Comprehensive Analysis of Genome-instability Related IncRNAs and Tumor-immune Microenvironment to Discriminate the Prognosis, Immunotherapy and PARP Inhibitor Responses in Patients with Low-grade Glioma

Zijian Zhou (zhouzijian@163.com)  
Qingdao Municipal Hospital Group  
https://orcid.org/0000-0001-8859-3741

Bin Lu  
Qingdao Municipal Hospital Group

JinHong Wei  
Southwest Medical University

Lei Guo  
Qingdao Municipal Hospital Group

ZhongMing He  
Southwest Medical University

Yini Liu  
Southwest Medical University

Weimin Wang  
Qingdao Municipal Hospital Group

Luo Li  
Qingdao Municipal Hospital Group

Research Article

**Keywords:** low-grade glioma, genome-instability related IncRNAs, tumor-immune microenvironment, prognosis, immunotherapy/PARP inhibitor

**DOI:** https://doi.org/10.21203/rs.3.rs-711385/v1

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

Background: Previous study revealed that Genome-instability was correlated with tumor-immune microenvironment in cancer. We try to discriminate the prognosis, immunotherapy and poly (ADP)-ribose polymerase (PARP) inhibitor responses through comprehensive analysis of genome-instability related IncRNAs and the tumor-immune microenvironment in patients with low-grade glioma (LGG).

Methods and Results: RNAseq data, genome variation profiling data and copy number variation (CNV) data were used to evaluate the genomic instability of LGG patients. Genomic unstable-like (GU-like) and genomic stable-like (GS-like) clusters were identified by hierarchical clustering analysis of 102 genome-instability related IncRNAs (GILncRNAs). GS-like cluster had a tendency to receive better clinical outcome. Patients in GU-like cluster were more likely to respond to immunotherapy, especially anti-PD-1/PD-L1 treatment. PARP inhibitors including Rucaparib and Olaparib will get better therapeutic effects for patients in GU-like cluster. Lasso and Cox regression analysis were utilized to construct the risk model based on GILncRNAs. As for the risk model constructed by 9 GILncRNAs, the overall survival, clinical outcome, immunotherapeutic response, and PARP inhibitor sensitivity were significantly different between patients of high and low-risk groups.

Conclusions: The genome-instability related IncRNAs signature involved in our risk model had great advantages in predicting prognosis, immunotherapy and PARP inhibitor response.

Introduction

Low-grade gliomas, composed of grade II-III according to the World Health Organization classification system, are the most common primary malignant tumors in central nervous system in children, and differ with high-grade glioma (glioblastoma, GBM) in biological and clinical features[1, 2]. The therapeutic outcome is not satisfactory for LGG patients received conventional treatment[3, 4]. Research focusing on LGG treatment still has a long way to go. Genomic instability is a basic and common feature in the development of malignant tumors, which were characterized as either chromosomal instability (CIN) or microsatellite instability (MSI)[5]. Due to rapid DNA replication and transcription, tumor cells are sensitive to DNA damage, especially damage from double-strand breaks (DSBs)[6]. In response to DNA damage, the DNA damage response (DDR) network was activated to be implicated in the repair of DNA damage, in which erroneous repair is correlated with genomic instability[7, 8]. Poly (ADP)-ribose polymerase (PARP) plays a critical role for DNA repair in cancer cells with DDR gene mutations[9] and inhibition of PARP as a synthetic lethal treatment has got promising therapeutic effect in breast cancer[10]. Previous studies have demonstrated an underlying correlation between genomic alterations and tumor-immune microenvironment (TIME)[11, 12]. Moreover, TIME has been reported to play a crucial role in the progression and aggressiveness of cancer. The abundance and type of infiltrating immune cell were correlated with cancer therapy and clinical outcome [13, 14]. In recent years, an increasing number of studies have been paying attention to immunotherapy such as immune checkpoint inhibitors, antitumor vaccines and cellular immunotherapy for glioma. However, the clinical outcome for patients received
immunotherapy can be diverse and only a minority of patients benefitted from it[15, 16]. Considering the many factors affecting the response to immunotherapy, a concise method for physicians to predict the immunotherapeutic response for individuals is on need. Long non-coding RNAs (lncRNAs) have been demonstrated to be involved in various biological process as they directly interact with the cell cycle, proliferation pathways and microbiome balance. Furthermore, lncRNAs are increasingly deemed to be important in the regulation of immune cell differentiation and function[17]. A large number of prognostic models based on lncRNAs have been established to predict the prognosis of patients with cancer[18-21]. Sun et al. identified that lncRNAs were related to tumor immune infiltration and built tumor immune infiltration-associated lncRNAs signature for improving prognosis and immunotherapy response of patients with non-small cell lung cancer[22]. However, there is rare study focusing on investigating a risk model for predicting prognosis, immunotherapy and chemotherapy response for patients with LGG from the respect of genome-instability related lncRNAs. In this manuscript, we aimed to construct a precise risk model to predict prognosis, immunotherapy and PARP inhibitor responses through comprehensive analysis of genome-instability related lncRNAs (GILncRNAs) and tumor-immune microenvironment in low-grade glioma.

Materials And Methods

Acquisition of Multi-omics and Clinical Data

The RNAseq (Fragments Per Kilobase Million [FPKM] value) data consisting of 529 samples and genome variation profiling data (maf format) consisting of 507 samples for LGG patients were obtained from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/). The copy number variation (CNV) data (n=527) for LGG patients was downloaded from the University of California Santa Cruz (UCSC) Xena browser[23](https://xena.ucsc.edu/). The corresponding clinical information was downloaded from cBioPortal for cancer genomics(https://www.cbioportal.org/). Four parts of the data were merged according to the sample names. The RNAseq and clinical data including 169 GBM observations were acquired from TCGA database.

Screening of lncRNAs Related to Genome Instability

The RNA annotation file of Genome Reference Consortium Human Build 38 (GRCh38) was downloaded from the Ensembl website (http://asia.ensembl.org/) for annotation of RNAseq data. The cumulative counts of somatic mutations in each sample were calculated and ranked in a decreasing manner, followed which the top 25% of the patients and the bottom 25% of the patients were designed as genomic unstable (GU) group and genomic stable (GS) group, separately. The mutation types and frequencies of the top 20 genes with the highest mutation frequency were analyzed and visualized using “maftools” package in R (version 4.1.0). To better exhibit the gain/loss alterations of chromosomes, CNV summary plots were visualized by Circos plots using the “RCircos” R package. LncRNA and mRNA expression data were all log2 transformed for differentially expression analysis using R project. The
differentially expressed IncRNAs between the two groups were analyzed using the Wilcoxon rank-sum test in “limma” package of R software and p values were regulated by false discovery rate (FDR).

Unsupervised Hierarchical Clustering Analysis

Unsupervised hierarchical clustering analysis was conducted through “sparcl”, “pheatmap” and “limma” packages of R software using the differentially expressed IncRNAs between GU/GS groups and samples were divided into two gene clusters. The gene cluster with higher mutation counts was identified as genomic unstable-like (GU-like) cluster, whereas the other was defined as genomic stable-like (GS-like) cluster.

Differential Analysis between GU-like and GS-like Clusters

We identified the differentially expressed mRNAs between the two gene clusters (|Fold Change| >4.0 and false discovery rate (FDR) adjusted P<0.01]. Top 20 differentially expressed genes (DEGs) were visualized in a heatmap. Subsequently, functional annotation and pathway enrichment analyses, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses, were carried out on the DEGs by using “clusterProfiler,” “org.Hs.eg.db,” “enrichplot,” and “ggplot2” packages of R software. FDR <0.001 was considered statistically significant in GO analysis and FDR <0.05 was considered significantly enriched in KEGG analysis.

Construction of Protein-Protein Interaction (PPI) Network

The PPI of DEGs between GU-like and GS-like clusters was explored by the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database online (https://string-db.org/). The PPI network was conducted by minimum required interaction score ≥ 0.900 with active interaction sources including textmining, experiment, databases, co-expression, neighborhood, gene fusion and co-occurrence. Top 30 key genes were identified according to the number of adjacent nodes.

Analysis of Tumor-Immune Microenvironment

The infiltration of 23 immune cells in tumor was estimated through single sample gene set enrichment analysis (ssGSEA) by using “GSEABase” and “GSVA” R package. We obtained the ssGSEA score for each sample according to the given immune gene set file, after which the scores were normalized for further analysis. The abundance of infiltrating immune cells and related immune functions between subgroups were evaluated and visualized in boxplot and heatmap. The correlation between risk score and infiltrating immune cells was analyzed and visualized using “limma” and “corrplot” R packages. Stromal and immune scores were calculated by the “ESTIMATE” R package based on the gene expression profiles of 141 immune genes and 141 stromal genes to determine the tumor purity. An unsupervised subclass
A method was used to predict the response to immunotherapy, such as anti-PD1 and anti-CTLA4 therapy, in the different subtypes.

**Construction and Verification of Prognostic Signature based on Genome-Instability related LncRNAs**

Univariate Cox regression analysis was firstly employed to evaluate the correlation between expression level of genome-instability related lncRNAs (GILncRNA) and overall survival of patients with LGG using “survival” R package, in which COX p value less than 1.0*10^-9 was considered as lncRNAs with prognostic value. The results were visualized by “forestplot” R package. The LGG samples were randomly divided into train and test group at a ratio of 1:1. The least absolute shrinkage and selection operator (LASSO) regression algorithm with penalty parameter tuning conducted by 1000-fold-crosss validation was used to further select the valuable GILncRNAs from the previous obtained prognostic IncRNAs. The Receiver Operating Characteristic (ROC) curve analysis was conducted and the Areas Under Curve (AUC) values were calculated to evaluate the accuracy of the risk model. R packages including "survival","caret","glmnet","survminer" and “timeROC” were involved above. The risk score for each sample was calculated by the following formula:

\[ \text{risk score} = \sum_{i=0}^{n} \text{coefIncRNA}i \times EXP \text{IncRNA}i \]

where the coefIncRNAi represents the coefficient of ith lncRNA, and the EXP IncRNAi represents the expression level of the ith lncRNA in the risk model. Univariate and multivariate Cox regression analysis were used to examine the independent prognostic value of risk score computed by the model. The distribution of clinical features among high and low risk groups was compared and visualized using “limma” and “ComplexHeatmap” packages in R. In addition, the drug sensitivity was predicted between the subtypes using “pRRophetic” R package, which was determined by the half maximal inhibitory concentration (IC50) of each sample based on the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.canncerrxgene.org/). Nomogram integrating risk score and clinical factors was utilized to optimize the prognostic model by using "rms" R package. Prediction for 1-, 3- and 5- years survival was identified based on multivariate Cox regression results. Calibration curves were used to determine the performance of the nomogram. PCA and t-SNE analysis were employed to examine whether the 9 genome-instability related IncRNAs involved in the risk model can distinguish the samples in our research by using “Rtsne” R package.

**Statistical Analyses**

The Kruskal-Wallis test was used to compare multi-groups, and the Wilcoxon test was used to compare two groups. Chi-square tests were implemented to evaluate the distribution of categorical variables between subgroups. The student’s t-test was employed to compare the continuous data between pairs of
subgroups. Correlations between normally distributed variables were analyzed by Pearson’s correlation test, while correlations between non-normally distributed variables were assessed with Spearman’s correlation test. The Kaplan-Meier plotter was employed to generate survival curves for the subgroups, and the log rank test was used to evaluate the statistically significant differences. Two-tailed $p < 0.05$ was considered as statistically significant unless specifically noted. R version 4.1.0 (Institute for Statistics and Mathematics, Vienna, Austria) was used to conduct the statistical analysis and visualization of the results.

Results

Screening and identification of genome-instability related IncRNAs in LGG patients

The clinical and pathological characteristics of the LGG cohort used in our study were displayed in Table 1. Patients were divided into genomic unstable (GU) group and genomic stable (GS) group based on the cumulative number of somatic mutations. In order to evaluate the difference of genomic instability, we investigated the mutation types, mutation frequencies and copy number variations of the two groups. Top 20 genes with the highest alteration frequencies were analyzed (Fig. 1a and b). The results of the oncoplots and bar plots revealed that the alteration frequencies of most of the genes were significantly higher in patients of GU group than those of GS group including EGFR, FLG, HMCN1, LRP2, NOTCH1, OBSCN, PTEN, RYR2, TTN (The bar plots for FLG, HMCN1, LRP2, NOTCH1, OBSCN, RYR2 and TTN were displayed in Online Resource 5A, $p < 0.05$). However, the proportion of patients with IDH1 mutations was significantly higher in GS group (80%) compared with GU group (54%, $p < 0.001$) suggesting that IDH1 may play a crucial role in the tumorigenesis and work synergistically with other genes involved in the same pathway\cite{24}. The analysis of CNVs of chromosomes was further performed to explore the distinct genomic variations between the two groups (Fig. 1c and d). Patients in GU group tended to bear a greater burden of copy number amplifications ($P = 8.9 \times e^{-13}$) and deletions ($P = 3.5 \times e^{-7}$) than those in GS group. All these findings indicated a significant difference in the genomic instability between the two groups. We compared the IncRNA expression profiles between GU and GS groups. 102 genome-instability-related IncRNAs were identified with $|\text{Fold Change}| > 2.0$ and adjusted $p < 0.05$. Heatmap of top 20 differentially expressed IncRNAs was shown in Fig. 1e.

Table 1

Clinical information for LGG patients in this study.
| Covariates               | Type       | Total (n=503) | Test (n=249) | Train (n=254) | P-value |
|--------------------------|------------|---------------|--------------|---------------|---------|
| Age                      | <=60       | 443 (88.07%)  | 225 (90.36%) | 218 (85.83%)  | 0.1524  |
|                          | >60        | 60 (11.93%)   | 24 (9.64%)   | 36 (14.17%)   |         |
| Sex                      | Female     | 224 (44.53%)  | 107 (42.97%) | 117 (46.06%)  | 0.5434  |
|                          | Male       | 279 (55.47%)  | 142 (57.03%) | 137 (53.94%)  |         |
| Disease Free Time        | <=24       | 284 (56.46%)  | 141 (56.63%) | 143 (56.3%)   | 0.9516  |
|                          | >24        | 184 (36.58%)  | 90 (36.14%)  | 94 (37.01%)   |         |
|                          | unknow     | 35 (6.96%)    | 18 (7.23%)   | 17 (6.69%)    |         |
| Disease Free Status      | Disease Free | 306 (60.83%) | 155 (62.25%) | 151 (59.45%)  | 0.5011  |
|                          | Recurred or progressed | 162 (32.21%) | 76 (30.52%)  | 86 (33.86%)   |         |
|                          | unknow     | 35 (6.96%)    | 18 (7.23%)   | 17 (6.69%)    |         |
| Grade                    | G2         | 242 (48.11%)  | 120 (48.19%) | 122 (48.03%)  | 1       |
|                          | G3         | 260 (51.69%)  | 129 (51.81%) | 131 (51.57%)  |         |
|                          | unknow     | 1 (0.2%)      | 0 (0%)       | 1 (0.39%)     |         |
| Histologic Type          | Astrocytoma | 189 (37.57%) | 95 (38.15%)  | 94 (37.01%)   | 0.6627  |
|                          | Oligoastrocytoma | 128 (25.45%) | 59 (23.69%)  | 69 (27.17%)   |         |
|                          | Oligodendroglioma | 186 (36.98%) | 95 (38.15%)  | 91 (35.83%)   |         |
| IDH1 Mutation            | YES        | 91 (18.09%)   | 48 (19.28%)  | 43 (16.93%)   | 0.201   |
|                          | NO         | 32 (6.36%)    | 12 (4.82%)   | 20 (7.87%)    |         |
|                          | unknow     | 380 (75.55%)  | 189 (75.9%)  | 191 (75.2%)   |         |
| Therapy Outcome          | Complete Remission or Response | 83 (16.5%) | 48 (19.28%)  | 35 (13.78%)   | 0.7811  |
|                          | Partial Remission or Response | 50 (9.94%) | 25 (10.04%)  | 25 (9.84%)    |         |
|                          | Stable Disease | 59 (11.73%) | 30 (12.05%)  | 29 (11.42%)   |         |
|                          | Progressive Disease | 36 (7.16%) | 20 (8.03%)   | 16 (6.3%)     |         |
|                          | unknow     | 275 (54.67%)  | 126 (50.6%)  | 149 (58.66%)  |         |
Identification of two LGG clusters based on genome-instability related lncRNAs with distinct survival outcomes and clinical features

As shown in the Fig. 2a, unsupervised hierarchical clustering analysis was conducted for LGG patients based on 102 genome-instability related lncRNAs (GILncRNAs). The somatic mutation counts were significantly different between the two clusters (Fig. 2b, \( p < 2.22 \times 10^{-16} \)). Thus, one cluster with higher mutation counts was defined as GU-like cluster, the other was defined as GS-like cluster. The expression level of UBQLN4, which is correlated with genomic instability, was significantly higher in patients of GS-like cluster than those of GU-like cluster (Fig. 2b, \( p = 1.4 \times 10^{-13} \)). The overall survival (OS) analysis was employed between the two clusters, indicating that patients in GS-like cluster displayed better prognosis compared with GU-like cluster (Fig. 2c, \( p < 0.001 \)). Heatmap of multi-clinicopathological characteristics was adopted to demonstrate the correlation between the clustering and clinical features (Fig. 2d). As shown in Fig. 2e, patients in GU-like cluster were significantly older than that in GS-like cluster (\( p < 0.001 \)). Tumors in patients of GU-like cluster were more likely to recur or progress compared with GS-like cluster (\( p < 0.001 \)). Histologic grade for patients in GU-like cluster was higher than that in GS-like cluster (\( p < 0.001 \)). Patients in GU-like cluster tended to receive postoperative radiotherapy (\( p < 0.001 \)). IDH1 for patients in GU-like cluster was more frequently mutated compared with GS-like cluster (\( p < 0.001 \)). Karnofsky Performance Scores for patients in GU-like cluster were higher than that in GS-like cluster (\( p = 0.002 \)), indicating self-care abilities for patients in GU-like cluster were more likely to be affected. Clinical outcomes for patients in GU-like cluster seemed to be worse than those in GS-like cluster (\( p = 0.01 \)), with more patients in tumor-bearing status. Differentially expressed mRNAs (DEGs) were identified between the two clusters (|Fold Change| >4.0, adjusted \( p < 0.01 \)). Heatmap exhibiting top 20 DEGs was shown in Fig. 2f. Protein-Protein interaction network was employed to further evaluate the interplay of the DEGs (Fig. 2g). 30 hub genes with the highest number of adjacent nodes were noted in Fig. 2h.

Enrichment analysis identified molecular mechanisms affected by DEGs between GS-like and GU-like clusters

To explore the underlying different molecular mechanisms between the two clusters, GO analysis was employed, regarding molecular function (MF), cellular component (CC), and biological process (BP). As shown in Online Resource 1A-C, the DEGs were significantly enriched in molecular functions associated with formation of genomic instability, including DNA-binding transcription activator activity, RNA polymerase II-specific. KEGG pathway analysis identified signaling pathways associated with tumorigenesis and tumor immunology including cytokine-cytokine receptor interaction pathway, ECM-receptor interaction pathway and transcriptional misregulation in cancer (Online Resource 1D-F). These results of functional enrichment analysis suggested that GU-like cluster differs in genome-instability and tumor-immunology associated functions compared with GS-like cluster, implying significant differences in tumor-immune environment and genomic instability.
Distinct tumor-immune microenvironment (TIME) and immunogenomic patterns were identified between GS-like and GU-like clusters

Based on the ESTIMATE algorithm, the compositions of the TIME for each sample were analyzed. Compared with those of GS-like cluster, the immune scores \( (p = 2.9 \times 10^{-13}) \), stromal scores \( (p < 2.22 \times 10^{-16}) \) and ESTIMATE scores \( (p = 3.3 \times 10^{-15}) \) were significantly higher in patients of GU-like cluster, indicating higher abundances of immune and stromal cells and lower tumor purity (Fig. 3a). Furthermore, comparisons of the abundance of 23 tumor infiltrating immune cells in different clusters were carried out using ssGSEA. Most of the immune cells were significantly more abundant in patients of GU-like cluster (Fig. 3b). The heatmap of the infiltration levels of the immune cells was displayed in the Fig. 3c. As shown in Fig. 3d, the associated immune functions were more active in patients of GU-like cluster.

Patients of GU-like clusters were more sensitive to immunotherapy and PARP inhibitors

An increasing number of researches revealed that immunotherapy targeting immune check-points brought hope for clinical treatment of malignant tumors[25]. The expressions of immune check-points including PDCD1(PD1), CD274(PD-L1), PDCD1LG2(PD-L2), CTLA4, CD80, CD86 were analyzed between the two clusters (Fig. 3e). The expression levels of all the immune genes mentioned above were elevated in patients of GU-like cluster compared with GS-like cluster. The results of unsupervised subclass mapping analysis demonstrated that patients of GU-like cluster were more likely to respond to the immunotherapy of anti-PD1 (Fig. 3f, Bonferroni corrected \( p = 0.002) \). However, no positive results were found in the response to anti-CTLA4 therapy between the clusters. The estimated IC50 values of Rucaparib and Olaparib were significantly lower in patients of GU-like cluster (Fig. 3g, \( p = 3.1 \times 10^{-7}, p = 1.7 \times 10^{-13} \), respectively), indicating that LGG patients of GS-like cluster tended to be more resistant to PARP-inhibitor therapy than those of GU-like cluster. However, there was no significant difference in sensitivity to Veliparib between the two clusters \( (p = 0.27) \).

Identification of genome-instability related lncRNA signature for prognosis

Univariate cox proportional risk regression was utilized to select the genome-instability related lncRNAs with prognostic values, combining the expression levels of genome-instability related lncRNAs and survival information of LGG patients. 33 genome-instability-related lncRNAs with predictive prognosis roles were determined with \( p \) value \(< 1.0 \times 10^{-9} \). The results were shown in the forest plot with hazard ratio (HR) of 95% confidence intervals (Fig. 4a). Patients with LGG were randomly divided into train and test group at a ratio of 1:1. Lasso regression analysis was applied to further evaluate and filter these
IncRNAs in the train set and 9 IncRNAs were identified as the key factors in the risk model with 1000 replicates of cross validation (Fig. 4b and c), including LINC01831, CRNDE, AC025171.5, AL390755.1, AC012073.1, AL355974.2, AC010273.3, AC131097.3, AC092718.3. LncRNSs with associated coefficients were listed in Fig. 4d. The risk score was calculated by the following formula:

$$\text{risk score} = \sum_{i=0}^{n} \text{coef}_i \text{IncRNA}_i \times EXP \text{IncRNA}_i.$$

LGG patients were further divided into high-risk group and low-risk group with the median risk score as the cutoff, in the train and test set, respectively.

**Evaluation of genome-instability related IncRNA signature for prognosis**

Kaplan-Meier analysis showed that overall survival of patients in low-risk group were significantly higher compared with those in high-risk group in the train and test set (Fig. 4e and l, p < 0.001). ROC curve analysis of the risk model based on genome-instability related IncRNAs signature (GILncSig) over time for 1, 2, and 3 years in the train set showed AUC values with 0.892, 0.912, and 0.905, respectively (Fig. 4f). In the test set, AUC values over time for 1, 2, 3 years were 0.819, 0.899 and 0.900, separately (Fig. 4m). In the train set, the AUC value of GILncSig over time for 1 year is 0.870, which is higher than that of other clinicopathological factors including diagnosis age (AUC = 0.704), sex (AUC = 0.509), histologic type (AUC = 0.354) and histologic grade (AUC = 0.669). Similar results were obtained regarding the ROC curve analysis over time of 2, 3 and 5 years (Online Resource 5B). The expression levels of 9 genome-instability-related IncRNAs in the risk model were significantly higher in patients of high-risk group than those of low-risk group (Fig. 4g and n). The scatter plots of survival time showed that the survival time decreased with the increasing of the risk score and the mortality rate was higher in patients of high-risk group (Fig. 4h and o). As for the scatter plots of somatic mutations, it seemed that the somatic mutation counts were also higher in patients of high-risk group (Fig. 4i and p) and it was validated in the boxplot with p value less than 0.01 (Fig. 4k and r). The expression level of UBQLN4 was significantly higher in patients of low-risk group compared with high-risk group (Fig. 4l and s). Univariate and multivariate cox regression were performed to analyze whether the risk score calculated based on the above 9 genome-instability related IncRNAs were independent prognostic factors of clinical or pathological factors (age, sex, histologic grade, histologic type, and mutation count). The results in the train set showed that risk score either in univariate cox regression analysis (hazard ratio = 4.008, 95% CI [1.876–8.562], p < 0.001) or multivariate cox regression analysis (hazard ratio = 2.937, 95% CI [1.261–6.844], p = 0.013) were significantly correlated with overall survival of LGG patients (Online Resource 2A and B). Similar results were obtained in the test set (Online Resource 2C and D). Nomogram integrating risk score, age, sex, grade as variables was employed to predict the 1-, 3-, 5- year overall survival in LGG cohort (Online Resource 2E). The calibration plots of 3-, 5- years overall survival indicated that the predictive curves were
close to the actual curves of overall survival, suggesting that the performance of the nomogram was good in predicting the overall survival of patients with LGG (Online Resource 2F).

Validation of genome-instability related IncRNA signature for prognosis

GBM dataset from TCGA database was used in this study to verify the performance of the risk model based on 9 genome-instability related IncRNAs. According to the expression levels of the 9 genome-instability related IncRNAs in GBM dataset, risk score for each sample was calculated, after which patients were divided into high and low-risk group according to the cutoff of median score. Kaplan-Meier curves showed that patients in the high-risk group had significantly lower overall survival probability compared with the low-risk group (Online Resource 2G, p < 0.001). As the standard chemotherapy for GBM postoperatively, temozolomide (TMZ) can significantly increase the overall survival of GBM patients according to the previous studies. The estimated IC50 values of TMZ were significantly lower in patients of low-risk group (p = 0.0071), indicating that GBM patients in high-risk group tended to be more resistant to TMZ therapy than those in low-risk group (Online Resource 2H). There was no statistical significance in the estimated IC50 values of 3 PARP inhibitors between patients in low- and high-risk groups (Online Resource 5C).

In addition, the predictive performance of our risk model based on genome-instability related IncRNAs signature (GILncSig) was compared with other glioma-associated prognostic IncRNAs signatures established in previous studies (Online Resource 2I). Although the predictive power of our GILncSig for 1 year rate was not better than Qiu’s model, the AUC values of our GILncSig were significantly higher than other three risk models over time for 2, 3 and 5 years, indicating that the risk model created by our group had better prognostic performance in predicting overall survival of LGG patients[18–20].

Distinct clinicopathological features, tumor-immune microenvironment (TIME), immunotherapy response and sensitivity to PARP inhibitors were identified between different risk groups

PCA and t-SNE analysis revealed that the 9 genome-instability related IncRNAs involved in our risk model can distinguish the samples with LGG (Online Resource 3A). As shown in the heatmap (Online Resource 3B), the distribution of diagnosis age, disease free time, disease free status, histologic grade, histologic type, IDH1 mutation status, Karnofsky performance score, postoperative radiotherapy status, person neoplasm status, and therapy outcome between the two groups were significantly different. The detailed results were demonstrated in boxplots to show the distribution of risk scores of patients with multi-
clinicopathological characteristics, indicating that the risk scores of patients with worse clinicopathological features tended to be significantly higher (Online Resource 3C).

The distribution of patients in the two gene-clusters corresponding with the risk groups and survival status was shown in alluvial diagram (Online Resource 4A). Comparison between patients of high- and low-risk groups demonstrated that the immune scores ($p = 8 \times e^{-5}$), stromal scores ($p = 0.00018$) and ESTIMATE scores ($p = 8.2 \times e^{-5}$) were significantly higher in patients of high-risk group, indicating higher abundance of immune and stromal cells, while lower tumor purity (Online Resource 4B). Furthermore, most of the immune cells were more abundant in patients of high-risk group (Online Resource 4C and D). The associated immune functions were more active in patients of high-risk group (Online Resource 4E). The expression levels of all the immune checkpoints mentioned above were elevated in patients of high-risk group compared with low-risk group (Online Resource 4F). In addition, we found a positive correlation between the expression of these immune check-points and risk score (Online Resource 4G). The results of unsupervised subclass mapping analysis demonstrated that patients in high-risk group were more likely to respond to the immunotherapy of anti-PD1 (Fig. 5a, Bonferroni corrected $p = 0.007$). However, no positive results were found regarding the response to anti-CTLA4 therapy between the two groups. The estimated IC50 values of Rucaparib, Olaparib and Veliparib were significantly lower in patients of high-risk group (Fig. 5b-d, $p = 1.4\times e^{-6}$, $p = 8.3\times e^{-9}$, $p = 0.0034$, respectively), suggesting that patients in low-risk group tended to be more resistant to PARP inhibitor therapy than those in high-risk group.

**Discussion**

As a common subtype of glioma, LGG which was composed of astrocytoma, oligoastrocytoma and oligodendroglioma, was not sensitive to current therapy[3]. Accumulating studies have placed special emphasis on the investigation of immunotherapy for glioma, which was revolutionizing cancer care in recent years[26, 27]. However, only a minority of patients can get a satisfactory therapeutic effect[28]. It is extremely important to select proper patients for immunotherapy. Taking anti-PD1/PDL1 treatment as an example, there are numerous factors affecting the response to immunotherapy including the expression level of PDL1, tumor mutation burden and abundance of infiltrating immune cells, which makes the clinical decision complex[29, 30]. There is an urgent need for an easy method to select proper potential candidates to receive immunotherapy. Recent researches had shown that a large number of long non-coding RNAs (lncNRAs) had important roles on the prediction of prognosis and immune microenvironment of gliomas[18-20, 22]. So far, there is no ideal classification or prognostic model to identify LGG patients with different response to immunotherapy and chemotherapy based on genome-instability related lncRNAs. Inhibition of PARP is a new strategy of chemotherapy for glioma patients[31]. Inhibition of PARP triggers a synthetically lethal effect in tumors with homozygous deletions, deleterious mutations, or both in DNA damage repair genes including BRCA1/2, ATM, PALB2, and the FANC gene family, which is correlated with genomic instability[32-35]. PARP inhibitors, such as Rucaparib, Olaparib and Veliparib, can inhibit the activity of PARP so as to prevent DNA replication and transcription as well as generate double-strand DNA breaks, and eventually lead to cell death[36]. UBQLN4, a member of the
ubiquilin family, which was correlated with genome instability, had been demonstrated to be associated with actionable poly (ADP)-ribose polymerase (PARP) inhibitor sensitivity[37]. Drug sensitivity regarding PARP inhibitors for LGG patients was involved in our study.

This study provided a comprehensive analysis of genomic instability and tumor-immune microenvironment in low-grade glioma through exploring the expression patterns of genome-instability related long non-coding RNAs (GILncRNAs), looking forward to finding a novel method for classification to distinguish subtypes with different prognosis, immunotherapy and PARP inhibitor response. Our results suggested that the GILncRNAs were significantly related to prognosis, immunotherapy and PARP inhibitor response of LGG patients than other established IncRNA signatures and had great advantages for prediction. Firstly, LGG samples of genomic unstable (GU) group and genomic stable (GS) group were extracted from TCGA dataset, followed which the genomic instability of the two groups were evaluated from the aspect of tumor mutation burdens, mutation frequencies of specific genes and copy number variations. Through the comparison of expression profiles of IncRNAs between GS and GU groups, we obtained genome-instability related IncRNAs (GILncRNAs), based on which 529 LGG samples were divided into two clusters, namely GS-like cluster and GU-like cluster. Then, we compared their clinicopathological parameters to clarify the correlation between the clustering and clinicopathological features. Distinct tumor-immune microenvironment (TIME) and immunogenomic patterns were identified between patients of GS-like and GU-like clusters. Patients in GU-like cluster were more likely to respond to immunotherapy, especially anti-PD-1/PD-L1 treatment, as predicted by unsupervised subclass mapping analysis. PARP inhibitors including Rucaparib and Olaparib will get better therapeutic effects for patients in GU-like cluster. In order to facilitate the clinical practice, genome-instability related IncRNAs signature was established as a powerful risk model for predicting the prognosis of LGG patients. According to the calculated risk score, patients were further divided into high-and low-risk groups. Patients in high-risk group tended to have worse overall survival, while immunotherapy and PARP inhibitors will get better therapeutic effects. In addition, the risk model was also applied to GBM patients to examine the performance in high-grade glioma. Unexpectedly, the overall survival of GBM patients with higher risk scores was significantly worse than those with lower risk scores. Moreover, TMZ will get better therapeutic effect for patients with lower risk scores.

In this study, LGG patients were separated into two clusters based on 102 genome-instability related IncRNAs (GILncRNAs). Patients of GS-like cluster had a tendency to receive better prognosis and clinical outcome. On the other hand, the TIME and immunogenomic patterns including stromal score, immune score, ESTIMATE score, abundance of infiltrating immune cells, immune functions and expression of immune checkpoints were significantly higher in patients of GU-like cluster, laying foundation for the response to immunotherapy. The expression level of UBQLN4 in patients of GS-like cluster was significantly higher than those of GU-like cluster and the somatic mutation counts which can be an indicator of genomic instability were significantly lower in patients of GS-like cluster, which were consistent with the results of Shiloh's group. However, the estimated IC50 values of PARP inhibitors including Rucaparib and Olaparib in patients of GS-like cluster were significantly higher compared with GU-like cluster, which indicated that patients of GS-like cluster were more resistant to PARP inhibitor
treatment. However, Shiloh's group demonstrated that UBQLN4 overexpression in tumors promotes PARP1 inhibitor sensitivity, which seemed differing with our results[37]. Considering they focused on the neuroblastoma cases, the difference of subtype of glioma may account for these results. In addition, in-depth investigation on this issue was demanded to clarify the relationship between UBQLN4 and PARP inhibitor sensitivity.

As for the risk model constructed by 9 genome-instability related lncRNAs (GILncRNAs), the overall survival, immunotherapeutic response, PARP inhibitor sensitivity and other specific clinical features were significantly different between high- and low-risk groups. As shown in the alluvial diagram, the majority of patients of GU-like cluster overlapped with patients of high-risk group. The distribution of high and low-risk groups as well as the difference of risk scores between the two clusters were analyzed (Online Resource 5D, p<0.001, p<2.22 x e\textsuperscript{-16}, respectively), which indicated that the risk score for one sample can be an indicator of the corresponding gene cluster to some extent. For example, one patient belonging to high-risk group with higher risk score was more likely to be in GU-like cluster, implying that the method of risk model based on 9 GILncRNAs can be substituted for the clustering method. The results of PCA and t-SNE analysis revealed that the method of classification based on the risk model was efficient and can well distinguish LGG samples in our research, indicating that the risk model was not only a model to predict prognosis but also a powerful method to distinguish patients with different clinical outcomes and therapy response.

**Conclusion**

Overall, based on 102 genome-instability related lncRNAs (GILncRNAs), we obtained two clusters with distinct prognosis, immunotherapeutic and PARP inhibitor response, demonstrating that the GILncRNAs were correlated with the above characteristics. Furthermore, the risk score calculated based on 9 genome-instability related lncRNAs in the risk model was an effective biomarker with powerful predictive function.

**Abbreviations**

AUC: Areas Under Curve; CNV: copy number variation; DDR: DNA damage response; DEGs: differentially expressed genes; DSBs: double-strand breaks; ECM: extracellular matrix; FDR: false discovery rate; GBM: glioblastoma; GDSC: Genomics of Drug Sensitivity in Cancer; GILncRNAs: genome-instability related lncRNAs; GO: Gene Ontology; GS: genomic stable; GU: genomic unstable; HR: hazard ratio; IC50: half maximal inhibitory concentration; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: least absolute shrinkage and selection operator; LGG: low-grade glioma; OS: overall survival; PARP: Poly (ADP)-ribose polymerase; PCA: principal component analysis; PPI: Protein-Protein Interaction; ROC: Receiver Operating Characteristic; ssGSEA: single sample gene set enrichment analysis; TCGA: The Cancer Genome Atlas; TIME: tumor-immune microenvironment; TMZ: temozolomide; t-SNE: t-Distributed Stochastic Neighbor Embedding.

**Declarations**
Acknowledgements

This work was supported by the National Natural Science Foundation of China (82002642). The authors would like to thank the TCGA databases, University of California Santa Cruz (UCSC) Xena browser, and cBioPortal for cancer genomics for data availability. I would like to thank my fiancée, Yuqi, for her support in my life.

Funding

This work was supported by the National Natural Science Foundation of China (82002642).

Conflicts of interest/Competing interests

The authors declare that there are no conflicts of interest and no financial and personal relationships with other people or organizations that could inappropriately influence our work.

Availability of data and material (data transparency)

We declare that the data supporting the findings of this study are available in the TCGA database (https://portal.gdc.cancer.gov/), the University of California Santa Cruz (UCSC) Xena browser (https://xena.ucsc.edu/) and cBioPortal for cancer genomics (https://www.cbioportal.org/).

Code availability (software application or custom code)

R (version 4.1.0) and Perl (5.32.1.1) software were applied in the study. The codes were custom codes.

Authors' contributions

Zijian Zhou and Luo Li contributed to the study conception and design. Material preparation and analysis were performed by Bin Lu, Jinhong Wei and Lei Guo. Data collection and visualization were conducted by ZhongMing He, Yini Liu and Weimin Wang. The first draft of the manuscript was written by Zijian Zhou and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Additional declarations for articles in life science journals that report the results of studies involving humans and/or animals
The data supporting the findings of this study were obtained from open-access database, so there is no additional declaration.

**Ethics approval**

Not applicable.

**Consent to participate**

All authors declare to approve to participate in this study.

**Consent for publication**

All authors declare to approve the publication of the findings in this study.

**References**

1. Hoshide R, Jandial R (2016) World Health Organization Classification of Central Nervous System Tumors: An Era of Molecular Biology. World Neurosurgery. 2016; 94: 561-2
2. Sidaway P (2020) Low-grade glioma subtypes revealed. Nature Reviews Clinical Oncology 17:335-336
3. Duffau H, Taillandier L (2014) New concepts in the management of diffuse low-grade glioma: Proposal of a multistage and individualized therapeutic approach. Neuro-Oncology 17:332–342
4. Darlix A, Mandonnet E, Freyschlag CF, Pinggera D, Forster M-T, Voss M et al (2018) Chemotherapy and diffuse low-grade gliomas: a survey within the European Low-Grade Glioma Network. Neuro-Oncology Practice 6:264–273
5. Lengauer C, Kinzler KW, Vogelstein B (1997) Genetic instability in colorectal cancers. Nature 386:623–627
6. Harper JW, Elledge SJ (2007) The DNA Damage Response: Ten Years After. Mol Cell 28:739–745
7. Reinhardt HC, Yaffe MB (2009) Kinases that control the cell cycle in response to DNA damage: Chk1, Chk2, and MK2. Curr Opin Cell Biol 21:245–255
8. Matsuno Y, Atsumi Y, Shimizu A, Katayama K, Fujimori H, Hyodo M et al (2019) Replication stress triggers microsatellite destabilization and hypermutation leading to clonal expansion in vitro. Nat Commun 10:3925
9. Javle M, Curtin NJ. The potential for poly (ADP-ribose) polymerase inhibitors in cancer therapy. 2011; 3: 257–67
10. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. 2017; 355: 1152–8
11. Huang T-X, Fu L (2019) The immune landscape of esophageal cancer. Cancer Commun 39:79
12. Rooney MS, Shukla Sachet A, Wu Catherine J, Getz G, Hacohen N (2015) Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. Cell 160:48–61

13. Barnes TA, Amir E (2017) HYPE or HOPE: the prognostic value of infiltrating immune cells in cancer. Br J Cancer 117:451–460

14. Jochems C, Schlom J. Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity. 2011; 236: 567–79

15. Weller M, Butowski N, Tran DD, Recht LD, Lim M, Hirte H et al (2017) Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. The Lancet Oncology 18:1373–1385

16. Reardon DA, Brandes AA, Omuro A, Mulholland P, Lim M, Wick A et al (2020) Effect of Nivolumab vs Bevacizumab in Patients With Recurrent Glioblastoma: The CheckMate 143 Phase 3 Randomized Clinical Trial. JAMA Oncology 6:1003–1010

17. Denaro N, Merlano MC, Lo Nigro C (2019) Long noncoding RNAs as regulators of cancer immunity. Mol Oncol 13:61–73

18. Maimaiti A, Wang X, Pei Y, Nuermammaiti N, Tuersunniyazi A, Abula Y et al. Identification and validation of a novel eight mutant-derived long non-coding RNAs signature as a prognostic biomarker for genome instability in low-grade glioma. Aging. 2021; 13

19. Qiu X, Tian Y, Xu J, Jiang X, Liu Z, Qi X et al (2021) Development and Validation of an Immune-Related Long Non-coding RNA Prognostic Model in Glioma. J Cancer 12:4264–4276

20. Tu Z, Wu L, Wang P, Hu Q, Tao C, Li K et al (2020) N6-Methyladenosine-Related IncRNAs Are Potential Biomarkers for Predicting the Overall Survival of Lower-Grade Glioma Patients. Front Cell Dev Biol 8:642

21. Xia P, Li Q, Wu G, Huang Y (2021) An Immune-Related IncRNA Signature to Predict Survival In Glioma Patients. Cell Mol Neurobiol 41:365–375

22. Sun J, Zhang Z, Bao S, Yan C, Hou P, Wu N et al. Identification of tumor immune infiltration-associated IncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. J Immunother Cancer. 2020; 8

23. Goldman MJ, Craft B, Hastie M, Repecka K, McDade F, Kamath A et al (2020) Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol 38:675–678

24. Bai H, Harmanci AS, Erson-Omay EZ, Li J, Coşkun S, Simon M et al (2016) Integrated genomic characterization of IDH1-mutant glioma malignant progression. Nat Genet 48:59–66

25. Olson B, Li Y, Lin Y, Liu ET, Patnaik A (2018) Mouse Models for Cancer Immunotherapy Research. Cancer Discov 8:1358

26. Price G, Bouras A, Hambardzumyan D, Hadjipanayis CJE. Current knowledge on the immune microenvironment and emerging immunotherapies in diffuse midline glioma. 2021; 69: 103453

27. Yang F, He Z, Duan H, Zhang D, Li J, Yang H et al. Synergistic immunotherapy of glioblastoma by dual targeting of IL-6 and CD40. 2021; 12: 3424
28. McGranahan T, Therkelsen KE, Ahmad S, Nagpal S (2019) Current State of Immunotherapy for Treatment of Glioblastoma. Curr Treat Options Oncol 20:24
29. Wang X, Guo G, Guan H, Yu Y, Lu J, Yu J (2019) Challenges and potential of PD-1/PD-L1 checkpoint blockade immunotherapy for glioblastoma. Journal of Experimental Clinical Cancer Research 38:87
30. Touat M, Li YY, Boynton AN, Spurr LF, Iorgulescu JB, Bohrson CL et al (2020) Mechanisms and therapeutic implications of hypermutation in gliomas. Nature 580:517–523
31. Karpel-Massler G, Nguyen T, Shang E, Siegelin MJ. Novel IDH1-Targeted Glioma Therapies. 2019; 33: 1155–66
32. Ashworth A, Lord CJ. Synthetic lethal therapies for cancer: what’s next after PARP inhibitors? 2018; 15: 564 – 76
33. Ahn D, Bekaii-Saab, TJ. Biliary tract cancer and genomic alterations in homologous recombinant deficiency: exploiting synthetic lethality with PARP inhibitors. 2020; 9: 6
34. Slade DJ, development. PARP and PARG inhibitors in cancer treatment. 2020; 34: 360 – 94
35. Patel P, Algouneh A, Hakem R. Exploiting synthetic lethality to target BRCA1/2-deficient tumors: where we stand. 2021; 40: 3001–14
36. Shen Y, Aoyagi-Scharber M, Wang B (2015) Trapping Poly(ADP-Ribose) Polymerase. J Pharmacol Exp Ther 353:446
37. Jachimowicz RD, Beleggia F, Isensee J, Velpula BB, Goergens J, Bustos MA et al. UBQLN4 Represses Homologous Recombination and Is Overexpressed in Aggressive Tumors. Cell. 2019; 176: 505 – 19 e22

Figures
Figure 1

Comparison of the genome instability between GS and GU group. (a) Oncoplots showing the top 20 mutated genes in GS and GU group. (b) The comparisons of TMB, mutation status of IDH1, PTEN, EGFR between the two groups. (c) The circos plots for each group showing the copy number variations of chromosomes, with red dots representing amplifications, blue dots representing deletions, and black dots representing no significant CNVs. (d) Boxplots exhibiting the differences of burdens of copy number
amplifications and deletions between GS and GU group. (e) The heatmap of expression patterns of top 20 differentially expressed lncRNAs between GU group and GS group, with red indicating high expressions and blue indicating low expressions.

Figure 2

Identification of clinicopathological characteristics and differentially expressed genes between GS-like and GU-like clusters. (a) Heatmap of unsupervised hierarchical clustering analyses of patients with LGG
according to the expression levels of 102 GILncRNAs. (b) Comparisons of somatic mutation count and expression level of UBQLN4 between the two clusters. (c) Overall survival curves of patients in GS-like and GU-like clusters. (d) The heatmap showing the distribution of clinicopathological factors in GS-like and GU-like clusters. * means $P < 0.05$, ** means $P < 0.01$, and *** means $P < 0.001$. (e) Barplots showing the proportion of related clinicopathological factors in the GS-like and GU-like clusters. (f) The heatmap of top 20 DEGs between GS-like and GU-like clusters. (g) The network of protein-protein interaction of the DEGs. (h) The barplot displaying 30 hub genes with highest number of adjacent nodes.

Figure 3
Comparisons of tumor-immune microenvironment (TIME), immunogenomic patterns, immunotherapy response and chemotherapy response between GS-like and GU-like clusters. (a) Comparisons of stromal scores, immune scores, and estimate scores between the two clusters. (b) Comparison of the abundance of infiltrated immune cells between the two clusters. (c) The heatmap showing the abundance of infiltrated immune cells between the two clusters. (d) Comparison of the associated immune functions between the two clusters. (e) Comparisons of the expression level of PDCD1/CD274/PDCD1LG2 and CTLA4/CD80/CD86 between the two clusters. (f) Subclass mapping analysis for predicting the likelihood of response to immunotherapy of GS-like and GU-like clusters. (g) Boxplot demonstrated the differences in the IC50 values estimated by pRRophetic algorithm between GS-Like and GU-Like clusters, regarding Veliparib, Rucaparib and Olaparib, respectively.
Figure 4

Construction and evaluation of the risk model based on GILncSig with prognostic values in the train (e-l) and test (m-s) set. (a) Univariate Cox proportional risk regression identified 33 genome-instability-related lncRNAs with predictive prognosis roles. (b) Lasso Cox analysis identified 9 genome-instability-related lncRNAs with prognostic values. (c) Determination of the optimal value of penalty parameters through 1000 replicates of cross-validation. (d) 9 lncRNAs involved in the risk model with corresponding
coefficients. (e) Overall survival curves in low- or high-risk patients by Kaplan-Meier estimates in the train set. (f) Time-dependent ROC curves of GLncSig at 1, 2 and 3 years in the train set. (g–j) Heatmap of 9 LncRNA expression profiles, scatter plot of somatic mutations, scatter plot of survival states and UBQLN4 expression in LGG patients. (k-l) Box plots for the comparisons of somatic mutations and UBQLN4 expression in highand low-risk groups. (m–s) Verification of the above results using the test set.

![Figure 5](image)

Identification of the sensitivities to immunotherapy and PARP inhibitors between different risk groups. (a) Unsupervised subclass mapping analysis of the response to the immunotherapy between patients in the two risk groups. (b-d) Boxplots showing the estimated IC50 values of Rucaparib, Olaparib and Veliparib, respectively, between patients in different risk groups.

### Supplementary Files

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

- ESM1.eps
- ESM2.eps
- ESM3.eps
