was generated on the basis of correlation of miRNA-mRNA which identified via miRWalk database (Fig. 6B). The heatmap of correlation biomarkers and miRNAs indicates that CDKN2A significantly correlate with hsa-miR-330-5P, hsa-miR-181d-5P, hsa-miR-181b-5P, and hsa-miR-1301-3P, and GLS significantly correlates with hsa-miR-34a-5P, hsa-miR-181d-5P, and hsa-miR-1301-3P, and LIAS significantly correlates with 6 miRNAs including hsa-miR-942-5P, hsa-miR-432-5P, hsa-miR-181d-5P, hsa-miR-181b-5P, hsa-miR-125b-5P, and hsa-miR-125a-5P (Fig. 6C). Besides, there are 8 miRNA are significantly correlated with OS of patients, which includes hsa-miR-185-5P, hsa-miR-320b, hsa-miR-181b-5P, hsa-miR-34b-5P, hsa-miR-370-3P, hsa-miR-490-3P, hsa-miR-942-5P, and hsa-miR-432-5P (Fig. 6D-K).

3.8 TFs-mRNA network of biomarkers
TFs can affect the function of biomarkers by affecting the transcription process of mRNA. Differentially expressed gene analysis was implemented to obtain 3489 DEGs between normal and GC samples, and then those DEGs were intersected with TFs identified through the NetworkAnalyst database (Fig. 7A, B). There are 5 differentially expressed TFs in the intersection and they were used for constructing a TFs-mRNA network (Fig. 7C). The heatmap demonstrates that CDKN2A significantly correlates with RCOR2, GLS significantly correlates with SOX5, RCOR2, EZH2, and ELF3, and LIAS is significantly correlated with SOX5, KLF9, EZH2, and ELF3 (Fig. 7D). The survival analysis of 5 TFs indicated that 3 TFs including EZH2, SOX5, and KLF9 have a significant correlation with OS of GC patients (Fig. 7E-G).

3.9 Co-expression Network of biomarkers

We constructed a Co-expression Network to indicate reciprocity between DEGs and biomarkers (Fig. 8A), and there are 82 genes co-express with GLS, 21 genes co-express with LIAS, and 19 genes co-express with CDKN2A in the network. Then, we performed enrichment analysis on the basis of the network, and we found that there are 20 functions of biomarkers and co-expressed genes, such as cardiac septum development, collagen fibril organization, and sensory organ morphogenesis (Fig. 8B).

4. Discussion

GC remains to be one of the most common epithelial cancers while most patients are diagnosed at an advanced stage and still cannot benefit from the developing comprehensive therapeutic strategies[25]. Therefore, developing effective prognostic models and methods will contribute to guide personalized medicine for GC patients. Cuproptosis is the recently identified pathway which may affect cancer development and progression[26]. However, the expression patterns and clinical significance of CRGs in cancers remain unclear. It is necessary to perform a full-scale investigation of CRGs in patients with GC. In this study, we used the large-scale public database of gastric cancer transcriptome data and clinical data to screen key CRGs that may have potential prognostic, diagnostic and other guidance implications. Moreover, the underlying mechanisms and transcriptomic data of CRGs in GC were further explored and elucidated. In previous studies, ten genes related to copper-induced cell death pathways were identified. Building on this result, the present study aims to further identify copper death genes that play a key role in gastric cancer. [15]A relatively low mutation frequency but high frequency of copy number alterations was found in CRGs, consistent with the previous study[27], which implicated that CRGs might be a potential treatment targets and served as a prognosis factors in GC patients.

We ultimately identified three molecular biomarkers, GLS, LIAS and CDKN2A, which play a critical role in the copper death pathways of gastric cancer patients. Further analysis was conducted on these markers using methods such as methylation analysis, enrichment analysis, immune infiltration analysis, and construction of multi-factor regulatory networks to explore their significance in guiding prognosis and diagnosis of gastric cancer patients. GLS is an essential substance for cellular energy metabolism, which is responsible for the conversion of glutamine to glutamate[28]. The high GLS expression related with poor prognosis in GC patients. LIAS is located in the mitochondrion and encodes the protein of the biotin and lipoic acid synthetases family[29]. Decreased LIAS expression is associated with diminished
hepatic alpha-lipoic acid and tissue oxidative stress[30]. A high LIAS expression was related to the good prognosis in patients with various cancers[31], which is consistent with the result of our study. CDKN2A is a tumor suppressor gene that encodes two distinct proteins to inhibit the cell cycle and promote apoptosis. [32] The main regulatory pathway for CDKN2A is via the p53 signaling pathway, in cancer, mutations or deletions of CDKN2A are common, leading to loss of its tumor suppressive functions and contributing to tumor growth and progression.[33] Our study showed that higher expression of CDKN2A was associated with lower DSS in gastric cancer patients, but there was no significant difference in OS or DFS. As shown in the Results section, we analyzed hundreds of enriched pathways for the three biomarkers mentioned above. Moreover, a nomogram of GC prognosis prediction was constructed using clinical features and the expression level of biomarkers, which has a great prediction accuracy of 1-year survival.

In addition to the above-mentioned biomarkers, we also analyzed other molecules related to prognosis. MTF1 is an essential metal-binding transcription factor that binds to conserved DNA sequence motifs in the heavy metal response, resulting in the loss of heavy metal response gene transcription and cellular protection[34]. Our study firstly revealed that the high expression of MTF1 resulted in better OS in GC patients, which was consistent with the previously reported role for MTF1 and Cu in cell differentiation and gene expression[35].

Immune infiltration analysis and the epigenetic regulation of immune response has been widely applied in clinical research on gastric cancer and provides useful guidance for patient treatment selection and prognostic evaluation. [36; 37; 38; 39] Considering the important position of immunotherapy in gastric cancer, immune infiltration analysis studies were also used to assess the distinct roles of the subclusters and to investigate immune cell dysregulation in GC. GLS has a negative correlation with Tregs, NK cells resting, and Neutrophils, LIAS has a positive correlation with T cells CD4 memory activated, and has a negative correlation with Macrophages M0 and Macrophages M1, CDKN2A has a negative correlation with B cells naive and Plasma cells. The results of immune infiltration were largely consistent with some other studies. For example, recent research indicated that GLS was involved in immune-related signaling pathways, such as T-cell receptor signaling pathway, chemokine signaling pathway and hypoxia-related pathways[28; 40; 41]. However, further research is needed to determine the optimal immune infiltration analysis method and how to apply it to personalized treatment decision-making.

Besides, a multifactorial regulatory network was constructed for key genes, and prognostic analysis of miRNAs and TFs was performed. In our study, multiple survival-related miRNAs and TFs were screened out. There were while 8 miRNA are significantly correlated with OS of patients. And the survival analysis of 5 TFs indicated that 3 TFs including EZH2, SOX5, and KLF9 have a significant correlation with OS of GC patients. Accumulating evidence indicates that copper are involved in the regulation of miRNAs and TFs in gastric cancer, thereby promoting the proliferation, invasion, and metastasis of cancer cells[42]. They contribute to GC as oncogenes or tumour suppressors by inhibiting either directly or indirectly the expression of target genes[43]. In addition to miRNAs, copper ions can also affect the expression and
function of some transcription factors, thereby affecting the development of various cancers[44; 45], while the role remains unclear in gastric cancer. Our research provided important reference information for the future development of targeted therapeutic strategies for GC patients.

Our study still has some limitations. First, the database of GC samples needs to be expanded for more comprehensive investigation. Second, recent molecular researches on GC including us are currently not integrated into clinical practice. Future validation studies still warrant to verify the specific mechanism of CRGs at the human, animal, and cellular levels. Even so, our results provide new insights into the diagnosis, prognosis, and therapy of GC patients.

5. Conclusion

Our study identified possible biomarkers through bioinformatics analysis and systematically investigated the interactive gene landscape, prognosis role and molecular changes of CRGs in GC patients. Multi-factor regulatory networks were constructed to screen out factors with regulating influences of biomarkers. The results indicated that these CRGs may play a key role in the tumor development and progression of GC and highlighted its potential for clinical applications to guide clinical care and improve treatment selection in patients.

Declarations

Funding: This work was supported by Guangzhou Basic and Applied Research Project (SL2024A04J01162, to RP,Z)

Conflicts of interest: The authors declare no competing interests.

Availability of data and material: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability: The code applied during and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions: WQ,C: Conceptualization; RP,Z,BW,Z and YQ,H: Methodology; FY,Z and ZK,L: Formal analysis and investigation; RP,Z and WQ,C Writing - original draft preparation; YH,L: Writing - review and editing; LL and ST,L: Technical support. All authors read and approved the final manuscript.

Ethics approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Consent for publication: Additional informed consent was obtained from all legal guardians for whom identifying information is included in this article.
Acknowledgments: None.

References


Figures
Figure 1

Expression and genetic alteration of CRGs. (A) The differential expression of 10 CRGs between tumor and normal tissues from GSCA database. (B-E) The expression level of CRGs between the normal and tumor samples from the TCGA cohort. (F) The normalized expression level of each CRG between normal stomach and STAD samples with pairing correlation. (G) The correlation heatmap of CRGs in STAD. (H-I) The CNV and mutation frequency and classification of 10 CRGs in STAD. * \( p < 0.05 \), *** \( p < 0.001 \), **** \( p < 0.0001 \).; CRGs: cuproptosis-related gene
Figure 2

Clinical application of CRGs in patients with STAD. (A-D) The diagnostic ROC curve of CRGs in STAD. (E-G) The expression level of key genes in different clinical stages. (H) The nomogram using prognostic factors identified by multivariate Cox analysis to predict the 1-year, 3-year and 5-year OS. (I) The calibration curve for determining the reliability of the nomogram to predict the 1-year OS.
Figure 3

Methylation analysis of CRGs in STAD patients. (A) The methylation heatmap of CRGs. (B-G). The scatter plot of correlation between CRGs and methylation sites. (H-J) Kaplan-Meier curves for the OS of STAD patients between high- and low-expressed groups from the TCGA cohort.
Figure 4

Functional enrichment analysis of CDKN2A, GLS, and LIAS. (A,C,E) The enriched item in the gene ontology analysis. (B,D,F) The enriched item in the aKyoto Encyclopedia of Genes and Genomes analysis.
Figure 5

Immune infiltration analyses. (A) The stacked bar chart of the ratio of the immune cells in STAD samples. (B) Association between the CRGs and immune cells by CIBERSORT. (C-E) The different distribution of immune cells in the low- or high-expression subgroups. (F-H) Correlation between CRGs and immune cells. (I) Correlation between CRGs and immune checkpoints.
Figure 6

MiRNA-mRNA network of key CRGs. (A) The volcano plot of DE-mRNAs in normal and STAD samples. (B) The miRNA-mRNA network of 22 screened miRNAs. (C) The co-expression heatmap between miRNAs and key CRGs. (D-K) Kaplan-Meier curves for the OS of 8 miRNA in STAD patients.
Figure 7

TFs-mRNA network of key CRGs. (A) The volcano plot of differentially expressed gene analysis between normal and STAD samples. (B) The intersection of DEGs with identified TFs. (C) The TFs-mRNA network of 5 differentially expressed TFs in the intersection. (D) The co-expression heatmap between TFs and key CRGs. (E-G) Kaplan-Meier curves for the OS of 3 TFs in STAD patients.
Figure 8

Co-expression Network of key CRGs. (A) A Co-expression Network to indicate reciprocity between DEGs and biomarkers. (B) Enrichment analysis of key CRGs and co-expressed genes.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.docx
- AdditionalFile2.tif
- AdditionalFile3.tif
- AdditionalFile4.tif
- AdditionalFile5.tif