Construction of an immune infiltration-associated ceRNA network and prognostic model for low-grade gliomas

Jing-Yuan Liu

Binzhou Medical University

Ming Jing

Binzhou Medical University

Yuan-Yu Liu

Binzhou Medical

Fan Zhang

Shantou University Medical College

Wen-Ying Liu

Binzhou Medical University

Sen Xu (✉ xuseneco@163.com)

Binzhou Medical University

Research Article

Keywords: low-grade glioma (LGG), immune infiltration, ceRNA network, prognostic model

Posted Date: February 23rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1371267/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Due to the low-grade glioma (LGG) infiltrative nature, it can be challenging to cure patients suffering from this condition only with surgery or multi-drug adjuvant chemotherapy. LGG patients usually have poor prognoses, making the exploration in therapeutic tools and accurate prognosis that are different from traditional ones a novel topic in the medical field. In recent years, the regulation of ceRNA networks in the tumor microenvironment (TME) by immune cell infiltration has generated a research boom, providing new perspectives for tumor treatment and predicting patient prognosis. However, the role of the ceRNA network in regulating the LGG tumor microenvironment is not effectively understood yet. In this study, we obtained LGG patient data with complete survival information from the TCGA database and constructed a novel immune infiltration pattern-related ceRNA regulatory network and a novel prognostic model, which was multi-dimensionally validated to have better predictive features in terms of LGG prognosis. Our study reveals the potential association and clinical significance of the ceRNA network with immune infiltration and survival prediction of these patients.

Introduction

Low-grade gliomas (LGG) are heterogeneous brain tumors of glial or precursor cell origin in the central nervous system (CNS) (1). Included in this group are the diffuse low-grade and moderate gliomas (classified by WHO as grade II and III, respectively) (2), which account for approximately 20% of all cases of CNS gliomas (3). In recent decades, despite many advances in surgical resection and multidrug adjuvant chemotherapy, most low-grade gliomas (LGG) progressively advance to high-grade gliomas (HGG). Due to the highly aggressive and infiltrative nature of LGG, recurrence and malignant progression of LGG is unavoidable, so the prognostic outcome of patients remains poor (4–7). Therefore, exploring the key molecular therapeutic mechanisms and targets of LGG is important for an accurate prognosis.

In addition to traditional LGG therapeutic strategies, emerging immunological therapies and therapeutic strategies targeting the tumor microenvironment (TME) have attracted a lot of interest from the medical community (8, 9). Tumor formation is the result of the combined action of tumor cells and the tumor microenvironment (TME) (10) in which immune and mesenchymal cells, as the major non-tumor components of the TME, are essential for the diagnosis and prognostic assessment of such clinical conditions (11).

Many scholars have reported that ceRNA networks can be used as a novel gene regulatory pathway to explain important regulatory mechanisms of complex post-transcriptional networks, including proliferation, invasion, migration, and drug resistance (12, 13). Recent studies have found that ceRNA regulatory networks and immune infiltration play a key role in the development of various tumors, including lung, gastric, and colon cancers (14–16). Currently, research focused on ceRNA networks in LGG is generally based on differential gene and biomarker screens between tumors and normal controls. However, the combined application of ceRNA networks and immune infiltration patterns has rarely been reported. Therefore, the construction of ceRNA regulatory networks in LGG tumor immune

microenvironment has greatly deepened our understanding of gene network regulation and molecular mechanisms, providing new therapeutic avenues and exploring directions for LGG treatment.

Meanwhile, along with modern medicine entering the era of big data and information technology, various high-throughput databases are used for target discovery and mechanism research. In this study, we constructed a novel immune infiltration pattern-related ceRNA regulatory network and prognostic model by combining the clinical features of LGG with in-depth information mining of TCGA and CGGA databases. We designed and validated a prognostic risk scoring model and a column line graph with predictive value. Our study reveals the clinical significance and application prospects of the regulation of ceRNA network with immune infiltration and prognosis prediction in LGG patients and provides new perspectives and research targets for the treatment of LGG.

Results

Immune infiltration score analysis

Tumor purity, immune cell infiltration, and stromal cell infiltration were scored using the ESTIMATE algorithm in 527 TCGA-LGG samples with complete survival information. ESTIMATE scores ranged from -3422.598988 to 3762.9069 (median = -1136.358097), immune scores ranged from -1676.001802 to 2477.025966 (median = -231.9068585), and stromal scores ranged from -1769.169966 to 1710.689096 (median = -919.0959986). Of the 527 LGG patients, 289 (54.8%) were male, and 238 (45.2%) were female, and age at first diagnosis ranged from 32 to 53 years (median = 40). A total of 427 patients had IDH status mutations (81%), 223 (42.3%) were WHO class II, and 357 had a total deletion of 1p/19q codeletion (67.7%) (table.1).

The 527 TCGA-LGG samples were divided into three high and low scoring subgroups by setting the median as the grouping value: (i) high/low ESTIMATE scoring group; (ii) high/low immune scoring group; and (iii) high/low stromal scoring group. The survival analysis showed significant differences ($P<0.05$) in survival status among all three high- and low-scoring subgroups (Figure.1A-C), with a higher probability of OS in the low-immune scoring group. We explored the association of the three scores concerning the clinical characteristics of LGG patients. Regarding the IDH status, we found that WHO Grade was positively correlated with the three scores, while 1p/19q codeletion was negatively correlated with the three scores (Figure.1D).

Table1. LGG patient data in TCGA database

Characteristic	levels	Overall
n		527
Gender, n (%)	Female	238 (45.2%)
	Male	289 (54.8%)
Age group, n (%)	≥40	263 (49.9%)
	≤40	264 (50.1%)
Grade, n (%)	G2	223 (42.3%)
	G3	243 (46.1%)
	unknown	61 (11.6%)
IDH status, n (%)	Mutant	427 (81%)
	unknown	3 (0.6%)
	WT	97 (18.4%)
1p/19q codeletion, n (%)	code1	170 (32.3%)
	non-code1	357 (67.7%)
Age, median (IQR)		40 (32, 53)

Identification of immune infiltration-related differential gene sets and WGCNA analysis

Gene expression differences in the immune scoring groups were analyzed using $|\log_2FC| > 1$ and FDR value < 0.05 as identification criteria. A total of 2033 differential genes were detected in the mRNA data; 1086 up-regulated and 947 down-regulated genes (Figure.2A). lncRNA data expressed 145 differential genes, being 51 up-regulated and 94 down-regulated (Figure.2B). In addition, other 192 genes were detected as significantly differentially expressed in the miRNA data; 67 were up-regulated and 125 were down-regulated (Figure.2C).

Differential lncRNAs and mRNAs from all samples were included in the WGCNA. No significant outliers in sample clustering (Figure.3A), including immune scores and trait information as trait variables, were used to calculate the optimal beta value of 8 in the WGCNA (Figure.3B). Gene co-expression modules were further generated using a dynamic tree cutting approach. The sheer height of module trait gene clustering was 0.3 (Figure.3C). Module-trait relationship plots showed that gray gene modules were highly correlated with immune scores ($R = -0.68$, $P = 9e-73$) (Figure.3D). Our analysis of the 160 genes of the grey gene module revealed a significant correlation ($R = 0.33$, $P = 2e-05$) between gene module membership (MM) and gene salience (GS) (Figure.3E). Therefore, we identified the 160 genes of the gray module as the core gene set as pivotal genes associated with immune infiltration.

Construction of ceRNA networks based on differentially expressed genes between immune scores

We designed a ceRNA network (Figure.4A) using queries from the database based on differentially expressed RNA (145 lncRNAs, 192 miRNAs, and 2033 mRNAs) interaction relationships between immune scoring groups. We have also screened 117 lncRNA-miRNAs and 265 mRNAs with interaction relationships with miRNAs, including 16 lncRNAs, 52 common miRNAs, and 189 mRNAs. A sub-network of ceRNAs (12 lncRNAs, 39 miRNAs, and 54 mRNAs) highly correlated with immune infiltration patterns was constructed by gene identification in the intersection of the ceRNA network and the WGCNA gray module (Figure.4 B).

Enrichment analysis and PPI network construction and optimization

The DEGs were intersected with those in the WGCNA gray module, and 96 genes were identified for enrichment analysis. In the KEGG analysis, the main enrichment was central to the carbon metabolism in cancer patients, PI3K-Akt signaling pathway, and other pathways (Figure.5B). the GO analysis's main enrichment is for biological processes such as cellular response to BMP stimulus and sensory perception of mechanical stimulus (Figure.5A). Additional string diagrams specifically show the relevant genes involved in the enrichment process (Figure.5C-D).

A PPI network was constructed through the STRING database to study protein-protein interactions, and three core clusters of MCODE inserts were further identified. The Mcode score of the first core PPI cluster was 6.923, and mainly consisted of immune response-related proteins, including DLX1, GSX1, SPHKAP, ETV4, CALB1, PDGFRA, SCG2, NCAN, NXPH1, VSTM2A, DLX5, EYA1, SFRP1, and TFAP2A (Figure.5E). The second Mcode score of 4.25 contained SHISA6, PCDHB15, COL2A1, PCDHB3, KIF26A, KIF26B, RHBDL3, PDZD2, C1QL2 (Figure.5F). The third Mcode score was 2.85 and contained COL9A1, PRICKLE1, ACAN, DAB1, NEUROD1, SLC17A8, GNB3, SEMA6C (Figure.5G).

Immune infiltration risk score model construction and prognostic validation

In the TCGA-LGG cohort, we integrated miRNAs, lncRNAs, and mRNAs from immune infiltration-associated ceRNAs and genes from the WGCNA gray module to construct the immune infiltration risk score model. Univariate Cox regression analysis was first performed to identify 66 of the genes significantly associated with OS (Figure.6B), and then Lasso regression was applied to produce a scoring model with the optimal number of genes (optimal $\lambda = -3.27$) (Figure.6A). A total of 15 potential modeled genes were identified and further analyzed with the stepwise use of multivariate Cox regression. Finally, at AIC=1213.43, the optimal with one lncRNAs (CRDNE) and 12 mRNAs (NT5C1A, LNX1, HSF4, FHDC1, AFAP1L2, SPHKAP, ETV4, FABP6, SLC10A4, GALNT14, ST8SIA2, TFAP2A) was identified in the model. By summarizing the above gene regression coefficients and gene expression (Figure.6C-D), a prognostic risk scoring model for predicting OS was constructed using the risk scoring formula. Then, the risk score for each patient of TCGA-LGG was calculated by the scoring model to use the median as the cut-off value between the high-risk and low-risk groups. Kaplan-Meier survival analysis showed (Figure.6E) that OS was significantly reduced in the high-risk group compared with the low-risk group ($P < 0.01$), while the risk

score, survival time, and expression levels of genes were discretely distributed among the low-risk groups (Figure.6G-I). As shown in Figure 6F, we analyzed the AUC of time-dependent ROC curves for different survival times, and the results were as follows: (i) the AUC for 1-year survival time was 0.906; (ii) the AUC for 3-year survival time was 0.899; and (iii) the AUC for 5-year survival time was 0.837 (Figure.6F). These results indicate that the risk scoring model we constructed had satisfactory OS prediction ability. To assess the prognostic significance of the scoring model, Cox regression analysis combined with LGG clinical characteristics showed that univariate analysis determined that risk score and clinical characteristics other than gender were associated with OS (Figure.S1A). In the multifactorial analysis, the risk score was further identified along with age, IDH status, WHO classification and 1p/19q codeletion as independent prognostic factors for OS in LGG patients (Figure.S1C).

Construction and evaluation of column line graphs

As shown in Figure. 7A, to construct a column line plot, we used the validated prognostic factors (risk score, age, IDH status, WHO classification, and 1p/19q codeletion) to construct a column line plot. To verify whether it could predict the probability of OS in LGG patients at 1, 3, and 5 years, we analyzed the AUC of the time-dependent ROC curves for different survival times. The results were as follows: the 1-year survival time AUC was 0.930, the 3-year survival time AUC was 0.920, and the 5-year survival time AUC was 0.845 (Figure.7C). The correlation calibration plots over the 1-year, 3-year, and 5-year periods show that the deviation-corrected lines of the column plots are close to the standard line (Figure.7B). The above results show that the column line plots constructed based on the risk score model have excellent predictive power in delimiting the probability of OS in LGG patients.

Joint validation of external databases

The China Glioma Genome Atlas (CGGA) cohort was used as an external validation of the risk score model, and we have used it together with the grouping and study methods described above for TCGA-LGG data. We analyzed the K-M survival curves of the CGGA cohort and showed that the probability of OS was also significantly lower in the high-scoring group of CGGA-LGG patients compared with the lower-scoring group (Figure.8A). The CGGA cohort was similar to the TCGA cohort in terms of risk score, survival time, and degree of gene expression (Figure.8B-D). We performed different analyses of ROC curves for 1-year, 3-year, and 5-year survival times in CGGA-LGG patients with AUC values of 0.694, 0.765, and 0.779, respectively (Figure.8E). Then, we further validated the column line plots constructed based on the TCGA cohort data in the CGGA cohort with the following ROC curve analyses: the AUC values for 1-year survival time were 0.801, 0.855 for 3-year survival time, and 0.858 for 5-year survival time, respectively (Figure.8F). Correlation calibration plots showed that the deviation-corrected line of the CGGA cohort column line plot was close to the standard line (Figure.8G). In addition, a decision curve analysis was performed to assess the diagnostic benefit of the risk score model by total LGG case counts (Figure.8H-J). The DCA analysis pointed out that the risk score model curve showed the highest net benefit. To explore the potential clinical value of the risk score model, we further performed the multidimensional analysis of TCGA-LGG and CGGA-LGG data (principal components were analyzed using log 2). The results showed that samples

from the low-risk scoring group and the high-risk scoring group differed significantly in gene expression diversity (Figure.9A). We further explored the genes' expression involved in model construction in various tumors in the TCGA joint GTEx database (Figure.S2) and observed that 14 mRNAs and one lncRNA were significantly different in most tumors.

GSEA and signaling pathway analysis

To understand the biological significance of the differences between high/low-risk groups, we performed GSEA analysis in these gene sets, such as `c2.cp.kegg.v7.4.symbols.gmt` and `c5.go.v7.4.symbols.gmt`. In the TCGA cohort, samples from the high-risk group were mainly enriched in immune-related pathways, such as B cell-mediated immunity, adaptive immune response, and other related signaling pathways (Figure.9B). The samples in the low-risk group were mainly enriched in signaling pathways related to cell-matrix adhesion and methylation, such as RIBOSOME and CELL-DIFFERENTIATION-IN-HINDBRAIN (Figure.9C). In the CGGA cohort, the high-risk group was mainly enriched in signaling pathways related to cell activation and signaling regulation pathways, such as ECM-RECEPTOR-INTERACTION, JAK-STAT, and other related pathways (Figure.9D). Samples from the low-risk group were mainly enriched in signaling pathways related to signaling processes, such as NEUROTRANSMITTER-SECRETION, REGULATION-OF-POSTSYNAPTIC-MEMBRANE-POTENTIAL, and others (Figure.9E).

Comprehensive immunological analysis of low-grade gliomas

In recent years, the influence of exosomes on the regulation of ceRNA networks in TME has attracted a lot of attention, which led to a surge of research around this topic. To this end, we selected low-grade glioma tumor tissue exosomes and exosomes from different tumor tissues and compared them using exoRBase database analysis. Our outcomes show that LGG patients had higher levels of FHDC1 mRNA expression in exosomes compared to other tumor patients (Figure.10A). The risk score assessed by the ESTIMATE algorithm with the immune score in immune infiltration intensity was significantly correlated ($P=2.3e-4$) (Figure.9F). Regarding the abundance of different immune cell infiltration types, we used the bioinformatics tool CIBERSORT to identify the six major immune cell infiltration types in TCGA-LGG patients (Figure.10C). The identification revealed relatively high levels of CD8⁺ T cells, M0-type macrophages, and T cell gamma delta infiltration in the high-risk group with poorer prognosis (Figure.10B).

Drugs targeting immune checkpoints can already act in TME by reversing inhibitory signaling pathways. Many scholars are currently working to screen which immune checkpoint blockade (ICB) therapies can benefit LGG patients (17). To better assess the response of LGG patients to immunotherapy, we explored the expression of 47 currently commonly used immune checkpoints between the high- and low-risk groups. It was found that most of the more classical checkpoints, such as PD-L1 and CD276, were expressed at significantly higher levels in the high-risk group than in the low-risk group ($P<0.05$) (Figure.10D).

Discussion

For the first time, we applied infiltration scores as a basis for grouping to construct immune-related ceRNA networks and combined them with core gene sets in WGCNA to construct ceRNA subnetworks with highly correlated immune infiltration patterns and optimal models of prognostic risk. We found that the probability of OS was significantly higher in the three low-scoring subgroups than in the high-scoring subgroups during the modeling process ($P < 0.05$). Gene set enrichment analysis revealed enhanced activation of tumor immune and inflammation-related signaling pathways and elevated immune response status in the high-risk group. Elevated immune scores were significantly associated with poor prognosis, consistent with the findings of Chen et al. (18), which prompted us to explore the potential association between immune scores and OS in LGG patients. Currently, elevated immune-related scores are associated with the survival and prognosis of tumor patients in breast, cervical, and prostate cancers (19–22). Immune scores play an important role in the staging of various tumors, including colon and gastric cancers (23, 24). Our study also found that risk score could remain as an independent prognostic factor after changes in clinical characteristics, suggesting that local immune status may have better potential in prognosis, which is consistent with the findings of Shi et al. in sarcoma (25). Zhang et al. found (26) that risk score was associated with tumor aggressiveness, tumor subtype, rapid proliferation, and shorter survival time, which matches our study results. This study provides a complementary immunological perspective for individualized tumor therapy and potential clues for the future of immunotherapy for LGG patients.

Tumor cells establish an immunosuppressive network through immune selection, which leads to immune escape and promotes tumor development in immunocompromised hosts (26). Analysis of the infiltration abundance of various immune cell types revealed that CD8 + T cells, M0-type macrophages, and T cell gamma delta had relatively high levels of infiltration in the high-risk group with poorer prognosis(27). The tumor microenvironment is closely related to immune cell infiltration, and high levels of immune cell infiltration imply that the complex immune microenvironment may induce severe alterations in tumor microenvironment adaptation. In turn, this has implications in tumor progression and prognosis (5, 28). Currently, glioma immune infiltration research is generally based on differential identification and prognostic marker screening between tumors and normal controls (18, 29). However, the mechanism of tumor immunity on the tumor microenvironment is still unclear. The related molecular mechanisms still need to be further investigated, which may provide new research perspectives for tumor targeting and immunotherapy in the future.

To better assess the response of LGG patients to immunotherapy, we explored 47 immune checkpoints currently in common use and we found that the expression levels of most classical checkpoints were significantly higher in the high-risk group than in the low-risk group ($P < 0.05$). In recent years, prognostic factors (IDH status, WHO Grade, 1p/19q co-deletion) have been found to correlate with immune checkpoint expression. Xiao et al. suggested that higher immune checkpoint expression in the high-risk group of LGG patients with the same IDH status may be associated with activation of immunosuppressive pathways (30). In a study by Lei Lv et al., it was noted that 1p/ 19q co-deletion could

lead to reduced expression levels of several immune checkpoints, including PD1 and PD-L1/2, and suggested that both IDH mutations and 1p/19q co-deletion are markers of better prognosis in LGG patients (31). Meanwhile, it was found that the expression of suppressor molecules such as CTLA-4, PD-1, PD-L1, and TIM-3 was relatively high in the high-risk group of LGG (32, 33). The above findings align with our results, and we have provided a new basis for targeted therapy for LGG. In addition, we explored the relationship between mRNA expression/mutation and immune cell infiltration. By analyzing the immune infiltration of the 12 mRNAs involved in the modeling, we found that most mRNA expression levels correlated mainly with CD8 + T cell and macrophage infiltration ($P < 0.05$). In contrast, mRNA mutations were associated with immune cell infiltration. This association reveals the potential prognostic value of CD8 + T cells and macrophages in LGG.

Our study provides insights into the role of exosomes of immune infiltrating cells in LGG in ceRNA networks and suppressive immune checkpoint expression. However, as these findings were based on analyses of database stored data, no in-depth excavation was performed in terms of LGG immune infiltration and clinical therapeutic tools (e.g., chemotherapy and drug resistance). Furthermore, the WHO CNS5(the fifth edition of the WHO Classification of Tumors of the Central Nervous System) in 2021 introduced new tumor types and subtypes (34), and the main population of this study was patients with adult LGG, which was not explored in depth in individuals with pediatric LGG. Therefore, some limitations emerged, and multidimensional studies in different areas are needed to expand and confirm our findings.

In summary, the immune-related ceRNA networks (including ceRNA sub-networks with highly correlated immune infiltration patterns) and risk scoring models we constructed have the potential to provide a more theoretical basis for the development of LGG immunotherapy and gene targeting. The column line graph containing the risk scoring model can improve physicians' prognosis judgment of LGG patients so they can make more accurate diagnoses and improve the quality of life of patients recovered from the condition.

Materials And Methods

Data set sources and data pre-processing LGG patient data were obtained from the TCGA and CGGA databases, incorporating gene expression data (mRNA, LncRNA, miRNA) and the corresponding clinical data of LGG samples for 529 TCGA-LGG and 1080 CGGA-LGG LGG patients. We screened out LGG patients without survival time and finally kept the data of 528 from the TCGA database and 929 patients from the CGGA database. Gene expression data from TCGA-LGG and CGGA-LGG sample sets were normalized by the SVA (Surrogate Variable Analysis)package.

Identification of DEGs and mRNA and WGCNA analysis The Estimate algorithm scored the infiltrating stromal and immune cell infiltration expression data of LGG tumor tissues, and samples were grouped according to the median infiltration score. Tumor purity was indicated by the ESTIMATE score, immune cell infiltration by the immune score, and stromal cell infiltration by the stromal score). The limma

package was used to identify differentially expressed miRNAs (DEmiRs) and DEGs, including *mRNA* and *lncRNA*, following the identification criteria $|\log_2FC| > 1$ and FDR value < 0.05 .

The weighted co-expression network (WGCNA) algorithm was used to explore the correlation between clinical features and gene sets. In this study, we used the WGCNA algorithm R package to analyze the core set of genes most associated with immune scores in TCGA-LGG samples.

ceRNA network construction To predict RNA molecule interactions, we used the miRDB, miRTarBase and TargetScan databases to screen and retain all mRNA-miRNA intersections identified by the three databases. The LncBasev2 dataset was used to screen and retain the identified lncRNA-miRNA interactions. Combining the results of WGCNA, all *mRNA* and *lncRNA* in the core module were used to construct ceRNA subnetworks, and Cytoscape software was used to visualize ceRNA and PPI networks.

Gene function enrichment analysis and PPI network construction The enrichment analysis of genes intersecting DEGs with the WGCNA core module was performed using the "ggplot2" and "clusterProfiler" R packages, including the Gene Ontology (GO), the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, and the KEGG enrichment analysis. Co-expressed gene PPI networks were constructed from the STRING database to study protein interactions, and the minimum interaction score required was set at 0.9 (highest confidence level). We used the MCODE plugin in Cytoscape software to explore the highly connected and related core clusters in the PPI network.

Construction and validation of the prognostic model One-way Cox regression analysis was applied to screen for genes associated with OS. Essential genes screened in univariate Cox regression analysis were selected by using the *glmnet* package using Lasso regression. Using multivariate Cox regression, risk score models were then constructed based on the risk score formula (risk score = coefficient * gene expression) and following the Akaike information criterion (AIC). K-M survival analysis was conducted to test the correlation between the level of risk score and prognosis. To assess the predictive accuracy of the risk score model, time-dependent 1-, 3-, and 5-year survival time (ROC) curve analyses were performed using the SURVIVAL package. We also performed univariate and multifactorial Cox regression analyses using risk scores and clinical characteristics. Then, all independent prognostic factors were retained to construct a prognostic column line graph. The discriminatory effect of the column line graph was assessed by analyzing time-dependent ROC curves. The consistency of the column line graphs was examined through time-dependent calibration plots. In addition, decision curve analysis was performed using the *rmda* package. We used all of the above methods to evaluate the risk score models and column line graphs in both the TCGA and the CGGA cohorts.

Comprehensive immunoassay To explore the role of exosomes in the ceRNA network, we accessed the exoRBase database and analyzed exosomes from a variety of tumor tissues. We used an estimation algorithm to calculate tumor purity and understand the differences in immune cell infiltration, along with a circular classification algorithm to estimate the abundance of immune cell infiltration and immune checkpoint expression levels in different risk groups. The TIMER database was employed to explore the relationship between immune infiltration and gene expression/mutation.

Statistical analysis All statistical analyses were performed using R software (v.4.0.1), and comparisons of log-transformed normalized expression data between the two groups were performed using independent t-tests, with $P < 0.05$ considered statistically significant differences.

Declarations

1. Ethical Approval and Consent to participate

Not applicable

2. Availability of data and materials

Not applicable

3. Competing interests

We state that there is no competing financial interest in relation to the work described and all the authors agree to its submission.

4. Funding

This study is financially supported by the Medical and Health Technology of Shandong, China (2016WS0007), Binzhou Medical University SARS-CoV-2 research program (BY2021XGFY04).

5. Authors' contributions

Jing-Yuan Liu, Bioinformatics analysis, Thesis writing

Ming Jing, Thesis writing

Yuan-Yu Liu, Thesis writing

Fan Zhang, Bioinformatics analysis

Wen-Ying Liu, Thesis writing

Sen Xu, Thesis proofreading

References

1. Lapointe S, Perry A, Butowski NA. Primary brain tumours in adults. *The Lancet*. 2018;392(10145):432–46.
2. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol*. 2016;131(6):803–20.

3. Shields LB, Choucair AK. Management of low-grade gliomas: a review of patient-perceived quality of life and neurocognitive outcome. *World Neurosurg.* 2014;82(1–2):e299–309.
4. Cancer Genome Atlas Research N, Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med.* 2015;372(26):2481–98.
5. Deng X, Lin D, Zhang X, Shen X, Yang Z, Yang L, et al. Profiles of immune-related genes and immune cell infiltration in the tumor microenvironment of diffuse lower-grade gliomas. *J Cell Physiol.* 2020;235(10):7321–31.
6. Qiu H, Li Y, Cheng S, Li J, He C, Li J. A Prognostic Microenvironment-Related Immune Signature via ESTIMATE (PROMISE Model) Predicts Overall Survival of Patients With Glioma. *Front Oncol.* 2020;10:580263.
7. Kiran M, Chatrath A, Tang X, Keenan DM, Dutta A. A Prognostic Signature for Lower Grade Gliomas Based on Expression of Long Non-Coding RNAs. *Mol Neurobiol.* 2019;56(7):4786–98.
8. Makkouk A, Weiner GJ. Cancer immunotherapy and breaking immune tolerance: new approaches to an old challenge. *Cancer Res.* 2015;75(1):5–10.
9. Yang L, Song X, Gong T, Jiang K, Hou Y, Chen T, et al. Development a hyaluronic acid ion-pairing liposomal nanoparticle for enhancing anti-glioma efficacy by modulating glioma microenvironment. *Drug Deliv.* 2018;25(1):388–97.
10. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature.* 2013;501(7467):346–54.
11. Jia D, Li S, Li D, Xue H, Yang D, Liu Y. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging (Albany NY).* 2018;10(4):592–605.
12. Peng Z, Liu C, Wu M. New insights into long noncoding RNAs and their roles in glioma. *Mol Cancer.* 2018;17(1):61.
13. Liu H, Wang S, Zhou S, Meng Q, Ma X, Song X, et al. Drug Resistance-Related Competing Interactions of lncRNA and mRNA across 19 Cancer Types. *Mol Ther Nucleic Acids.* 2019;16:442–51.
14. Liang ZZ, Guo C, Zou MM, Meng P, Zhang TT. circRNA-miRNA-mRNA regulatory network in human lung cancer: an update. *Cancer Cell Int.* 2020;20:173.
15. Zheng X, Wang X, Zheng L, Zhao H, Li W, Wang B, et al. Construction and Analysis of the Tumor-Specific mRNA-miRNA-lncRNA Network in Gastric Cancer. *Front Pharmacol.* 2020;11:1112.
16. Zhou H, Xiong Y, Liu Z, Hou S, Zhou T. Expression and prognostic significance of CBX2 in colorectal cancer: database mining for CBX family members in malignancies and vitro analyses. *Cancer Cell Int.* 2021;21(1):402.
17. Abe S, Nagata H, Crosby EJ, Inoue Y, Kaneko K, Liu CX, et al. Combination of ultrasound-based mechanical disruption of tumor with immune checkpoint blockade modifies tumor microenvironment and augments systemic antitumor immunity. *J Immunother Cancer.* 2022;10(1).

18. Chen J, Wang Z, Wang W, Ren S, Xue J, Zhong L, et al. SYT16 is a prognostic biomarker and correlated with immune infiltrates in glioma: A study based on TCGA data. *Int Immunopharmacol.* 2020;84:106490.
19. Long X, Hou H, Wang X, Liu S, Diao T, Lai S, et al. Immune signature driven by ADT-induced immune microenvironment remodeling in prostate cancer is correlated with recurrence-free survival and immune infiltration. *Cell Death Dis.* 2020;11(9):779.
20. Li SJ, Zhao L, Wang HY, Zhou HN, Ju J, Du H, et al. Gustave Roussy Immune Score based on a three-category risk assessment scale serves as a novel and effective prognostic indicator for surgically resectable early-stage non-small-cell lung cancer: A propensity score matching retrospective cohort study. *Int J Surg.* 2020;84:25–40.
21. Pan JW, Zabidi MMA, Ng PS, Meng MY, Hasan SN, Sandey B, et al. The molecular landscape of Asian breast cancers reveals clinically relevant population-specific differences. *Nat Commun.* 2020;11(1):6433.
22. Qi X, Fu Y, Zhang M, Lu C, Wang Y, Peng J, et al. An Innovative Immune Score-Based Prognostic Nomogram for Patients with Cervical Cancer. *Biomed Res Int.* 2020;2020:8882576.
23. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou F-S, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet.* 2018;391(10135):2128–39.
24. Wang H, Wu X, Chen Y. Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer. *Front Oncol.* 2019;9:1212.
25. Shi D, Mu S, Pu F, Zhong B, Hu B, Liu J, et al. Development of a Novel Immune Infiltration-Related ceRNA Network and Prognostic Model for Sarcoma. *Front Cell Dev Biol.* 2021;9:652300.
26. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
27. Feng P, Li Z, Li Y, Zhang Y. Phosphatase and Tensin Homolog Mutation in Immune Cell Infiltration and Clinicopathological Features of Low-Grade Gliomas. *Front Mol Biosci.* 2020;7:562416.
28. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol.* 2002;3(11):991–8.
29. Zhang C, Wang L, Liu H, Deng G, Xu P, Tan Y, et al. ADPRH is a prognosis-related biomarker and correlates with immune infiltrates in low grade glioma. *J Cancer.* 2021;12(10):2912–20.
30. Xiao Y, Cui G, Ren X, Hao J, Zhang Y, Yang X, et al. A Novel Four-Gene Signature Associated With Immune Checkpoint for Predicting Prognosis in Lower-Grade Glioma. *Front Oncol.* 2020;10:605737.
31. Lv L, Zhang Y, Zhao Y, Wei Q, Zhao Y, Yi Q. Effects of 1p/19q Codeletion on Immune Phenotype in Low Grade Glioma. *Front Cell Neurosci.* 2021;15:704344.
32. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol.* 2015;15(8):486–99.

33. Zhang M, Wang X, Chen X, Zhang Q, Hong J. Novel Immune-Related Gene Signature for Risk Stratification and Prognosis of Survival in Lower-Grade Glioma. *Front Genet.* 2020;11:363.

34. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol.* 2021;23(8):1231–51.

Figures

Figure 1

Infiltration score analysis: (A-C) Survival analysis plots for the three high and low-scoring subgroups. (D) Heat map of the association between the three scores and the five clinical features.

Figure 2

Volcano and heat map of differentially expressed genes between immune scoring groups: (A) mRNA; (B) lncRNA, and (C) miRNA.

Figure 3

Identification of immune infiltration-associated gene sets: (A) Sample clustering tree diagram; (B) Calculation and evaluation of optimal β values in WGCNA; (C) Module-trait gene clustering tree diagram; (D) Heat map of the module-trait relationship; (E) Correlation plot of gray module immune score genes with gene module members ($R = 0.33$, $P = 2e-05$).

Figure 4

ceRNA network construction: (A) Construction of ceRNA network based on differential expression between immune scoring groups; (B) Immune infiltration-related ceRNA sub-networks.

Figure 5

Enrichment analysis and PPI network optimization: (A) GO analysis visualization bar graph; (B) KEGG analysis and visualized dot plots; (C, D) String plots showing the pathway-related genes involved in enrichment; (E) First Mcode score of 6.923 for the core PPI cluster; (F) Second Mcode score of 4.25 for the core PPI clusters. (G) The third Mcode score of 2.85 for the core PPI clusters.

Figure 6

Risk score model construction and validation:(A, B) Lasso Cox regression to achieve dimensionality reduction and construct feature sets with an optimal number of genes by parameter cross-validation. (C) One-dimensional histogram of regression coefficients for the optimal model genes (D) Multivariate Cox regression analysis of the optimal model forest plot. (E) Analytical plots of survival K-M survival analysis for high-risk and low-risk groups; (G-H) Discrete distribution plots of risk scores and survival status for

patients in both groups. (I) Heat map of gene expression levels of patients in the high and low-risk group. (F) Time-dependent ROC curve plots validate that the scoring model can accurately predict OS.

Figure 7

Column line graphs for the evaluation of (A) column line graphs based on independent prognostic factors; (B) time-dependent calibration plots of column line graphs; and (C) time-dependent ROC plots of column line plots.

Figure 8

Joint validation of external datasets: (A) K-M survival curves for the CGGA cohort; (B,C) Discrete distribution of risk scores, survival status of patients in the high/low-risk groups of the CGGA cohort; (D) Heat map of gene expression levels of patients in the high/low-risk group of the CGGA cohort; (E) Time-dependent ROC plot of the risk score model for the CGGA cohort; (F) Columnar line graph of the time-dependent ROC plot of the CGGA cohort; (G) Time-dependent calibration plots of column line graphs for the CGGA cohort; (H-G) DCA analysis plots for the CGGA cohort at 1, 3, and 5 years.

Figure 9

GSEA and signaling pathway analysis: (A) PCA plots of gene expression diversity. (B) Plot of gene set enrichment analysis for the high-risk group in the TCGA cohort. (C) Plot of gene set for the enrichment analysis of the low-risk group in the TCGA cohort. (D) Gene set for the enrichment analysis of the high-risk group in the CGGA cohort. € Gene set for the enrichment analysis of the low-risk group in the CGGA cohort. (F) Box plot of correlation between risk score and immune score on the intensity of immune infiltration.

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

- [Supplementarymaterial.docx](#)