

Identification of natural killer cell-related characteristics to predict the clinical prognosis and immune microenvironment of patients with low-grade glioma

Fei Sun^{1,2,*}, Hongtao Lv^{1,*}, Baozhi Feng¹, Jiaao Sun³, Linyun Zhang¹, Bin Dong¹

¹Department of Neurosurgery, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China

²Department of Neurosurgery, Xinhua Hospital Affiliated to Dalian University, Dalian, Liaoning, China

³Department of Urology, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China

*Equal contribution

Correspondence to: Bin Dong, Linyun Zhang; email: stocktondb@163.com, <https://orcid.org/0000-0002-0349-446X>; zly542644190@163.com, <https://orcid.org/0000-0002-6665-0634>

Keywords: low-grade glioma (LGG), natural killer (NK) cells, tumor immune microenvironment, RiskScore model, bioinformatics

Received: March 8, 2023

Accepted: June 15, 2023

Published: July 5, 2023

Copyright: © 2023 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Individuals with low-grade glioma (LGG) have a dismal prognosis, and most patients will eventually progress to high-grade disease. Therefore, it is crucial to accurately determine their prognoses.

Methods: Seventy-nine NK cell genes were downloaded from the LM22 database and univariate Cox regression analysis was utilized to detect NK cell-related genes affecting prognosis. Molecular types were established for LGG using the “ConsensusClusterPlus” R package. The results from a functional enrichment analysis and the immune microenvironment were intensively explored to determine molecular heterogeneity and immune characteristics across distinct subtypes. Furthermore, a RiskScore model was developed and verified using expression profiles of NK cells, and a nomogram consisting of the RiskScore model and clinical traits was constructed. Moreover, pan-cancer traits of NK cells were also investigated.

Results: The C1 subtype included the greatest amount of immune infiltration and the poorest prognosis among well-established subtypes. The majority of enriched pathways were those involved in tumor progression, including epithelial-mesenchymal transition and cell cycle pathways. Differentially expressed genes among distinct subtypes were determined and used to develop a novel RiskScore model. This model was able to distinguish low-risk patients with LGG from those with high-risk disease. An accurate nomogram including the RiskScore, disease grade and patient’s age was constructed to predict clinical outcomes of LGG patients. Finally, a pan-cancer analysis further highlighted the crucial roles of NK cell-related genes in the tumor microenvironment.

Conclusions: An NK cell-related RiskScore model can accurately predict the prognoses of patients with LGG and provide valuable insights into personalized medicine.

INTRODUCTION

Brain tumors are the tenth most fatal human malignancy, accounting for around 3 percent of all cancer deaths, according to recent data from cancer centers [1]. Glioma is the most common malignant

intracranial tumor [2, 3]. The World Health Organization (WHO) has classified gliomas by grades based on their pathological characteristics, with grades I and II representing low-grade glioma (LGG) and grades III and IV representing high-grade glioma (HGG). LGG accounts for approximately 15% of all gliomas. Most

LGGs eventually progress to HGGs, with poor prognosis [4]. At present, LGG tumors are often surgically resected, accompanied by radiotherapy and chemotherapy [5, 6]. Due to the invasive growth characteristics of gliomas, surgical resection or radiotherapy cannot avoid tumor recurrence [7]. Therefore, it is necessary to determine an efficient treatment plan. Currently, when treating LGG, it is best to surgically remove as much of the tumor as possible, but the treatment plan for individuals who have unresectable disease is highly controversial. Therefore, an in-depth understanding of the pathogenesis of LGG is essential to manage and treat patients with LGG to improve their prognoses.

In order for cancer cells to form and spread, they need a specific environment known as the tumor microenvironment (TME) [8, 9]. The infiltration of immune cells into the TME has been shown to influence cancer progression and has strong prognostic value for LGG [10]. The onset and progression of LGG are inextricably linked to the dysfunction of the body's immune system, particularly natural killer (NK) cells, which perform non-specific immune responses [11]. These cells preferentially eliminate major histocompatibility complex class I downregulated neoplastic tumors [12]. During homeostasis, NK cells reside in the brain parenchyma and circulate through the blood-brain barrier (BBB) [13]. The presence of NK cells in brain tumors and the surrounding brain parenchyma microenvironment has been previously identified [14, 15]. It is unclear, however, which factors facilitate NK cell passage across the BBB and subsequent activation in the brain. Further evidence suggests that *NKp46*, *NKp44*, and *NKp3* can synergistically inhibit or activate the antitumor function of NK cells [16]. In addition, NK cells recognize a variety of cell- and tumor-associated surface ligands and may control tumor cell function in a positive or negative manner [17]. *In vitro*, glioma cell lines can be killed by NK cells [18]. LGG is a highly vascularized tumor that can alter the immune system and impair immune system function [19]. The relationship between NK cells and LGG is still not completely understood. Hence, for the development of new immunotherapies, understanding immune surveillance mechanisms in LGG patients with greater survival rates is critical.

Through using the Chinese Glioma Genome Atlas (CGGA) and The Cancer Genome Atlas (TCGA) databases [20], this study carried out molecular typing of LGG based on LM22 data and transcriptomic data using NK cell-related genes [21]. In addition, a risk model was constructed, and immune cells in the LGG tumor microenvironment (TME) were also assessed. Discrepancies in drug sensitivity among different

clusters were also predicted. Immune profiles and therapeutic sensitivities significantly differed among LGG molecular subtypes. Finally, this study summarized a pan-cancer overview of NK cell-related genes, thereby elucidating the relationship between LGG and NK cells, and informing the personalized treatment of LGG.

METHODS

Data acquisition and processing

Similar to the method used in previous studies [22, 23], clinical samples and mRNA transcriptome data of LGG were accessed using the CGGA (<http://www.cgga.org.cn/>) and TCGA (<https://portal.gdc.cancer.gov/>) databases. Exclusion criteria were as follows: samples with incomplete clinical follow-up data including survival time and status. Two datasets, “mRNAseq_325 (batch 2)” and “mRNAseq_693 (batch 1)”, from the CGGA were merged using the “RTCGAToolbox” package and named as the “CGGA dataset”. Subsequently, batch effects were eliminated using the “sva” package. Clinical information of the TCGA and CGGA datasets is shown in Supplementary Table 1. Gene mutation and DNA methylation data of LGG patients were downloaded from the TCGA database [24, 25]. Transcriptome data for NK cell-related genes, containing 79 NK cell genes, were downloaded from the LM22 database (Supplementary Table 2) [21]. The LM22 database contains 547 genes that distinguish 22 human hematopoietic cell phenotypes, including T cells, naïve and memory B cells, plasma cells, NK cells, and myeloid subsets.

Molecular typing

Univariate Cox regression analysis was implemented to determine NK cell genes associated with LGG prognoses in the TCGA and CGGA datasets. The LGG molecular subtypes were obtained using the “ConsensusClusterPlus” R package based on prognosis-related NK cell genes [26]. In total, 500 bootstraps were conducted using the “K-M” algorithm and “canberra” as the metric distance, with 80% and 20% of patients in the respective training and validation sets involved in each bootstrapping process. Clusters ranged from 10 to 2, and molecular subtypes were determined using the optimum classification system established by calculating the consistency cumulative distribution function (CDF) and the consistency matrix [27].

In 2020, Zheng et al. stated that patients with LGG were categorized into six molecular subtypes, including Codel, G-CIMP-high, G-CIMP-low, Classic-like,

Mesenchymal-like and PA-like [28]. To comprehensively uncover the molecular traits of our NK cell-based typing, we also investigated the close association of NK cell subtypes with these six molecular subtypes.

Gene set enrichment analysis (GSEA)

The optimal molecular subtyping was determined, and subsequent differential gene analysis was performed on the different LGG molecular types using the “limma” package with thresholds set at $P < 0.05$ and $|\log \text{fold change (FC)}| > 1$ [29]. Candidate gene sets in The Molecular Signatures Database (MSigDB) (<https://ngdc.cncb.ac.cn/databasecommons/database/id/1077>) [30] were subjected to “GSEA” enrichment analysis to explore possible pathways linked to DEGs to screen significant pathways using a FDR < 0.05 [30]. In addition, gene sets related to angiogenesis and inflammatory characteristics were collected from previously published studies [31, 32].

Tumor immune microenvironment and clinical characteristics in different clusters

The following immune prediction algorithms were utilized to explore discrepancies in the tumor immune microenvironment among distinct clusters: “CIBERSOFT” and “ESTIMATE” algorithms. The “CIBERSOFT” algorithm assessed the infiltration abundance of various immunocytes in the TCGA and CGGA datasets, while the “ESTIMATE” algorithm calculated the “ImmuneScore”. The “ImmuneScore” reflects the cumulative content of immune cells in the TME. Importantly, immune-associated signaling pathways were also explored, including interferon and MHC.

The TCGA and CGGA datasets provided a series of clinical information, including age, grade, and IDH mutation. These indicators are key factors in clinical decision-making and therapeutic strategies for LGG patients. Moreover, discrepancies in clinical information among these four clusters were also investigated.

Immunotherapy/chemotherapy prediction in different clusters

A recent study by Palmeri et al. stated that the tumor mutational burden (TMB) could serve as an accurate biomarker for immunotherapy outcome prediction [33]. Davoli and colleagues further reported that tumor aneuploidy could also be used as an immunotherapy biomarker [34]. Thus, we compared TMB levels and aneuploidy scores among different clusters. Importantly, the expression of various typical immune

checkpoint genes among different clusters was also explored. The IC50 of chemotherapy drugs in the GDSC (Genomics of Drug Sensitivity in Cancer) database, which contained a large panel of cancer cell lines, was calculated using the “pRRophetic” package [35]. Finally, IC50 values of each drug among different clusters were also compared.

Risk model construction

The “Limma” package was implemented to determine differentially expressed genes (DEGs) among different clusters. Subsequently, these DEGs were selected to develop a risk model using least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression analysis. Specifically, the number of model genes was reduced using a LASSO regression to further screen the valid genes. LASSO regression solved the multicollinearity during the regression study by compressing coefficients and setting some of them to zero. Alongside the gradual increase in lambda, selection of the optimum number of factors was carried out when the coefficients of independent variables tended to zero. In addition, the best model was developed by performing stepwise regression applying the Akaike information criterion (AIC) in the “MASS package”. This process of regression takes into consideration the best statistical fit of the model and parameter number. Model development was conducted by initially establishing the most complex model, followed by a successive reduction in the number of variables involved to decrease the AIC values. The lower the AIC values, the better the model, as these low values indicated the least number of parameters used to achieve a sufficient model fit. After determining valid genes, a risk model was constructed by the following equation: $\text{RiskScore} = \text{Exp}_i \times \sum \beta_i$, where β is the multivariate Cox regression coefficient of a corresponding gene, and Exp_i refers to the expression of NK cell-related gene characteristics [36]. Subsequently, a z-score was obtained, and patients were classified into low-RiskScore and high-RiskScore groups with the criteria of z-score = 0. Furthermore, the Kaplan-Meier method was used to plot survival curves for prognostic analysis. Moreover, the log-rank test was performed to examine whether differences were considered significant.

Immune microenvironment, molecular characteristics, and drug prediction in the risk model

In LGG, the CIBERSORT (<https://cibersort.stanford.edu/>) algorithm calculated the relative abundance of 22 immune cells to analyze the differences in immune cell composition and function between high- and low-RiskScore groups. Additionally, to measure immune cell amounts, the “ESTIMATE” R

package was applied to calculate the ImmuneScore, the StromalScore and the ESTIMATEScore [37]. Based on previous studies, scores of 10 oncogenic pathways (cell cycle, HIPPO, MYC, NOTCH, NRF1, PI3K, TGF-Beta, RAS, TP53, WNT), 7 metagenes pathways (HCK, IgG, Interferon, LCK, MHC_I, MHC_II, STAT1), T cell inflamed GEP, and cytolytic activity were calculated by “ssGSEA” using the “GSVA” package [38]. Discrepancies in the above pathway activities between different risk groups were comprehensively investigated. Based on the results of the “pRRophetic” package, chemotherapy drug sensitivities between high- and low-RiskScore groups were also further explored.

Nomogram

Univariate and multivariate Cox regression analysis using clinicopathological variables were performed to determine if the RiskScore may be used as an independent prediction measure [39]. The ‘rms’ software was used to construct a nomogram containing all independent prognostic factors to estimate 1-, 3-, and 5-year overall survival (OS) probability. The nomogram’s discriminative power was measured using the concordance index (C-index) and calibration.

Pan-cancer characterization of NK cell-related genes

Based on the pan-cancer cohort samples of the TCGA, we gathered and compiled genomic data, transcriptome data, and clinical follow-up information for dozens of human malignancies. It is currently recognized that information on promoter methylation levels and gene mutations may have a significant effect on gene expression. Hence, we first investigated the pan-tumor mutation and methylation data of NK cell-related genes. Of note, we interpreted the mutation spectrum primarily from the viewpoints of CNV and SNV. Subsequently, we conducted a complete analysis of the differential expression of each NK cell-related gene in distinct malignant and precancerous tissues and evaluated the prognostic relevance of each gene across tumor types. Importantly, based on the Msigdb platform, we collected the classical metabolic pathways and immune pathways and, using the GSEA concept, and evaluated the potential correlation between the NK cell gene set and these pathways, thereby laying the groundwork for subsequent mechanistic studies.

Statistical analysis

Data analysis was conducted using the R 4.0.2 statistical tool (<https://www.r-project.org/>). The “limma” package was used for group differential gene analysis, the “MASS package” for stepAIC analysis, the “ssGSEA” for enrichment pathway analysis, and the “survminer”

package for survival analysis. Furthermore, the examination and comparison of survival differences between the aforementioned groups were carried out using the Kaplan-Meier (K-M) method and the log-rank test. The “rpart” and “rpart.plot” packages were utilized to develop a survival decision tree. Statistical tests were two-sided unless otherwise specified, and a significant difference was indicated by a P value < 0.05 .

Availability of data and materials

The datasets analyzed in this work may be found in the Supplementary Materials or contact with the corresponding author.

RESULTS

Molecular typing based on NK cell-related genes

LGG-related prognostic genes in the TCGA and CGGA datasets were analyzed separately using univariate Cox regression. Genes associated with NK cells were examined for correlation with OS in individuals with LGG, which revealed a link between 42 genes in the CGGA dataset ($P < 0.05$) and 51 out of 79 genes in the TCGA dataset ($P < 0.05$). Significant prognostic genes are displayed in Supplementary Table 3. Subsequently, prognostic genes in both datasets were intersected to obtain 34 prognosis-related significant genes (Figure 1A, 1B). Subsequently, the optimum number of clusters in the TCGA dataset was identified using consensus clustering for the aforementioned prognostic-related genes. Cluster number was defined according to CDF, which was reflected by the CDF delta area curves, and was determined as a stable cluster when the number of clusters was 4 (Figure 1C, 1D). Therefore, clustering was optimal when $k = 4$ (Figure 1E). Clusters of each LGG in the TCGA and CGGA datasets are displayed in Supplementary Table 4. Survival and K-M curve analysis suggested a better prognosis for cluster C4, whereas cluster C1 had the worst prognosis (Figure 1F). Similarly, LGG subtypes in the CGGA dataset were classified into four categories ($k = 4$), and the survival curves for each subtype were examined (Figure 1G). Variations between the expression of 34 NK cell-linked genes among various subtypes were compared. In the two independent datasets, risk genes exhibited increased expression in the C1 cluster, whereas protective genes depicted increased expression in the C4 cluster (Figure 1H, 1I).

Genomic landscape and pathway characteristics among molecular subtypes

In the TCGA-LGG cohort, molecular characteristics were compared among the four subtypes to investigate

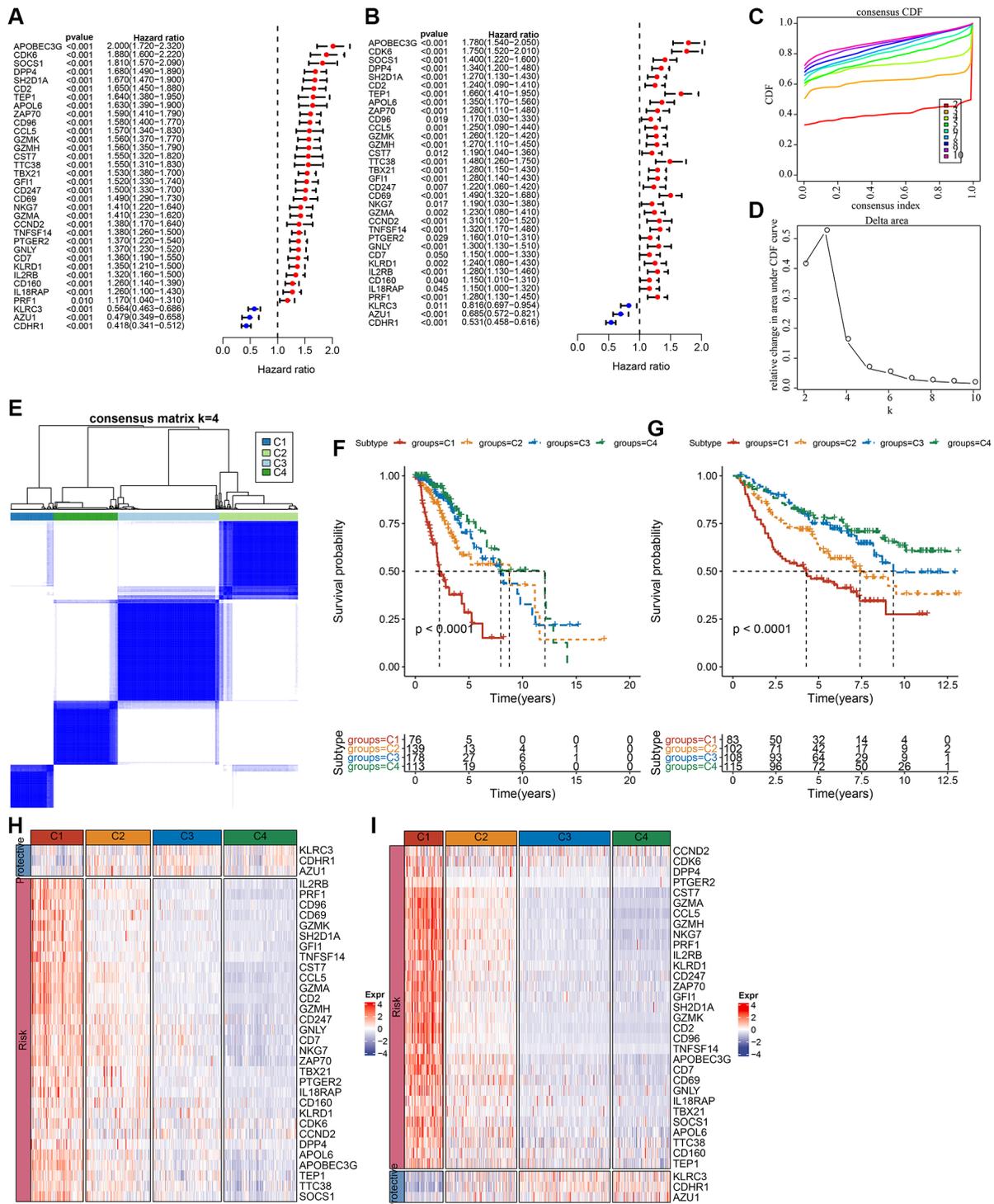


Figure 1. Molecular subtyping of LGG based on NK cell-related genes. (A) Univariate Cox forest plot of NK cell-related genes in the TCGA-LGG cohort. **(B)** Multivariate Cox forest plot of NK cell-related genes in the CGGA-LGG cohort. **(C)** CDF curve of samples in the TCGA-LGG cohort. **(D)** CDF delta area curve of TCGA-LGG cohort samples area curve, Delta area curve of consensus clustering indicated the relative variation in the area under the CDF curve for each category number k in comparison with k-1. The vertical axis represents the relative change in area under the CDF curve, the horizontal axis represents the category number k. **(E)** A heat map of sample clustering at consensus k = 4. **(F)** K-M survival curves showing the prognosis of the four subtypes in the TCGA-LGG cohort. **(G)** K-M survival curves for the prognosis of the four subtypes in the CGGA cohort. **(H)** A heat map showing the expression of prognostically significant NK cell-related genes in various subtypes in TCGA-LGG. **(I)** A heat map of the expression of prognostically significant NK cell-associated genes in various subtypes in CGGA. Abbreviations: CDF: cumulative distribution function; expr: expression; LGG: low-grade glioma; NK: natural killer; TCGA: The Cancer Genome Atlas; CGGA: Chinese Glioma Genome Atlas; K-M: Kaplan-Meier.

the differences in genomic alterations among them. Analysis revealed that the C1 cluster had a high TMB score and a high aneuploidy score (Supplementary Figure 1A). In addition, six previously established molecular subtypes [28] were compared with the currently defined four molecular subtypes. The results highlighted more “G-CIMP-high” and “Codel” molecular subtypes in clusters C3 and C4 (Supplementary Figure 1B). Subsequently, significant mutation differences among subtypes were determined. After visualizing the top 20 mutated genes, mutation frequencies of gene, such as isocitrate dehydrogenase (*IDH1*) and capicua (*CIC*), varied considerably among the four molecular subtypes (Supplementary Figure 1C).

Additionally, among the four subtypes, a distribution comparison of clinical characteristics was carried out in the TCGA dataset (Supplementary Figure 2A), and no significant difference was found between genders. A larger proportion of patients with the C1 subtype had a G3 grade and more patients in subtypes C3 and C4 had a G2 grade. As for *IDH* mutations, the highest proportion of patients that had mutations were in subtypes C4, C3, and C2. In addition, the majority of patients in subtypes C4 and C3 had *IDH* mutations and *IDH*mut-non-codel.

O-6-methylguanine-DNA methyltransferase (MGMT) promoter methylation level was also considerably increased in C3 and C4 subtypes rather than in C1 (Supplementary Figure 2A). Age, MGMT promoter methylation, *IDH* mutation, disease grade, and 1p19q co-deletion in the CGGA dataset were also examined (Supplementary Figure 2B). Gender did not show a significant difference among subtypes, while the 1p19q co-deletion level and *IDH* mutation were considerably increased in C3 and C4 subtypes when compared to C1, which was similar to the observations reported in the above analysis of the TCGA dataset.

Furthermore, differentially activated pathways among subtypes were categorized by the “GSEA” enrichment analysis. Findings based on the TCGA cohort suggested considerable enrichment of cancer-associated pathways in subtypes C1 and C2, such as glycolysis, PI3K-AKT-mTOR, angiogenesis, hypoxia, P53, and apoptosis (Figure 2A). Moreover, the cell cycle, HIPPO, and TP53 pathway activities were also noticeably upregulated in the C1 subtype (Figure 2B). In short, we found that several signaling pathways were closely related to tumorigenesis and development which showed a significantly up-regulated trend in the C1 and C2 subtypes, and a significant downward trend in the C3 and C4 subtypes. This phenomenon may be potentially contributing to the observed difference in prognosis.

Differences in immunological characteristics and treatment efficacy among molecular subtypes

Variations in the immune microenvironment among subtypes was analyzed by determining the relative frequency of the 22 immune cells using “CIBERSORT”. Most immune cells differed significantly among subtypes. For example, the C1 subtype had a higher infiltration of CD8⁺T cells and macrophages in the TCGA dataset. Additionally, the ESTIMATE algorithm revealed that the C1 subtype had the highest “ImmuneScore” and “StromalScore”, followed by the C2, C3, and C4 subtypes (Figure 3A, 3B). Importantly, similar findings were also observed in the CGGA dataset (Figure 3C, 3D). In addition, we analyzed the inflammatory activity of four molecular subtypes, and the enrichment scores of all the metagene clusters (i.e., HCK, IgG, Interferon, LCK, MHC-I, MHC-II, STAT1) were significantly different among the four molecular subtypes. Overall, the C1 subtype had higher inflammatory activity, as shown in Figure 3E; similarly, this phenomenon was also observed in the CGGA cohorts, as per Figure 3F.

Since immune cell checkpoints are key targets for cancer treatment using immune checkpoint blockade (ICB), several checkpoint molecules, including cytotoxic T-lymphocyte-associated protein 4 (*CTLA-4*), programmed death 1 (*PD-1*), and programmed cell death-ligand 1 (*PD-L1*) were evaluated. These molecules showed significantly high expression in subtype C1 (Figure 4A). Given that IFN- γ plays a key role in immune modulation and anticancer immunity, we carried out ssGSEA analysis of the GOBP_RESPONSE_TO_INTERFERON_GAMMA gene set loaded from the GO database and found that the IFN- γ response is significantly enhanced in the C1 subtype (Figure 4B). Simultaneously, we also compared differences in the expression of the IFNG gene among the four subtypes and found that IFNG was significantly higher in the C1 subtype (Figure 4C). In addition, the cytotoxic (CYT) score, which characterized cytotoxic activity, was considerably increased in the C1 subtype (Figure 4D). In addition, we used the “T-cell-inflamed score” to evaluate the predictive potential of different molecular subtypes in cancer immunotherapy. As depicted in Figure 4E, the C1 subtype showed a higher T-cell-inflamed score. Taken together, these findings suggested that LGG patients with subtype C1 might benefit from immunotherapy.

Moreover, we analyzed the response of different molecular subtypes in the TCGA-LGG cohort to traditional chemotherapeutic drugs. C1 subtype patients were more sensitive to bleomycin, as evidenced by the biochemical half-maximal inhibitory concentration (IC₅₀) and their response (Figure 4F).

Establishment and validation of a clinical prognostic model

We identified a total of 739 DEGs among C1 vs. other, C2 vs. other, C3 vs. other, and C4 vs. other using the limma analysis (Supplementary Figure 3A–3C). There were no significant differentially expressed genes between C2 and

other subtypes. Subsequently, the prognostic performances of the above 739 DEGs were further explored, with the results showing that a total of 719 genes (including 466 “risk” and 253 “protective” genes) were significantly associated with LGG prognoses (Supplementary Figure 4A, Supplementary Table 5, $P < 0.001$). LASSO and multivariate Cox regression algorithms

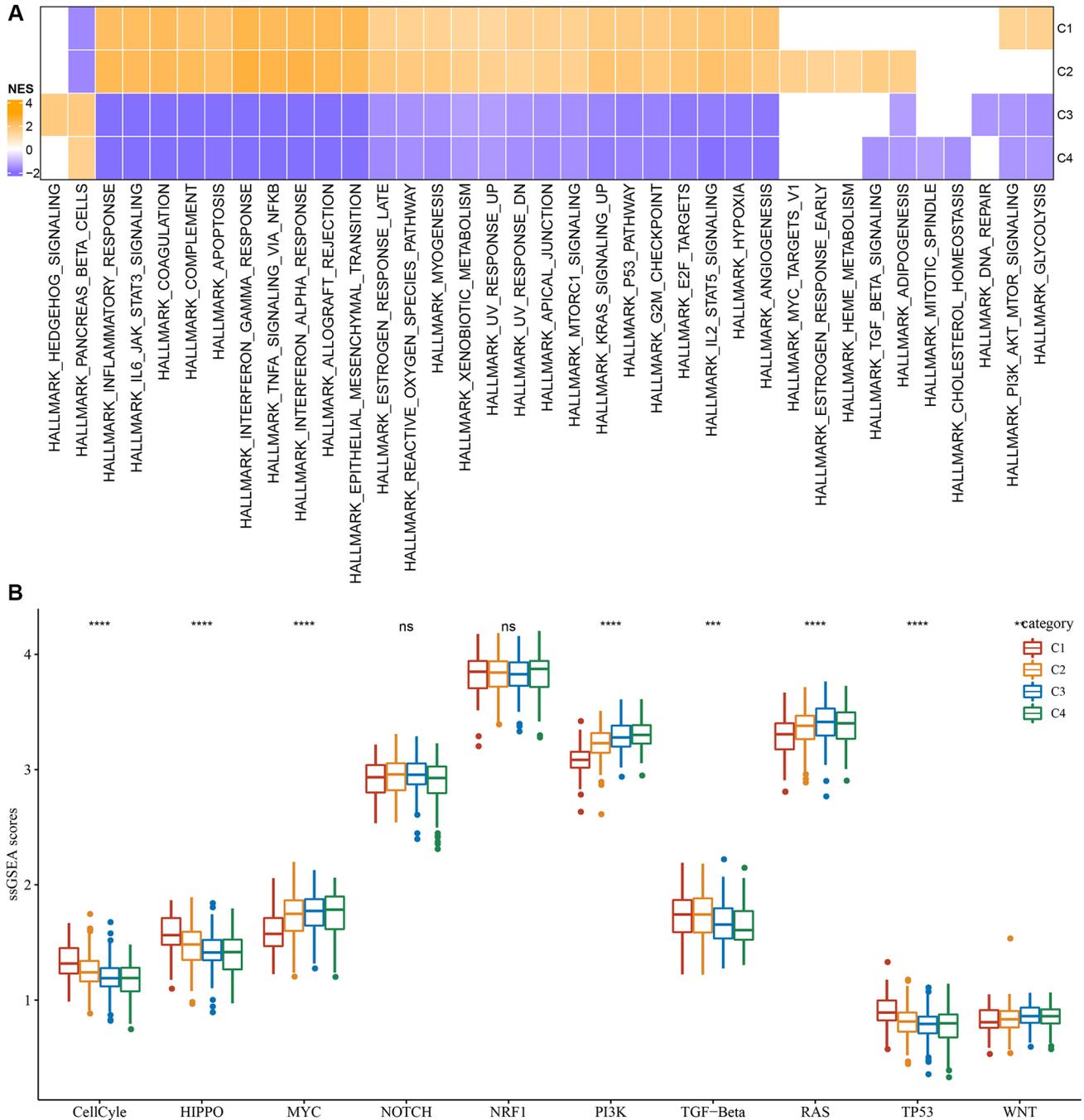


Figure 2. Significantly activated pathways in different molecular subtypes. (A) GSEA analysis results in the TCGA-LGG cohort. (B) Variation in the scores of 10 tumor abnormality-related pathways in various TCGA-LGG molecular subtypes in the $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$. Abbreviations: ns: no significance; NES: normalized enrichment scores; ssGSEA: single-sample GSEA; GSEA: gene set enrichment analysis; TCGA: The Cancer Genome Atlas; LGG: low-grade glioma; TP53: tumor protein p53; PI3K: phosphatidylinositol 3-kinase; NRF1: nuclear respiratory factor-1; TGF- β : transforming growth factor- β .

were utilized to construct a NK cell-based RiskScore. LASSO regression highlighted the crucial roles of 14 genes in the prognostic model (Supplementary Figure 4B, 4C). Finally, multivariate Cox regression analysis identified 6 genes to calculate the RiskScore (Supplementary Figure 4D). Using the z-score value = 0

as the dividing line, LGG patients were divided into high-risk group and low-risk groups, and the expression and distribution of model genes in each LGG patient were characterized (Figure 5A). For the TCGA and CGGA cohorts, KM survival analysis and ROC curves validated the accurate prediction performances of the

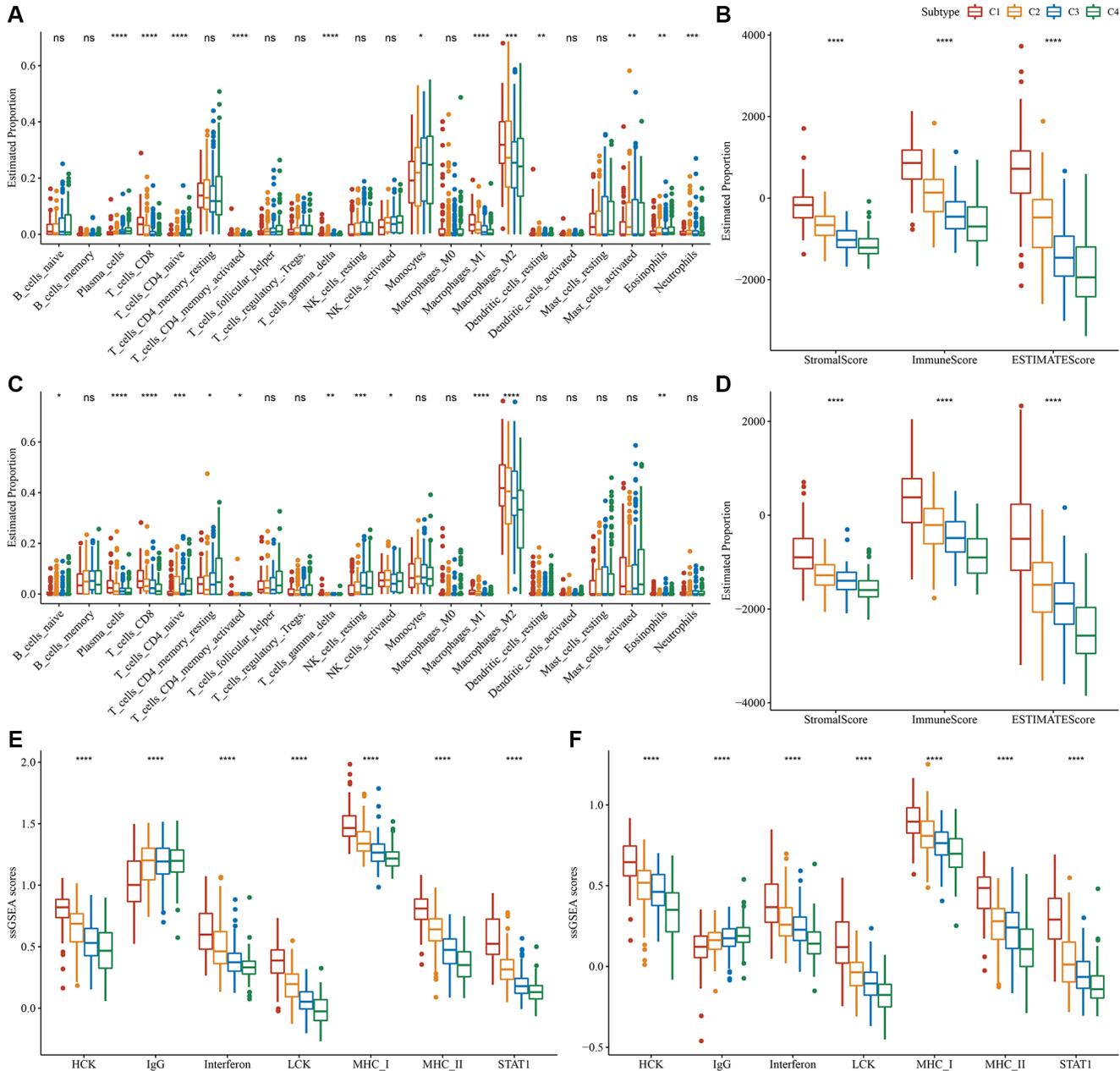


Figure 3. Degree of immune cell infiltration in different molecular subtypes. (A) Variation in 22 immune cell scores among various TCGA-LGG molecular subtypes. (B) Variation in ESTIMATE immune infiltration among various TCGA-LGG molecular subtypes. (C) Variation in 22 immune cell scores among different CGGA cohort molecular subtypes. (D) Variation in ESTIMATE immune infiltration among CGGA cohort molecular subtypes. (E) Variation in the gene cluster scores for the seven inflammation-related genesets among molecular subtypes in the TCGA-LGG cohort. (F) Differences in the gene cluster scores for the seven inflammation-related genesets among molecular subtypes in the CGGA cohort. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$. Abbreviations: ns: no significance; ssGSEA: single-sample GSEA; TCGA: The Cancer Genome Atlas; GSEA: gene set enrichment analysis; CGGA: Chinese Glioma Genome Atlas; LGG: low-grade glioma; HCK: hematopoietic cell kinase; IgG: Immunoglobulin G; LCK: lymphocyte-specific protein tyrosine kinase; MCH: melanin-concentrating hormone; STAT1: signal transducer and activator of transcription 1.

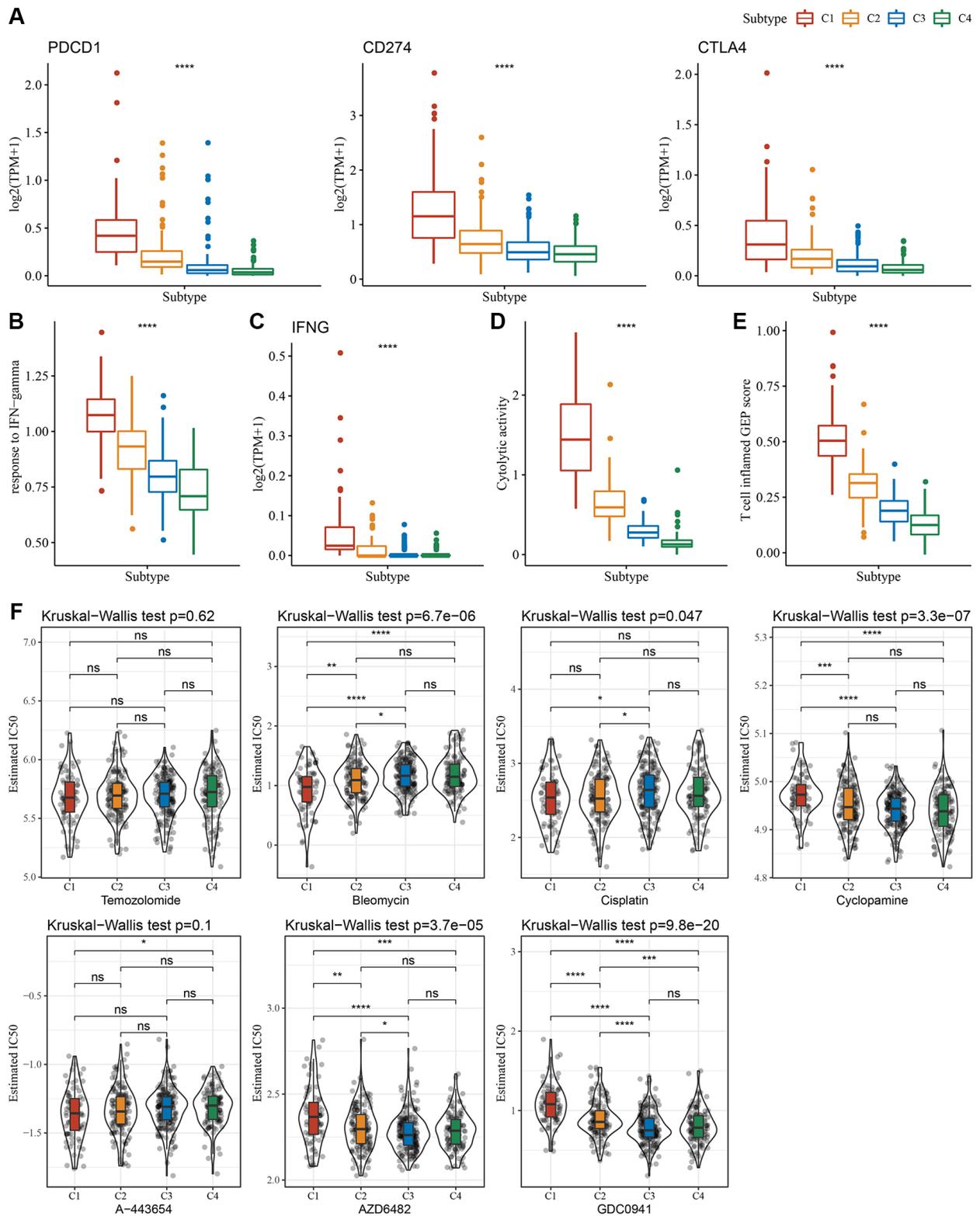


Figure 4. Immunological characteristics scores characterizing the effect of immunotherapy in different subtypes. (A) Expression differences of immune checkpoint-related genes among molecular subtypes. **(B)** Variation in response to IFN- γ among molecular subtypes. **(C)** The difference in expression of IFNG gene among molecular subtypes. **(D)** Differences in “Cytolytic activity” among molecular subtypes. **(E)** Variation in the T cell inflamed GEP score among molecular subtypes. **(F)** A box plot of the estimated IC₅₀ values for temozolomide, bleomycin, cisplatin, cyclopamine, A-443654, AZD6482, and GDC0941 in TCGA-LGG. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$. Abbreviations: ns: no significance; TPM: transcripts per million; IFN: interferon; GEP: gene expression profile; IC₅₀: half-maximal inhibitory concentration; TCGA: The Cancer Genome Atlas; LGG: low-grade glioma.

RiskScore (Figure 5B–5E). Specifically, LGG patients with high RiskScore were associated with worse clinical outcomes.

Association of the RiskScore with different clinicopathological features and molecular subtypes

To examine the relationship between the RiskScore and clinical features of LGG, we analyzed the differences in

RiskScores across age, gender, grade, IDH mutation, IDH/code subtype, and MGMT promoter methylation in the TCGA-LGG dataset. Results showed that higher RiskScores were associated with increased age and disease grade (Figure 6A). Furthermore, we compared RiskScores among different molecular subtypes (Figure 6B) and found that the C1 subtype had the highest RiskScores. In addition, we also compared whether there are prognostic differences in the high- and low-risk

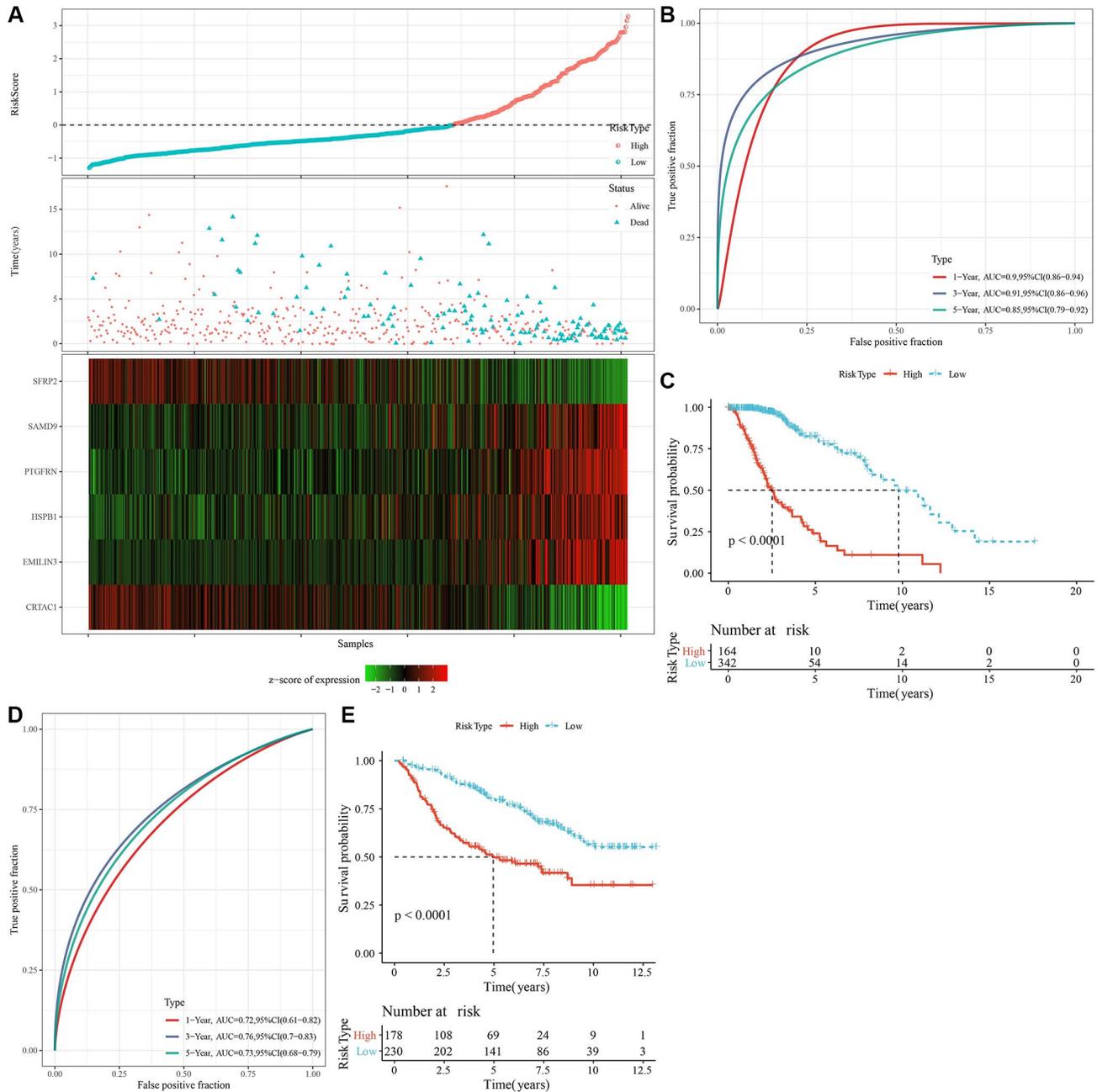


Figure 5. Calculation of RiskScore and determination of its robustness by ROC. (A) The RiskScore, survival status, survival time, and expression of oxidative stress-related prognostic genes in the TCGA dataset. (B) The ROC curve and AUC of RiskScore classification in the TCGA dataset. (C) The K-M survival curve distribution of RiskScore in the TCGA dataset. (D, E) The ROC curve and K-M survival curve of RiskScore in the CGGA cohort. Abbreviations: AUC: area under the ROC curve; CI: confidence interval; ROC: receiver operating characteristic; TCGA: The Cancer Genome Atlas; K-M: Kaplan-Meier; CGGA: Chinese Glioma Genome Atlas.

groups of RiskScores defined here among different clinicopathological characteristic groups in the TCGA-LGG cohort. Results showed that our risk groups also performed well in different clinical groups, validating the reliability of our defined risk groups (Figure 6C). The findings in the CGGA dataset showed similar results (Supplementary Figure 5).

Tumor immune microenvironment and molecular characteristics in both RiskScore groups

By employing the “CIBERSORT” algorithm, variation in the immune microenvironment of patients in both RiskScore groups was investigated by calculating the relative abundance of 22 immune cells. Of note, infiltration abundance of eight immune cells, including CD8⁺T cells, was found to be significantly different between the two groups (Figure 7A). Results of the

correlation analysis demonstrated that RiskScores were closely related to immunocyte infiltration (Figure 7B). In addition, we also used ESTIMATE to evaluate immune cell infiltration, as shown in Figure 7C. We observed that the “ImmuneScore” in the “High” group was significantly higher than that in the “Low” group, and that LGG patients with a high RiskScore had higher immune cell infiltration.

In order to explore discrepancies in the molecular characteristics between high- and low-RiskScore groups, the ssGSEA algorithm was utilized to evaluate the activity of signaling pathways. As shown in Figure 7D, the activity of the following pathways in the high-RiskScore group were noticeably higher than those in the low-RiskScore group: interferon, glycolysis, apoptosis, and angiogenesis pathways, among others. Correlation analysis further validated the close

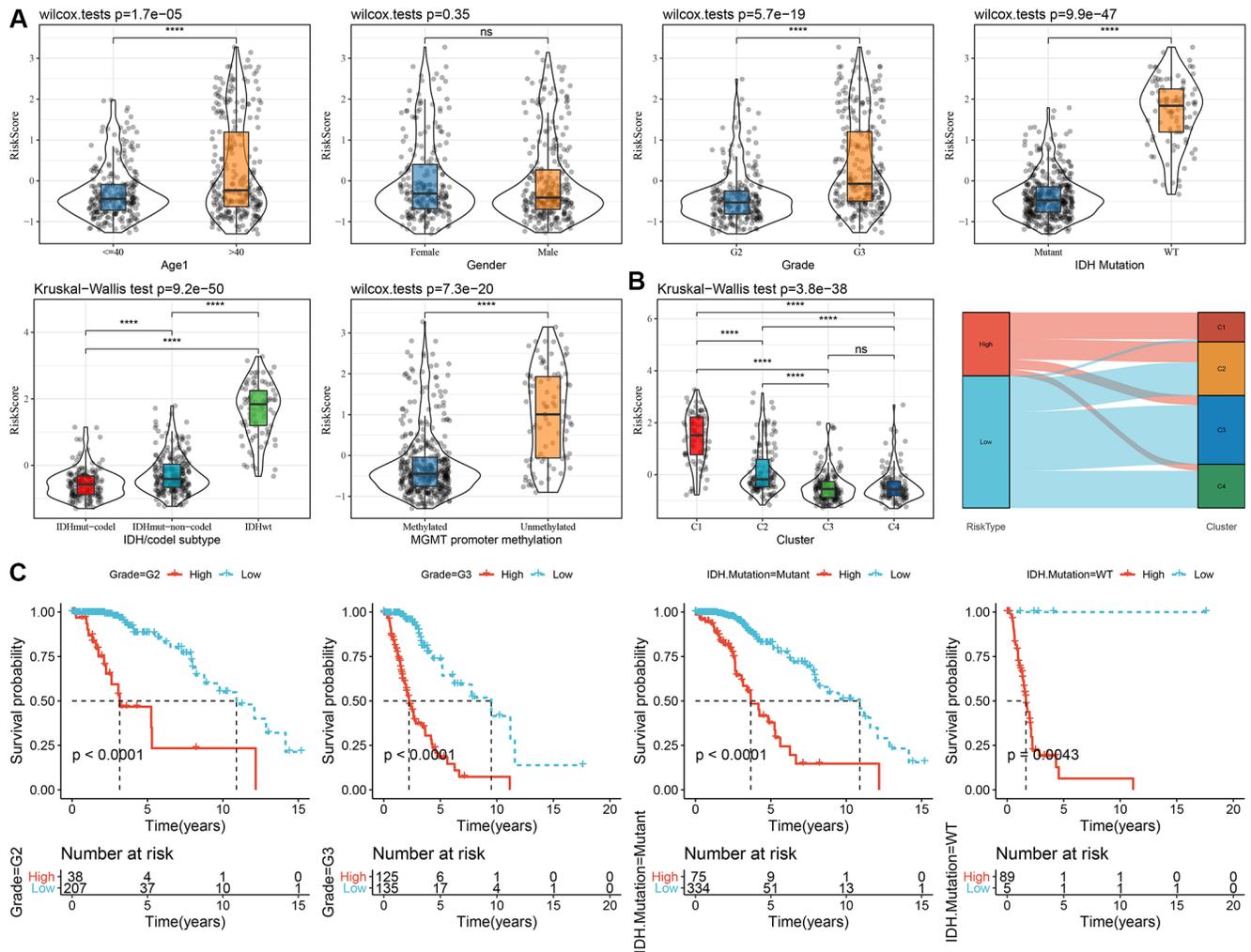


Figure 6. Association of RiskScore with clinical information in TCGA dataset. (A) Differences in RiskScore among different clinicopathological groups in the TCGA-LGG cohort; (B) Differences in RiskScore among different molecular subtypes in the TCGA-LGG cohort; (C) KM curves between high and low RiskScore groups in different clinical subgroup. Abbreviations: TCGA: The Cancer Genome Atlas; LGG: low-grade glioma; K-M: Kaplan-Meier.

association of the RiskScore with the above pathways (Figure 7E). Hence, dysfunction of the immune microenvironment and molecular signaling might be responsible for the different prognoses observed between high- and low-RiskScore groups.

Chemotherapy and immunotherapy efficacy differences across RiskScore groups

Initially, the “T-cell-inflamed GEP score” was used to assess the predictive potential of the RiskScore in ICB across different groups. The results suggested that, in the high-RiskScore group, the “T-cell-inflamed GEP score” and the “IFN- γ response” showed a considerable

increase, and that the CYT score was also higher (Figure 8A–8C). Moreover, the expression of some checkpoint molecules (CTLA-4, PD-1, PD-L1) was noticeably higher (Figure 8D). Overall, patients with a high-RiskScore are likely to be more sensitivity to ICB.

In addition, patients with a high-RiskScore had lower IC₅₀ values for bleomycin and A-443654, whereas patients with low-RiskScore had lower IC₅₀ values for cyclophosphamide, AZD6482, and GDC0941. These findings suggested that high-RiskScore patients might benefit from treatment with bleomycin and A-443654; while low-RiskScore patients might benefit from cyclophosphamide, AZD6482, and GDC0941 (Figure 8E).

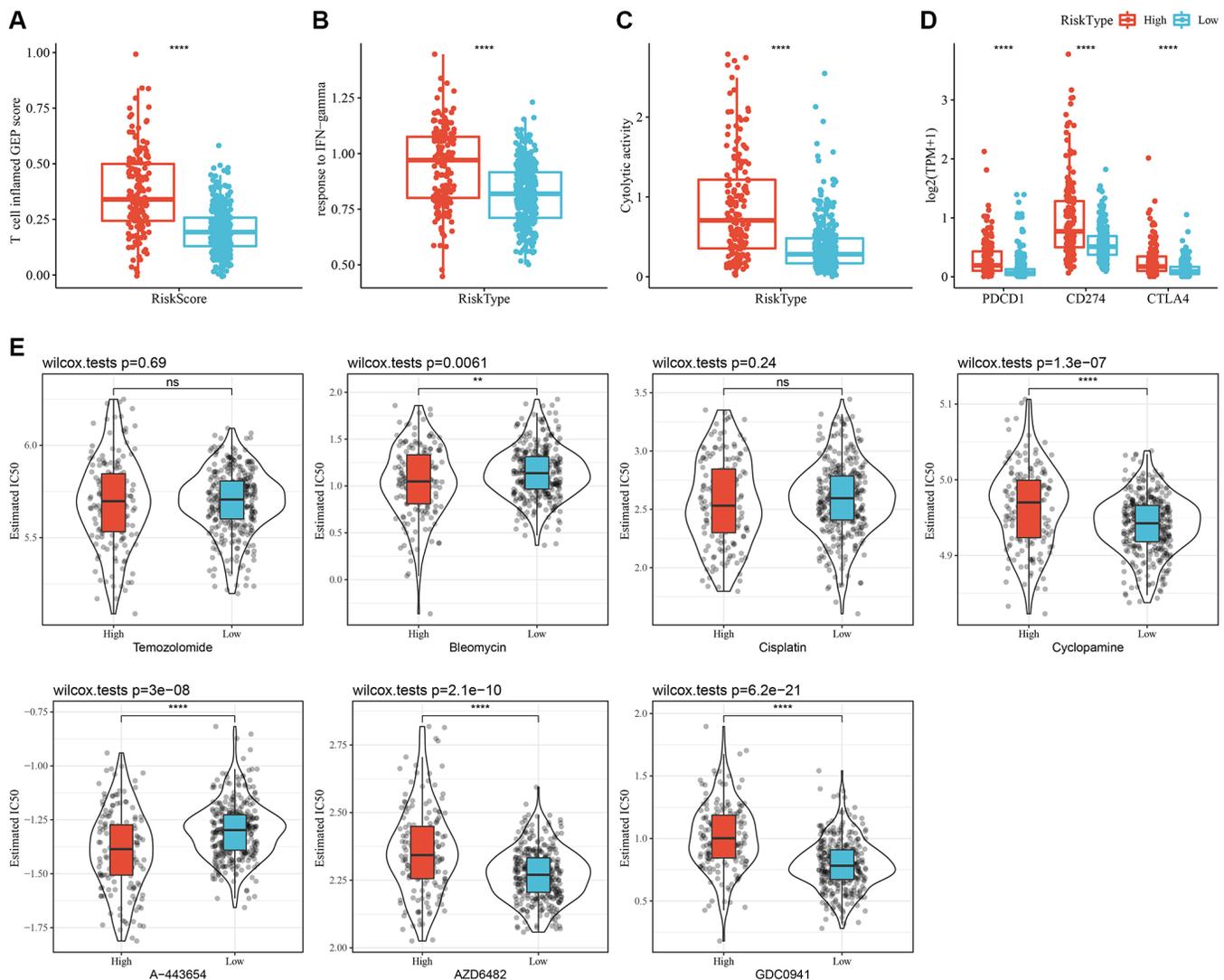


Figure 8. Immunological characteristics and drug sensitivities between different RiskScore groups. (A) The difference in “T cell inflamed GEP score” among various molecular subtypes. (B) The difference in “response to IFN- γ ” among various molecular subtypes. (C) The difference in “Cytolytic activity” among various molecular subtypes. (D) Expression differences of immune checkpoint-associated genes among various molecular subtypes. (E) A box plot of the estimated IC₅₀ values for temozolomide, bleomycin, cisplatin, cyclophosphamide, A-443654, AZD6482, and GDC0941 in TCGA-LGG. ** $P < 0.01$; **** $P < 0.0001$. Abbreviations: ns: no significance; GEP: gene expression profile; IFN: interferon; TPM: transcripts per million; IC₅₀: half-maximal inhibitory concentration; TCGA: The Cancer Genome Atlas; LGG: low-grade glioma.

RiskScore incorporates clinicopathological features to improve the prognosis models and survival prediction

To further improve the prognostic model, clinical information of patients (age, gender, tumor-node-

metastasis (TNM) stage, *IDH* mutation, and RiskType) and RiskScores were integrated, and a decision tree model was constructed. Ultimately, only two parameters, *IDH* mutation and RiskType, were retained. Considering the maximum weight of RiskType, three subtypes were identified (Figure 9A) and their

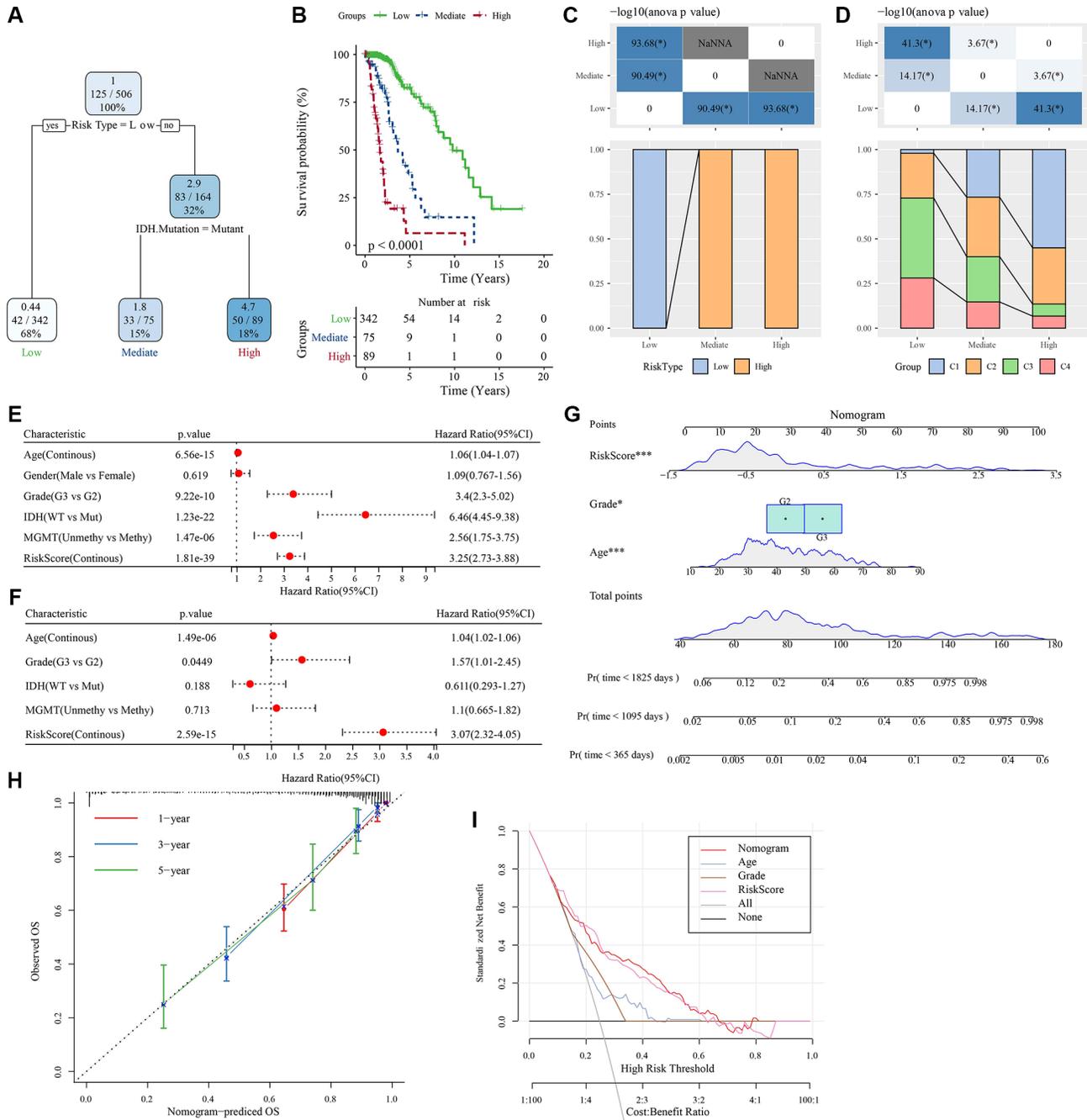


Figure 9. Determination of optimal prognostic factors by decision tree and their reliability. (A) Patients with full-scale annotations, including gender, RiskScore, age, and TNM stage, were employed to develop a survival decision tree for optimizing the risk stratification. (B) Significant differences in OS were found among the 4 risk subgroups. (C, D) Comparative analysis on different subgroups. (E, F) Univariate and multivariate Cox analyses of clinicopathological characteristics and RiskScore. (G) Nomogram model. (H) 1-, 3-, and 5-year calibration curves of the nomogram. (I) Decision curves of the nomogram. * $P < 0.05$; *** $P < 0.001$. Abbreviations: ANOVA: analysis of variance; IDH: isocitrate dehydrogenase; WT: wild-type; Mut: mutant; MGMT: O-6-methylguanine-DNA methyltransferase; CI: confidence interval; OS: overall survival; TNM: tumor-node-metastasis; Pr: predicted.

prognosis was examined (Figure 9B). The association of the decision tree model with the LASSO-Cox model is shown in Figure 9C, and the association of the decision tree model with consensus clusters is shown in Figure 9D. Univariate and multivariate Cox regression analyses highlighted the crucial role of the RiskScore in predicting clinical prognoses (Figure 9E, 9F). To quantify the risk assessment and survival probability of patients with LGG, we combined RiskScore and other clinicopathological features to establish a nomogram (Figure 9G). The calibration curve was used for model accuracy evaluation. The 1-, 3-, and 5-year predictive calibration curves nearly overlapped with the standard curve, suggesting a strong prediction performance (Figure 9H). In addition, the reliability of the model was assessed using the decision curve analysis (DCA). The RiskScore and nomogram benefits were both noticeably elevated when compared to the extreme curves, exhibiting strong survival prediction power over other clinicopathological characteristics (Figure 9I).

Pan-cancer characterization of NK cell-related genes

The CNV profiles of NK cell-related genes in the pan-cancer cohort were summarized and plotted in Figure 10A, 10B. KLRK1, KLRF1, KLRD1, KLRC3, and KLRB1 showed relatively noticeable CNV amplification; while, GZMM, GZMH, GZMB, GZMA, and GRAP2 showed CNV deletion (Figure 10A, 10B). The majority of NK cell-related genes were accompanied by low methylation levels in cancers when compared to precancerous tissues. Of note, the methylation levels of PTGDR and ZNF135 genes were substantially higher in cancers when compared to precancerous tissues (Figure 10C). In addition, we also investigated other genomics traits (i.e., SNV). Results showed that mutation frequencies of MGAM and TEP1 were relatively high among all NK cell-related genes (Figure 11A, 11B). Importantly, we explored the correlation among NK cell-related genes, the tumor immune microenvironment, and tumor metabolic

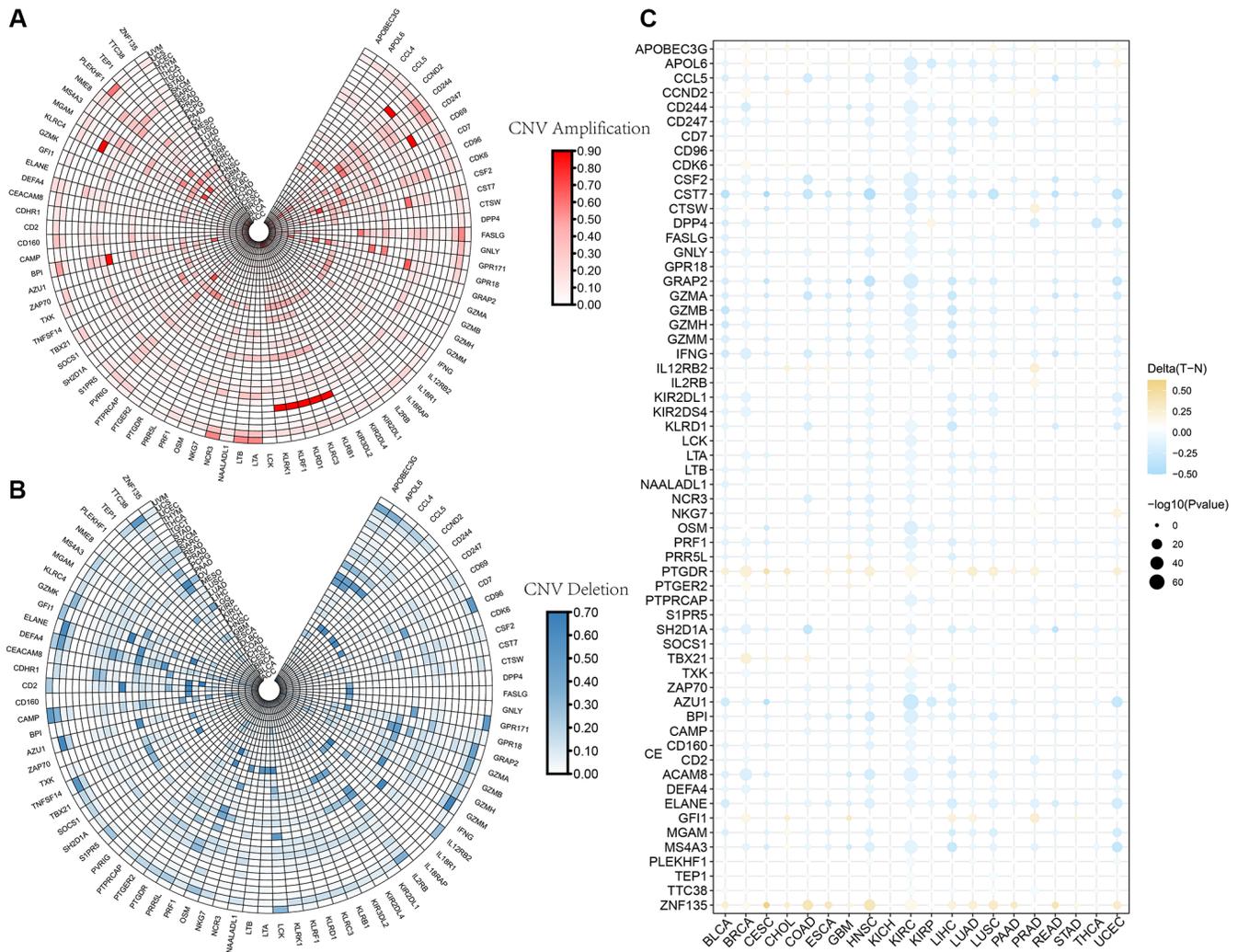


Figure 10. The CNV and methylation profiles of NK cell-related genes in pan-cancer. (A) CNV amplification of NK cell-related genes. **(B)** CNV deletion of NK cell-related genes. **(C)** DNA methylation traits of NK cell-related genes.

EMT-related pathways) [50]. Based on these findings, NK cells may contribute to LGG via both of these mechanisms. Moreover, the pan-characteristic of NK cell-related genes presented a wide range of CNVs and mutation statuses, confirming that NK cell gene mutations also contribute to the pathogenic mechanisms of tumor immune escape [51].

T cells are the primary immune cells targeted in ICB, and the balance between their activating and inhibiting signals plays a crucial role in ICB [52]. CTLA-4 and PD-1 are the two primary immunological checkpoint molecules on the surface of T cells. CTLA-4 modulates T cell activation during the first phase of immunological activation, while PD-1 is activated throughout the

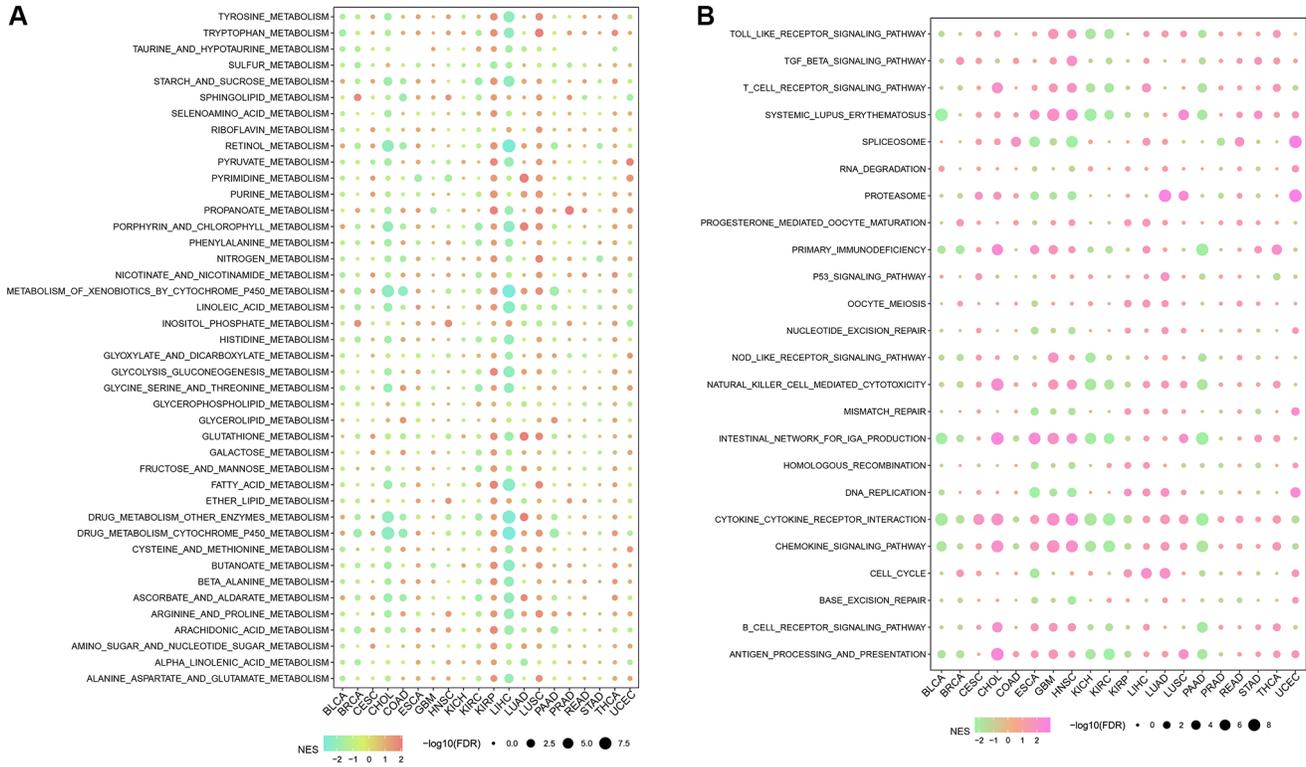


Figure 12. The correlation among NK cell-related genes, tumor metabolic reprogramming, and immune microenvironment in pan-cancer. (A) Correlation between NK cell-related genes and tumor metabolism-related pathways **(B)** Correlation between NK cell-related genes and immune-related pathways.

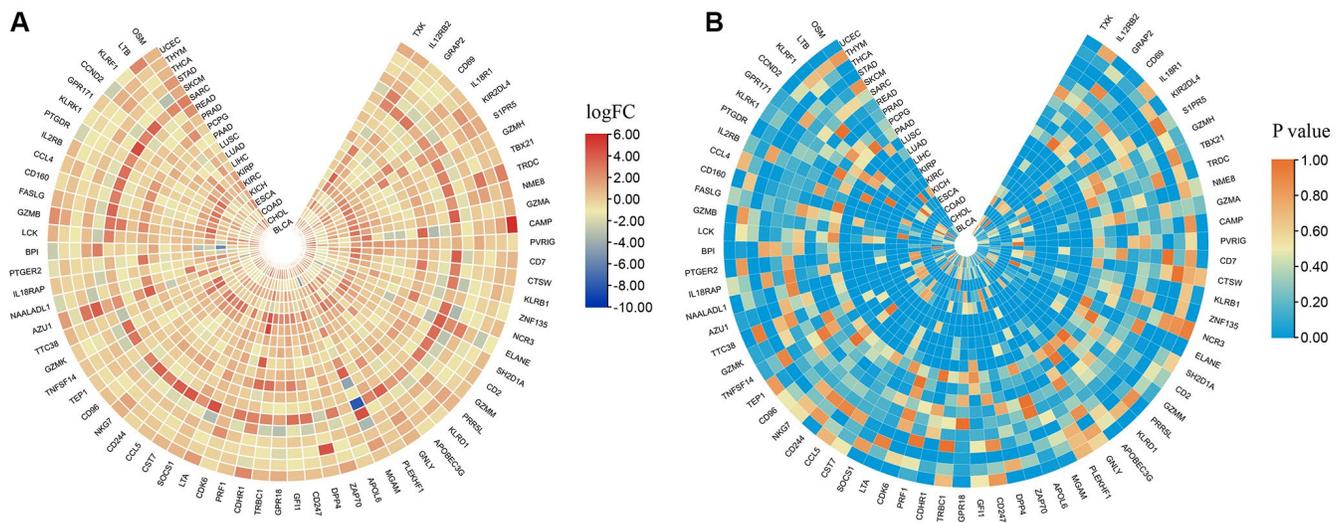


Figure 13. The expression levels of NK cell-related genes in pan-cancer. (A) $\log_{2}(\text{FC})$ values of differential expression analysis. **(B)** p values of differential expression analysis.

immune effector phase and is abundantly produced when tumor antigens are presented [53, 54]. In the present study, the expression of representative immune checkpoint molecules was evaluated in different LGG subtypes wherein *PD-1*, *PD-L1*, and *CTLA-4* were considerably highly expressed in the C1 subtype, which was consistent with a study by Ghoulzani et al. [55]. Therefore, patients in the C1 subtype may be potential candidates for ICB.

Most clinical trials investigating the use of *CTLA-4* and *PD-1* inhibitors in glioma are currently ongoing, with only a few published findings, and none achieving significant efficacy [56–58]. The reason for this may be attributed to drug resistance. There are many mechanisms associated with the resistance of ICB in the treatment of glioma, which can be broadly categorized as endogenous or exogenous. Endogenous mechanisms include alterations in immune recognition sites, abnormal cellular signaling pathways, alterations in gene expression, and DNA damage repair [59]. On the other hand, exogenous mechanisms include all factors related to immune cell activation [60]. Common factors used to determine patient prognoses and prediction of treatment outcomes include *PD-1/PD-L1* expression levels, TMB, tumor-infiltrating lymphocytes (TILs), and microsatellite instability (MSI) [61]. In the present study, ssGSEA revealed that *PD-1/PD-L1*, TMB, and TILs were significantly different among LGG subtypes, highlighting the diversity of gliomas and providing new indicators for predicting ICB efficacy.

Construction of a risk model enables the more exact identification of high-risk groups. In the present study, a prediction model was successfully constructed based on 6 NK cell-related genes, and patients were classified into high- and low-risk groups. Prostaglandin F2 receptor negative regulator (*PTGFRN*), a tumorigenesis-related gene, was associated with interleukin (IL)-12-mediated tumor recognition and killing efficacy for drugs [62, 63]. The role of *HSPB1* in tumor immunity has been previously demonstrated, and its mechanism of action is related to the direct immunosuppression of Ym1 produced by macrophages and T cell suppression [63]. The sterile alpha motif domain-containing protein 9 (*SAMD9*) is a potential antigen for the development of messenger RNA (mRNA) vaccines against diffuse glioma. Its expression is associated with tumor immune subtypes and determines immune-related processes of tumor-associated genes [64]. In addition, the expression of secreted frizzled-related protein 2 (*SFRP2*), a member of the secretory glycoprotein family, correlates with the degree of immune infiltration of tumor cells and plays a synergistic role in tumor progression. Therefore, *SFRP2* is a promising prognostic biomarker and therapeutic target [65]. Although less research has

been done on elastin Microfibril Interfacer 3 (*EMILIN3*) and cartilage acidic protein 1 (*CRTAC1*) in tumor immunity, previous findings suggest that *CRTAC1* may influence the onset and progression of tumors by binding to calcium ions and innate immune pathways [66].

Given the poor prognosis of individuals with LGG, it is vital to discover clinical variables that influence their prognosis and to implement appropriate therapies. This research successfully divided LGG patients into two groups by calculating RiskScores and determining patient prognoses, which were found to be significantly different. In addition, a nomogram was constructed and validated to determine the ability to accurately predict the prognosis of patients with LGG. Overall, the nomogram exhibited good prognostic ability and was a powerful tool in predicting the survival status of individuals with LGG across clinical settings. This study is the first to identify molecular subtypes and prognostic models associated with LGG based on NK cells. In addition, this new six-gene prognostic model has not been previously reported. Nonetheless, there are some limitations to this study, including the validation of these findings using PCR or immunohistochemical experiments. Moreover, some clinical factors were not considered due to the lack of necessary clinical follow-up information, especially diagnostic details. These limitations may influence the effectiveness of the clinical application of our RiskScore.

CONCLUSIONS

In this study, molecular subtypes were categorized using consistent clustering based on NK cell-correlated genes and showed different immunological, pathological, prognostic, and pathway features. Subsequently, from these molecular subtypes, 6 key NK cell-related genes were screened based on DEGs, and a clinical prognostic model was developed. This model was independent of other clinicopathological features, demonstrating stable predictive performance in independent datasets and strong robustness. We finally integrated clinicopathological characteristics with RiskScores and constructed a decision tree model to further improve its survival prediction. Overall, this tool may be used for patient stratification, particularly for ICB. Moreover, our study may provide critical information to overcome immune resistance by targeting NK cell-related genes.

AUTHOR CONTRIBUTIONS

The first author (Fei Sun and Hongtao Lv) is responsible for the LGG-related data collection, data analysis, and draft writing. The second author (Baozhi

Feng) is responsible for the data collection and analysis about pan-cancer omics. The third author (Jiaao Sun) participated in the genomics and transcriptomics pan-cancer analysis. The corresponding author (Linyun Zhang and Bin Dong) revised and finally approved the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

FUNDING

The authors wish to thank the National Natural Science Foundation of China (No. 81672968) for supporting this research.

REFERENCES

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022; 72:7–33. <https://doi.org/10.3322/caac.21708> PMID:35020204
2. Wang TJC, Mehta MP. Low-Grade Glioma Radiotherapy Treatment and Trials. *Neurosurg Clin N Am.* 2019; 30:111–8. <https://doi.org/10.1016/j.nec.2018.08.008> PMID:30470398
3. Wang J, Ren J, Liu J, Zhang L, Yuan Q, Dong B. Identification and verification of the ferroptosis- and pyroptosis-associated prognostic signature for low-grade glioma. *Bosn J Basic Med Sci.* 2022; 22:728–50. <https://doi.org/10.17305/bjbms.2021.6888> PMID:35276059
4. Jooma R, Waqas M, Khan I. Diffuse Low-Grade Glioma - Changing Concepts in Diagnosis and Management: A Review. *Asian J Neurosurg.* 2019; 14:356–63. https://doi.org/10.4103/ajns.AJNS_24_18 PMID:31143247
5. Bell EH, Zhang P, Fisher BJ, Macdonald DR, McElroy JP, Lesser GJ, Fleming J, Chakraborty AR, Liu Z, Becker AP, Fabian D, Aldape KD, Ashby LS, et al. Association of MGMT Promoter Methylation Status With Survival Outcomes in Patients With High-Risk Glioma Treated With Radiotherapy and Temozolomide: An Analysis From the NRG Oncology/RTOG 0424 Trial. *JAMA Oncol.* 2018; 4:1405–9. <https://doi.org/10.1001/jamaoncol.2018.1977> PMID:29955793
6. Buckner JC, Shaw EG, Pugh SL, Chakravarti A, Gilbert MR, Barger GR, Coons S, Ricci P, Bullard D, Brown PD, Stelzer K, Brachman D, Suh JH, et al. Radiation plus Procarbazine, CCNU, and Vincristine in Low-Grade Glioma. *N Engl J Med.* 2016; 374:1344–55. <https://doi.org/10.1056/NEJMoa1500925> PMID:27050206
7. de Blank P, Bandopadhyay P, Haas-Kogan D, Fouladi M, Fangusaro J. Management of pediatric low-grade glioma. *Curr Opin Pediatr.* 2019; 31:21–7. <https://doi.org/10.1097/MOP.0000000000000717> PMID:30531227
8. Pottier C, Wheatherspoon A, Roncarati P, Longuespée R, Herfs M, Duray A, Delvenne P, Quatresooz P. The importance of the tumor microenvironment in the therapeutic management of cancer. *Expert Rev Anticancer Ther.* 2015; 15:943–54. <https://doi.org/10.1586/14737140.2015.1059279> PMID:26098949
9. Zhang C, Zhang Y, Tan G, Mi W, Zhong X, Zhang Y, Zhao Z, Li F, Xu Y, Zhang Y. Prognostic Features of the Tumor Immune Microenvironment in Glioma and Their Clinical Applications: Analysis of Multiple Cohorts. *Front Immunol.* 2022; 13:853074. <https://doi.org/10.3389/fimmu.2022.853074> PMID:35677045
10. van den Bent MJ, Wefel JS, Schiff D, Taphoorn MJ, Jaeckle K, Junck L, Armstrong T, Choucair A, Waldman AD, Gorlia T, Chamberlain M, Baumert BG, Vogelbaum MA, et al. Response assessment in neuro-oncology (a report of the RANO group): assessment of outcome in trials of diffuse low-grade gliomas. *Lancet Oncol.* 2011; 12:583–93. [https://doi.org/10.1016/S1470-2045\(11\)70057-2](https://doi.org/10.1016/S1470-2045(11)70057-2) PMID:21474379
11. Sun Y, Sedgwick AJ, Palarasah Y, Mangiola S, Barrow AD. A Transcriptional Signature of PDGF-DD Activated Natural Killer Cells Predicts More Favorable Prognosis in Low-Grade Glioma. *Front Immunol.* 2021; 12:668391. <https://doi.org/10.3389/fimmu.2021.668391> PMID:34539622
12. Karlhofer FM, Ribaldo RK, Yokoyama WM. MHC class I alloantigen specificity of Ly-49+ IL-2-activated natural killer cells. *Nature.* 1992; 358:66–70. <https://doi.org/10.1038/358066a0> PMID:1614533
13. Korin B, Ben-Shaan TL, Schiller M, Dubovik T, Azulay-Debby H, Boshnak NT, Koren T, Rolls A. High-dimensional, single-cell characterization of the brain's immune compartment. *Nat Neurosci.* 2017; 20:1300–9. <https://doi.org/10.1038/nn.4610> PMID:28758994
14. Holl EK, Frazier VN, Landa K, Beasley GM, Hwang ES, Nair SK. Examining Peripheral and Tumor Cellular

- Immunome in Patients With Cancer. *Front Immunol*. 2019; 10:1767.
<https://doi.org/10.3389/fimmu.2019.01767>
PMID:[31417550](https://pubmed.ncbi.nlm.nih.gov/31417550/)
15. Ren F, Zhao Q, Huang L, Zheng Y, Li L, He Q, Zhang C, Li F, Maimela NR, Sun Z, Jia Q, Ping Y, Zhang Z, et al. The R132H mutation in IDH1 promotes the recruitment of NK cells through CX3CL1/CX3CR1 chemotaxis and is correlated with a better prognosis in gliomas. *Immunol Cell Biol*. 2019; 97:457–69.
<https://doi.org/10.1111/imcb.12225>
PMID:[30575118](https://pubmed.ncbi.nlm.nih.gov/30575118/)
16. Bryceson YT, March ME, Ljunggren HG, Long EO. Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion. *Blood*. 2006; 107:159–66.
<https://doi.org/10.1182/blood-2005-04-1351>
PMID:[16150947](https://pubmed.ncbi.nlm.nih.gov/16150947/)
17. Barrow AD, Martin CJ, Colonna M. The Natural Cytotoxicity Receptors in Health and Disease. *Front Immunol*. 2019; 10:909.
<https://doi.org/10.3389/fimmu.2019.00909>
PMID:[31134055](https://pubmed.ncbi.nlm.nih.gov/31134055/)
18. Castriconi R, Daga A, Dondero A, Zona G, Poliani PL, Melotti A, Griffiero F, Marubbi D, Spaziante R, Bellora F, Moretta L, Moretta A, Corte G, Bottino C. NK cells recognize and kill human glioblastoma cells with stem cell-like properties. *J Immunol*. 2009; 182:3530–9.
<https://doi.org/10.4049/jimmunol.0802845>
PMID:[19265131](https://pubmed.ncbi.nlm.nih.gov/19265131/)
19. Appolloni I, Alessandrini F, Ceresa D, Marubbi D, Gambini E, Reverberi D, Loiacono F, Malatesta P. Progression from low- to high-grade in a glioblastoma model reveals the pivotal role of immunoediting. *Cancer Lett*. 2019; 442:213–21.
<https://doi.org/10.1016/j.canlet.2018.10.006>
PMID:[30312732](https://pubmed.ncbi.nlm.nih.gov/30312732/)
20. Wang G, Cen Y, Wang C, Xiang W, Li S, Ming Y, Chen L, Zhou J. Elevated RECQL1 expression predicts poor prognosis and associates with tumor immune infiltration in low-grade glioma. *Transl Cancer Res*. 2022; 11:1552–64.
<https://doi.org/10.21037/tcr-21-2762>
PMID:[35836526](https://pubmed.ncbi.nlm.nih.gov/35836526/)
21. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12:453–7.
<https://doi.org/10.1038/nmeth.3337>
PMID:[25822800](https://pubmed.ncbi.nlm.nih.gov/25822800/)
22. Liu H, Wang J, Luo T, Zhen Z, Liu L, Zheng Y, Zhang C, Hu X. Correlation between ITGB2 expression and clinical characterization of glioma and the prognostic significance of its methylation in low-grade glioma(LGG). *Front Endocrinol (Lausanne)*. 2023; 13:1106120.
<https://doi.org/10.3389/fendo.2022.1106120>
PMID:[36714574](https://pubmed.ncbi.nlm.nih.gov/36714574/)
23. Miao Y, Liu J, Liu X, Yuan Q, Li H, Zhang Y, Zhan Y, Feng X. Machine learning identification of cuproptosis and necroptosis-associated molecular subtypes to aid in prognosis assessment and immunotherapy response prediction in low-grade glioma. *Front Genet*. 2022; 13:951239.
<https://doi.org/10.3389/fgene.2022.951239>
PMID:[36186436](https://pubmed.ncbi.nlm.nih.gov/36186436/)
24. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018; 28:1747–56.
<https://doi.org/10.1101/gr.239244.118>
PMID:[30341162](https://pubmed.ncbi.nlm.nih.gov/30341162/)
25. Liu H, Liu X, Zhang S, Lv J, Li S, Shang S, Jia S, Wei Y, Wang F, Su J, Wu Q, Zhang Y. Systematic identification and annotation of human methylation marks based on bisulfite sequencing methylomes reveals distinct roles of cell type-specific hypomethylation in the regulation of cell identity genes. *Nucleic Acids Res*. 2016; 44:75–94.
<https://doi.org/10.1093/nar/gkv1332>
PMID:[26635396](https://pubmed.ncbi.nlm.nih.gov/26635396/)
26. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010; 26:1572–3.
<https://doi.org/10.1093/bioinformatics/btq170>
PMID:[20427518](https://pubmed.ncbi.nlm.nih.gov/20427518/)
27. Yuan Q, Zhang W, Shang W. Identification and validation of a prognostic risk-scoring model based on sphingolipid metabolism-associated cluster in colon adenocarcinoma. *Front Endocrinol (Lausanne)*. 2022; 13:1045167.
<https://doi.org/10.3389/fendo.2022.1045167>
PMID:[36518255](https://pubmed.ncbi.nlm.nih.gov/36518255/)
28. Zheng C, Berger NA, Li L, Xu R. Epigenetic age acceleration and clinical outcomes in gliomas. *PLoS One*. 2020; 15:e0236045.
<https://doi.org/10.1371/journal.pone.0236045>
PMID:[32692766](https://pubmed.ncbi.nlm.nih.gov/32692766/)
29. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015; 43:e47.
<https://doi.org/10.1093/nar/gkv007>
PMID:[25605792](https://pubmed.ncbi.nlm.nih.gov/25605792/)

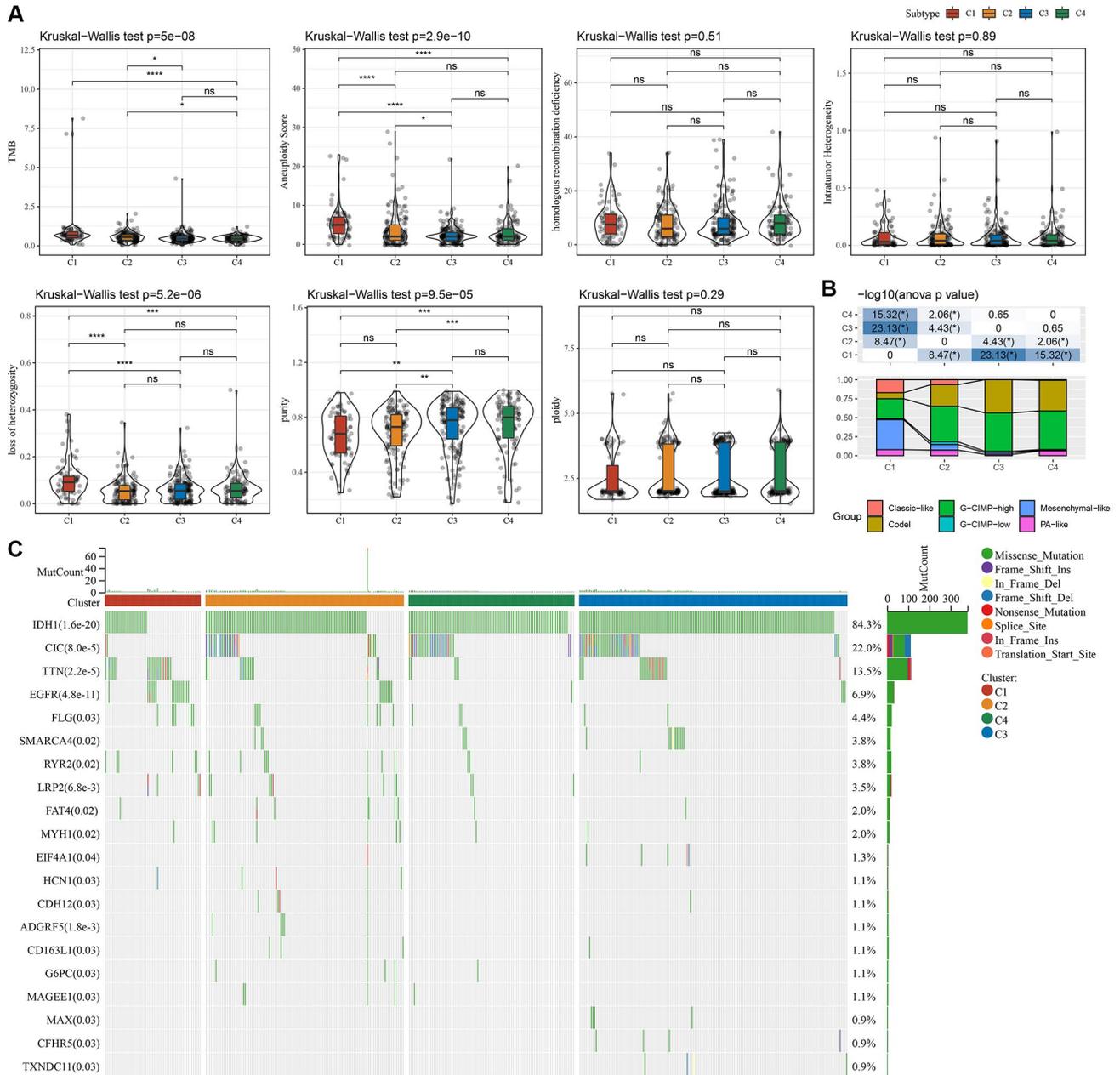
30. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015; 1:417–25.
<https://doi.org/10.1016/j.cels.2015.12.004>
PMID:26771021
31. Liu Q, Cheng R, Kong X, Wang Z, Fang Y, Wang J. Molecular and Clinical Characterization of PD-1 in Breast Cancer Using Large-Scale Transcriptome Data. *Front Immunol.* 2020; 11:558757.
<https://doi.org/10.3389/fimmu.2020.558757>
PMID:33329517
32. Masiero M, Simões FC, Han HD, Snell C, Peterkin T, Bridges E, Mangala LS, Wu SY, Pradeep S, Li D, Han C, Dalton H, Lopez-Berestein G, et al. A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell.* 2013; 24:229–41.
<https://doi.org/10.1016/j.ccr.2013.06.004>
PMID:23871637
33. Palmeri M, Mehnert J, Silk AW, Jabbour SK, Ganesan S, Popli P, Riedlinger G, Stephenson R, de Meritens AB, Leiser A, Mayer T, Chan N, Spencer K, et al. Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. *ESMO Open.* 2022; 7:100336.
<https://doi.org/10.1016/j.esmoop.2021.100336>
PMID:34953399
34. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science.* 2017; 355:eaaf8399.
<https://doi.org/10.1126/science.aaf8399>
PMID:28104840
35. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One.* 2014; 9:e107468.
<https://doi.org/10.1371/journal.pone.0107468>
PMID:25229481
36. Yuan Q, Ren J, Chen X, Dong Y, Shang D. Contributions and prognostic performances of m7G RNA regulators in pancreatic adenocarcinoma. *Chin Med J (Engl).* 2022; 135:2101–3.
<https://doi.org/10.1097/CM9.0000000000002179>
PMID:36201641
37. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013; 4:2612.
<https://doi.org/10.1038/ncomms3612>
PMID:24113773
38. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013; 14:7.
<https://doi.org/10.1186/1471-2105-14-7>
PMID:23323831
39. Yuan Q, Ren J, Li L, Li S, Xiang K, Shang D. Development and validation of a novel N6-methyladenosine (m6A)-related multi-long non-coding RNA (lncRNA) prognostic signature in pancreatic adenocarcinoma. *Bioengineered.* 2021; 12:2432–48.
<https://doi.org/10.1080/21655979.2021.1933868>
PMID:34233576
40. Ostrom QT, Bauchet L, Davis FG, Deltour I, Fisher JL, Langer CE, Pekmezci M, Schwartzbaum JA, Turner MC, Walsh KM, Wrensch MR, Barnholtz-Sloan JS. The epidemiology of glioma in adults: a "state of the science" review. *Neuro Oncol.* 2014; 16:896–913.
<https://doi.org/10.1093/neuonc/nou087>
PMID:24842956
41. Chen R, Smith-Cohn M, Cohen AL, Colman H. Glioma Subclassifications and Their Clinical Significance. *Neurotherapeutics.* 2017; 14:284–97.
<https://doi.org/10.1007/s13311-017-0519-x>
PMID:28281173
42. Ruff M, Kizilbash S, Buckner J. Further understanding of glioma mechanisms of pathogenesis: implications for therapeutic development. *Expert Rev Anticancer Ther.* 2020; 20:355–63.
<https://doi.org/10.1080/14737140.2020.1757440>
PMID:32301635
43. Cahill D, Turcan S. Origin of Gliomas. *Semin Neurol.* 2018; 38:5–10.
<https://doi.org/10.1055/s-0037-1620238>
PMID:29548046
44. Kaminska B, Cyranowski S. Recent Advances in Understanding Mechanisms of TGF Beta Signaling and Its Role in Glioma Pathogenesis. *Adv Exp Med Biol.* 2020; 1202:179–201.
https://doi.org/10.1007/978-3-030-30651-9_9
PMID:32034714
45. Wang Z, Su G, Dai Z, Meng M, Zhang H, Fan F, Liu Z, Zhang L, Weygant N, He F, Fang N, Zhang L, Cheng Q. Circadian clock genes promote glioma progression by affecting tumour immune infiltration and tumour cell proliferation. *Cell Prolif.* 2021; 54:e12988.
<https://doi.org/10.1111/cpr.12988>
PMID:33442944
46. Sokratous G, Polyzoidis S, Ashkan K. Immune infiltration of tumor microenvironment following

- immunotherapy for glioblastoma multiforme. *Hum Vaccin Immunother.* 2017; 13:2575–82.
<https://doi.org/10.1080/21645515.2017.1303582>
PMID:[28362548](https://pubmed.ncbi.nlm.nih.gov/28362548/)
47. Kang K, Xie F, Wu Y, Wang Z, Wang L, Long J, Lian X, Zhang F. Comprehensive exploration of tumor mutational burden and immune infiltration in diffuse glioma. *Int Immunopharmacol.* 2021; 96:107610.
<https://doi.org/10.1016/j.intimp.2021.107610>
PMID:[33848908](https://pubmed.ncbi.nlm.nih.gov/33848908/)
48. Duygu B, Olieslagers TI, Groeneweg M, Voorter CEM, Wieten L. HLA Class I Molecules as Immune Checkpoints for NK Cell Alloreactivity and Anti-Viral Immunity in Kidney Transplantation. *Front Immunol.* 2021; 12:680480.
<https://doi.org/10.3389/fimmu.2021.680480>
PMID:[34295330](https://pubmed.ncbi.nlm.nih.gov/34295330/)
49. Ryan JC, Naper C, Hayashi S, Daws MR. Physiologic functions of activating natural killer (NK) complex-encoded receptors on NK cells. *Immunol Rev.* 2001; 181:126–37.
<https://doi.org/10.1034/j.1600-065x.2001.1810110.x>
PMID:[11513134](https://pubmed.ncbi.nlm.nih.gov/11513134/)
50. Friede ME, Leibelt S, Dudziak D, Steinle A. Select Clr-g Expression on Activated Dendritic Cells Facilitates Cognate Interaction with a Minor Subset of Splenic NK Cells Expressing the Inhibitory Nkrp1g Receptor. *J Immunol.* 2018; 200:983–96.
<https://doi.org/10.4049/jimmunol.1701180>
PMID:[29263211](https://pubmed.ncbi.nlm.nih.gov/29263211/)
51. Polprasert C, Takeuchi Y, Makishima H, Wudhikarn K, Kakiuchi N, Tangnuntachai N, Assanasen T, Sitthi W, Muhamad H, Lawasut P, Kongkiatkamon S, Bunworasate U, Izutsu K, et al. Frequent mutations in HLA and related genes in extranodal NK/T cell lymphomas. *Leuk Lymphoma.* 2021; 62:95–103.
<https://doi.org/10.1080/10428194.2020.1821011>
PMID:[32964767](https://pubmed.ncbi.nlm.nih.gov/32964767/)
52. Lim M, Xia Y, Bettegowda C, Weller M. Current state of immunotherapy for glioblastoma. *Nat Rev Clin Oncol.* 2018; 15:422–42.
<https://doi.org/10.1038/s41571-018-0003-5>
PMID:[29643471](https://pubmed.ncbi.nlm.nih.gov/29643471/)
53. McGranahan T, Therkelsen KE, Ahmad S, Nagpal S. Current State of Immunotherapy for Treatment of Glioblastoma. *Curr Treat Options Oncol.* 2019; 20:24.
<https://doi.org/10.1007/s11864-019-0619-4>
PMID:[30790064](https://pubmed.ncbi.nlm.nih.gov/30790064/)
54. Medikonda R, Dunn G, Rahman M, Fecci P, Lim M. A review of glioblastoma immunotherapy. *J Neurooncol.* 2021; 151:41–53.
<https://doi.org/10.1007/s11060-020-03448-1>
PMID:[32253714](https://pubmed.ncbi.nlm.nih.gov/32253714/)
55. Ghouzlani A, Kandoussi S, Tall M, Reddy KP, Rafii S, Badou A. Immune Checkpoint Inhibitors in Human Glioma Microenvironment. *Front Immunol.* 2021; 12:679425.
<https://doi.org/10.3389/fimmu.2021.679425>
PMID:[34305910](https://pubmed.ncbi.nlm.nih.gov/34305910/)
56. Schalper KA, Rodriguez-Ruiz ME, Diez-Valle R, López-Janeiro A, Porciuncula A, Idoate MA, Inogés S, de Andrea C, López-Díaz de Cerio A, Tejada S, Berraondo P, Villarroel-Espindola F, Choi J, et al. Neoadjuvant nivolumab modifies the tumor immune microenvironment in resectable glioblastoma. *Nat Med.* 2019; 25:470–6.
<https://doi.org/10.1038/s41591-018-0339-5>
PMID:[30742120](https://pubmed.ncbi.nlm.nih.gov/30742120/)
57. Omuro A, Vlahovic G, Lim M, Sahebjam S, Baehring J, Cloughesy T, Voloschin A, Ramkissoon SH, Ligon KL, Latek R, Zvirtes R, Strauss L, Paliwal P, et al. Nivolumab with or without ipilimumab in patients with recurrent glioblastoma: results from exploratory phase I cohorts of CheckMate 143. *Neuro Oncol.* 2018; 20:674–86.
<https://doi.org/10.1093/neuonc/nox208>
PMID:[29106665](https://pubmed.ncbi.nlm.nih.gov/29106665/)
58. Reardon DA, Brandes AA, Omuro A, Mulholland P, Lim M, Wick A, Baehring J, Ahluwalia MS, Roth P, Bähr O, Phuphanich S, Sepulveda JM, De Souza P, et al. Effect of Nivolumab vs Bevacizumab in Patients With Recurrent Glioblastoma: The CheckMate 143 Phase 3 Randomized Clinical Trial. *JAMA Oncol.* 2020; 6:1003–10.
<https://doi.org/10.1001/jamaoncol.2020.1024>
PMID:[32437507](https://pubmed.ncbi.nlm.nih.gov/32437507/)
59. Riaz N, Morris L, Havel JJ, Makarov V, Desrichard A, Chan TA. The role of neoantigens in response to immune checkpoint blockade. *Int Immunol.* 2016; 28:411–9.
<https://doi.org/10.1093/intimm/dxw019>
PMID:[27048318](https://pubmed.ncbi.nlm.nih.gov/27048318/)
60. Genoud V, Marinari E, Nikolaev SI, Castle JC, Bukur V, Dietrich PY, Okada H, Walker PR. Responsiveness to anti-PD-1 and anti-CTLA-4 immune checkpoint blockade in SB28 and GL261 mouse glioma models. *Oncoimmunology.* 2018; 7:e1501137.
<https://doi.org/10.1080/2162402X.2018.1501137>
PMID:[30524896](https://pubmed.ncbi.nlm.nih.gov/30524896/)
61. Zhang N, Wei L, Ye M, Kang C, You H. Treatment Progress of Immune Checkpoint Blockade Therapy for Glioblastoma. *Front Immunol.* 2020; 11:592612.
<https://doi.org/10.3389/fimmu.2020.592612>
PMID:[33329578](https://pubmed.ncbi.nlm.nih.gov/33329578/)

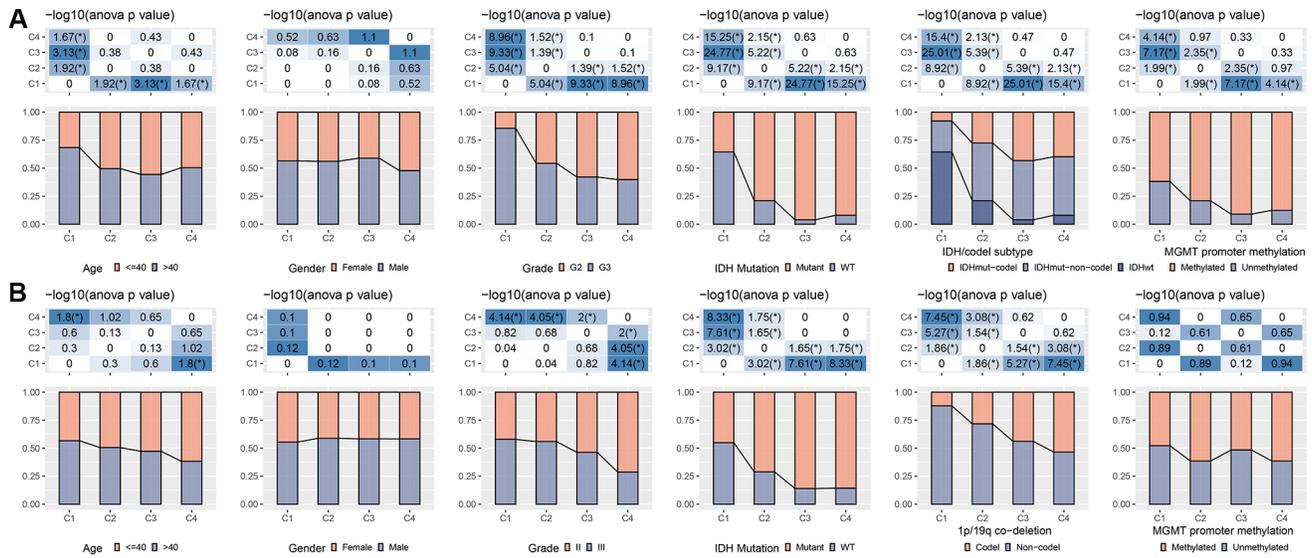
62. Hotta K, Kitamoto T, Kitamoto A, Ogawa Y, Honda Y, Kessoku T, Yoneda M, Imajo K, Tomeno W, Saito S, Nakajima A. Identification of the genomic region under epigenetic regulation during non-alcoholic fatty liver disease progression. *Hepatol Res.* 2018; 48:E320–34.
<https://doi.org/10.1111/hepr.12992>
PMID:[29059699](https://pubmed.ncbi.nlm.nih.gov/29059699/)
63. Lewis ND, Sia CL, Kirwin K, Haupt S, Mahimkar G, Zi T, Xu K, Dooley K, Jang SC, Choi B, Boutin A, Grube A, McCoy C, et al. Exosome Surface Display of IL12 Results in Tumor-Retained Pharmacology with Superior Potency and Limited Systemic Exposure Compared with Recombinant IL12. *Mol Cancer Ther.* 2021; 20:523–34.
<https://doi.org/10.1158/1535-7163.MCT-20-0484>
PMID:[33443094](https://pubmed.ncbi.nlm.nih.gov/33443094/)
64. Zhou Q, Yan X, Zhu H, Xin Z, Zhao J, Shen W, Yin W, Guo Y, Xu H, Zhao M, Liu W, Jiang X, Ren C. Identification of three tumor antigens and immune subtypes for mRNA vaccine development in diffuse glioma. *Theranostics.* 2021; 11:9775–90.
<https://doi.org/10.7150/thno.61677>
PMID:[34815785](https://pubmed.ncbi.nlm.nih.gov/34815785/)
65. Wu ZH, Zhang YJ, Yue JX, Zhou T. Comprehensive Analysis of the Expression and Prognosis for *SFRPs* in Breast Carcinoma. *Cell Transplant.* 2020; 29:963689720962479.
<https://doi.org/10.1177/0963689720962479>
PMID:[32990024](https://pubmed.ncbi.nlm.nih.gov/32990024/)
66. Katzeff JS, Bright F, Lo K, Kril JJ, Connolly A, Crossett B, Ittner LM, Kassiou M, Loy CT, Hodges JR, Piguet O, Kiernan MC, Halliday GM, Kim WS. Altered serum protein levels in frontotemporal dementia and amyotrophic lateral sclerosis indicate calcium and immunity dysregulation. *Sci Rep.* 2020; 10:13741.
<https://doi.org/10.1038/s41598-020-70687-7>
PMID:[32792518](https://pubmed.ncbi.nlm.nih.gov/32792518/)

SUPPLEMENTARY MATERIALS

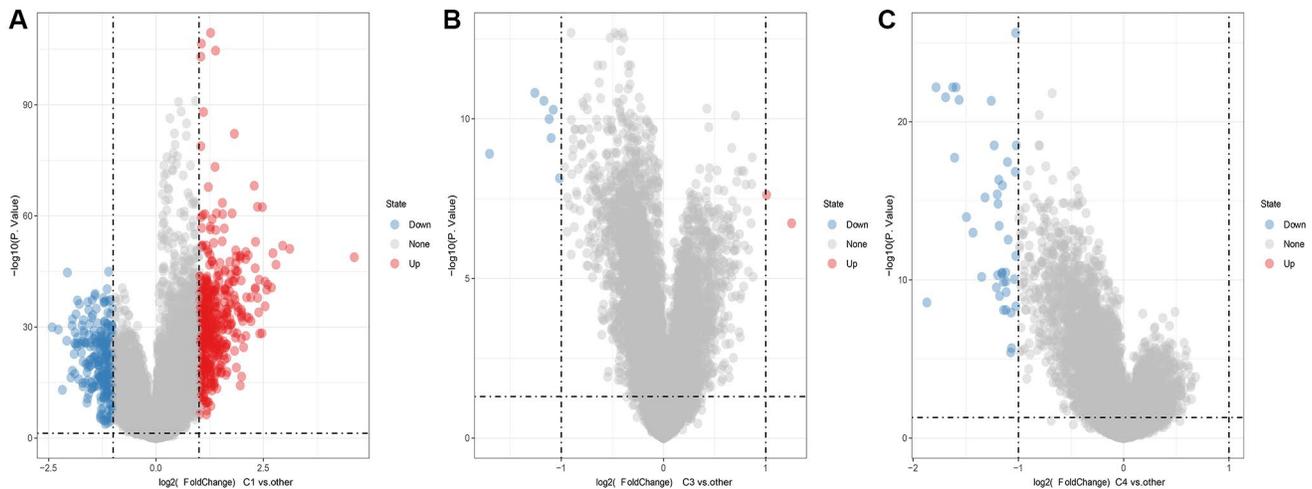
Supplementary Figures



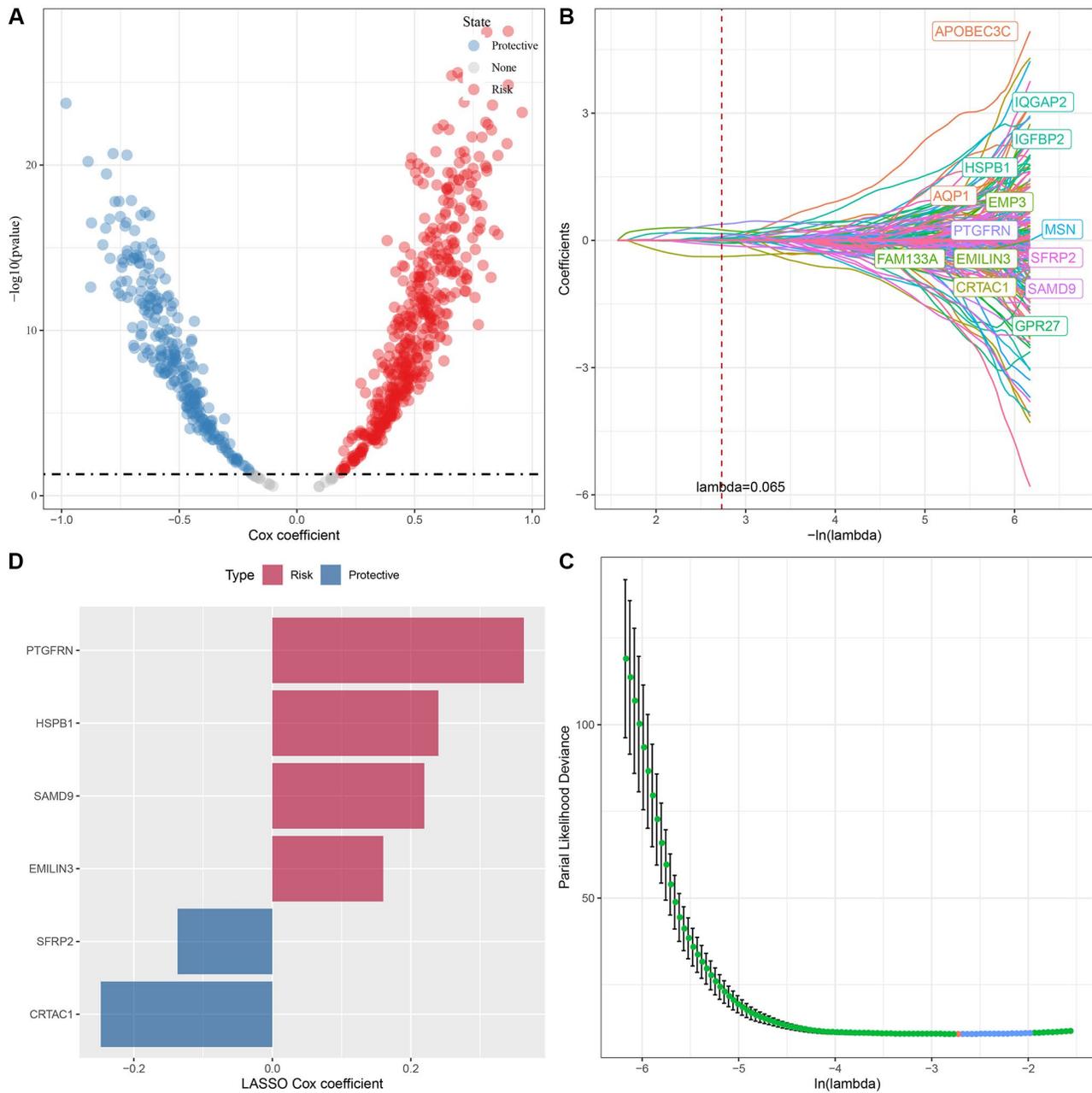
Supplementary Figure 1. Mutation burden in different molecular subtypes. (A) Comparison of TMB, aneuploidy score, homologous recombination defects, intratumor heterogeneity, LOH, purity, and ploidy among different molecular subtypes in the TCGA-LGG cohort. (B) Comparison of the four molecular subtypes with immune molecular subtypes. (C) Somatic mutations in the four molecular subtypes (chi-square test). * $P < 0.05$. ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$. Abbreviations: ns: no significance; ANOVA; analysis of variance; IDH: isocitrate dehydrogenase; WT: wild-type; Mut: mutant; MGMT: O-6-methylguanine-DNA methyltransferase; TMB: tumor mutation burden; LOH: loss of heterozygosity; TCGA: The Cancer Genome Atlas; LGG: low-grade glioma.



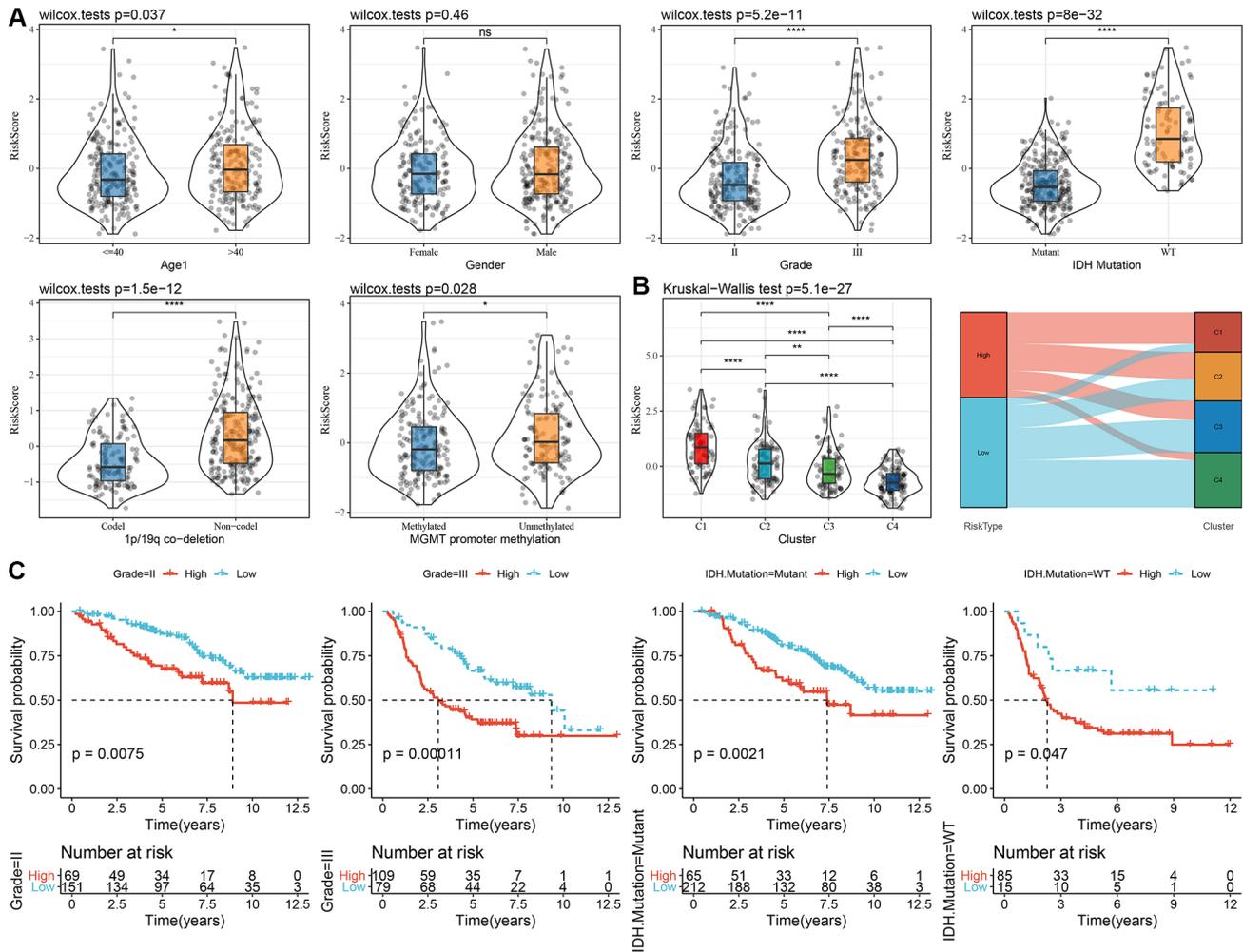
Supplementary Figure 2. Distribution patterns of various molecular subtypes in each clinical variable. (A) Clinicopathological characteristics of molecular subtypes in the TCGA-LGG cohort. **(B)** Clinicopathological characteristics of molecular subtypes in the CGGA cohort; the upper part is the statistical significance of the pairwise distribution difference, the lower part is the proportion. Abbreviations: TMB: tumor mutation burden; ANOVA: analysis of variance; TCGA: The Cancer Genome Atlas; LGG: low-grade glioma; CGGA: Chinese Glioma Genome Atlas.



Supplementary Figure 3. Identification of DEGs among different clusters. (A) C1 vs. other. **(B)** C3 vs. other. **(C)** C4 vs. other. Abbreviation: DEGs: Differentially expressed genes.



Supplementary Figure 4. Construction of a prognostic model with key genes screened by LASSO algorithm. (A) A total of 719 promising candidates were filtered from all the DEGs. **(B)** Changed trajectory of each independent variable with lambda. **(C)** Confidence interval under lambda. **(D)** Distribution of LASSO-Cox coefficients of the NK cell-associated prognostic gene signature. Abbreviations: LASSO: least absolute shrinkage and selection operator; DEGs: differentially expressed genes; NK: natural killer.



Supplementary Figure 5. Association of RiskScore with clinical information in CGGA dataset. (A) Differences in RiskScore among different clinicopathological groups in the CGGA-LGG cohort. (B) Differences in RiskScore among different molecular subtypes in the CGGA-LGG cohort. (C) KM curves between high and low RiskScore groups in different clinical subgroup. Abbreviations: CGGA: Chinese Glioma Genome Atlas; LGG: low-grade glioma; K-M: Kaplan-Meier.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 2–5.

Supplementary Table 1. The clinical information of TCGA and CGGA dataset.

Characteristics	TCGA (N = 506)	CGGA (N = 408)	Total (N = 914)
Age			
≤40	249 (27.24%)	214 (23.41%)	463 (50.66%)
>40	257 (28.12%)	193 (21.12%)	450 (49.23%)
NA	0 (0.0e + 0%)	1 (0.11%)	1 (0.11%)
Gender			
Female	226 (24.73%)	172 (18.82%)	398 (43.54%)
Male	280 (30.63%)	236 (25.82%)	516 (56.46%)
Grade			
G2	245 (26.81%)	220 (24.07%)	465 (50.88%)
G3	260 (28.45%)	188 (20.57%)	448 (49.02%)
NA	1 (0.11%)	0 (0.0e + 0%)	1 (0.11%)
IDH.Mutation			
Mutant	409 (44.75%)	277 (30.31%)	686 (75.05%)
WT	94 (10.28%)	100 (10.94%)	194 (21.23%)
NA	3 (0.33%)	31 (3.39%)	34 (3.72%)
IDH.codel.subtype			
IDHmut-codel	165 (18.05%)	0 (0.0e + 0%)	165 (18.05%)
IDHmut-non-codel	244 (26.70%)	0 (0.0e + 0%)	244 (26.70%)
IDHwt	94 (10.28%)	0 (0.0e + 0%)	94 (10.28%)
NA	3 (0.33%)	408 (44.64%)	411 (44.97%)
MGMT.promoter.methylation			
Methylated	418 (45.73%)	191 (20.90%)	609 (66.63%)
Unmethylated	88 (9.63%)	150 (16.41%)	238 (26.04%)
NA	0 (0.0e + 0%)	67 (7.33%)	67 (7.33%)
1p.19q.co. deletion			
Codel	0 (0.0e + 0%)	131 (14.33%)	131 (14.33%)
Non-codel	0 (0.0e + 0%)	243 (26.59%)	243 (26.59%)
NA	506 (55.36%)	34 (3.72%)	540 (59.08%)

Supplementary Table 2. Identification of 79 genes associated with NK cells.

Supplementary Table 3. The results of univariate cox analysis in the CGGA cohort.

Supplementary Table 4. The results of clustering analysis in the CGGA cohort.

Supplementary Table 5. The information of univariate Cox regression analysis.