Molecular Therapy
Oncolytics
Original Article

Genomic instability in lower-grade glioma: Prediction of prognosis based on lncRNA and immune infiltration

Kai Kang,[1,4] Fucun Xie,[2,4] Yijun Wu,[1,4] Chang Han,[1] Yi Bai,[3] Junyu Long,[2] Xin Lian,[1] and Fuquan Zhang[1]

[1]Department of Radiation Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China; [2]Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China; [3]Department of Hepatobiliary Surgery, First Central Hospital, Tianjin 300190, China

Glioma is the most common type of primary malignant tumor in the central nervous system. Tumor recurrence and progression are common in lower-grade glioma (LGG). Immune checkpoint blockade (ICB), as an emerging immunotherapy, is expected to improve the prognosis of patients undergoing conventional treatment, but it currently performs poorly in glioma. We divided patients into genome-stable and -unstable groups according to the somatic mutation count and then found that the expression of *CDC20* was positively correlated with genomic instability. We compared the differences in long non-coding RNA (lncRNA) expression and immune infiltration between the two groups. Five lncRNAs and three immune cell types were identified to construct risk models and a nomogram combing clinical features. Through internal and external validation, the models exhibited sufficient ability to predict the prognosis and the possible response to ICB therapy of patients. This study provided a potential predictive approach for the precise application of ICB and support for improving the prognosis of LGG patients.

INTRODUCTION

Glioma is the most common primary malignant brain and other central nervous system tumor, accounting for 80.8% of patients.[1] Gliomas are conventionally classified into World Health Organization (WHO) grades I to IV based on histopathological features, among which WHO grades II and III are considered as lower-grade gliomas (LGGs). According to the latest WHO classification combined with molecular parameters in 2016, LGGs are divided into isocitrate dehydrogenase (IDH)-wild-type astrocytomas, IDH mutant astrocytomas, IDH mutant and 1p19q-codeleted oligodendrogliomas, and not otherwise specified (NOS) categories.[2] Despite well-developed conventional treatment, tumor recurrence and progression to WHO grade IV glioblastomas are common in LGG due to the highly invasive nature and incomplete surgical resection, causing significant differences in survival outcomes among patients.[3–5] Therefore, accurate prediction of LGG patients' prognosis is essential for treatment guidance and comprehensive management.

Genomic instability (GI), as a ubiquitous characteristic of human cancers, represents an elevated trend of acquiring new mutations in the genome, thereby driving tumor progression and resistance to therapy.[6] Previous studies have shown that higher GI is associated with worse prognosis for several cancer types, such as ovarian, colorectal, and breast cancers.[7–10] On the other hand, GI drives tumors to accumulate point mutation burden, which increases the production of neoantigens.[11] Neoantigens trigger immune responses and modulate immune infiltration in the tumor microenvironment (TME), often making tumors more sensitive to immune checkpoint blockade (ICB) and leading to better prognosis.[12,13] The specific prognostic role of GI in LGG needs further exploration.

Long non-coding RNA (lncRNA) is defined as a group of various RNA molecules with more than 200 nucleotides that do not encode proteins. lncRNAs regulate gene expression through various mechanisms, thereby affecting biological activities of cells, including glioma cells naturally, such as GI, proliferation, migration, and apoptosis.[14,15] Regarding the link between lncRNA and GI, recent studies have revealed that *NORAD* could maintain genomic stability by sequestering proteins that bind to mRNAs and inhibit their critical activity such as mitosis, DNA repair, and DNA replication.[16,17] *DINO* has been reported to bind to and promote p53 protein stabilization and plays an important role in the process of cell-cycle arrest and apoptosis in response to DNA damage.[18] Besides, increasing studies have revealed the functional roles of lncRNAs in the modulation of TME.[19,20] Compared with the total amount, the lncRNAs in which mechanisms have been revealed are only a drop in the ocean. The current research results can only partially explain the impact of lncRNA on GI and immune infiltration.

In the current study, based on the somatic mutation profiles and transcriptome profiles from public databases, we explored the impact of

Received 18 March 2021; accepted 17 July 2021;
https://doi.org/10.1016/j.omto.2021.07.011.

[4]These authors contributed equally

Correspondence: Fuquan Zhang, Department of Radiation Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 1 Shuaifuyuan, Dongcheng District, Beijing 100730, China.
E-mail: zhangfq_pumc@163.com

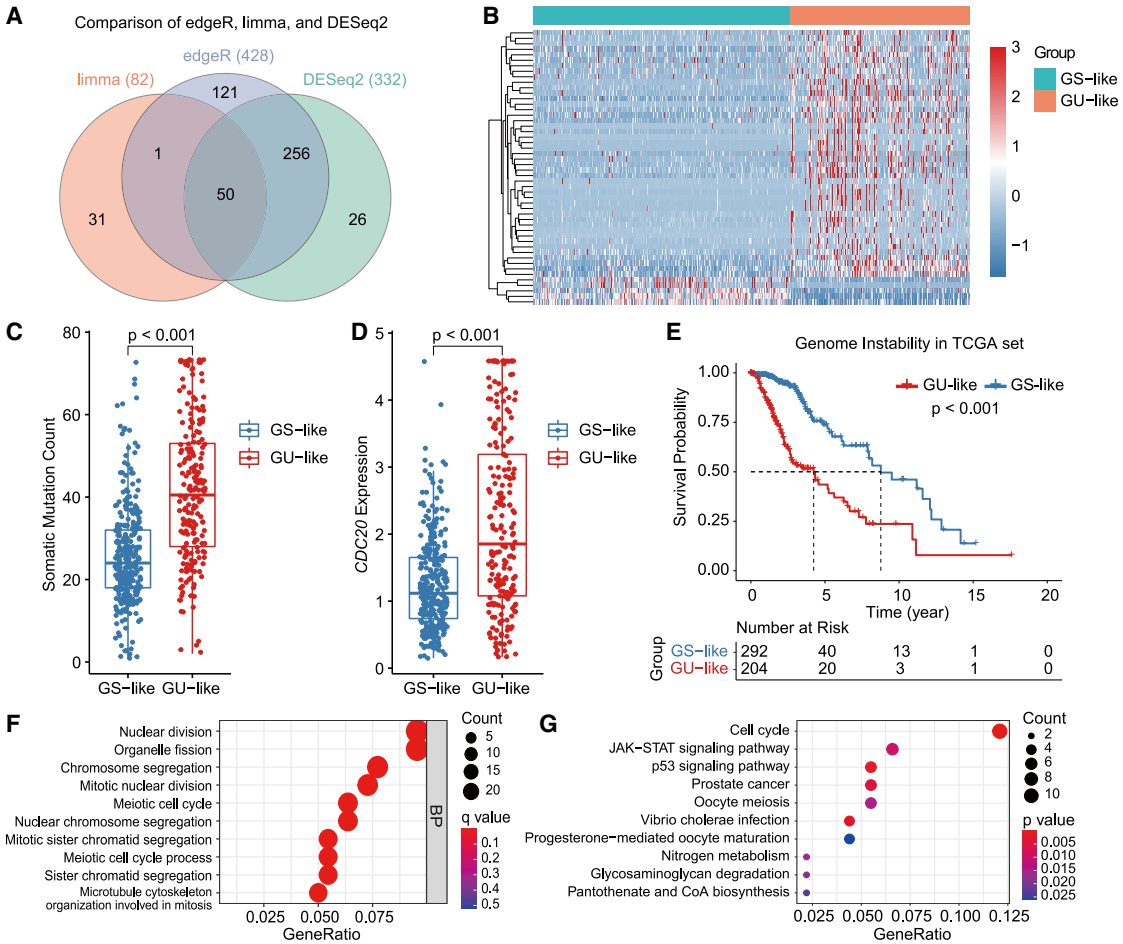

Figure 1. Identification and function enrichment analyses of GI-related lncRNAs

(A) Comparison of three differential expression analysis methods, including limma, edgeR, and DESeq2. (B) Based on 50 differentially expressed lncRNAs, patients were divided into two groups using hierarchical clustering. (C) The two groups were distinguished by the number of cumulative somatic mutations. (D) The expression of *CDC20* was higher in the GU-like group. (E) K-M survival analysis exhibited that the GS-like group had significantly better prognosis. (F) GO analysis of the biological process based on the GI-related lncRNAs. (G) KEGG analysis based on the GI-related lncRNAs.

GI on LGG from the level of lncRNA expression and immune infiltration. Corresponding risk models and a nomogram were developed and validated to predict the prognosis and the possible response to ICB therapy of patients. The results of this study may help identify patients suitable for ICB therapy and improve the current plight of LGG treatment. The conclusions need to be further confirmed in biological experiments and clinical trials.

RESULTS

Identification of GI-related lncRNAs

Based on the somatic mutation profiles of LGG samples (n = 496) downloaded from The Cancer Genome Atlas (TCGA), we calculated the number of cumulative point mutations for each sample.

Samples with mutations in the top 25% (n = 130) and bottom 25% (n = 128) were defined as genome-unstable (GU) group and genome-stable (GS) group, respectively. Subsequently, the GS group was set as the control group, and the differential expression analysis of lncRNAs between GU and GS groups was performed using three methods separately, limma, edgeR, and DESeq2. After calculating the fold change (FC) and false discovery rate (FDR) of lncRNAs, we set $|\log_2 FC| > 1.5$ and FDR < 0.05 as the filter criteria and then took the intersection of results from three methods. Finally, 50 differentially expressed lncRNAs were selected for subsequent analyses (Figure 1A).

With the aid of hierarchical clustering, 496 LGG samples were separated into two groups in terms of the differential expression pattern of 50 lncRNAs, which were exhibited in a heatmap (Figure 1B). The cumulative somatic mutations in the two groups were significantly different, and we defined the group with fewer mutations as the GS-like group, whereas the other group as the GU-like group

Table 1. Comparison of clinical characteristics of patients in the training set and testing set

Variables	Training set	Testing set	TCGA set	p value
Gender				
Female	163 (46.8%)	59 (39.9%)	222 (44.8%)	0.183
Male	185 (53.2%)	89 (60.1%)	274 (55.2%)	
Age				
≤40	165 (47.4%)	76 (51.4%)	241 (48.6%)	0.481
>40	183 (52.6%)	72 (48.6%)	255 (51.4%)	
Histological type				
Astrocytoma	130 (37.4%)	57 (38.5%)	187 (37.7%)	0.193
Oligoastrocytoma	96 (27.6%)	30 (20.3%)	126 (25.4%)	
Oligodendroglioma	122 (35.1%)	61 (41.2%)	183 (36.9%)	
WHO grade				
Grade II	174 (50.0%)	65 (43.9%)	239 (48.2%)	0.281
Grade III	174 (50.0%)	82 (55.4%)	256 (51.6%)	
Unknown	0 (0%)	1 (0.7%)	1 (0.2%)	
Molecular feature				
IDH mutant and 1p19q codeleted	52 (14.9%)	31 (21.0%)	83 (16.7%)	0.430
IDH mutant	95 (27.3%)	39 (26.4%)	134 (27.0%)	
IDH wild type	37 (10.6%)	16 (10.8%)	53 (10.7%)	
Unknown	164 (47.1%)	62 (41.9%)	226 (45.6%)	

TCGA, The Cancer Genome Atlas; WHO, World Health Organization; IDH, isocitrate dehydrogenase.

(Wilcoxon test, $p < 0.001$; Figure 1C). Besides, the expression of *CDC20* in the GU-like group was much higher (Wilcoxon test, $p < 0.001$; Figure 1D). *CDC20* is a newly identified gene related to genomic stability. The overexpression of *CDC20* may forcibly activate the unregulated mitosis cycle of cells that have undergone prolonged mitotic arrest, which in turn leads to cell proliferation in an unregulated manner.[21] According to the Kaplan-Meier (K-M) survival analysis between the two groups, the GS-like group had better overall survival (OS) (log-rank test, $p < 0.001$; Figure 1E).

Furthermore, based on the transcriptome profiles downloaded from TCGA, we analyzed the co-expression relationship between lncRNAs and protein-coding genes. A co-expression network was constructed based on the top 10 most related protein-coding genes for each differentially expressed lncRNA (Figure S1). Functional enrichment analyses of lncRNA-related protein-coding genes were conducted to explore the prospective pathways of the association between GI and lncRNAs. In terms of the Gene Ontology (GO) analysis, the function was enriched in the biological process such as nuclear division and chromosome segregation (Figure 1F). As for the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, pathways such as cell cycle and p53 signaling pathway were identified (Figure 1G). These enriched GI-related pathways provided a possible explanation for the effects of lncRNAs on GI in LGG samples.

Association between GI-related lncRNAs and prognosis

To explore the impact of GI-related lncRNAs on prognosis, further survival-related analyses were performed for screening modeling lncRNAs. Based on a ratio of seven to three, 496 LGG patients from TCGA were separated randomly into the training set (n = 348) and the testing set (n = 148). Patients' clinical characteristics from the two sets were compared to ensure the differences between the two sets were insignificant (Table 1). In the training set, through univariate Cox regression, 49 of the 50 GI-related lncRNAs met the criteria that the p value was less than 0.05. Subsequently, in the least absolute shrinkage and selector operation (LASSO) regression, eight lncRNAs met the requirement to be included more than 990 times in all 1,000 repetitions. Finally, through further filtration of stepwise regression, a model based on the expression of five lncRNAs in the training set was established as follows:

$$\textit{lncRNA risk score} = 0.4069 \times \textit{H19 expression level} - 0.6600 \times \textit{AC109439.1 expression level} + 0.3623 \times \textit{LINC02587 expression level} + 1.1741 \times \textit{AC015909.3 expression level} + 1.6457 \times \textit{AC003986.2 expression level}$$

.

Based on the median risk score of the training set (−0.431), patients from the training set, testing set, and total TCGA set were divided into high- and low-risk groups, respectively. In the total TCGA set, we visualized the relationship among lncRNA risk score, expression level, and survival outcome (Figure 2A). The higher expression level of *AC109439.1* indicated better prognosis, whereas the association in the other four lncRNAs was opposite. The somatic mutation count and the expression of *CDC20* in the high-risk group were both higher than those in the low-risk group (Wilcoxon test, $p < 0.001$; Figure 2B).

As for the lncRNA risk model's predictive ability, K-M analyses and receiver operating characteristic (ROC) curves were performed in the training set, testing set, and total TCGA set, respectively. Patients in the high-risk groups had worse prognosis than the low-risk groups (log-rank test, $p < 0.001$, Figure 2C; $p = 0.002$, Figure 2D; $p < 0.001$, Figure 2E). According to the ROC curves of 1-, 3-, and 5-year OS prediction, the area under the curve (AUC) ranged between 0.769 and 0.942 (Figures 2F–2H). The lncRNA risk model exhibited a quite good prediction performance in the internal validation. As for the external validation, since modeling lncRNAs other than *H19* lacked enough expression data, we only explored the prognostic role of *H19* in the Chinese Glioma Genome Atlas (CGGA) set (n = 592) and Gene Expression Omnibus (GEO): GSE16011 set (n = 107). Similar to the performance in TCGA set, higher expression of *H19* in external sets also indicated worse prognosis (log-rank test, CGGA: $p < 0.001$; GEO: GSE16011: $p = 0.117$; Figure S2).

Association between GI-related immune infiltration and prognosis

To explore the association between GI and immune infiltration, we compared the enrichment scores of 28 immune cell types, stromal

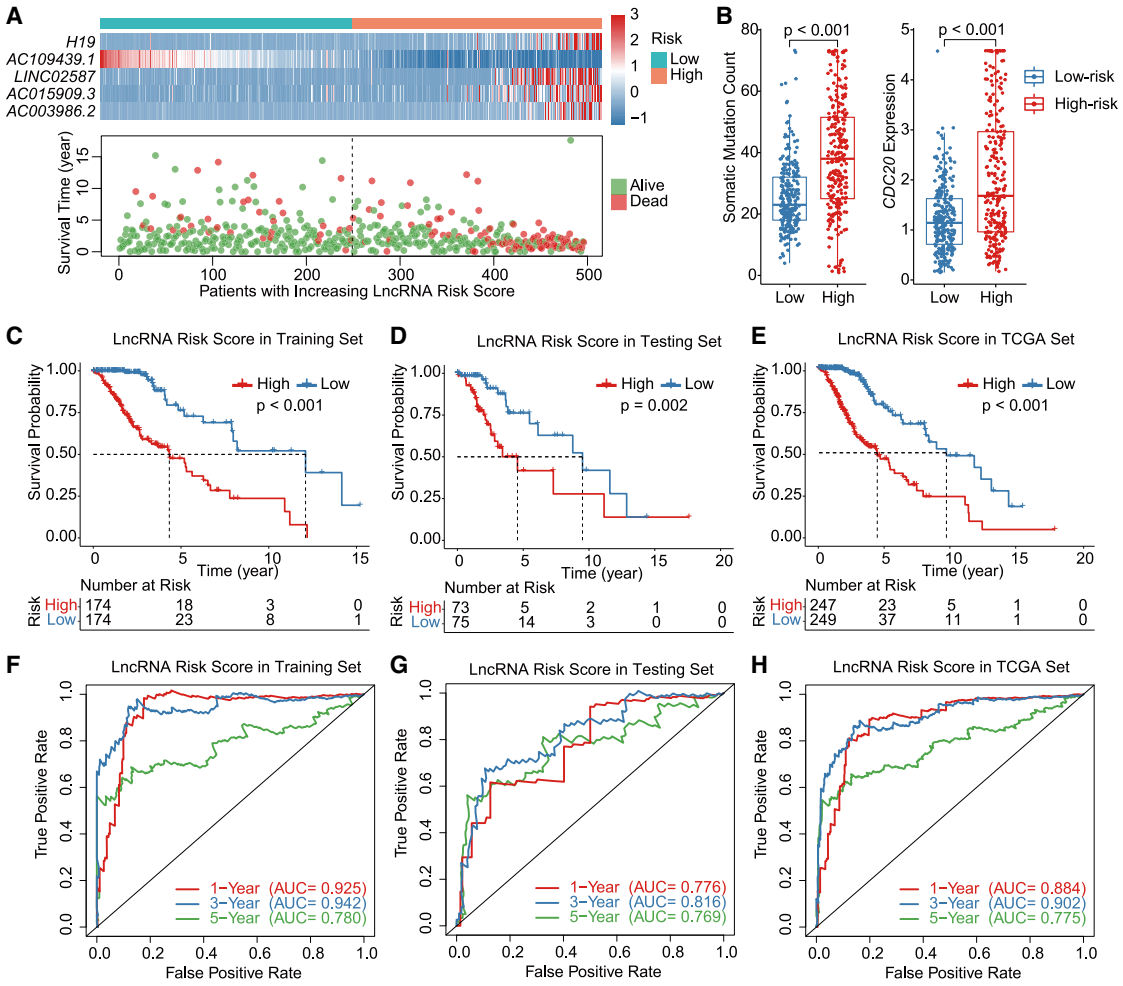

Figure 2. Construction and validation of the lncRNA risk model

(A) With the increase of lncRNA risk score, the differences in lncRNA expression level and survival outcome. (B) Comparison of somatic mutation count and *CDC20* expression between low and high lncRNA risk groups. (C–E) Groups with higher lncRNA risk scores were related to worse prognosis in the training set, testing set, and TCGA set. (F–H) ROC curves for 1-, 3-, and 5-year OS prediction of the lncRNA risk model were performed in the training set, testing set, and TCGA set.

score, and immune score between GS- and GU-like groups. Since the distribution of each score was quite different, the comparison between the two groups was exhibited after consecutive mean subtraction and standardization (Figure 3A). The differences between the two groups of effector memory CD8 T cell and eosinophil were not significant, and the score of the CD56-dim natural killer cell in the GS-like group was significantly higher than the GU-like group, whereas all other scores in the GU-like group were significantly higher than the GS-like group.

To explore the relationship between prognosis and immune cell scores, we conducted univariate Cox regression analyses for 28 immune cell types. In the training set, 22 cell types met the criteria that the p value was less than 0.05. LASSO regression was performed for further filtration and found that four cell types were included more than 990 times in all 1,000 repetitions. With the aid of stepwise regression, an immune risk score based on enrichment scores of three cell types was constructed as follows: $\textit{immune risk score} = 13.610 \times \textit{type 2 T helper cell score} + 4.202 \times \textit{memory B cell score} + 10.580 \times \textit{plasmacytoid dendritic cell score}$.

Like the lncRNA risk score, patients were separated into high- and low-risk groups in terms of the median risk score of the training set (18.657). In the total TCGA set, the relationship among immune risk score, enrichment score, and survival outcome was exhibited (Figure 3B). Higher scores of all three cell types were correlated with worse prognosis. The somatic mutation count and the expression of *CDC20* were both elevated in the high-risk group (Wilcoxon test, $p < 0.001$; Figure 3C). Besides, with combining mutation profiles, we compared the difference in *IDH1* and *TP53* mutation rates between the high- and low-risk groups. It was found that fewer *IDH1* mutations but more *TP53*

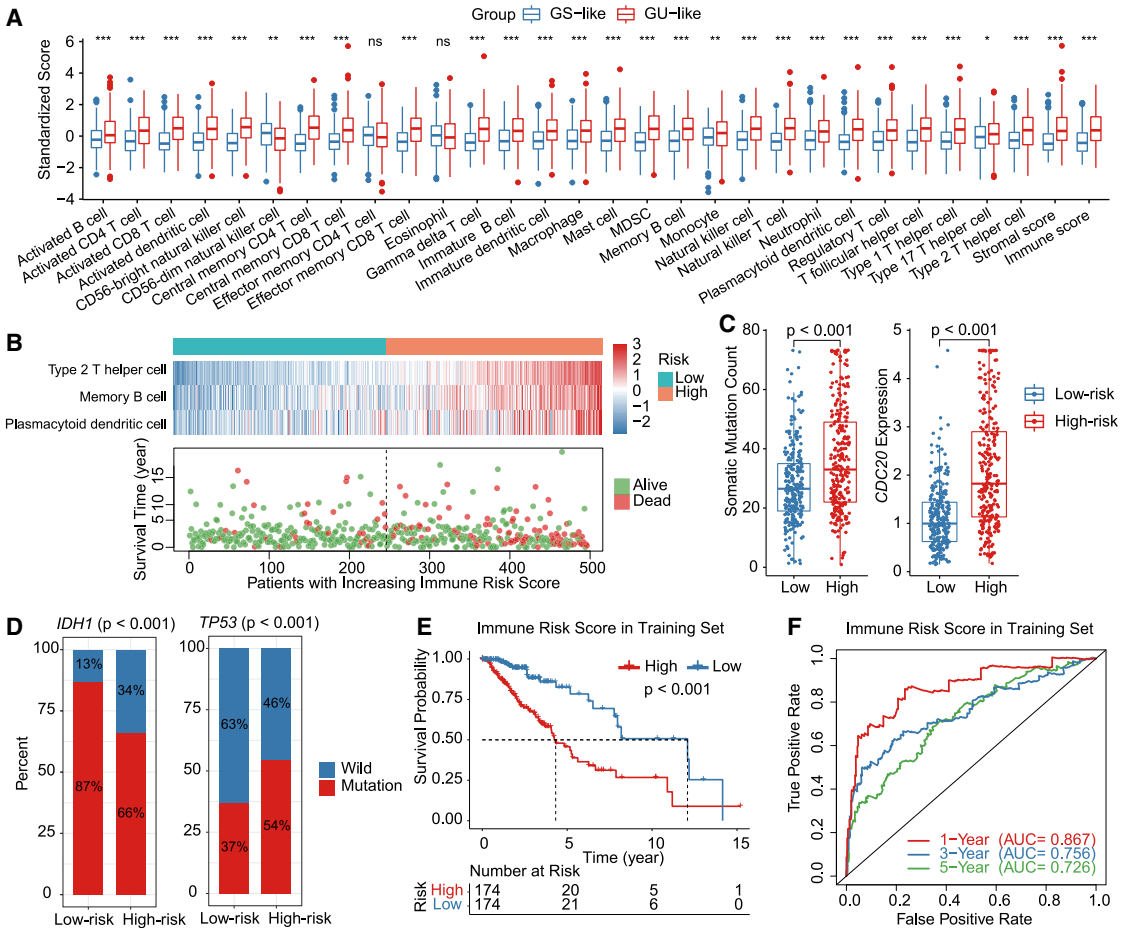

Figure 3. Evaluation of the immune infiltration and construction of the immune risk model

(A) Comparison of enrichment scores of 28 immune cell types between GS- and GU-like groups. (B) With the increase of immune risk score, the differences in immune enrichment level and survival outcome. (C) Comparison of somatic mutation count and *CDC20* expression between low and high immune risk groups. (D) Differences in *IDH1* and *TP53* mutation rates between low and high immune risk groups. (E) Groups with higher immune risk scores were related to worse prognosis in the training set. (F) ROC curves for 1-, 3-, and 5-year OS prediction of the immune risk model were performed in the training set. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

mutations occurred in the high-risk group (chi-square test, $p < 0.001$; Figure 3D).

K-M analyses and ROC curves were also performed for the immune risk model to assess the prediction accuracy. In the training set, patients' survival outcomes in the high-risk group were significantly worse (log-rank test, $p < 0.001$; Figure 3E). The AUCs of the ROC curves of 1-, 3-, and 5-year OS were 0.867, 0.756, and 0.726, respectively (Figure 3F). As for three validation sets, similar to the training set, high-risk scores were also associated with poor prognosis in the testing set, CGGA set, and GEO: GSE16011 set (log-rank test, $p = 0.012$, Figure 4A; log-rank test, $p < 0.001$, Figure 4B; log-rank test, $p = 0.003$, Figure 4C). According to the ROC curves of three validation sets, the model showed better effectiveness in the 3-year OS prediction, and the AUCs were 0.747, 0.702, and 0.778, respectively (Figures 4D–4F). Compared with the lncRNA risk model, the immune risk model was much easier to promote in the external validation sets due to the higher availability of the enrichment score.

With the aid of the Tumor Immune Estimation Resource (TIMER), we analyzed the relationship between *CDC20* and immune infiltration in LGG samples, involving B cell, $CD8^+$ T cell, $CD4^+$ T cell, macrophage, neutrophil, and dendritic cell (DC). The expression level of *CDC20* was positively correlated with the level of immune infiltration when it was relatively low, especially B cell and DC (Figure 4G). Besides, compared with diploid copy numbers of *CDC20*, the copy number of *CDC20* was significantly positively correlated with the level of immune infiltration (Figure 4H).

To further improve the immune risk model's prediction accuracy, we developed a nomogram combined with the model and several clinical features. Through univariate and multivariate Cox regression analyses, immune risk score, age, WHO grade, and molecular feature were

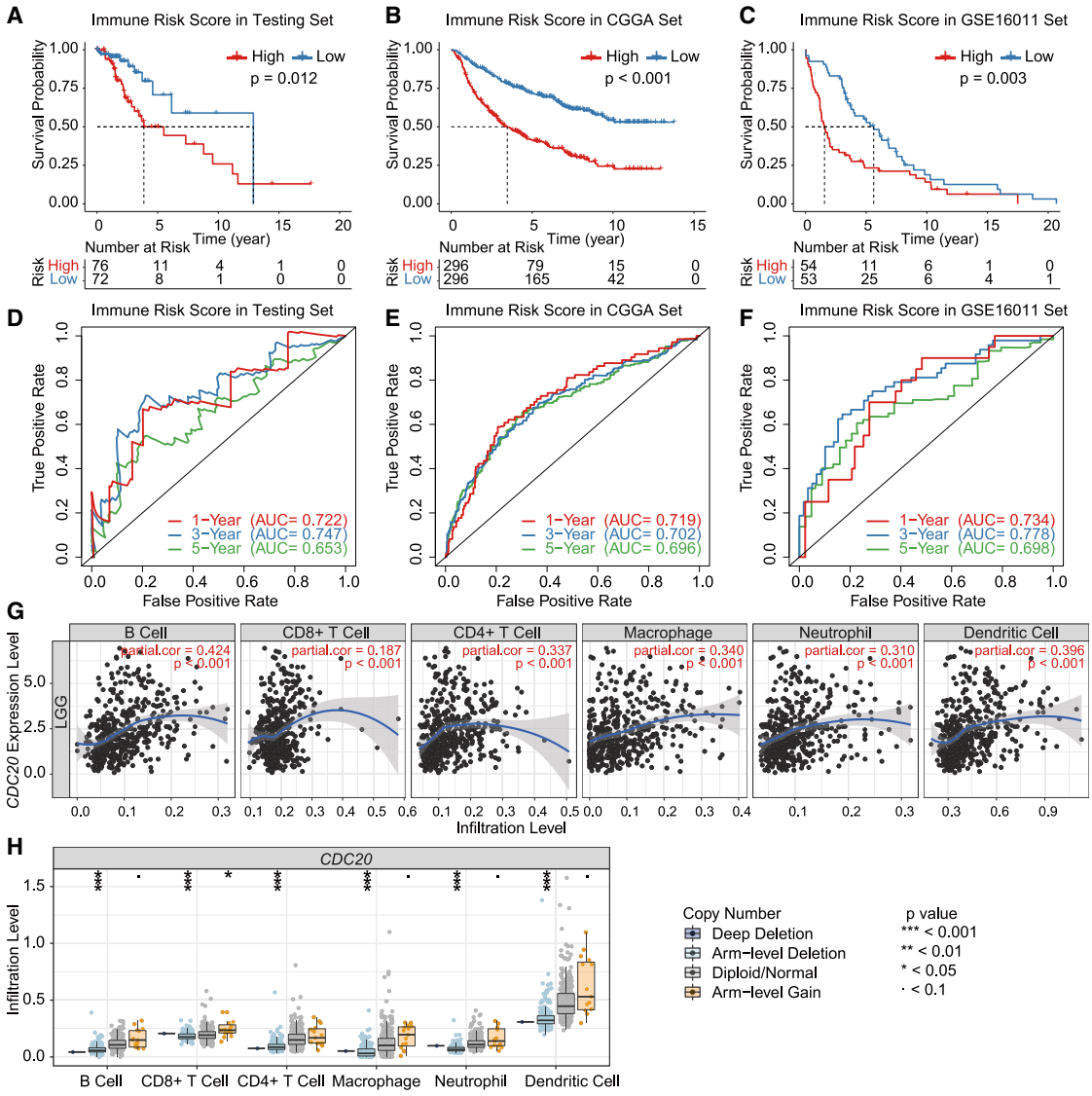

Figure 4. Validation of the immune risk model

(A–C) Groups with higher immune risk scores were related to worse prognosis in the testing set, CGGA set, and GEO: GSE16011 set. (D–F) ROC curves for 1-, 3-, and 5-year OS prediction of the immune risk model were performed in the testing set, CGGA set, and GEO: GSE16011 set. (G) Association between immune infiltration and the expression level of *CDC20*. (H) Association between immune infiltration and the copy number of *CDC20*.

singled out as factors in the nomogram (Table 2; Figure 5A). The Harrell's consistency index (C-index) of the nomogram was 0.853 in TCGA set. Subsequently, the nomogram was validated in both TCGA set and the CGGA set. The calibration curves of the OS prediction at 1, 3, and 5 years exhibited a pretty good prediction ability of the nomogram (Figures 5B and 5C). The AUCs of the ROC curves of 1-, 3-, and 5-year OS prediction were 0.893, 0.905, and 0.825 in TCGA set, whereas AUCs in the CGGA set were 0.760, 0.800, and 0.787 (Figures 5D and 5E). Decision curve analysis (DCA) curves showed that the net benefits of clinical decisions based on the nomogram were better than those based on individual included factors (Figures 5F and 5G).

Higher immune risk score indicated better response to immunotherapy

Based on the immune risk model, LGG samples were separated into high-risk and low-risk groups, and we compared the expression levels of several immune checkpoints between the two groups, including PD-L1 (*CD274*), CTLA-4 (*CTLA4*), PD-1 (*PDCD1*), LAG-3 (*LAG3*), TIM-3 (*HAVCR2*), VISTA (*VSIR*), and CD161 (*KLRB1*). The expression levels of above immune checkpoints in the high-risk group were significantly higher than the low-risk group (Wilcoxon test, PD-L1: $p < 0.001$; CTLA-4: $p < 0.001$; PD-1: $p < 0.001$; LAG-3: $p = 0.015$; TIM-3: $p < 0.001$; VISTA:

Table 2. Univariate and multivariate Cox regression for the development of the nomogram

	Univariate analysis		Multivariate analysis	
Variables	HR (95% CI)	p value	HR (95% CI)	p value
Gender				
Female	1 (ref)			
Male	0.95 (0.60−1.53)	0.846		
Age	1.06 (1.04−1.08)	<0.001	1.05 (1.03−1.07)	<0.001
WHO grade				
Grade II	1 (ref)		1 (ref)	
Grade III	3.80 (2.20−6.54)	<0.001	2.38 (1.28−4.41)	0.006
Molecular feature				
IDH mutant and 1p19q codeleted	1 (ref)		1 (ref)	
IDH mutant	1.30 (0.68−2.49)	0.419	1.41 (0.72−2.76)	0.321
IDH wild type	9.60 (4.95−18.62)	<0.001	4.48 (1.97−10.16)	<0.001
Immune risk score	3.18 (2.27−4.44)	<0.001	1.53 (1.01−2.32)	0.046

HR, hazard ratio; CI, confidence interval; ref, reference.

$p < 0.001$; CD161: $p < 0.001$; Figure 6A). Besides, with the aid of the Immune Cell Abundance Identifier (ImmuCellAI), we predicted the response to ICB of patients in TCGA and CGGA sets. The response rates in the high-risk groups were higher than the low-risk groups in both sets (chi-square test, $p < 0.001$; Figure 6B). Based on the predicted response to ICB, we divided patients from TCGA and CGGA sets into no-response and response groups, respectively. The differences of three modeling immune cell scores between the two groups were explored. The scores in the response groups were higher than the no-response groups in both sets (Wilcoxon test, $p < 0.001$; Figure 6C). Similarly, the differences of modeling lncRNAs between the two groups were also compared (Figure S3).

Furthermore, we examined the immune risk model in ICB therapy datasets. For cancer types with many somatic mutations and many neoantigens, such as melanoma and urothelial cancer, the model was validated in a total set of GEO: GSE35640 (n = 56), GEO: GSE115821 (n = 11), and IMvigor210 (n = 348). The immune risk score of the response group was significantly higher than the no-response group (Wilcoxon test, $p < 0.001$; Figure 6D). Besides, for clear cell renal cell carcinoma with fewer somatic mutations, the verification was performed in only one dataset (n = 33). Similarly, the risk score of the clinical benefit group was higher than the no clinical benefit group (t test, $p = 0.046$; Figure 6E). Moreover, we explored the association between the immune risk model and targets of chimeric antigen receptor (CAR) T cell therapy. The expression levels of interleukin-13 receptor (IL-13R)α2 and erythropoietin-producing hepatocellular receptor A2 (EphA2) were higher in the high-risk group compared to the low-risk group (Wilcoxon test, IL-13Rα2: $p < 0.001$; EphA2: $p < 0.001$; Figure 6F).

Higher ratio of neoantigen to mutation indicated better prognosis

Based on the previous publication of TCGA research network, we obtained neoantigens predicted from a single-nucleotide variant (SNV) of 494 patients and neoantigens predicted from insertion-deletion (InDel) of 391 patients in TCGA set. We explored the association between immune risk scores and predicted neoantigens. The SNV neoantigen counts in the high-risk group were significantly higher than the low-risk group, whereas the difference of InDel neoantigen counts was not significant (Wilcoxon test, SNV neoantigen: $p < 0.001$; InDel neoantigen: $p = 0.61$; Figure 7A).

Furthermore, we calculated the ratio of the neoantigen counts to the mutation counts. There was no significant difference in the ratios between high- and low-risk groups (Wilcoxon test, SNV neoantigen: $p = 0.14$; InDel neoantigen: $p = 0.13$; total neoantigen: $p = 0.14$; Figure S4). Nonetheless, according to the K-M analysis, a higher ratio of total neoantigen to mutation was correlated for a better prognosis (log-rank test, $p = 0.011$; Figure 7B).

DISCUSSION

In this study, we identified GI-related lncRNAs and divided patients into GS-like and GU-like groups based on the expression levels of 50 lncRNAs. Compared to the GS-like group, patients in the GU-like group had significantly more somatic mutations and higher expression levels of *CDC20*. *CDC20* plays a crucial role in cell division and genomic stability, and the expression is elevated in various cancers, indicating an oncogenic function in tumorigenesis.[22] Besides, it has been observed that high expression of a gene module co-expressed with *CDC20* correlated severe GI in glioma.[23] The marked difference of somatic mutation and expression level of *CDC20* between GS-like and GU-like groups indicated that the hierarchical clustering method was reliable enough for subsequent analyses.

Similar to previously reported cancer types, possibly due to the driving effect of GI on tumor progression and therapy resistance, LGG patients in the GU-like group had a worse prognosis than those in the GS-like group. However, patients with a higher ratio of neoantigen count to mutation count had a better prognosis, although not receiving immunotherapy. According to functional enrichment analyses, protein-coding genes co-expressed with 50 GI-related lncRNAs were involved in pathways such as nuclear division, chromosome segregation, cell cycle, and p53 signaling pathway. Among those GI-related pathways, the p53 signaling pathway induces tumor-suppressive mechanisms such as cell-cycle arrest and apoptosis, thereby avoiding DNA damages and leading to GI.[24] In addition to the impact of lncRNAs on GI, we also explored the prognostic role of lncRNAs in LGG. Five lncRNAs were finally singled out for the construction of the lncRNA risk model. The expression levels of *H19*, *LINC02587*, *AC015909.3*, and *AC003986.2* were negatively correlated with prognosis, whereas the expression level of *AC109439.1* was positively correlated with prognosis. It had been reported that *H19* promoted the proliferation and metastasis of glioma through the interaction

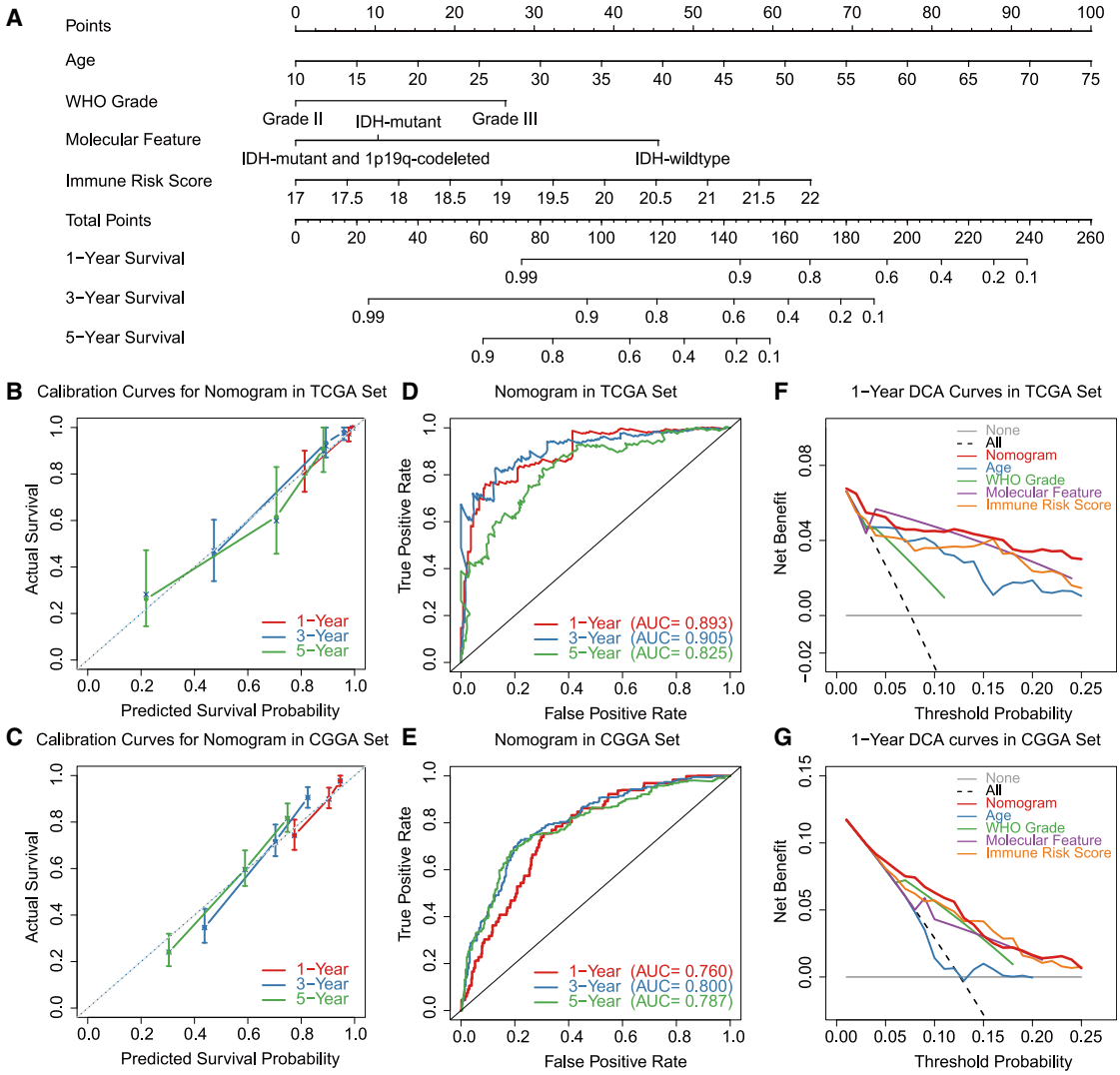

Figure 5. Development and evaluation of the nomogram

(A) The nomogram was developed based on immune risk score, age, WHO grade, and molecular feature. (B and C) Calibration curves of the nomogram on 1-, 3-, and 5-year survival probability. (D and E) ROC curves for 1-, 3-, and 5-year OS prediction of the nomogram were performed in both TCGA and CGGA sets. (F and G) DCA curves exhibited that clinical decisions based on the nomogram had the highest net benefits.

with several microRNAs.[25,26] Besides, *H19* can destabilize the genome and induce new mutations in tumor cells.[27] The verified association of *H19* with prognosis and GI indicates that plausible mechanisms may also exist in the association of other modeling lncRNAs, which needs to be explored by further research.

The difference in immune infiltration between GS- and GU-like groups also played a crucial role in the prognosis of LGG patients. Among 28 types of immune cells, the enrichment scores of most types were significantly higher in the GU-like group, which may be related to the production of neoantigens caused by GI. We also observed that the expression level and copy number of *CDC20* were both positively correlated with the immune infiltration level in LGG, indicating that *CDC20* may cause higher immune infiltration by destructing the genomic stability. Three cell types were finally singled out to construct the immune risk model, and higher enrichment scores of all three types were correlated with worse prognosis. Typically, a type 2 T helper (Th2) cell mediated immune responses related to parasitic infections and chronic inflammatory diseases.[28] Moreover, compared to healthy subjects, the immune response in patients with glioblastomas exhibited a significantly elevated level of Th2 cells, and the trend was more remarkable for recurrent glioblastomas.[29] As for the memory B cells, the negative correlation between enrichment levels and prognosis still lacked sufficient explanation. Besides, it has been reported that the infiltration level of memory B cells in tumors of responders to ICB was significantly higher than tumors of non-responders.[30] This

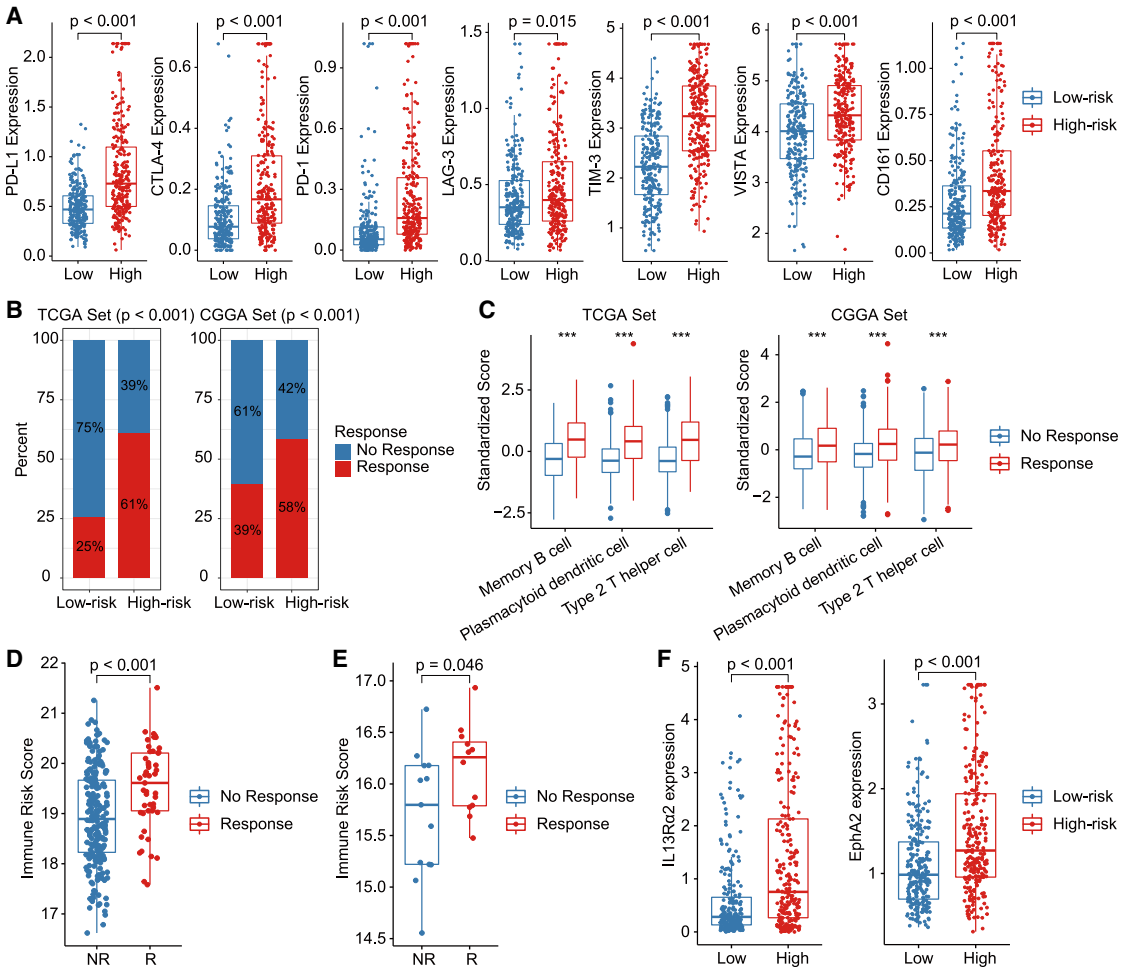

Figure 6. Prediction of the response to immunotherapy

(A) The expression levels of seven immune checkpoints were higher in the high immune risk group. (B) The rate of response to ICB was higher in the high immune risk group in both TCGA and CGGA sets. (C) The scores of three modeling immune cells were higher in the response group in both TCGA and CGGA sets. (D) The immune risk score of the response group was higher than the no-response group in melanoma and urothelial cancer. (E) The immune risk score of the response group was higher than the no-response group in clear cell renal cell carcinoma. (F) The expression levels of two potential targets of CAR T cell therapy were higher in the high immune risk group.

finding indicates that patients with higher immune risk scores may also have better responses to ICB. The plasmacytoid DC (pDC), as the major antigen presenting cell in glioma, was involved in multiple immune-suppression mechanisms, such as inducing an increased expression of indoleamine 2,3 dioxygenase to recruit immunosuppressive regulatory T cells and binding the PD-1 of T cells via the PD-L1 expressed by pDCs.[31–33] The immune suppression in glioma was consistent with the relationship between pDCs and prognosis, and it indicated that higher immune risk scores might be associated with better prognosis of patients undergoing ICB therapy.

After exploring the prognostic role of infiltrating immune cells, we also evaluated the accuracy of the immune risk score to predict the response to ICB. In addition to PD-L1, which is widely recognized as a predictive biomarker in immunotherapy in other cancers,[34] the expression levels of several other immune checkpoints were also positively related to the immune risk score. Among them, CD161 is a newly discovered immunosuppressive receptor expressed in glioma-infiltrating T cells, and the blockade of CD161 enhances the anti-tumor effect.[35] The pathway of CD161 shares similar mechanisms with the PD-1 pathway and is expected to be a novel target of immunotherapy for gliomas. Besides, the expression levels of IL-13Rα2 and EphA2 were positively correlated to the immune risk score. IL-13Rα2 has long been recognized as an available target for CAR T cell therapy.[36] CAR T cells targeting EphA2 have exhibited anti-tumor effects in animal experiments, and related clinical trials are already underway.[37,38]

Based on the experience in other cancers, higher expression levels of immune checkpoints may indicate a better response to ICB, which to some extent reflected the predictive value of the immune risk score. However, the conclusions were not based on the direct impact of

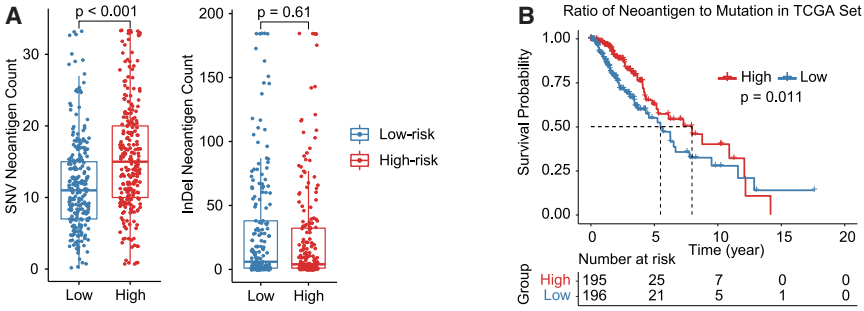

Figure 7. Evaluation of the predicted neoantigen
(A) The SNV neoantigen counts were higher in the high immune risk group. (B) The group with a higher ratio of neoantigen to mutation had a better prognosis.

these immune checkpoints on biology, so further experiments are indispensable to support the results. Besides, with ImmuCellAI predictions, we found that patients with higher immune risk had a higher potential response rate to ICB. Furthermore, we validated the immune risk model in several ICB treatment cohorts and found that the model's predictive ability was not limited to cancer types and may also be applicable to glioma, which needs to be validated in clinical trials of glioma. Perhaps the immune risk model we constructed can help improve the plight of ICB therapy in glioma.

Although the current study exhibited the significant role of GI at the level of lncRNA and immune infiltration in LGG, several shortcomings need to be improved in further studies. The risk models and the nomogram exhibited good predictive performance in predicting prognosis. In other cancers, these immune checkpoints were related to the response to ICB, but further research is required in glioma. We have verified them in the CGGA, GEO, and ICB treatment sets but still need to be verified in more LGG sets, especially in the LGG cohort under ICB therapy. Furthermore, we identified pathways, lncRNAs, and immune cells that were closely related to GI. However, due to the limitations of bioinformatics methods, the correlation was mainly understood through speculation, and the specific mechanisms on the level of molecular still need to be further explored. In addition, the difference in the format of sequencing data due to various sequencing processes may cause the distribution of risk score to vary with research cohorts, and the employment of a consistent cut-off value for grouping may bring errors to the analysis. Besides, the use of a merged CGGA set for external validation may lead to reduced reliability of the conclusions.

In conclusion, with the aid of bioinformatics analysis methods, this study interpreted the influence of GI on LGG from the perspective of lncRNA and immune infiltration. We identified that higher GI in LGG was associated with worse prognosis. Besides, several lncRNAs and infiltrating immune cells were closely correlated to GI and survival outcomes. The risk models and nomogram based on the expression level of lncRNAs and the infiltration level of immune cells exhibited a good predictive performance. Among them, the immune risk model, after further verification based on biological experiment and clinical feasibility, may help identify patients suitable for ICB therapy and improve the current plight of LGG treatment.

MATERIALS AND METHODS

Data source

From the LGG project of TCGA (https://portal.gdc.cancer.gov/), we downloaded somatic mutation profiles, gene-expression data, and clinical features of patients. After excluding patients with incomplete information and survival time = 0, a total of 496 patients were retained for subsequent analyses. The gene-expression data in the "counts" type was used for differential expression analysis, whereas the "FPKM" (fragments per kilobase per million mapped reads) type after the $\log_2$(FPKM + 1) transformation was used for other analyses. As for the external validation sets, transcriptome profiles and clinical features of datasets "mRNAseq_693" and "mRNAseq_325" from the CGGA database (http://www.cgga.org.cn/) were downloaded, and a total of 592 LGG patients with complete survival information were retained for validation.[39,40] The CGGA set was merged from two datasets, thus the "ComBat" algorithm in the "sva" R package was employed to eliminate the batch effect.[41] The validation was also conducted in the GEO: GSE16011 set downloaded from the GEO database, which contained 107 patients with complete survival information.[42] Furthermore, the results were also validated in several ICB treatment datasets. Two datasets of melanoma treatment, GEO: GSE35640 and GSE115821, were downloaded from the GEO database.[43,44] Besides, a dataset of urothelial cancer, namely IMvigor210, and a dataset of clear cell renal cell carcinoma were obtained according to Supplemental materials of respective studies.[45,46] If necessary, the validation sets were also log transformed to ensure that they were similar in the order of magnitude to TCGA set.

Differential expression analysis and hierarchical cluster analysis

From the Ensembl database (http://useast.ensembl.org/index.html), we downloaded the gene transfer format file, namely "Homo_sapiens.GRCh38.102.gtf," which contained annotations of human genes. Then we singled out the expression data of lncRNAs from the complete transcriptome profiles according to the annotations. According to the mutation data, samples with cumulative point mutations in the top 25% and bottom 25% were defined as GU and GS groups. With the aid of "limma," "edgeR," and "DESeq2" R packages, the differential expression analysis of lncRNAs between the two groups was conducted using three different methods based on the gene-expression data in the "counts" type. After the differences in the expression of all lncRNAs between the two groups were calculated, $|\log_2 FC| > 1.5$ and FDR < 0.05 were set as the standard, and lncRNAs that met the standards of all three methods were retained for subsequent analyses. In terms of the expression levels of lncRNAs, using Euclidean distance and the complete linkage method for the hierarchical cluster analysis, patients in TCGA set were separated into GU-like and GS-like groups, and the result was exhibited using the

"pheatmap" R package. K-M survival analysis was performed to explore the prognostic role of GI.

Functional enrichment analysis

The co-expression relationship between differentially expressed lncRNAs and protein-coding genes was explored using Pearson correlation analysis. The top 10-related protein-coding genes based on correlation coefficients for each lncRNA were retained for functional enrichment analysis. With the use of "clusterProfiler" and "enrichplot" R packages, GO and KEGG analyses were performed based on the Entrez Gene ID corresponding to each related protein-coding gene.

Risk-score model and nomogram

Before constructing risk-score models, TCGA set was divided into the training set (70%) and the testing set (30%) randomly. A chi-square test was employed to compare two sets' clinical features and assess whether the differences were significant. Univariate Cox regression, LASSO regression, and stepwise regression were used in sequence to single out candidates of variables in the model. The criteria of the LASSO regression was to be retained in the model more than 990 times in all 1,000 repetitions. Based on coefficients calculated by the stepwise regression, the risk-score model was constructed as follows: $risk\ score = \sum_{i=1}^{n}(coef_i \times \alpha_i)$. The α_i represented the expression levels of lncRNAs in the gene risk model, while representing the enrichment scores of immune cells in the immune risk model. In the training set, K-M survival analyses and ROC curves were performed to evaluate the predictive accuracy of models, also in the testing set and external validation sets.

The nomogram was constructed based on the immune risk model. Univariate and multivariate Cox regression analyses were performed successively to assess the importance of the immune risk score and several clinical features in the prediction of OS, with $p < 0.05$ as the criteria. The nomogram was developed using "rms" R package, and the C-index was calculated using the "rcorr.cens" function from the "Hmisc" R package. Besides, in both TCGA and CGGA sets, calibration curves for different years were plotted to exhibit the disparity between actual and predicted survival possibility. ROC curves were plotted to assess the prediction accuracy. DCA curves were also plotted to evaluate the net benefits of different clinical decisions.

Evaluation of immune infiltration and neoantigen

With the aid of "GSVA" and "GSEABase" R packages, the enrichment scores of 28 types of infiltrating immune cells were calculated using single-sample Gene Set Enrichment Analysis (ssGSEA).[47,48] Besides, the scores of stromal cells and immune cells in TME pre-calculated by the ESTIMATE (Estimation of Stromal and Immune Cells in Malignant Tumors Using Expression Data) algorithm were downloaded (https://bioinformatics.mdanderson.org/estimate/). TIMER (https://cistrome.shinyapps.io/timer/) was used to explore the relationship between immune infiltration and the expression level and copy number of the GI-related gene.[49] Based on the transcriptome profiles of TCGA and CGGA sets, ImmuCellAI (http://bioinfo.life.hust.edu.cn/ImmuCellAI/#!/) was used to predict the response of patients to ICB therapy.[50] The predicted neoantigen, including prediction from SNVs and InDel, was obtained according to supplementary materials of a previously published study from TCGA research network.[51]

Statistical analysis

R software (version 4.0.2) and several packages mentioned above were used for statistical analyses. Survival analyses were performed using the K-M method, with the log-rank test to assess the significance. Wilcoxon rank-sum test, t test, and chi-square test were used for the comparison between two groups. All statistical tests were two sided, and $p < 0.05$ was considered statistically significant.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omto.2021.07.011.

ACKNOWLEDGMENTS

This study was funded by the National Key Research and Development Plan and the Ministry of Science and Technology of the People's Republic of China (grant numbers 2016YFC0105207 and 2017YFC1311004).

AUTHOR CONTRIBUTIONS

K.K., F.X., Y.W., and F.Z. designed the study. K.K., Y.W., C.H., and Y.B. performed the bioinformatics analyses. K.K., F.X., and Y.W. drafted the manuscript. K.K., J.L., X.L., and F.Z. critically revised the study. All authors have confirmed the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Ostrom, Q.T., Patil, N., Cioffi, G., Waite, K., Kruchko, C., and Barnholtz-Sloan, J.S. (2020). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013-2017. Neuro-oncol. *22* (12, Suppl 2), iv1–iv96.

2. Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., and Ellison, D.W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. *131*, 803–820.

3. Claus, E.B., Walsh, K.M., Wiencke, J.K., Molinaro, A.M., Wiemels, J.L., Schildkraut, J.M., Bondy, M.L., Berger, M., Jenkins, R., and Wrensch, M. (2015). Survival and low-grade glioma: the emergence of genetic information. Neurosurg. Focus *38*, E6.

4. Cancer Genome Atlas Research Network, Brat, D.J., Verhaak, R.G.W., Aldape, K.D., Yung, W.K.A., Salama, S.R., Cooper, L.A.D., Rheinbay, E., Miller, C.R., Vitucci, M., et al. (2015). Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N. Engl. J. Med. *372*, 2481–2498.

5. Schomas, D.A., Laack, N.N., Rao, R.D., Meyer, F.B., Shaw, E.G., O'Neill, B.P., Giannini, C., and Brown, P.D. (2009). Intracranial low-grade gliomas in adults: 30-year experience with long-term follow-up at Mayo Clinic. Neuro-oncol. *11*, 437–445.

6. Lee, J.K., Choi, Y.L., Kwon, M., and Park, P.J. (2016). Mechanisms and Consequences of Cancer Genome Instability: Lessons from Genome Sequencing Studies. Annu. Rev. Pathol. *11*, 283–312.

7. Gemoll, T., Auer, G., Ried, T., and Habermann, J.K. (2015). Genetic Instability and Disease Prognostication. Recent Results Cancer Res. *200*, 81–94.
8. Habermann, J.K., Doering, J., Hautaniemi, S., Roblick, U.J., Bündgen, N.K., Nicorici, D., Kronenwett, U., Rathnagiriswaran, S., Mettu, R.K., Ma, Y., et al. (2009). The gene expression signature of genomic instability in breast cancer is an independent predictor of clinical outcome. Int. J. Cancer *124*, 1552–1564.
9. Walther, A., Houlston, R., and Tomlinson, I. (2008). Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. Gut *57*, 941–950.
10. Åkeson, M., Jakobsen, A.-M., Zetterqvist, B.-M., Holmberg, E., Brännström, M., and Horvath, G. (2009). A population-based 5-year cohort study including all cases of epithelial ovarian cancer in western Sweden: 10-year survival and prognostic factors. Int. J. Gynecol. Cancer *19*, 116–123.
11. Andor, N., Maley, C.C., and Ji, H.P. (2017). Genomic Instability in Cancer: Teetering on the Limit of Tolerance. Cancer Res. *77*, 2179–2185.
12. Hegde, P.S., and Chen, D.S. (2020). Top 10 Challenges in Cancer Immunotherapy. Immunity *52*, 17–35.
13. Picard, E., Verschoor, C.P., Ma, G.W., and Pawelec, G. (2020). Relationships Between Immune Landscapes, Genetic Subtypes and Responses to Immunotherapy in Colorectal Cancer. Front. Immunol. *11*, 369.
14. Huarte, M. (2015). The emerging role of lncRNAs in cancer. Nat. Med. *21*, 1253–1261.
15. Peng, Z., Liu, C., and Wu, M. (2018). New insights into long noncoding RNAs and their roles in glioma. Mol. Cancer *17*, 61.
16. Lee, S., Kopp, F., Chang, T.C., Sataluri, A., Chen, B., Sivakumar, S., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. Cell *164*, 69–80.
17. Munschauer, M., Nguyen, C.T., Sirokman, K., Hartigan, C.R., Hogstrom, L., Engreitz, J.M., Ulirsch, J.C., Fulco, C.P., Subramanian, V., Chen, J., et al. (2018). The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. Nature *561*, 132–136.
18. Schmitt, A.M., Garcia, J.T., Hung, T., Flynn, R.A., Shen, Y., Qu, K., Payumo, A.Y., Peres-da-Silva, A., Broz, D.K., Baum, R., et al. (2016). An inducible long noncoding RNA amplifies DNA damage signaling. Nat. Genet. *48*, 1370–1376.
19. Zhou, L., Zhu, Y., Sun, D., and Zhang, Q. (2020). Emerging Roles of Long non-coding RNAs in The Tumor Microenvironment. Int. J. Biol. Sci. *16*, 2094–2103.
20. Chen, D., Lu, T., Tan, J., Li, H., Wang, Q., and Wei, L. (2019). Long Non-coding RNAs as Communicators and Mediators Between the Tumor Microenvironment and Cancer Cells. Front. Oncol. *9*, 739.
21. VanGenderen, C., Harkness, T.A.A., and Arnason, T.G. (2020). The role of Anaphase Promoting Complex activation, inhibition and substrates in cancer development and progression. Aging (Albany NY) *12*, 15818–15855.
22. Kapanidou, M., Curtis, N.L., and Bolanos-Garcia, V.M. (2017). Cdc20: At the Crossroads between Chromosome Segregation and Mitotic Exit. Trends Biochem. Sci. *42*, 193–205.
23. Zhang, Y., Li, J., Yi, K., Feng, J., Cong, Z., Wang, Z., Wei, Y., Wu, F., Cheng, W., Samo, A.A., et al. (2019). Elevated signature of a gene module coexpressed with CDC20 marks genomic instability in glioma. Proc. Natl. Acad. Sci. USA *116*, 6975–6984.
24. Eischen, C.M. (2016). Genome Stability Requires p53. Cold Spring Harb. Perspect. Med. *6*, a026096.
25. Chen, L., Wang, Y., He, J., Zhang, C., Chen, J., and Shi, D. (2018). Long Noncoding RNA H19 Promotes Proliferation and Invasion in Human Glioma Cells by Downregulating miR-152. Oncol. Res. *26*, 1419–1428.
26. Zhao, H., Peng, R., Liu, Q., Liu, D., Du, P., Yuan, J., Peng, G., and Liao, Y. (2016). The lncRNA H19 interacts with miR-140 to modulate glioma growth by targeting iASPP. Arch. Biochem. Biophys. *610*, 1–7.
27. Lecerf, C., Le Bourhis, X., and Adriaenssens, E. (2019). The long non-coding RNA H19: an active player with multiple facets to sustain the hallmarks of cancer. Cell. Mol. Life Sci. *76*, 4673–4687.
28. Walker, J.A., and McKenzie, A.N.J. (2018). T_H2 cell development and function. Nat. Rev. Immunol. *18*, 121–133.
29. Shimato, S., Maier, L.M., Maier, R., Bruce, J.N., Anderson, R.C., and Anderson, D.E. (2012). Profound tumor-specific Th2 bias in patients with malignant glioma. BMC Cancer *12*, 561.
30. Helmink, B.A., Reddy, S.M., Gao, J., Zhang, S., Basar, R., Thakur, R., Yizhak, K., Sade-Feldman, M., Blando, J., Han, G., et al. (2020). B cells and tertiary lymphoid structures promote immunotherapy response. Nature *577*, 549–555.
31. Dey, M., Chang, A.L., Miska, J., Wainwright, D.A., Ahmed, A.U., Balyasnikova, I.V., Pytel, P., Han, Y., Tobias, A., Zhang, L., et al. (2015). Dendritic Cell-Based Vaccines that Utilize Myeloid Rather than Plasmacytoid Cells Offer a Superior Survival Advantage in Malignant Glioma. J. Immunol. *195*, 367–376.
32. Wainwright, D.A., Balyasnikova, I.V., Chang, A.L., Ahmed, A.U., Moon, K.S., Auffinger, B., Tobias, A.L., Han, Y., and Lesniak, M.S. (2012). IDO expression in brain tumors increases the recruitment of regulatory T cells and negatively impacts survival. Clin. Cancer Res. *18*, 6110–6121.
33. Buchbinder, E.I., and Desai, A. (2016). CTLA-4 and PD-1 Pathways: Similarities, Differences, and Implications of Their Inhibition. Am. J. Clin. Oncol. *39*, 98–106.
34. Patel, S.P., and Kurzrock, R. (2015). PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. Mol. Cancer Ther. *14*, 847–856.
35. Mathewson, N.D., Ashenberg, O., Tirosh, I., Gritsch, S., Perez, E.M., Marx, S., Jerby-Arnon, L., Chanoch-Myers, R., Hara, T., Richman, A.R., et al. (2021). Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. Cell *184*, 1281–1298.e26.
36. Sharma, P., and Debinski, W. (2018). Receptor-Targeted Glial Brain Tumor Therapies. Int. J. Mol. Sci. *19*, 3326.
37. Yi, Z., Prinzing, B.L., Cao, F., Gottschalk, S., and Krenciute, G. (2018). Optimizing EphA2-CAR T Cells for the Adoptive Immunotherapy of Glioma. Mol. Ther. Methods Clin. Dev. *9*, 70–80.
38. Bagley, S.J., Desai, A.S., Linette, G.P., June, C.H., and O'Rourke, D.M. (2018). CAR T-cell therapy for glioblastoma: recent clinical advances and future challenges. Neuro-oncol. *20*, 1429–1438.
39. Wang, Y., Qian, T., You, G., Peng, X., Chen, C., You, Y., Yao, K., Wu, C., Ma, J., Sha, Z., et al. (2015). Localizing seizure-susceptible brain regions associated with low-grade gliomas using voxel-based lesion-symptom mapping. Neuro-oncol. *17*, 282–288.
40. Zhao, Z., Meng, F., Wang, W., Wang, Z., Zhang, C., and Jiang, T. (2017). Comprehensive RNA-seq transcriptomic profiling in the malignant progression of gliomas. Sci. Data *4*, 170024.
41. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics *28*, 882–883.
42. Gravendeel, L.A., Kouwenhoven, M.C., Gevaert, O., de Rooi, J.J., Stubbs, A.P., Duijm, J.E., Daemen, A., Bleeker, F.E., Bralten, L.B., Kloosterhof, N.K., et al. (2009). Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. Cancer Res. *69*, 9065–9072.
43. Ulloa-Montoya, F., Louahed, J., Dizier, B., Gruselle, O., Spiessens, B., Lehmann, F.F., Suciu, S., Kruit, W.H., Eggermont, A.M., Vansteenkiste, J., and Brichard, V.G. (2013). Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. J. Clin. Oncol. *31*, 2388–2395.
44. Auslander, N., Zhang, G., Lee, J.S., Frederick, D.T., Miao, B., Moll, T., Tian, T., Wei, Z., Madan, S., Sullivan, R.J., et al. (2018). Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. Nat. Med. *24*, 1545–1549.
45. Mariathasan, S., Turley, S.J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., Kadel, E.E., III, Koeppen, H., Astarita, J.L., Cubas, R., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature *554*, 544–548.
46. Miao, D., Margolis, C.A., Gao, W., Voss, M.H., Li, W., Martini, D.J., Norton, C., Bossé, D., Wankowicz, S.M., Cullen, D., et al. (2018). Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. Science *359*, 801–806.
47. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 7.

48. Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. *18*, 248–262.

49. Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J.S., Li, B., and Liu, X.S. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res. *77*, e108–e110.

50. Miao, Y.R., Zhang, Q., Lei, Q., Luo, M., Xie, G.Y., Wang, H., and Guo, A.Y. (2020). ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. Adv. Sci. (Weinh.) *7*, 1902880.

51. Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al.; Cancer Genome Atlas Research Network (2018). The Immune Landscape of Cancer. Immunity *48*, 812–830.e14.