Identification and Validation of an 11-Ferroptosis Related Gene Signature and Its Correlation With Immune Checkpoint Molecules in Glioma

Zhuohui Chen†, Tong Wu†, Zhouyi Yan and Mengqi Zhang*

Department of Neurology, Xiangya Hospital, Central South University, Changsha, China

Background: Glioma is the most common primary malignant brain tumor with significant mortality and morbidity. Ferroptosis, a novel form of programmed cell death (PCD), is critically involved in tumorigenesis, progression and metastatic processes.

Methods: We revealed the relationship between ferroptosis-related genes and glioma by analyzing the mRNA expression profiles from The Cancer Genome Atlas (TCGA), Chinese Glioma Genome Atlas (CGGA), GSE16011, and the Repository of Molecular Brain Neoplasia Data (REMBRANDT) datasets. The least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed to construct a ferroptosis-associated gene signature in the TCGA cohort. Glioma patients from the CGGA, GSE16011, and REMBRANDT cohorts were used to validate the efficacy of the signature. Receiver operating characteristic (ROC) curve analysis was applied to measure the predictive performance of the risk score for overall survival (OS). Univariate and multivariate Cox regression analyses of the 11-gene signature were performed to determine whether the ability of the prognostic signature in predicting OS was independent. Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were conducted to identify the potential biological functions and pathways of the signature. Subsequently, we performed single sample gene set enrichment analysis (ssGSEA) to explore the correlation between risk scores and immune status. Finally, seven putative small molecule drugs were predicted by Connectivity Map.

Results: The 11-gene signature was identified to divide patients into two risk groups. ROC curve analysis indicated the 11-gene signature as a potential diagnostic factor in glioma patients. Multivariate Cox regression analyses showed that the risk score was an independent predictive factor for overall survival. Functional analysis revealed that genes were enriched in iron-related molecular functions and immune-related biological processes. The results of ssGSEA indicated that the 11-gene signature was correlated...
with the initiation and progression of glioma. The small molecule drugs we selected showed significant potential to be used as putative drugs.

**Conclusion:** We identified a novel ferroptosis-related gene signature for prognostic prediction in glioma patients and revealed the relationship between ferroptosis-related genes and immune checkpoint molecules.

**Keywords:** ferroptosis, tumor immunity, prognosis, gene signature, glioma

**INTRODUCTION**

Glioma is the most common primary malignant intracranial tumor. Glioblastoma, the most malignant form (WHO grade IV glioma, GBM), has a 5-year survival rate of less than 5% (Ostrom et al., 2014; Gusyatiner and Hegi, 2018). Though low-grade gliomas have a better prognosis than glioblastomas, 70% of the patients inevitably develop into glioblastomas within 10 years, posing the importance of early diagnosis and risk assessment to improve the prognosis of gliomas (Kirang et al., 2019). The rapid progression, along with the highly heterogeneous nature of gliomas, makes prognostic prediction challenging. Standard treatment of gliomas involves observation, surgery, chemotherapy, and radiotherapy (Wang and Mehta, 2019). Despite significant advances in glioma management over the past decades leading to remarkable improvements in overall survival, treatment of gliomas remains a challenge because of heterogeneity, highly proliferative rate and the infiltrative nature of the tumor cells (Delgado-López et al., 2017; Li and Ding, 2017; Esparragosa et al., 2018). All these malignant biological features make gliomas highly recurrent and drug-resistant. Recently discovered biomarkers indicate improved survival and specific antitumor treatment (Ostrom et al., 2014; Esparragosa et al., 2018). Numerous clinical trials targeting these molecule markers for glioma therapies have been carried out, but few have ultimately succeeded. Therefore, identifying novel and effective prognostic models and drug targets is an urgent and critical task not only for glioma management, but also for drug discovery.

Ferroptosis, first proposed by Dixon in 2012, is a newly discovered type of programmed cell death (PCD) that occurs through Fe(II)-dependent lipid peroxidation due to insufficient cellular reducing capacity (Dixon et al., 2012; Fearnhead et al., 2019; Doll et al., 2019; Hassannia et al., 2015) and TCGA-GBM (Brennan et al., 2013) combined (n = 703) were used to find aberrantly expressed genes between cancer and normal tissue, based on previously published ferroptosis related genes in the CGGA dataset. We further performed the functional annotation to explore the underlying mechanisms. Finally, we selected several small molecule drugs as potential therapeutic target for glioma.

**MATERIALS AND METHODS**

The flow chart of data collection and analysis is shown in Figure 1.

**Data Acquisition**

In our current study, we collected four public cohorts containing RNA-seq data (expression matrix has been transformed to TPM) and clinical information of patients obtained from The Cancer Genome Atlas (TCGA)\(^1\), Chinese Glioma Genome Atlas (CGGA)\(^2\) database, GEO\(^3\), and GlioVis\(^4\). TCGA-LGG (Brat et al., 2015) and TCGA-GBM (Brennan et al., 2013) combined (n = 703) were used to find aberrantly expressed genes between cancer and normal tissue, based on previously published ferroptosis related genes, and to construct the prognostic signature model (Stockwell et al., 2017; Bersuker et al., 2019; Doll et al., 2019; Hassannia et al., 2019). Basically, we chose the mentioned protein-coding

---

**Abbreviations:** TCGA, The Cancer Genome Atlas; CGGA, Chinese Glioma Genome Atlas; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; OS, overall survival; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; ssGSEA, single sample gene set enrichment analysis; GBM, glioblastoma; PCD, programmed cell death; DEGs, differentially expressed genes; PRS, primary/recurrent/secondary; C-index, concordance index; CMap, Connectivity Map; PPI, protein–protein interaction; PCA, Principal component analysis; t-SNE, t-distributed stochastic neighbor embedding; AUC, area under the curve; HR, Hazard ratio; CI, Confidence interval; dDC, Activated dendritic cell; DC, dendritic cell; iDC, Immature dendritic cell; NK, natural killer; pDC, Plasmacytoid dendritic cell; Th, T helper cell; Th1, T follicular helper cell; Th2, T helper; TIL, Tumor Infiltrating Lymphocyte; Treg, T regulation; APC, Antigen presenting cell; CCR, Cytokine–cytokine receptor; HLA, Human leukocyte antigen; MHC, major histocompatibility complex; IFN, interferon; OSCC, oral squamous carcinoma; MT, metallothionein; HCC, hepatocellular carcinoma.

---

\(^1\)https://portal.gdc.cancer.gov/
\(^2\)http://www.cgga.org.cn/
\(^3\)https://www.ncbi.nlm.nih.gov/geo/
\(^4\)http://gliovis.bioinfo.cnio.es/
genes from these studies and combined them as ferroptosis related genes. The retrieved genes were listed in Supplementary Table 1. CGGA dataset (Zhao et al., 2017) \((n = 325)\), GSE16011 (Gravendeel et al., 2009) \((n = 276)\) and Repository of Molecular Brain Neoplasia Data (REMBRANDT) (Madhavan et al., 2009) \((n = 444)\) were used as validation cohorts to assess the efficacy of our gene signature model.

**Identification of Differentially Expressed Genes (DEGs)**

DEGs were identified by comparing the mRNA expression of 60 ferroptosis related genes between tumor and normal tissue in TCGA datasets using Limma package version 3.44.3 in R software version 4.0.3. The information of sample sources (TCGA-LGG or TCGA-GBM) was included as covariates during the analysis. Genes with \(p < 0.05\) were selected for further analysis. The DEGs between low- and high-risk group were also identified in both TCGA and CGGA cohorts after the calculation of the risk score using Limma package version 3.44.3 in R software version 4.0.3. Genes with \( \text{fdr} < 0.05 \) and \(|\log \text{ FC}| > 1\) were selected for further analysis.

**Gene Correlation Analysis**

The protein-protein association analysis of DEGs was performed based on the STRING database version 11.0 (von Mering et al., 2005)\(^5\). Genes having no predictive interaction with other DEGs were not presented in the final figure.

\(^5\)http://string.embl.de/
Identification and Validation of Prognostic Gene Signature

Univariate Cox regression analysis was performed to determine the genes significantly associated with overall survival (OS) in TCGA datasets. Due to the existence of missing value in overall survival and survival status, only 665 patients were included in the model construction in TCGA datasets. The overlapping genes between DEGs and clinically associated genes were identified using the Venn diagram. After that, the overlapping genes were sent to further construct a ferroptosis-associated gene signature by using the least absolute shrinkage and selection operator (LASSO) Cox regression analysis with the aid of glmnet package version 4.0-2 (Tibshirani et al., 2012) in R software version 4.0.3. Based on the prognostic gene signature, the risk score for each patient can be calculated as followed:

\[
\text{riskScore} = \sum_{i=1}^{n} \text{Coef}(X_i) * \text{Exp}(X_i)
\]

In formula (1), \(\text{Coef}(X_i)\) represented the coefficient of each ferroptosis-related genes \(X_i\), and \(\text{Exp}(X_i)\) represented the expression levels of these genes. The calculated risk score divided all the patients into low- or high-risk using median risk score as a cutoff. The risk score of patients from CGGA datasets can also be calculated to validate the efficacy of the prognostic gene signature.

Gene Enrichment Analysis

To functionally annotate differentially expressed gene sets during the analysis, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed in R software version 4.0.3 using ClusterProfiler version 3.16.1 (Yu et al., 2012) package. PFAM Protein Domains, INTERPRO Protein Domains and Features enrichment were performed directly in the STITCH database version 5.0 (Szklarczyk et al., 2016). To estimate the immune cell infiltration and immune function status in high-risk patients vs. low-risk ones, single sample gene set enrichment analysis (ssGSEA) was performed using GSVA version 1.36.3 (Hänzelmann et al., 2013) package in R software version 4.0.3. In short, the immune-related gene set enrichment score of each patient was first calculated. Then the patients were divided into high- or low-risk groups based on the formerly mentioned cutoff, after which the immune status was compared between the two groups.

Nomogram Construction

The independent clinical factor validated by univariate and multivariate Cox regression analysis were enrolled to construct a nomogram for prognosis prediction, which included primary/recurrent/secondary (PRS) type of glioma, grade, 1p19q codeletion status and risk score. Patients with missing data were excluded from the analysis and thus only 275 patients were included in the univariate and multivariate Cox regression analysis. Package Rms version 6.1-0 was utilized to perform the construction and calculate the concordance index (C-index) to evaluate the model efficacy in prognosis prediction. The closer its value is to 1, the better the performance. In addition, calibration curves for 1-, 2-, and 3-year prediction were plotted to assess the consistency between predicted and actual survival.

Candidate Small Molecule Drugs Analysis and Downstream Target Molecule Identification

The Connectivity Map (CMap) database version build 02\(^7\) was used to predict the putative drugs targeting DEGs in our present analysis. CMap database (Lamb et al., 2006) can be utilized to explore functional links between disease, genetic interference and drug action. The enrichment scores ranging from –1 to 1 were calculated for each putative drugs. The negative enrichment score of a drug represented the reversing effects on the input gene set, thus indicating its anti-tumor capacity when it comes to cancer-related gene set. Besides, percent non-null represented the percentage of meaningful results obtained in the whole n times experiments conducted by CMap database. The small molecule compounds were chosen with \(p < 0.05\) and enrichment scores < –0.85. Subsequently, all the small molecule compounds with \(p < 0.05\) were collected and analyzed in the STITCH database version 5.0 (see text footnote 6) to identify the target proteins and mechanism of action. The STITCH database (Szklarczyk et al., 2016) is a platform for searching known and predicted interactions between drug compounds and proteins. The interactions between drug compounds and proteins are verified through experiments, databases, and studies in the literature.

Statistical Analysis

All the data were analyzed using the R software version 4.0.3\(^8\). Kaplan-Meier analysis was performed to compare the overall survival curves between different patient groups, with a log-rank test to evaluate the statistical significance. Spearman’s rank correlation coefficient was calculated to evaluate the linear correlation between risk score and the expression of immune checkpoint related genes. Kruskal-Wallis H-test was used to compare difference between groups. \(p < 0.05\) was considered statistically significant for all the analyses.

RESULTS

Identification of 25 Prognostic Ferroptosis-Related DEGs in the TCGA Cohort

A total of 703 GBM patients from two TCGA cohorts (including TCGA-GBM and TCGA-LGG) and 325 GBM patients from CGGA cohort were included in the analysis. To evaluate the expression differences of ferroptosis-related genes between tumor tissues and adjacent normal tissues, we analyzed RNA-seq data from TCGA dataset. Subsequently, 27 of 60 ferroptosis-related

\(^7\)https://portals.broadinstitute.org/cmap
\(^8\)https://www.R-project.org/
genes were selected, and 25 of them were related to patient survival (Figure 2A). Among the 25 overlapping genes, HSBP1, FANCD2, PGD, SAT1, CD44, SLC1A5, LPCAT3, NFE2L2, ACO1, ALOX12, ZEB1, TP53, KEAP1, PEBP1, FADS2, AKR1C3, and CRYAB were upregulated, and AKR1C2, ACSI4, CISD1, GLS2, GOT1, MT1G, CHAC1, and PTGS2 were downregulated in tumor tissues (Figure 2B). The results of univariate Cox regression analysis between gene expression and OS is shown in Figure 2C (all p < 0.05). The protein-protein interaction (PPI) network and functional analysis of these genes indicated that TP53 and PTGS2 were the hub genes (Figure 2D). The correlation of candidate genes is presented in Figure 2E.

Construction of a Gene-Based Prognostic Model in the TCGA Cohort

To evaluate the risk of each patient, LASSO-Cox regression analysis was applied to establish a gene-based prognostic model using the expression profile of the 25 genes mentioned above. We identified an 11-gene signature based on the LASSO regression with the optimal value of λ and performed survival analyses according to the optimal cut-off expression value of each gene (Supplementary Figure 1). The results indicated that high expression of CD44, FANCD2, HSBP1, MT1G, NFE2L2, SAT1, and low expression of AKR1C3, ALOX12, CRYAB, FADS2, and ZEB1 correlated with a poor prognosis. The calculated coefficient was exhibited in Table 1. The distribution of risk scores in the TCGA dataset is presented in Figure 3A. The patients were divided into a high-risk group and a low-risk group based on the median cut-off value. Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) analysis were used to separate between the two different risk groups (Figures 3B,D). The result indicated that patients in high-risk

| Gene     | Coefficient |
|----------|-------------|
| AKR1C3   | −0.120098473470357 |
| ALOX12   | −0.0122009713910145 |
| CD44     | 0.214177710154476 |
| FANCD2   | 0.684183489406273 |
| CRYAB    | −0.104571744619555 |
| MT1G     | 0.0523002991814781 |
| SAT1     | 0.17208169682249 |
| HSBP1    | 0.0301696571931318 |
| ZEB1     | −0.37149858625237 |
| FADS2    | −0.037897482917486 |
| NFE2L2   | 0.452805076174773 |
group and low-risk group were distributed in discrete directions. The OS status in the high-risk group was significantly poorer than that in the low-risk group (Figure 3C). Consistently, Kaplan-Meier survival curve in the TCGA cohort indicated that a high-risk score was correlated with a worse prognosis (Figure 3E, $p < 0.001$). Time-dependent ROC curve analysis was applied to measure the predictive performance of the risk score for OS. The signature of the 11 ferroptosis-related genes exhibited remarkable prognostic validity, with the area under the curve (AUC) reaching 0.879 ($0.843-0.915$, 95%CI) at 1 year, 0.903 ($0.871-0.935$, 95%CI) at 2 years, and 0.919 ($0.884-0.954$, 95%CI) at 3 years (Figure 3F).
Validation of the Signature of 11 Ferroptosis-Related Genes in the CGGA, GSE16011, and REMBRANDT Cohorts

To test the robustness of the model constructed by the TCGA cohort, we performed prognostic analyses in the CGGA, GSE16011, and REMBRANDT validation cohorts. In the CGGA cohort, the patients were also divided into high-risk and low-risk groups according to the median value with the same calculation formula as the TCGA cohort (Figure 4A). Likewise, PCA and t-SNE analysis in the CGGA cohort confirmed that patients in high-risk group and low-risk group were distributed in two directions (Figures 4B,D). Similar to the results from the TCGA cohort, patients with a high-risk score were more likely to have a significantly shorter OS and poorer prognosis in the CGGA dataset (Figures 4C,E). The AUC of the signature was 0.790 (0.741–0.839, 95%CI) at 1 year, 0.875 (0.835–0.915, 95%CI) at 2 years, and 0.878 (0.836–0.919, 95%CI) at 3 years (Figure 4F). The GSE16011 cohort (Supplementary Figure 2) and the REMBRANDT cohort (Supplementary Figure 3) exhibited a pattern similar to the CGGA cohort. Heat maps

![Heat maps](image-url)
showed clinical and molecular features and different expression levels of 25 selected genes using hierarchical clustering in the TCGA dataset (Supplementary Figure 4) and CGGA dataset (Figure 5). Consistent with the calculated risk score, patients were roughly clustered into two groups by risk score. In the CGGA dataset, with an increase in risk score, the expression levels of CRYAB, AKR1C3, FADS2, AKR1C2, PEBP1, CISD1, GLS2, and GOT1 were downregulated; the expression levels of SAT1, SLC1A5, CD44, NFE2L2, ACSL4, PTGS2, CHAC1, MT1G, ALOX12, HSBP1, KEAP1, FANCD2, PGD, TP53, ZEB1, LPCAT3, and ACO1 were upregulated. Clinical and molecular features, such as methylated MGMTp, 1p19q non-codeletion, IDH wild types, recurrent/secondary tumor types were enriched in high-risk-score samples. In the TCGA dataset, there are only 443 samples with both clinical data and RNAseq data, so blank and missing data existed when clinical parameters were added to the heatmap. The incomplete clinical data may not be completely random, leading to bias in the clinical correlation analysis. The result in the TCGA dataset was basically consistent with CGGA dataset. These results indicated that the risk score of the ferroptosis-related gene signatures positively correlated with glioma.

Independent Prognostic Value of the 11-Gene Signature

Univariate and multivariate Cox regression analyses of the 11-gene signature were performed in the TCGA, CGGA, GSE16011, and REMBRANDT datasets to determine whether the ability of the prognostic signature in predicting OS was independent. The results of univariate Cox regression analyses determined that the risk score was significantly related to OS in both the TCGA cohort and the CGGA cohort (TCGA cohort: $HR = 3.107$, $95\% CI = 2.506$–$3.853$, $p < 0.001$; CGGA cohort: $HR = 1.943$, $95\% CI = 1.737$–$2.174$, $p < 0.001$) (Figures 6A,B). In multivariate Cox regression analyses, the risk score still proved to be an independent predictive factor for OS (TCGA cohort: $HR = 1.568$, $95\% CI = 1.100$–$2.235$, $p < 0.05$; CGGA cohort: $HR = 1.651$, $95\% CI = 1.415$–$1.926$, $p < 0.001$) (Figures 6A,B). These consistent results were also validated in the GSE16011 and REMBRANDT datasets (Supplementary Figure 5). Based on the independent prognostic parameters for the OS in the TCGA dataset, we constructed a nomogram to predict 1, 2, and 3-year survival (Figure 6C). Besides, the calibration curve for the probability of 1, 2, and 3-year OS showed an optimal
FIGURE 6 | Independent prognostic value of the 11-gene signature in the TCGA and CGGA cohorts and construction of the predictive nomogram from the CGGA cohort. Univariate and multivariate Cox regression analyses of the signature in the TCGA derivation cohort (A) and the CGGA validation cohort (B). (C) A nomogram of the 11-gene signature for predicting 3-year survival in the TCGA dataset.

agreement between observation and prediction in the TCGA dataset (Supplementary Figure 6). The C-index was calculated as 0.83 after bias correction, showing relatively high performance in clinical diagnosis. Therefore, these results indicated that the 11-gene signature may serve as a potential diagnostic factor in glioma patients. Meanwhile, the above findings also provided us with useful information that upregulated expression of ferroptosis-related genes has an impact on the prognosis of glioma.

**Functional Annotation of the 11-Gene Signature**

To identify the potential biological functions and pathways of the 11-gene signature, the DEGs between the high-risk group and the low-risk group were used to conduct GO analysis and KEGG pathway analysis. Based on GO analysis, the DEGs were enriched in iron-related molecular functions, such as gated channel activity in the TCGA and CGGA cohorts, and channel activity and ion channel activity in
the CGGA cohort. In addition, DEGs in the TCGA cohort were highly enriched in several immune-related biological processes, such as adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains, leukocyte migration, interferon-gamma-mediated signaling pathway, response to interferon-gamma, regulation of immune effector process, neutrophil activation and neutrophil-mediated immunity (Figure 7A). Two immune-related biological processes were validated in the CGGA cohort, including leukocyte migration and adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains (Figure 7C). KEGG pathway analysis indicated that DEGs were enriched in immune-related pathways, which was consistent with the results of GO analysis. The terms included phagosome, complement and coagulation cascades, allograft rejection, human T-cell leukemia virus 1 infection, antigen processing and presentation, cell adhesion molecules and graft-vs.-host disease in both the TCGA dataset (Figure 7B) and the CGGA dataset (Figure 7D). Besides, several cancer-related terms were included in the two datasets, such as proteoglycans in cancer and Epstein-Barr virus infection in the TCGA and CGGA datasets, and human papillomavirus infection in the TCGA dataset. The aforementioned results indicate that the 11-gene signature highly correlates with cancer progression, particularly by affecting immune-related functions.

The Ferroptosis Related Gene Signature Is Highly Correlated With Immune Function and Immune Checkpoint Molecules in Glioma

To further explore the correlation between risk scores and immune status, ssGSEA was performed to quantify the enrichment scores of different immune cell subpopulations and related functions or pathways in the TCGA (Figures 8A,B), CGGA (Figures 8C,D), GSE16011 (Figures 8E,F), and REMBRANDT (Figures 8G,H) cohorts. As expected, contents of immune response had higher scores in the high-risk group, including macrophages, pDCs, T helper cells, TIL, and Treg in all datasets. The score of NK cells was lower in the high-risk group in the TCGA, GSE16011, and REMBRANDT datasets, while exhibited the opposite in the CGGA dataset. Moreover, the scores of DEGs in immune-related biological processes or molecule functions were statistically different between two risk groups in all cohorts, including APC co-inhibition, APC co-stimulation, CCR, check-point, cytolytic activity, HLA, inflammation-promoting, MHC class I, parainflammation, T cell co-inhibition, T cell co-stimulation, type I IFN response, type II IFN response. Subsequently, we calculated Spearman’s rank correlation coefficient to evaluate the linear correlation between the expression of immune checkpoint related genes (CD274, CD276, CTLA4, HAVCR2, LAG3, and PDCD1) and risk scores in the TCGA, CGGA (Supplementary Figure 7), GSE16011, and REMBRANDT (Supplementary Figure 8) datasets. The results
indicated that the expression of immunosuppression-related genes had positive correlation with risk scores in all cohorts. Furthermore, in order to preliminarily determine which of the 11 genes may be the most related to immune response, we performed the correlation analysis for each ferroptosis-related gene and immune checkpoint gene (| r | > 0.4) (Supplementary Figure 9). The results revealed that key genes in the TCGA dataset were CD44, FANCD2 and SAT1, and key genes in the CGGA dataset were AKR1C3, ALOX12, CD44, SAT1, and NFE2L2. The common critical genes in the two datasets were CD44 and SAT1. A previous study has revealed that CD44 was identified as a key positive regulator of PD-L1 expression in triple-negative breast cancer and non-small cell lung cancer (Kong et al., 2020). SAT1 has been reported to contribute to p53-mediated reactive oxygen species (ROS) response and ferroptosis (Ou et al., 2016). Therefore, it can be speculated that CD44 and SAT1 may regulate immune checkpoints to involve in tumor immunity. Collectively, the results above indicated that the 11-gene signature was correlated with the initiation and progression of glioma.

### Drug Screening Using the 11-Gene Signature With CMap and STITCH Database

In order to test whether the selected 11 ferroptosis-related genes are good candidates for therapy target, CMap (the connectivity map) analysis was performed to screen small molecule drugs using the selected prognostic genes as up signature or down signature. Of the 27 prognostic genes, 18 genes with positive cox coefficient were set as the up-regulated signature and the other nine genes were set as the down-regulated signature. Seven candidate small molecule drugs (butacaine, CAY-10397, BAS-012416453, PHA-00816795, STOCK1N-35696, piperlongumine, sanguinarine) with potential value were identified with p < 0.05 and enrichment score <−0.85 (Table 2). Among the seven small molecule drugs, butacaine and CAY-10397 (p < 0.001) showed significantly negative correlation with selected mRNAs and potential to be used as putative drugs. To understand the correlation between candidate drugs and immune checkpoints, the network of interacting proteins and candidate drugs with p < 0.05 was constructed in the STITCH database (see text footnote 6) (Supplementary Figure 10). The results showed that the direct downstream target genes of candidate drugs included HDAC1, ESR2, CYP19A1, ESR1, AR, ADRA1A, PGR, ESRRA, NR3C1, and KCNJ11. We further performed GO, PFAM Protein Domains, INTERPRO Protein Domains and Features enrichment analyses on the above genes, and found that steroid hormone related functions were enriched, especially estrogen receptor (ER) related functions (Supplementary Table 2).

A previous study has revealed that in cells with low ER expression, ferroptosis was easier to be triggered by sulfasalazine in breast cancer cells (Ou et al., 2016). Besides, the correlation between ER and tumor immunity has been also extensively researched (Kurozumi et al., 2019; Romero et al., 2020; Smida et al., 2020). Therefore, the candidate drugs targeting estrogen receptor are likely to involve in the downstream tumor immunity through ferroptosis.

### Table 2 | Drugs selected by CMap.

| CMap name         | n | Enrichment | p     | Percent non-null |
|-------------------|---|------------|-------|------------------|
| Butacaine         | 4 | −0.881     | 0.00048 | 100              |
| CAY-10397         | 3 | −0.92      | 0.00092 | 100              |
| BAS-012416453     | 3 | −0.886     | 0.00288 | 100              |
| PHA-00816795      | 2 | −0.956     | 0.00421 | 100              |
| STOCK1N-35696     | 2 | −0.884     | 0.02696 | 100              |
| Piperlongumine    | 2 | −0.876     | 0.03078 | 100              |
| Sanguinarine      | 2 | −0.871     | 0.03306 | 100              |

FIGURE 8 | The ssGSEA results of different risk groups in the TCGA cohort (A,B), CGGA cohort (C,D), GSE16011 cohort (E,F), and REMBRANDT cohort (G,H). The scores of 16 immune cells (A,C,E,G) and 13 immune-related functions (B,D,F,H) were shown in boxplots. Adjusted p were showed as: ns, not significant; *p < 0.05; **p < 0.01; ***p < 0.001.
DISCUSSION

Glioma is the most common primary malignant brain tumor characterized by rapid progression and treatment resistance, causing significant mortality and morbidity (Gusyatiner and Hegi, 2018; Liu et al., 2020). Ferroptosis is a novel form of PCD. Previous studies have identified the critical role of ferroptosis in tumorigenesis and therapies (Liu et al., 2020). In the current study, we used comprehensive bioinformatics analysis to investigate variations in expression profiling of 60 ferroptosis-related genes in glioblastoma and their association with OS. We identified the signature of 11 ferroptosis-related genes associated with progression and prognosis of GBM patients and validated the novel prognostic model in an external cohort. Functional analyses indicated that immune-related biological processes were highly enriched. In addition, we selected several small molecule compounds as potential therapeutic drugs.

The 11 prognostic genes consist of AKR1C3, ALOX12, CD44, CRYAB, FADS2, FANCD2, HSBP1, MT1G, NFE2L2, SAT1, and ZEB1. Previous studies have illustrated that most of these ferroptotic genes are critically involved in tumorigenesis, including glioma. AKR1C3 involves in steroids, prostaglandins and lipid aldehydes metabolism and plays a role in tumorigenesis of breast carcinoma, endometrial carcinoma and prostate carcinoma (Park et al., 2010). ALOX12, a hotspot of monoallelic deletion in cancers, plays an important role in p53-mediated ferroptosis. ALOX12 missense mutations diminishes polyunsaturated fatty acids oxidation and p53-mediated ferroptosis (Chu et al., 2019). CD44 is a single-pass type I transmembrane protein that has shown to be closely related to tumor development. Ferroptotic responses correlate with the transcriptional regulation of SLC7A11, a key component of the cystine-glutamate antiporter. The level of CD44 can regulate the sensitivity of tumor cells to ferroptosis and the stability of SLC7A11, a key component of the cystine-glutamate antiporter related to ferroptosis regulation (Liu et al., 2019). CRYAB, secreted via exosomes, has been reported to be upregulated in GBM and exert anti-apoptotic activity (Kore and Abraham, 2014). FADS2 has been shown to overexpress in colorectal cancer and facilitate cancer cell proliferation by increasing the metabolism of PGE2, an oncogenic molecule associated with colorectal cancer tumorigenesis (Tian et al., 2020). Previous studies have demonstrated that FANCD2 is overexpressed in high-grade gliomas and depletion of FANCD2 may serve as a potential strategy for the treatment high-grade gliomas (Metselaar et al., 2019). HSBP1 expression has reported to be elevated in oral squamous carcinoma (OSCC) and increased HSBP1 expression enhances the sensitivity of OSCC cells in radiation (Shen et al., 2014). MT1G, a member of metallothioneins (MTs), is frequently downregulated in HCC, which can be regarded as an early event in HCC progression (Ji et al., 2014). NFE2L2 is a redox-sensitive transcriptional factor mainly located in cytoplasm. The expression of NFE2L2 has been shown to positively correlate with the expression of immune checkpoint markers in brain lower grade glioma (Ju et al., 2020). A previous study has demonstrated that SAT1 causes resistance to radiation in GBM through an shRNA screen. SAT1 also involves in cell migration, proliferation and tumor growth (Thakur et al., 2019). ZEB1 has been reported to increase in gliomas and positively correlate with tumor progression (Chen et al., 2017).

We further demonstrated that a high-risk score was associated with worse prognosis. Time-dependent ROC curve of the signature of 11 ferroptosis-related genes predicted patient OS. To investigate the correlation between the risk signature and glioma grade, we further studied the levels of risk score stratified by the grade in the TCGA, CGGA, GSE16011, and REMBRANDT cohorts (Supplementary Figure 11). The risk score increased as the grade of glioma increased. In the TCGA and CGGA datasets, WHO grade IV patients had the highest increase in the risk score, while WHO grade II patients had the lowest increase in the risk score. Patients with WHO grade III were assigned a medium risk score in both the TCGA and CGGA datasets ($p < 0.001$). In the GSE16011 and REMBRANDT datasets, due to the limited samples of patients with WHO grade I and II, the results of WHO grade I and II exhibited no statistical difference. Therefore, we reassigned the samples to WHO grade I-III and WHO grade IV, and the results were statistically different. Besides, we also plotted Kaplan-Meier curves for glioma patients with low and high risk scores classified as WHO grade II to IV in the TCGA, CGGA, GSE16011, and REMBRANDT datasets (Supplementary Figure 12). In consideration of the limited samples of patients with Grade I and II in the GSE16011 and REMBRANDT datasets, we also plotted Kaplan-Meier curves of Grade I-III. The results showed that patients with high risk had significantly shorter OS than patients with low risk in WHO grade II, WHO grade III, and WHO grade IV groups. Exceptionally, the survival plot between low- and high-risk in WHO IV patients from TCGA and REMBRANT cohorts showed no significance because there were almost no patients with WHO IV were low-risk. Similarly, the survival plot between low- and high-risk in WHO II patients from CGGA and GSE16011 cohorts showed no significance, also because there were almost no patients with WHO II were high-risk. In addition, the prognostic values based on the 11-gene signature were independent of other clinical variables, including grade, age, radiotherapy status, IDH mutation status. It can be noticed that differences existed in the prognostic value for grade, age, and IDH status in the four datasets. It is probably due to the data of the four datasets coming from people in different regions. To adjust the differences caused by populations, it requires a wider range of multi-center clinical verification. Besides, the clinical data in the TCGA cohort are incomplete, some of which have missing value. The lack of data may not be completely random, leading to bias in the clinical correlation analysis. This is also the limitation of using public data to conduct analysis. Among the 11 genes, functional analysis showed that these genes were involved in immune-related biological processes and pathways. It can be reasonably assumed that ferroptosis may be critically involved in tumor immunity. To further explore potential mechanism of the signature of 11 ferroptosis-related genes, ssGSEA analysis was performed. The enrichment scores of macrophages, pDCs, T helper cells, TIL and Treg were statistically different between two risk groups and exhibited a similar pattern in the TCGA, CGGA, GSE16011, and REMBRANDT datasets. The scores of macrophages were
studies have demonstrated that B7-H3 is highly overexpressed transcriptional factor activity (Zhang et al., 2009). Numerous inhibitor associated with diminished NFAT, NF-κB families (Picarda et al., 2016). Induced on antigen-presenting and 0.59 in the REMBRANDT cohort. B7-H3 (CD276) is a cohort, 0.81 in the CGGA cohort, 0.57 in the GSE16011 cohort with the regression coefficient reaching 0.73 in the TCGA correlation with risk scores in all cohorts. Notably, the expression GSE16011, and REMBRANDT cohorts. The results indicated that to immunosuppression and the risk scores in the TCGA, CGGA, analyzed the correlation between the expression of genes related play a dominant role in the immune response. Therefore, we possible speculation is that the effect of immunosuppression may inhibit in both cohorts are higher in the high-risk group. One scores of Treg cells and the activity of T cell co- exhaustion and inhibit pDC-derived IFN-I. In addition to the protective functions of T helper populations, they are also involved in the pathogenesis of chronic inflammatory disorders (Cosmi et al., 2014). Tumor infiltrating lymphocytes (TIL) were previously reported to play a predictive role in mediating response to chemotherapy and increasing overall survival, while the increase in immunosuppressive regulatory T-cells (Tregs) has correlation with poor prognosis (Santoemma and Powell, 2015; Stanton and Disis, 2016). Given that high-risk group exhibited poorer prognosis, we speculated that patients with high risk may suffer pDC functional exhaustion and weakened anti-tumor immunity. Besides, the inflammatory effect of T helper cells and immunosuppression of Tregs play a dominant role in the tumor immunity, rather than the protective functions of T helper populations or TIL. Moreover, the high-risk groups in the TCGA, GSE16011, and REMBRANDT cohorts had lower fractions of NK cells, indicating that impaired antitumor immunity in patients of high-risk group may contribute to their poor prognosis. The score of NK cells between two groups in the CGGA cohort was the opposite. We speculated that the result may be due to limited samples and individual differences of the patients in the CGGA cohort. Interestingly, although high-risk groups have higher scores of functions related to antitumor immunity in both cohorts, including the activity of the type I IFN response and type II IFN response, the prognosis of the high-risk group still proves poorer than that of the low-risk group. We also noticed that cells or functions associated with immunosuppression such as the fractions of Treg cells and the activity of T cell co-inhibition in both cohorts are higher in the high-risk group. One possible speculation is that the effect of immunosuppression may play a dominant role in the immune response. Therefore, we analyzed the correlation between the expression of genes related to immunosuppression and the risk scores in the TCGA, CGGA, GSE16011, and REMBRANDT cohorts. The results indicated that the expression of immunosuppression-related genes (CD274, CD276, CTLA4, HAVCR2, LAG3, and PDCD1) had a positive correlation with risk scores in all cohorts. Notably, the expression of B7-H3 exhibited a significant correlation with risk scores, with the regression coefficient reaching 0.73 in the TCGA cohort, 0.81 in the CGGA cohort, 0.57 in the GSE16011 cohort and 0.59 in the REMBRANDT cohort. B7-H3 (CD276) is a critical immune checkpoint molecule belonging to B7-CDS8 families (Picarda et al., 2016). Induced on antigen-presenting cells, B7-H3 was previously shown to act as a T cell co-inhibitor associated with diminished NFAT, NF-κB and AP-1 transcriptional factor activity (Zhang et al., 2009). Numerous studies have demonstrated that B7-H3 is highly overexpressed in human malignancies and correlates with negative prognosis and poor clinical outcome, including glioma (Wang et al., 2018; Zheng et al., 2019), HCC (Zheng et al., 2019), pancreatic cancer (Inamura et al., 2018), ovarian carcinoma (Zang et al., 2010), colorectal cancer (Ingebritsen et al., 2014), and bone cancer (He and Li, 2019). Therefore, B7-H3 may serve as an attractive target for immunotherapy against cancers. The immune checkpoint molecules, CTLA4 and PDCD1 (PD-1), send a negative signal to T cells, thereby suppressing effector T-cell responses (Dyck and Mills, 2017). Immunotherapy targeting immune checkpoints CTLA4 and PD-1/PD-L1 (CD274) have been used in cancer patients (Chen and Mellman, 2017; Lee et al., 2017). However, only a small portion of patients exhibit durable responses and a large number of cancers such as colorectal cancer remain largely refractory to the therapy (Chen and Mellman, 2017; Das et al., 2017). HAVCR2 (TIM-3), highly expressed on innate cells, acts as a negative regulator of Th1 and CTL responses and critically involves in tumor growth (Das et al., 2017; Joller and Kuchroo, 2017). LAG3 is also a potential target for cancer immunotherapy due to its negative regulatory role on T cells (Andrews et al., 2017). Therefore, we might suppose that the signature of 11 ferroptosis-related genes was critically involved in tumorigenesis and progression of glioma probably by regulating immune-related biological processes and pathways, which correlated with the poor prognosis of glioma. In addition, we performed a CMap analysis to screen small molecule drugs using the selected prognostic genes. Some of the candidate drugs in our results have been proven to have anti-cancer effects, including butacaine (Mizuno and Ishida, 1982), piperlongumine (Duan et al., 2016; Liu et al., 2017; Chen et al., 2019), and sanguinarine (Ma et al., 2017; Zhang et al., 2017, 2018). Piperlongumine has been reported to rapidly induce the death of human pancreatic cancer cells through, at least in part, the induction of ferroptosis (Yamaguchi et al., 2018). A previous study has revealed that sanguinarine can block PPM1A phosphatase activity to restore c-Jun N-terminal kinase (JNK) activation, resulting in increased apoptosis of M. tuberculosis (Mt)-infected macrophages (Schaaf et al., 2017). To explore the potential mechanisms of these putative drugs, we further constructed the network of protein and drug compound interactions from the STITCH database, and found that the candidate drugs were likely to involve in tumor immunity through ferroptosis by targeting estrogen receptors. Considering that all these drugs may share similar mechanisms because of the principles of CMap database, drugs with no previous publications may also have the anti-tumor effects via targeting estrogen receptors and ferroptosis. Therefore, the 11-gene signature in the current study has the potential to be used as drug targets for therapy.

This study exists some limitations. First, all the data used to construct and validate the prognostic model in the current study were obtained from publicly available datasets. This is a retrospective study. A prospective study is needed to assess the potential application of the signature we have built because of the principles of CMap database, drugs with no previous publications may also have the anti-tumor effects via targeting estrogen receptors and ferroptosis. Therefore, the 11-gene signature in the current study has the potential to be used as drug targets for therapy.
CONCLUSION

Our present study identified a ferroptosis related 11-genes prognostic model for glioma patients. This model proved to have relatively high efficacy and to be clinically independent of other factors, providing insight into the prediction of glioma patients' prognosis. We also noticed that the differentially expressed ferroptosis genes were highly correlated with both pro-tumor and anti-tumor pathways, yet the latter seemed to be suppressed during the progression of glioma, indicating the putative effects of immunotherapy on glioma. Several potential drugs were also predicted based on the differentially expressed ferroptosis genes. The curative effect of the drugs and the underlying mechanisms between ferroptosis and tumor immunity in glioma remained lack of research and warranted further investigation.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: TCGA repository, https://portal.gdc.cancer.gov/; CGGA repository, http://www.cgga.org.cn/; GEO data can be found here: TCGA repository, https://portal.gdc.cancer.gov/; CGGA repository, http://www.cgga.org.cn/; GEO.

REFERENCES

Andrews, L. P., Marciscano, A. E., Drake, C. G., and Vignali, D. A. (2017). LAG3 (CD223) as a cancer immunotherapy target. *Immunol. Rev.* 276, 80–96. doi: 10.1111/imr.12539

Belavgeni, A., Bornstein, S. R., von Mässenhausen, A., Tonnis, W., Stumpf, J., Meyer, C., et al. (2019). Exquisite sensitivity of adrenocortical carcinomas to induction of ferroptosis. *Proc. Natl. Acad. Sci. U.S.A.* 116, 22269–22274. doi: 10.1073/pnas.1912700116

Bersuker, K., Hendricks, J. M., Li, Z., Magtanong, L., Ford, B., Tang, P. H., et al. (2019). The CoQ oxidoreductase FSP1 acts parallel to GXPD to inhibit ferroptosis. *Nature* 575, 688–692. doi: 10.1038/s41586-019-1705-2

Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., Cooper, L. A., et al. (2015). Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *Nat. Eng. J. Med.* 362, 2481–2498. doi: 10.1056/NEJMoa1402121

Brennan, C. W., Verhaak, R. G., McKenna, A., Campos, B., Noshmehr, H., Salama, S. R., et al. (2013). The somatic genomic landscape of glioblastoma. *Cell* 155, 462–477. doi: 10.1016/j.cell.2013.09.034

Cao, J. Y., and Dixon, S. J. (2016). Mechanisms of ferroptosis. *Cell. Mol. Life Sci.* 73, 2195–2209. doi: 10.1007/s00018-016-2194-1

Chen, B., Lei, Y., Wang, H., Dang, Y., Fang, P., Wang, J., et al. (2017). Repression of the expression of TET2 by ZEB1 contributes to invasion and growth in glioma cells. *Mol. Med. Rep.* 15, 2625–2632. doi: 10.3892/mmr.2017.6288

Chen, D. S., and Mellman, I. (2017). Elements of cancer immunity and the cancer-immune set point. *Nature* 541, 321–330. doi: 10.1038/nature21349

Chen, W., Lian, W., Yuan, Y., and Li, M. (2019). The synergistic effects of oxaliplatin and piperlongumine on colorectal cancer are mediated by oxidative stress. *Cell Death Dis.* 10:600. doi: 10.1038/s41419-018-1824-6

Chu, B., Kon, N., Chen, D., Li, T., Liu, T., Jiang, L., et al. (2019). ALOX12 is required for p53-mediated tumour suppression through a distinct ferroptosis pathway. *Nat. Cell Biol.* 21, 579–591. doi: 10.1038/s41566-019-0305-6

Cosmi, L., Maggi, L., Santarlasci, V., Liotta, F., and Annunziato, F. (2014). T-helper cells plasticity in inflammation. *Cytometry A* 85, 36–42. doi: 10.1002/cyto.a.22348

Das, M., Zhu, C., and Kuchroo, V. K. (2017). Tim-3 and its role in regulating anti-tumor immunity. *Immunol. Rev.* 276, 97–111. doi: 10.1111/imr.12520

AUTHOR CONTRIBUTIONS

ZC: conceptualization and software and formal analysis. TW: writing-original draft preparation and visualization. ZY: data curation and writing-review and editing. MZ: conceptualization, supervision, and funding acquisition and validation. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the Key Research and Development Program of Hunan Province (No. 2020SK2063), the Natural Science Foundations of Hunan Province (No. 2020JJ4134), and the National Natural Science Foundation of China (No. 81501025).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.652599/full#supplementary-material
and AP002754.2. *Cancer Res.* 80, 1804–1818. doi: 10.1158/0008-5472.Can-19-2389

Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., et al. (2012). Strong rules for discarding predictors in lasso-type problems. *J. R. Stat. Soc. Series B Stat. Methodol.* 74, 245–266. doi: 10.1111/j.1467-9868.2011.01004.x

von Mering, C., Jensen, L. J., Snel, B., Hooper, S. D., Krupp, M., Fogliertini, M., et al. (2005). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* 33, D433–D437. doi: 10.1093/nar/gki005

Wang, T. J. C., and Mehta, M. P. (2019). Low-grade glioma radiotherapy treatment and trials. *Neurosurg. Clin. N. Am.* 30, 111–118. doi: 10.1016/j.nec.2018.08.008

Wang, Z., Wang, Z., Zhang, C., Liu, X., Li, G., Liu, S., et al. (2018). Genetic and clinical characterization of B7-H3 (CD276) expression and epigenetic regulation in diffuse brain glioma. *Cancer Sci.* 109, 2697–2705. doi: 10.1111/cas.13744

Xia, X., Fan, X., Zhao, M., and Zhu, P. (2019). The relationship between ferroptosis and tumors: a novel landscape for therapeutic approach. *Curr. Gene Ther.* 19, 117–124. doi: 10.2174/1566523219666190628152137

Yamaguchi, Y., Kasukabe, T., and Kumakura, S. (2018). Piperlongumine rapidly induces the death of human pancreatic cancer cells mainly through the induction of ferroptosis. *Int. J. Oncol.* 52, 1011–1022. doi: 10.3892/ijo.2018.4259

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). Clusterprofiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zang, X., Sullivan, P. S., Soslow, R. A., Waitz, R., Reuter, V. E., Wilton, A., et al. (2010). Tumor associated endothelial expression of B7-H3 predicts survival in ovarian carcinomas. *Mod. Pathol.* 23, 1104–1112. doi: 10.1038/modpathol.2010.95

Zhang, G., Xu, Y., Lu, X., Huang, H., Zhou, Y., Lu, B., et al. (2009). Diagnosis value of serum B7-H3 expression in non-small cell lung cancer. *Lung Cancer* 66, 245–249. doi: 10.1016/j.lungcan.2009.01.017

Zhang, R., Wang, G., Zhang, P. F., Zhang, J., Huang, Y. X., Lu, Y. M., et al. (2017). Sanguinarine inhibits growth and invasion of gastric cancer cells via regulation of the DUSP4/ERK pathway. *J. Cell. Mol. Med.* 21, 1117–1127. doi: 10.1111/jcmm.13043

Zhang, S., Leng, T., Zhang, Q., Zhao, Q., Nie, X., and Yang, L. (2018). Sanguinarine inhibits epithelial ovarian cancer development via regulating long non-coding RNA CASC2-EIF4A3 axis and/or inhibiting NF-κB signaling or PI3K/AKT/mTOR pathway. *Biomed. Pharmacother.* 102, 302–308. doi: 10.1016/j.biopha.2018.03.071

Zhao, Z., Meng, F., Wang, W., Wang, Z., Zhang, C., and Jiang, T. (2017). Comprehensive RNA-seq transcriptomic profiling in the malignant progression of gliomas. *Sci. Data* 4:170024. doi: 10.1038/sdata.2017.24

Zheng, Y., Liu, A., Wu, Y., Gao, J., Li, Z., Liu, W., et al. (2019). High expression of B7-H2 or B7-H3 is associated with poor prognosis in hepatocellular carcinoma. *Mol. Med. Rep.* 19, 4315–4325. doi: 10.3892/mmr.2019.10080

**Conflict of Interest:** 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.

*Copyright © 2021 Chen, Wu, Yan and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*