Identification and validation of disulfidptosis-related subtypes and the prognostic model in hepatocellular carcinoma

Qiming Huang
Department of Gastroenterology, First Affiliated Hospital of Anhui Medical University

Tiewen Li
Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine

Sisi Liu
Department of Gastroenterology, First Affiliated Hospital of Anhui Medical University

Difei Chen
Department of Gastroenterology, First Affiliated Hospital of Anhui Medical University

Yemei Du
Department of Gastroenterology, First Affiliated Hospital of Anhui Medical University

Lei Jie
Department of Gastroenterology, First Affiliated Hospital of Anhui Medical University

Yang Ma
The First Affiliated Hospital of Bengbu Medical College

Ru-Tao Hong (✉ hongruutaoah@aliyun.com)
Department of Gastroenterology, First Affiliated Hospital of Anhui Medical University

Research Article

Keywords: disulfidptosis, hepatocellular carcinoma, prognosis, immunotherapy, tumor immune microenvironment

Posted Date: June 13th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3029518/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Disulfdptosis is a newly discovered form of cell death. The latest reports have revealed potential mechanisms for disulfide-mediated cell damage, including actin cytoskeleton proteins and cellular scaffold proteins, which are susceptible to disulfide stress. Accumulation of disulfides in cells results in the breakdown of the actin network and cell death. However, the prognostic role and regulatory mechanism of disulfidptosis-related genes in hepatocellular carcinoma remain unclear.

Methods

The differential expression, copy number variation, and prognostic relevance of 10 genes associated with disulfidptosis were analyzed. Based on the expression levels of the disulfidptosis-related genes, unsupervised clustering analysis was performed to classify the samples into three subtypes, and their gene expression, biological functions, and immune cell infiltration were analyzed. Subsequently, the common differentially expressed genes were screened and a gene prognostic model was established. Samples with high-risk scores showed higher immune cell infiltration and expression of immune checkpoint genes.

Results

Firstly, we found that four genes (LRPPRC, NCKAP1, RPN1, SLC7A11) showed significant prognostic ability for overall survival in HCC patients. Subsequently, using consensus clustering analysis, the samples were divided into three clusters (A, B and C cluster) based on the expression levels of the four prognostic disulfidptosis genes, where the prognostic of B cluster was worse, and the cell cycle pathway activation was significantly enriched. Furthermore, the immune cell infiltration abundance was calculated, significant differences in TME were observed among the three subtypes. Additionally, among the common differentially expressed genes among the three subtypes, lasso regression was used to screen six survival-related genes, and a prognostic model was established. Finally, our results suggest that the risk score may serve as a useful tool for predicting sensitivity to immunotherapy and drug treatment in HCC patients.

Conclusion

These findings highlight the significant clinical implications of the sulfide death pathway and provide new insights into guiding personalized immune therapy strategies for patients with hepatocellular carcinoma.
1 Introduction

Globally, primary liver cancer is the fourth most frequent cause of cancer-related fatalities (1), with hepatocellular carcinoma (HCC) accounting for around 90% of cases. To effectively treat HCC, the patient's liver condition, the stage of the tumor, and any associated complications must be considered. Treatment options include surgical resection, liver transplantation, radiofrequency ablation, transarterial chemoembolization, immunotherapy, etc. Discovering new molecular markers associated with HCC prognosis and treatment remains a priority, as most patients are ineligible for surgical or liver transplantation treatment at the time of diagnosis (2).

In contrast to ferroptosis and cuproptosis, disulfidptosis is caused by disulfide stress. In comparison to regular cells, tumor cells possess a more rapid glucose metabolism (3). Disulfide accumulation from cellular metabolism, such as cysteine, γ-glutamylcysteine, and glutathione-cysteine disulfide, leads to oxidative stress and cell death (4). Tumor cells with high expression of the disulfidptosis-related gene SLC7A11 accelerate this process, and the use of glucose transporter inhibitors (GLUTs) has been shown to kill SLC7A11-overexpressing tumor cells and inhibit their growth (3). Recent reports have uncovered a potential mechanism for disulfidptosis, which involves the susceptibility of actin cytoskeleton proteins and cellular scaffold proteins to disulfide stress; the accumulation of disulfides in cells results in the disintegration of the actin network and eventually leads to cell death (5). Moreover, further investigations have revealed that clearing the buildup of disulfides in cells can postpone the process of cell death. However, the prognostic role and regulatory mechanism of disulfidptosis-related genes in hepatocellular carcinoma remain unknown.

This research aimed to investigate the influence of 10 genes related to disulfidptosis on the prognosis and treatment of hepatocellular carcinoma. Firstly, the expression, prognostic value, and copy number of these 10 genes were analyzed in hepatocellular carcinoma, combined with single-cell sequencing data. Secondly, univariate Cox regression analysis was performed on these 10 genes, and the 4 identified prognostic genes were further analyzed by consistency clustering. In addition, bioinformatics methods were used to explore their functional enrichment differences, immune cell infiltration, tumor immune microenvironment, and prognosis. Finally, a gene prognostic model and drug sensitivity analysis was established by screening common differential genes between clusters. These studies revealed the role and potential regulatory mechanisms of disulfidptosis-related genes in hepatocellular carcinoma.

2 Materials and Methods

2.1 Ten disulfidptosis-related genes and sample data sources.

Sample information was obtained from the TCGA-LIHC cohort (including 50 tumor and 374 liver cancer samples) as well as RNA-Seq data (in FPKM and Counts format), copy number variations, clinical data, and survival information using the TCGA data base(https://portal.gdc.cancer.gov/). The GSE162616
The dataset was procured from the GEO database (Home - GEO - NCBI (nih.gov)) for single-cell sequencing analysis. The 10 disulfidptosis-related genes were identified as SLC7A11, SLC3A2, RPN1, NCKAP1, NUBPL, NDUFA11, LRPPRC, OXSM, NDUFS1, and GYS1.

2.2 Differential analysis of the 10 disulfidptosis genes

Using R and the limma package, differential expression analysis was conducted on the samples from the TCGA-LIHC database (50 normal liver tissue samples and 369 liver cancer tissue samples). The analysis was performed with an FDR threshold of < 0.05, and the results were visualized using boxplots. Gene copy information was summarized using perl, and copy number variation frequency was calculated and visualized using circular plots.

2.3 Clustering analysis of prognosis-related disulfidptosis genes

In this study, 369 liver cancer patients were randomly divided into the training group and the validation group for univariate Cox regression analysis. Subsequently, a PCA plot was generated to demonstrate the disparities in sample distribution among the clusters, utilizing ConsensusClusterPlus package in R as the input information for consistency clustering analysis. The R package "survival" was used to analyze survival differences between each cluster. Furthermore, a heat map was generated to illustrate the disparities in clinical data and disulfidptosis gene expression between clusters.

2.4 Functional enrichment and differences in tumor immune microenvironment among disulfidptosis clusters

The Kyoto Encyclopedia of Genes and Genomes (KEGG) C2 dataset was downloaded to explore differential pathways among the clusters of disulfidptosis. The dataset contained gene sets from the Molecular Signatures Database (MSigDB) from the gsea-msigdb (http://www.gsea-msigdb.org) website. Subsequently, gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA) were performed. The ssGSEA algorithm was applied to ascertain disparities in the tumor immune microenvironment between the clusters, and the ESTIMATE algorithm was used to compute the matrix score, immune score, and tumor cell purity for each cluster.

2.5 Acquisition of public differential genes and establishment of prognostic models

Using the Limma package, differentially expressed genes between each pair of clusters were identified based on the criteria of |log2 Fold Change (FC)| > 0 and FDR < 0.05, and a volcano plot was generated. Differentially expressed genes that were common to the three clusters related to disulfidptosis were selected based on the criterion of |log2 Fold Change (FC)| > 1. Subsequently, the samples were divided into the training and validation sets. Using the glmnet package, LASSO regression analysis was conducted on the common differentially expressed genes in the training set. Six genes with significant prognostic ability were chosen to construct a prognostic model, and the risk score was calculated.
model was then validated using the validation set, and Kaplan-Meier curves and ROC curves were plotted to assess its predictive capability.

### 2.6 Consistent clustering and risk score correlation analysis

To investigate the relationship between three disulfidptosis clusters and risk scores, the ggplot and ggalluvial packages were used to draw Sankey diagrams and box plots. Furthermore, a heatmap was generated to display the expression of prognostic-related differential genes and the relationship between high-risk/low-risk groups and clinical data of samples.

### 2.7 Analysis of immune cell infiltration, differences in immune checkpoint gene expression and drug sensitivity between the high- and low-risk groups

The immune cell infiltration of each sample was calculated between the high- and low-risk groups, and differential and correlation analysis of immune cell infiltration was conducted. The expression of immune checkpoint genes between the high- and low-risk groups was analyzed and the results were presented in a boxplot. Moreover, the R package oncoPrdict was utilized for drug sensitivity analysis to investigate the association between risk score and drug sensitivity.

### 2.8 Single-cell data analysis

Single-cell data analysis was performed using the GSE166635 dataset through the online website http://tisch.comp-genomics.org/.

### 2.9 Software and statistics

This study utilized R (R version 4.2.1) and Strawberry Perl (Perl version 5.32.1.1) for data processing and analysis. A significance threshold of $P < 0.05$ was used for all statistical differences.

### 3 Results

#### 3.1 Differential expression, copy number variation, and prognostic correlation of genes related to disulfidptosis

As shown in the Fig. 1, the box plots display the expression levels of disulfidptosis-associated genes in normal liver tissue samples and HCC samples (Fig. 1A). Seven genes (GYS1, LRPPRC, NCKAP1, DUUFA11, RPN1, SLC3A2, and SLC7A11) exhibited differential expression in HCC and were significantly upregulated. In addition, univariate COX analysis was performed on the ten disulfidptosis-associated genes and a forest plot was generated (Fig. 1B). Four genes (LRPPRC, NCKAP1, RPN1, SLC7A11) showed significant prognostic ability for overall survival in HCC patients.
Next, the survival data of HCC patients were analyzed, and Kaplan-Meier survival analysis was performed based on the expression of disulfidptosis-associated genes (Fig. 1C). The results showed significant differences in the survival curves for all four disulfidptosis genes.

Additionally, the frequency of copy number variations (CNVs) of disulfidptosis-associated genes was calculated. The genes GYS1, NDUFA11, and RPN1 had decreased copy numbers, while NUBPL, NDUFS1, NCKAP1, and LRPPRC had increased frequencies of CNVs (Fig. 1D). The circular plot of gene CNVs also revealed the location of disulfidptosis-associated genes on human chromosomes.

### 3.2 The three disulfidome subtypes and their gene expression patterns

Consensus clustering analysis of the expression levels of the four prognostic disulfidoptosis genes was performed, and the samples were divided into three clusters (Fig. 2A). The PCA results of all HCC samples (Fig. 2B) showed almost no overlap between the three HCC clusters, and the components within each HCC cluster showed good correlation. The expression patterns of disulfidoptosis genes also differed significantly among the three clusters. Cluster B had relatively higher expression levels of disulfidoptosis-related genes, and the expression levels of SLC7A11 and SLC3A2 were significantly higher than those in clusters A and C.

### 3.3 Survival and gene expression differences among three subtypes

Survival analysis and gene expression differences were assessed on three subtypes of disulfidptosis, revealing that although there was no significant difference in survival between cluster A and cluster C, cluster B still exhibited a significantly poorer prognosis (Fig. 3A), which is consistent with the overall high expression of disulfidptosis genes in cluster B. Thereafter, HCC patient subtyping and clinical information were integrated, and volcano plots and heatmaps were generated to display the differential expression patterns of genes among the three clusters (Figs. 3B-C).

### 3.4 Functional and immune microenvironment differences among three subtypes of disulfidptosis in HCC

Furthermore, the functional and tumor immune microenvironment differences among the three subtypes of disulfidptosis were analyzed using GSVA and GSEA algorithms. The KEGG pathways enriched among the three subgroups were plotted into a heatmap, showing that the C subtype was more active in olfactory transduction, arachidonic acid metabolism, and neuroactive ligand receptor interaction pathways compared to the A and B subgroups (Fig. 4A). The GSEA analysis results (Fig. 4B) were consistent with the GSVA analysis, showing cell cycle pathway activation in the A and B subgroups, which was inhibited in the C subgroup.
Furthermore, the immune cell infiltration abundance was calculated using the ssGSEA algorithm and the StromalScore, ImmuneScore, and ESTIMATEScore using the ESTIMATE algorithm. The C subtype has higher ImmuneScore and ESTIMATEScore scores. The boxplot (Fig. 4C) showed the infiltration of immune cells in each cluster, with a higher proportion of immune cell infiltration in the C subtype, including activated B cells, activated CD8 T cells, eosinophils, MDSC, macrophages, and neutrophils. Significant differences in TME were observed among the three subtypes (Fig. 4D), with the C subgroup demonstrating higher immune scores and ESTIMATEScores.

3.5 Prognostic value and prognostic models of the hub DEGs among three disulfidptosis subtypes

Among the common differentially expressed genes among the three disulfidptosis subtypes, lasso regression was used to screen six survival-related genes, and a prognostic model was established (Fig. 5A). The Receiver Operating Characteristic curve was plotted (Fig. 5B), with the area under the curve (AUC) for both the training and validation groups greater than 0.65. These findings indicated that the prognostic model had good predictive ability in both groups. The risk scores for all samples were calculated, and significant differences in gene expression were found between the high-risk and low-risk groups (Fig. 5C). SAA4 was significantly up-regulated in the low-risk group, while PLCH1, SPP1, FLNC, SLC7A11, and DIRAS2 were significantly down-regulated compared to the high-risk group.

3.6 Association between risk scores and survival differences, gene expression patterns, and three disulfidptosis subgroups

In this study, Kaplan-Meier survival analyses were conducted for the training set, validation set, and overall samples. Figure 6A demonstrated that patients in the high-risk group had a significantly worse prognosis in the training set, validation set, and overall samples. The differences in gene expression patterns between the high- and low-risk groups in the training set, validation set, and overall samples were presented using a heatmap, which revealed consistent expression patterns of prognostic-related genes among these samples. Moreover, significant differences in risk scores were present among the three subtypes of disulfidptosis (Fig. 6C), with higher risk scores in subtype B, indicating a worse prognosis, which was consistent with the results of PCA survival analysis. Furthermore, the relationship between clustering, risk score, and prognosis status was illustrated by a Sankey diagram (Fig. 6D).

3.7 Relationship between risk score and immune cell infiltration, immune cell correlation, and immune checkpoints
Figure 7A demonstrates the immune cell infiltration in the different risk samples. The samples in the high-risk group showed significantly higher infiltration of activated CD4+ T cells, activated dendritic cells, natural killer cells, MDSC, and regulatory T cells than the low-risk group. In contrast, only monocyte and neutrophil infiltration were decreased. The high-risk score was positively correlated with most immune cell infiltrates (Fig. 7B). The expression of immune checkpoint genes in samples from different risk groups was calculated (Fig. 7C), and the correlation analysis results were shown in the box plot. The expression of immune checkpoint genes was mostly higher in samples from the high-risk group, which may predict that patients in the high-risk group are more sensitive to immune checkpoint inhibitors.

3.8 Drug sensitivity analysis of patients in high- and low-risk group

The sensitivity of patients from different risk groups to the chemical drugs was analyzed and box plots were generated (Fig. 8). Patients in the high-risk group were less sensitive to trametinib, dasatinib, and bortezomib and more likely to benefit from irinotecan, oxaliplatin, and niraparib.

4 DISCUSSION

The incidence of primary liver cancer rises year by year, and by 2025, more than one million people are predicted to be affected by liver cancer each year (6). Studies have shown that the immune system plays an essential role in controlling the development of hepatocellular carcinoma, with intrinsic and specific immunity interactions leading to tumor immune surveillance, and dysfunctional tumor-immune system interactions resulting in immune escape (7). Current treatment modalities for hepatocellular carcinoma mainly include surgical resection, radiofrequency ablation, transarterial embolization chemotherapy, and liver transplantation (8). In the past 10 years, significant advances in the development of immune checkpoint inhibitors (ICIs) for the treatment of progressive tumors (9) have revolutionized traditional treatment modalities. However, due to individual differences, some patients are not sensitive to immunotherapy, and the sensitivity of immunotherapy drugs and the clinical prognosis of such patients requires further study.

Seven disulfidptosis-related genes were differentially expressed and upregulated in hepatocellular carcinoma, and only four genes showed an association with survival prognosis, but the 10 disulfidptosis genes CNV were equally important in the regulatory mechanism of hepatocellular carcinoma. The analysis suggested that disulfidptosis-related genes may influence hepatocellular carcinoma prognosis by regulating the cell cycle, arachidonic acid metabolism, and participating in immune cell infiltration. Solute carrier family 3 member 2 (SLC3A2) and solute carrier family 7 member 11 (SLC7A11) are members of the solute carrier family (SLCs) and mediate the transmembrane transport of cellular metabolites and solutes. SLC3A2 and SLC7A11 interact mainly through polar and hydrophobic interactions in extracellular and transmembrane regions (10–13). Previous reports have shown that overexpression of SLC7A11 is associated with the development of various malignancies, such as esophageal, hepatocellular, gastric, and lung adenocarcinomas (14–19). Compared to normal liver
tissues, hepatocellular carcinomas are more dependent on SLC7A11-mediated cystine uptake and thus maintain intracellular cysteine redox homeostasis due to more rapid glucose metabolism (20, 21). Related studies confirmed that the knockdown of SLC7A11 expression in mice inhibited tumor cell growth. In addition, specific inhibition of SLC7A11 enhanced the efficacy of the immune checkpoint inhibitor anti-ctla-4 (22).

The samples were classified into three subtypes by unsupervised clustering based on four prognosis-related disulfidptosis genes. The three clusters differed significantly in gene expression patterns, survival prognosis, immune cell infiltration, and biological function. As previously described, increased expression of SLC7A11 and SLC3A2 might be associated with higher risk scores and poor prognosis in patients with subtype B. Moreover, immune cell infiltration was more abundant in subtype C. In terms of biological function, subtype C was significantly downregulated compared to subtypes A and B cell cycle pathways, which is consistent with the generally low expression of bisulfide death genes in subtype C and better clinical prognosis.

A prognostic model associated with disulfidptosis was developed and a risk score was calculated. Immune cell infiltration of activated CD4 T cells, MDSCs, and regulatory T cells was significantly higher in the high-risk group than in the low-risk group. Furthermore, MDSCs can exert pro-tumor and immunosuppressive effects in cancer by inducing differentiation and expansion of Treg and suppressing DCs, T cells and NK cells (23). The expression of SPP1, PLCH1, FLNC, SLC7A11, and DIRAS2 was also higher in samples from the high-risk groups. SPP1 encodes Osteopontin (OPN), which plays a key role in the immune response of different types of immune cells (24). Upregulation of the SPP1 gene expression is usually associated with inflammatory responses such as infections, allergic reactions, and tissue damage (25, 26). The results of single-cell data analysis were combined, which observed high SPP1 expression in dendritic cells. SPP1 enhances the antigen presentation efficacy of dendritic cells, thereby stimulating allogeneic T cells and Th1 polarization (27–29). Based on the significantly higher expression of immune checkpoint genes in samples from the high-risk group, the relationship between risk scores and patients' sensitivity to therapeutic agents was further analyzed. These findings may assist in developing new drug treatment modalities for hepatocellular carcinoma.

In summary, this study provides a multifaceted analysis of the expression differences, survival prognosis, immune cell infiltration, TME, and chemotherapeutic relevance of disulfidptosis-related genes in hepatocellular carcinoma based on the TCGA database. This study is the first to elucidate the regulatory mechanisms and biological roles of disulfide death-related genes in hepatocellular carcinoma. The model was validated using ROC curves, and K-M survival curves confirmed significant differences in patient survival prognosis between the high and low risk score groups. Nevertheless, the present study lacks in vivo and ex vivo experimental validation and included data from a single database source, which introduces limitations. Still, this retrospective bioinformatics-based study may provide new ideas to guide personalized immunotherapy strategies for HCC patients.

5 Conclusion
In this study, we identified the differential expression of disulfidptosis-related genes in hepatocellular carcinoma. By unsupervised clustering analysis we demonstrated significant survival differences in patients with three different subtypes of hepatocellular carcinoma. In addition, we developed a 6-gene prognostic model to predict the prognosis of HCC patients based on the common differential genes. Finally we showed the correlation of risk score with the expression of immune checkpoints and drug sensitivity.

**Declarations**

**Ethics approval and consent to participate:**

Not applicable

**Consent for publication:**

Not applicable

**Availability of data and materials**

The datasets generated and analysed during the current study are available in the TCGA database. (https://portal.gdc.cancer.gov/). The GSE162616 dataset was procured from the GEO database (Home - GEO - NCBI (nih.gov)) for single-cell sequencing analysis.

**Competing interests**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Funding**

Not applicable

**Author Contributions**

Ru-Tao Hong designed this work and revised this manuscript. Qiming Huang and Tiewen Li prepared, integrated and analyzed the data. Qiming Huang wrote this manuscript. Sisi Liu, Yemei Du, Yang Ma and Difei Chen collected the data. All authors have read, revised and approved the final manuscript.

**Acknowledgements**

Not applicable

**References**


Figure 1

The expression, prognostic value and CNV of ten Disulfidptosis related genes. (A) Differential expression of ten Disulfidptosis related genes in hepatocellular carcinoma; (B) Prognostic value of Disulfidptosis related genes in HCC; (C) KM survival curve for each Disulfidptosis related genes; (D) CNV of ten Disulfidptosis related genes in HCC.
Figure 2

Identification of clusters using the unsupervised clustering analysis. (A) Unsupervised clustering analysis; (B) PCA diagram showing the distribution of different clusters; (C) Differential expression of Disulfidptosis related genes between different clusters.
Figure 3

The features of clusters. (A) KM survival curve between each disulfidptosis clusters; (B) Differential gene expression between each disulfidptosis clusters; (C) heatmap showing the relationship between clinical features, genes expression and clusters.
Figure 4

Functional enrichment and differences in tumor immune microenvironment. (A-B) KEGG pathway were downloaded separately from the Msigdb database and the pathways were scored using the R package GSVA and GSEA; (C) Differences in immune cell infiltration between different clusters; (D) Differences between Stromal Score, Immune Score and ESTIMATE Score in different clusters.* represents p<0.05, ** represents p<0.01, *** represents p<0.001, ns represents p>0.05.
Figure 5

Build a risk model. (A) lasso regression analysis of the training set public difference genes using the R language glmnet package; (B) ROC curve of training group and test group; (C) Differential expression of Prognosis-related genes in different risk clusters.
Figure 6

Risk scores, survival differences, gene expression patterns in three clusters. (A) KM survival curve, Risk Curve, Survival scatterplot and heatmap of training group, test group and overall sample; (C) Risk score for three subtypes; (D) Sankey diagram showing the relationship between cluster, risk score and prognostic status.
Figure 7

The immune cell infiltration, immune checkpoint and single-cell sequencing. (A) Differences in immune cell infiltration between high and low risk groups; (B) Correlation of immune cell infiltration and risk scores; (C) Correlation of immune checkpoint gene expression and risk scores; (D) Expression of 6 prognosis-related genes in single-cell sequencing.
Figure 8

Analysis of drug sensitivity associated with risk score. Predicting IC50 values for multiple anti-cancer drugs.