



## OPEN ACCESS

## EDITED BY

Anton A. Buzdin,  
European Organisation for Research and  
Treatment of Cancer, Belgium

## REVIEWED BY

Yuhao Zhang,  
Zhejiang Provincial People's Hospital,  
China  
Weiren Luo,  
The Second Affiliated Hospital of  
Southern University of Science and  
Technology, China

## \*CORRESPONDENCE

Shiyin Fu,  
✉ fu\_shiyin@126.com

†These authors have contributed equally  
to this work and share first authorship

## SPECIALTY SECTION

This article was submitted to Cancer  
Genetics and Oncogenomics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 31 October 2022

ACCEPTED 27 March 2023

PUBLISHED 07 April 2023

## CITATION

Hu Y, Yang Q, Cai S, Wang W and Fu S  
(2023), The integrative analysis based on  
super-enhancer related genes for  
predicting different subtypes and  
prognosis of patient with lower-  
grade glioma.

*Front. Genet.* 14:1085584.

doi: 10.3389/fgene.2023.1085584

## COPYRIGHT

© 2023 Hu, Yang, Cai, Wang and Fu. This  
is an open-access article distributed  
under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#).  
The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# The integrative analysis based on super-enhancer related genes for predicting different subtypes and prognosis of patient with lower-grade glioma

Yungang Hu<sup>1†</sup>, Qingqing Yang<sup>2†</sup>, Shuzhou Cai<sup>1</sup>, Wei Wang<sup>1</sup> and Shiyin Fu<sup>3\*</sup>

<sup>1</sup>Department of Neurosurgery, Wuhan University of Science and Technology Affiliated Xiaogan Central Hospital, Xiaogan, Hubei, China, <sup>2</sup>Department of Thyroid and Breast Surgery, Wuhan University of Science and Technology Affiliated Xiaogan Central Hospital, Xiaogan, Hubei, China, <sup>3</sup>Department of Pediatric, Jinchu University of Technology Affiliated Central Hospital, Jingmen, Hubei, China

**Objective:** Emerging evidence revealed that super-enhancer plays a crucial role in the transcriptional reprogramming for many cancers. The purpose aimed to explore how the super-enhancer related genes affects the prognosis and tumor immune microenvironment (TIME) of patients with low-grade glioma (LGG).

**Methods:** In this study, the differentially expressed genes (DEGs) between LGG cohorts and normal brain tissue cohort were identified by the comprehensive analysis of the super-enhancer (SE) related genes. Then non-negative matrix factorization was performed to seek the optimal classification based on the DEGs, while investigating prognostic and clinical differences between different subtypes. Subsequently, a prognostic related signature (SERS) was constructed for the comprehensive evaluation in term of individualized prognosis, clinical characteristics, cancer markers, genomic alterations, and immune microenvironment of patients with LGG.

**Results:** Based on the expression profiles of 170 DEGs, we identified three SE subtypes, and the three subtypes showed significant differences in prognostic, clinicopathological features. Then, nine optimal SE-related genes were selected to construct the SERS through the least absolute shrinkage and selection operator Cox regression analysis. Survival analysis showed that SERS had strong and stable predictive ability for the prognosis of LGG patients in the The Cancer Genome Atlas, China Glioma Genome Atlas, and Remdrandt cohorts, respectively. We also found that SERS was highly correlated with clinicopathological features, tumor immune microenvironment, cancer hallmarks, and genomic alterations in LGG patients. In addition, the predictive power of SERS for immune checkpoint inhibitor treatment is also superior. The qRT-PCR results and immunohistochemical results also confirmed the difference in the expression of four key genes in normal cells and tumors, as well as in normal tissues and tumor tissues.

**Conclusion:** The SERS could be suitable to utilize individualized prognosis prediction and immunotherapy options for LGG patients in clinical application.

## KEYWORDS

super-enhancer, lower-grade glioma, prognostic signature, tumor immune microenvironment, immunotherapy

## 1 Introduction

Gliomas are the most common intracranial malignant tumor, accounting for more than 80% of primary malignant tumor in central nervous system (Ostrom et al., 2022). Low-grade gliomas (LGG), also known as WHO grade II and III tumors defined by the World Health Organization, are composed of diffuse low-grade and intermediate-grade gliomas (Brat et al., 2015). Compared with patients with glioblastoma (GBM), LGG patients have a relatively low degree of malignancy (Chen et al., 2022). However, even with comprehensive treatment including surgical resection, radiotherapy and chemotherapy, some LGG patients still have the characteristics of high recurrence and progression rates (Liu et al., 2018; Jiang et al., 2021). Significant heterogeneity in patient outcomes and treatment response remains a major clinical challenge for neurosurgeons. Traditionally, WHO grade II gliomas were considered to have a better prognosis than WHO grade III gliomas, but since the WHO reclassification of gliomas in 2016, molecular alterations have been considered more objective and precise than grading (Gittleman et al., 2020). Although there have been some progress in the onlooker research on LGG in recent years (Xu et al., 2021a), few drugs are currently approved for the treatment of LGG patients, and the prognosis has not been significantly improved (Ye et al., 2021). Therefore, there is an urgent need to explore new biomarkers to predict the prognosis of LGG patients and find potential therapeutic targets.

Gene regulation plays a major role in tumor pathogenesis, and the regulation of long non-coding RNAs (lncRNAs) on tumors is the hotspot of current research (Lou et al., 2020). Aberrant gene expression promotes tumorigenesis, progression and metastasis (Mansour et al., 2014). Enhancers in gene regulatory elements can bidirectionally transcribe enhancer RNA, a non-coding RNA transcribed by enhancers, that not only drives tumorigenesis, but also regulates genes and immune checkpoints (Lee et al., 2020). Super-enhancers (SE) are clusters of enhancers formed by contiguously arranged enhancers in tandem. SE usually appear near most of the key genes that determine cell identity and function, and play a more effective role than typical enhancers (Hnisz et al., 2013; Whyte et al., 2013). They have the ability to flexibly regulate, by combining unneeded regions to form highly concentrated regional transcriptional machinery, thereby affecting epigenetics and regulating tumorigenesis and progression (Chen et al., 2018). The researchers also found that SE operate covertly in a particularly latent manner, but control across multiple cancer lineages, with cancer cells assembling their own super-enhancers, thereby overproducing malignant oncogenes, exhibiting cancer hallmarks of hyperplasia, invasion and metastasis (Whyte et al., 2013). Yang believed that identifying, mapping out, and disrupting SE has the potential to transform how clinical cancer is managed (Whyte et al., 2013). Hence, as we concentrated on personalized therapy for patients with cancers, SE can serve as the potential biomarkers to track and understand the evolution of individual cancers, and ultimately may become important targets in therapeutic interventions.

The lncRNA HCCL5 in human tissue cells was identified as a SE-driven oncogenic factor that promotes the malignant development of hepatocellular carcinoma by promoting HCC cell viability, migration, and classical epithelial-mesenchymal transition (Peng et al., 2019). TCOF1 depletion in triple-negative breast cancer patients significantly inhibited the growth and invasiveness of triple-negative breast cancer cells (Hu et al., 2022a). Heparanase (HPSE) is a cancer metastasis protein that is regulated by the hnRNP/p300/EGR1/HPSE axis, promotes high expression of HPSE enhancer RNA, is an independent prognostic factor for poor prognosis in cancer patients (Jiao et al., 2018). Regarding whether SE affect progression and overall survival in patients with LGG, the jury is still out.

Therefore, to solve the above problems, this study investigated the effect of different types of SE-related genes on the survival of patients with LGG by collecting data from The Cancer Genome Atlas (TCGA), China Glioma Genome Atlas (CGGA) and Rembrandt Database for LGG. At the same time, we constructed and evaluated prognostic score (SERS) based on 8 SE-related genes for patients with LGG. On top of that, the relationship was also explored between SERS and prognosis, clinicopathological features, tumor immune microenvironment, cancer hallmarks, genomic alterations and immunotherapy efficacy in patients with LGG. We provided a new strategy for predicting the prognosis of and assessing treatment effects for patients with LGG, and thus the findings of this study will help individualized prognosis prediction and immunotherapy decisions in patients with LGG.

## 2 Materials and methods

### 2.1 Data collection and study population

The RNA sequencing data and clinical information of LGG patients were extracted from TCGA (<https://portal.gdc.cancer.gov/>), CGGA (<http://www.cgga.org.cn/>) and Rembrandt (<http://gliovis.bioinfo.cnio.es/>) databases. A total of 5 LGG cohorts were gathered in this study, namely, the TCGA, CGG693, CGGA325, CGGA301 and Rembrandt cohorts, respectively. Patients with no survival data or overall survival (OS) < 30 days were excluded from further analysis. Zakharova et al. (Zakharova et al., 2022) had reclassified the TCGA sampling according to the updated WHO CNS Tumor Classification in 2021, we therefore used the updated glioma diagnoses for analysis in this study. The transcriptome data with normal brain tissue were also obtained from Genotype-Tissue Expression (GTEx; <https://gtexportal.org/home/>). Furthermore, SE-related gene can be downloaded from the SEA v. 3.0 database (<http://sea.edbc.org>). The clinicopathological characteristics of LGG patients in five cohorts are generalized in Table 1.

The differential expression analysis firstly performed based on GTEx dataset and TCGA dataset, and finally 1,672 differentially expressed genes (DEGs) were extracted with the cutoff values of log<sub>2</sub> fold-change |log<sub>2</sub>FC|>2 and *p*-value < 0.05. Then, in the same way, 285 DEGs were extracted between GTEx dataset and TCGA

**TABLE 1** Characteristics of glioma patients in training and validation cohorts.

Clinicopathological characteristics	Training cohort	Validation cohorts			
	TCGA	CGGA693	CGGA325	CGGA325	Rembrandt
Number of patients	331	420	170	158	119
Age (mean $\pm$ SD; years)	41.3 $\pm$ 13.2	40.3 $\pm$ 10.4	40.4 $\pm$ 10.9	39.6 $\pm$ 10.6	NA
Gender					
Female	146	185	65	68	37
Male	185	235	105	90	59
NA	0	0	0	0	0
Survival status					
Alive	272	223	82	85	34
Dead	59	197	88	73	85
Preoperative KPS					
<80	50	NA	NA	NA	NA
$\geq$ 80	93	NA	NA	NA	NA
NA	188	NA	NA	NA	NA
Histology					
Astrocytoma	193	254	110	102	80
Oligoastrocytoma	0	29	0	18	0
Oligodendroglioma	138	137	60	38	34
NA	0	0	0	0	0
WHO grade					
II	179	172	97	105	63
III	152	248	73	53	56
NA	0	0	0	0	0
IDH status					
Mutant	331	288	125	104	NA
Wild type	0	94	44	1	NA
NA	0	38	1	53	NA
1p19q codeletion					
Codeletion	138	125	55	16	8
Non-codeletion	193	257	113	33	13
NA	0	38	2	109	98
MGMT promoter status					
Methylated	271	200	84	43	NA
Unmethylated	60	129	70	106	NA
NA	0	38	16	9	NA
TERT status					
Mutant	116	NA	NA	NA	NA
Wild type	137	NA	NA	NA	NA
NA	78	NA	NA	NA	NA

dataset. Eventually, the differentially expressed SE-related genes were shared by two cohorts were considered eligible.

## 2.2 Identification of SE subtypes of LGG patients

Based on the above DEGs, non-negative matrix factorization (NMF) consensus clustering analysis was performed to obtain the optimal SE subtypes of LGG patients (Hillman et al., 2018). The commonality, dispersion and contour indicators are used to judge the optimal number of subtypes. The t-distributed stochastic neighbor

embedding (tSNE) algorithm we applied to confirm the reliability of clustering results by naked eyes. The Kaplan-Meier survival curves were then used to identify differences in survival difference among different SE subtypes. In addition, we compared differences in clinicopathological features among different SE subtypes.

## 2.3 Construction and validation of a prognostic SERS

The univariate Cox regression was conducted to select the prognostic SE-related DEGs. Then the least absolute shrinkage

and selection operator (LASSO) Cox regression analysis was performed to identify the SE-related prognostic signature (SERS) in the TCGA cohort (Friedman et al., 2010). The prognostic risk score of each LGG patient was calculated with the regression coefficient and the expression of the corresponding gene. The calculation formula of SERS was shown below:

$$\text{Risk score} = \sum_{i=1}^n (\text{Coe}f_i * X_i)$$

where  $n$  represents the number of all the selected gene;  $i$  represents the serial number of each gene;  $X_i$  and  $\text{Coe}f_i$  refer to the expression level of each selected gene and corresponding coefficient, respectively. The cut-off value, defined as the median risk score was divide the patients into high- or low-risk group. The Kaplan-Meier survival curve analysis were conducted to evaluate the accuracy of prognosis of LGG patients between the high- and low-risk groups. The receiver operating characteristic (ROC) curves and the area under the ROC curves (AUC) were plotted and calculated to describe the accuracy of predicting OS. The above analyses were performed simultaneously in the TCGA cohort and four independent validation cohorts. What is more, we finally conducted meta-analysis to calculate the pooled hazard ratio of SERS.

## 2.4 Development of a nomogram

Initially, the univariate Cox regression analysis were performed based on SERS and clinicopathological features, and then multivariate Cox regression analysis was used to identify independent predictors in the TCGA cohort. The nomogram was developed in the TCGA cohort to individually predict 1, 3, and 5-year survival probabilities in LGG patients. And the predicted outcomes for LGG patients were presented in the form of ROC curves. To evaluate the stability of this nomogram, a 10-fold cross-validation algorithm was performed in the TCGA cohort for the internal validation, and the external validation was conducted in the other four independent cohorts. In addition, calibration curves and C-index were performed in the TCGA and validation cohorts to evaluate the usability of this nomogram.

## 2.5 Evaluation of genomic alterations

Tumor mutational burden (TMB) was calculated as the total number of somatic, coding, base substitution, and indel mutations examined per megabase of genome (Mayakonda et al., 2018). The somatic mutation profile ordered in Mutation Annotation Format (MAF) was obtained from the TCGA database. The mutation spectrum and frequency differences were analyzed between high and low risk genes (Bi et al., 2020). In addition, copy number alteration (CNA) data in LGG patients were obtained from the TCGA database. We used GISTIC2.0 to identify significant amplifications or deletions genome-wide. CNA burden was defined as the total number of genes with copy number changes at the focal and arm levels (Shen et al., 2019).

## 2.6 Assessment of TIME and immunotherapeutic responses

For the purpose of better understanding the underlying biological functions of DEGs between high-risk and low-risk groups, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to identify annotated functions and gene enrichment pathways (Hu et al., 2022b). DEGs between high- and low-risk groups were set the cutoff values of  $|\log_2\text{FC}| > 2$  and the BH method adjusted  $p < 0.05$ .

There has been an increasing recognition that the interaction of cancer cells and tumor microenvironment may best be conceptualized as an ecological process (Kenny et al., 2006). Hence, the ESTIMATE algorithm was used to calculate the immune score, stromal score, ESTIMATE score and tumor purity in LGG patients (Yoshihara et al., 2013) for assessing the difference of stromal and immune cells in LGG. Simultaneously, CIBERSORT was performed to calculate the proportions of 22 immune cells from LGG based on gene expression (Newman et al., 2015). In addition, Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was also applied to assess potential response to immune checkpoint inhibitions (ICI) therapy for LGG patients (Jiang et al., 2018).

## 2.7 Quantitative real-time polymerase chain reaction (qRT-PCR) and immunohistochemistry (IHC)

The normal human astrocyte line HA1800 and glioma cell lines U87, U251, A172 and LN229, were purchased from the Cell Bank of the Chinese Academy of Sciences. The clinical specimens of 10 LGG patients were collected in the Department of Neurosurgery of Wuhan Union Hospital from June 2021 to December 2021. Ten non-tumor brain tissues were obtained from patients with brain tissue resection due to craniocerebral injury from June 2021 to December 2021. The study was approved by the Medical Ethics Committee of our hospital, and the informed consent was obtained from each patient. Total RNA was extracted from cell lines and sample tissues using RNAiso Plus (Takara 9109). According to the instruction, cDNA was synthesized by reverse transcription through using HiScript<sup>®</sup> III RT SuperMix for qPCR (+gDNA wiper) (Vazyme R323-01). The qRT-PCR analyses were performed using the AceQ<sup>®</sup> qPCR SYBR Green Master Mix (Vazyme Q111-02) with PCR LightCycler480 (Roche Diagnostics, Basel, Switzerland). All expression data was normalized to GAPDH as an internal control using the  $2^{-\Delta\Delta\text{CT}}$  method. All primers used were synthesized by GeneCreate Biological Engineering Co., Ltd. (Wuhan, China). The protein levels of the selected genes were then verified by IHC experiments. In addition, the relations between the selected gene and tumor immune features also analyzed in LGG patients.

## 2.8 Statistical analysis

The PERL language (version, 5.30.2, <http://www.perl.org>) was used to preprocess RNA-seq transcriptome information. The R software (version 4.0.1, <http://www.R-project.org>) were conducted

for statistical analyses and graph visualization. Continuous variables are described as mean  $\pm$  standard deviation, and categorical variables are described as frequency (n) and proportion (%). Chi-square test or Fisher's exact test was performed to compare categorical variables between two groups. Student's *t*-test or one-way ANOVA was used to compare continuous variables with normal distribution between two or among more groups. The Mann-Whitney *U* test was used to compare non-normally distributed continuous variables between two groups, while the Kruskal Wallis test was used to compare non-normally distributed continuous variables among three or more groups. Survival differences between groups were assessed using Kaplan-Meier curves. Univariate and multivariate cox proportional hazards models were applied to estimate hazard ratios for variables and to identify independent prognostic factors. The cutoff value with statistical significance was set at two-tailed  $p < 0.05$ .

## 3 Results

### 3.1 Overall structure of this study

First of all, the GTEs between LGG and normal brain tissues were screened out. Based on the expression profiles of these selected GTEs, NMF consensus clustering was performed to construct SE subtypes of LGG patients. Then, we explored the heterogeneities of prognosis and clinicopathological features for SE subtypes. Subsequently, the Univariate Cox regression analysis LASSO Cox algorithm were combined to screen for robust SERS and presented as a nomogram. The effectiveness of SERS was assessed in multiple dimensions. The overall flow diagram of this study was presented in [Figure 1](#).

### 3.2 Identification of SE subtypes in TCGA cohort based on the DEGs

The differential expression analysis based on GTEx dataset and TCGA dataset was shown as the volcano in [Figure 2A](#), and differential expression analysis based on GTEx dataset and CGGA693 dataset was also shown as the volcano in [Figure 2B](#). Then, a total of 170 DEGs ([Figure 2C](#)) shared by two cohorts were used for subsequent analysis, they can be found in [Supplementary Table S1](#).

Based on the expression profiles of 170 SE associated DEGs, the NMF was performed in the TCGA cohort to identify SE subtypes. As shown in [Figure 2D](#), we chose 3 as the optimal number of clusters based on common, scatter, and contour metrics. Then, a total of 469 LGG patients were divided into three subtypes ([Figure 2E](#)), named SE1 ( $n = 125$ ), SE2 ( $n = 75$ ), and SE3 ( $n = 125$ ). The heatmap of the consensus matrix exhibits clear boundaries, indicating the accuracy and robustness of the clustering results. t-SNE plot showed clear differences in the distribution between the three SE subtypes ([Figure 2F](#)). Significant differences in the expression of 170 prognostic SE-related DEGs can also be observed in the heatmap in [Figure 2G](#). Kaplan-Meier survival curves showed obvious survival differences among the three SE subtypes ([Figure 2H](#)). The LGG patients in SE1 subtype had the best survival outcome, while SE2 had the worst survival outcome. At

the same time, the heterogeneity of clinicopathological characteristics of these three subtypes were analyzed and found interestingly no significant differences among these clinicopathological characteristics ([Supplementary Figure S1](#)).

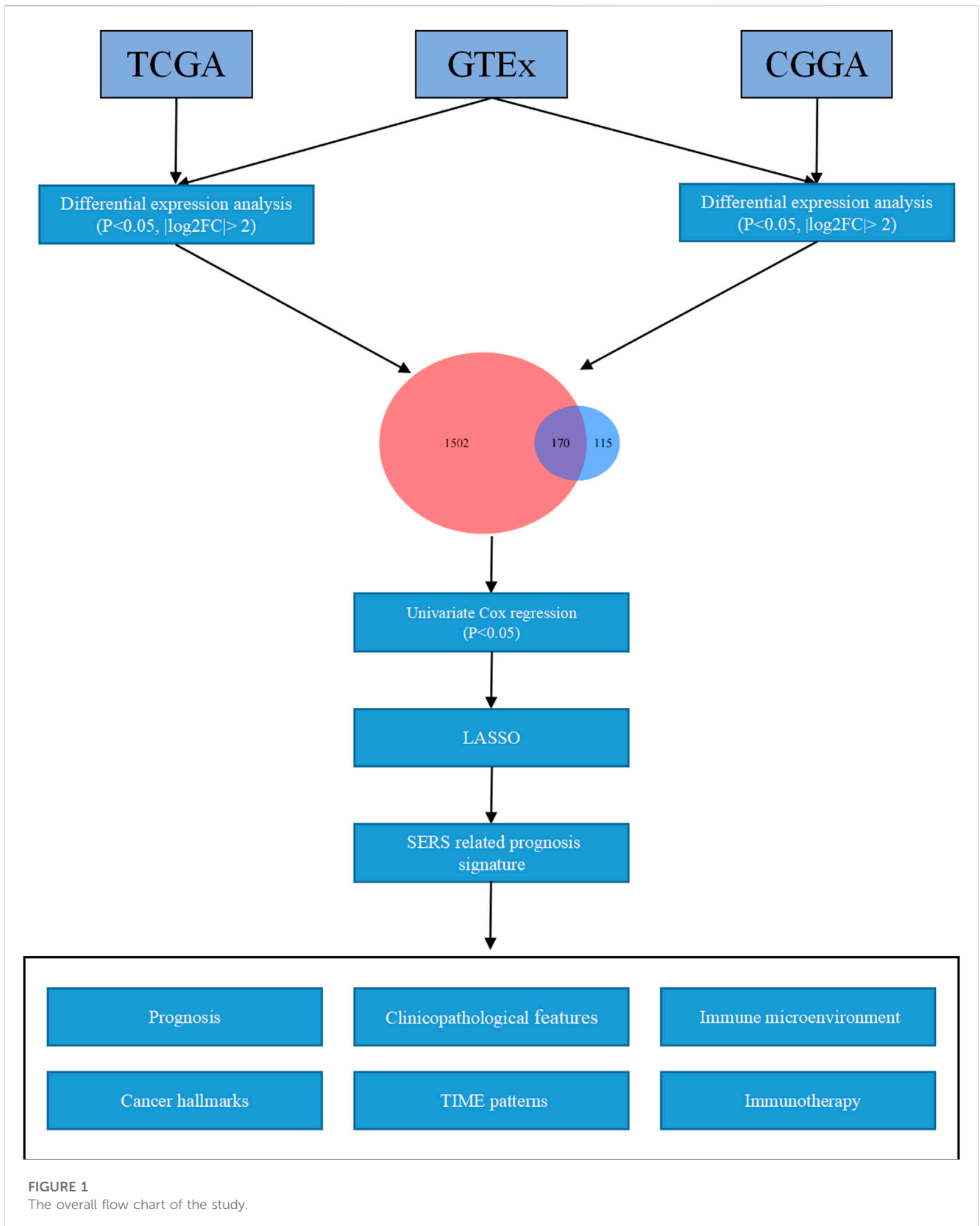
### 3.3 Development and validation of the SERS

The Univariate Cox regression analyses were conducted based on DEGs to identify prognostic SERG. The results of the analysis indicated 33 genes were obviously related to the OS of LGG, and detailed information for these prognostic SERG was shown in [Supplementary Table S2](#). Then, the LASSO analysis was performed on above 33 prognostic SERG in the TCGA cohort to explore simplest and most accurate model. Finally, a total of 9 optimal prognostic SERG (AQP7, MYOD1, CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2) were stood out and incorporated in the SERS ([Figures 3A, B](#)). [Figure 3C](#) exhibited the LASSO coefficients of each selected gene in this signature. Among them, the coefficients of seven genes (CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2) are positive number, which means that they are related to bad prognosis for LGG patients, whereas the coefficients AQP7 and MYOD1 are negative number, indicating a good prognosis. The Kaplan-Meier survival curves of these nine optimal genes were shown in [Supplementary Figure S2](#). The risk score of each patient was calculated as follows: SERS score =  $(-0.408 \times \text{AQP7}) + (-0.107 \times \text{MYOD1}) + 0.186 \times \text{CDCA2} + 0.625 \times \text{FAM92B} + 0.163 \times \text{HOXA11-AS} + 0.454 \times \text{E2F7} + 0.344 \times \text{KIF18A} + 0.130 \times \text{MC5R} + 0.201 \times \text{SKOR2}$ . Subsequently, the median SERS score in was set as the cut-off value to stratified the 325 LGG patients into the high- and low-risk groups. Heatmap analysis of nine genes showed markedly different distributions between high- and low-risk groups, the risky genes were upregulated in the high-risk group and the protective genes were upregulated in the low-risk group ([Figure 3D](#)).

The SERS was calculated with LASSO coefficients obtained from the TCGA cohort to stratified into with the median score high- and low-risk groups in other 4 cohorts. The Kaplan-Meier survival curves demonstrated that patients with high-SERS showed worse OS than low-SERS in the TCGA cohort (log-rank test  $p < 0.001$ ; [Figure 3E](#)). Consistent results were also observed in four other independent validation cohorts (log-rank test  $p < 0.001$ ; [Figures 3F-I](#)). The distribution plot of the risk score and survival status showed that the SERS had the positively correlation with the deaths of LGG patients ([Supplementary Figures S3A-E](#)). Furthermore, the ROC curves confirmed the satisfactory predictive performance of the of SERS in predicting 1-, 3-, and 5-year OS ([Supplementary Figures S3F-I](#)). Thus, SERS were sufficiently discriminative on both the validation cohorts. In addition, a meta-analysis was performed to assess the overall predicting accuracy, and the results indicated that the overall pooled HR for SERS was 3.2 (95% CI = 1.69–6.08; [Figure 3J](#)).

### 3.4 Relationship between SERS and clinicopathological characteristics, genomic alterations

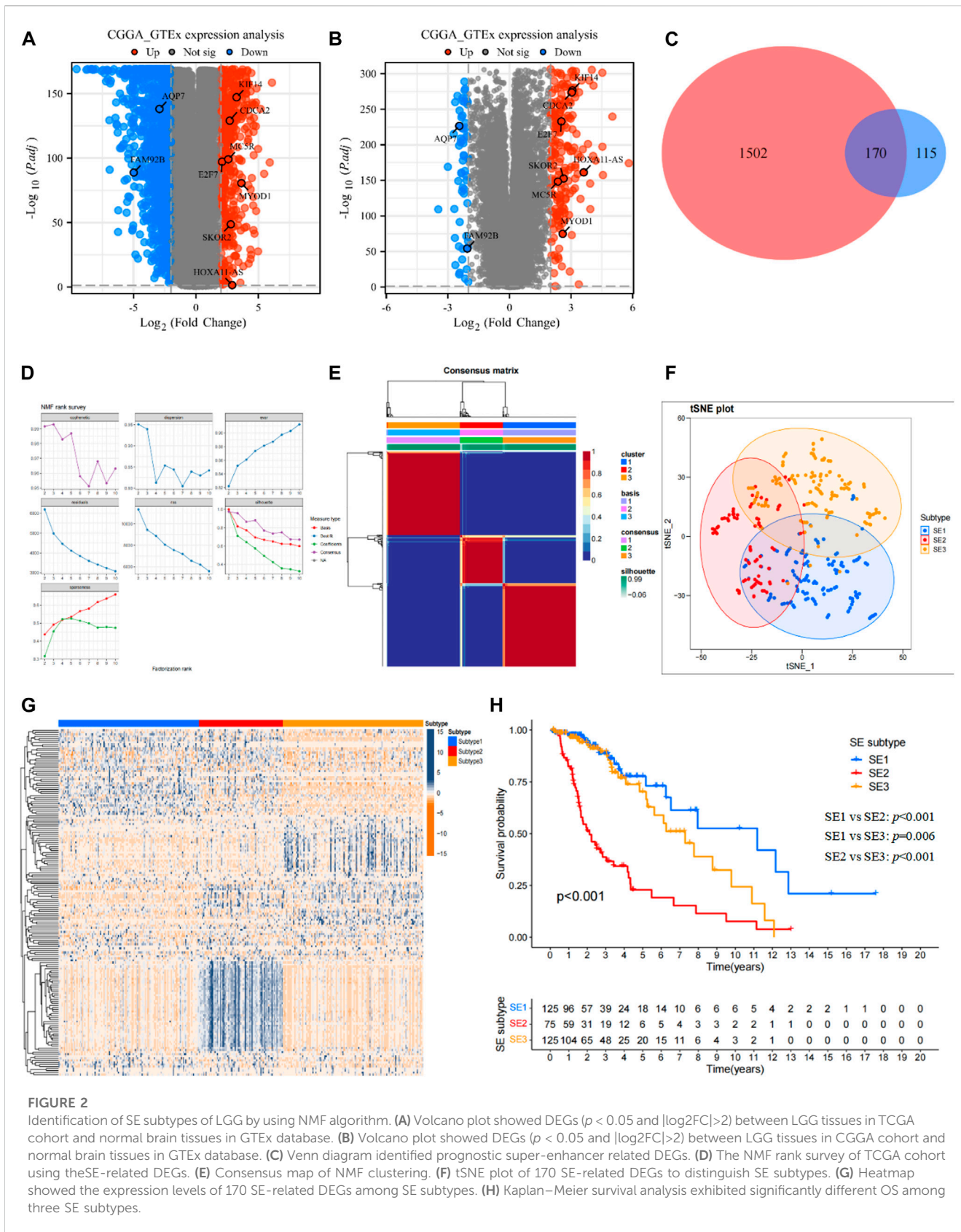
The clinical relationships of SERS were attempted to explore in the TCGA cohort. As shown in [Figure 4A](#), SERS were arranged



**FIGURE 1**  
The overall flow chart of the study.

from low to high to show the correlation between SERS and clinicopathological characteristics. There were significant differences in, survival status, Histology, 1p19q status, and SE

subtype between high and low SERS groups, but no significant differences in -age, gender, WHO grade, MGMT status, TERT status and Transcriptome subtype. Furthermore, SERS levels



between LGG patients stratified by various clinicopathological features were compared. The results of the analyses showed that LGG patients with, death status, Oligodendroglioma and

SE2 subtype showed significant higher SERS, while no significant differences of SERS were observed in other subgroups (Figures 4B–K). Likewise, the relationship between

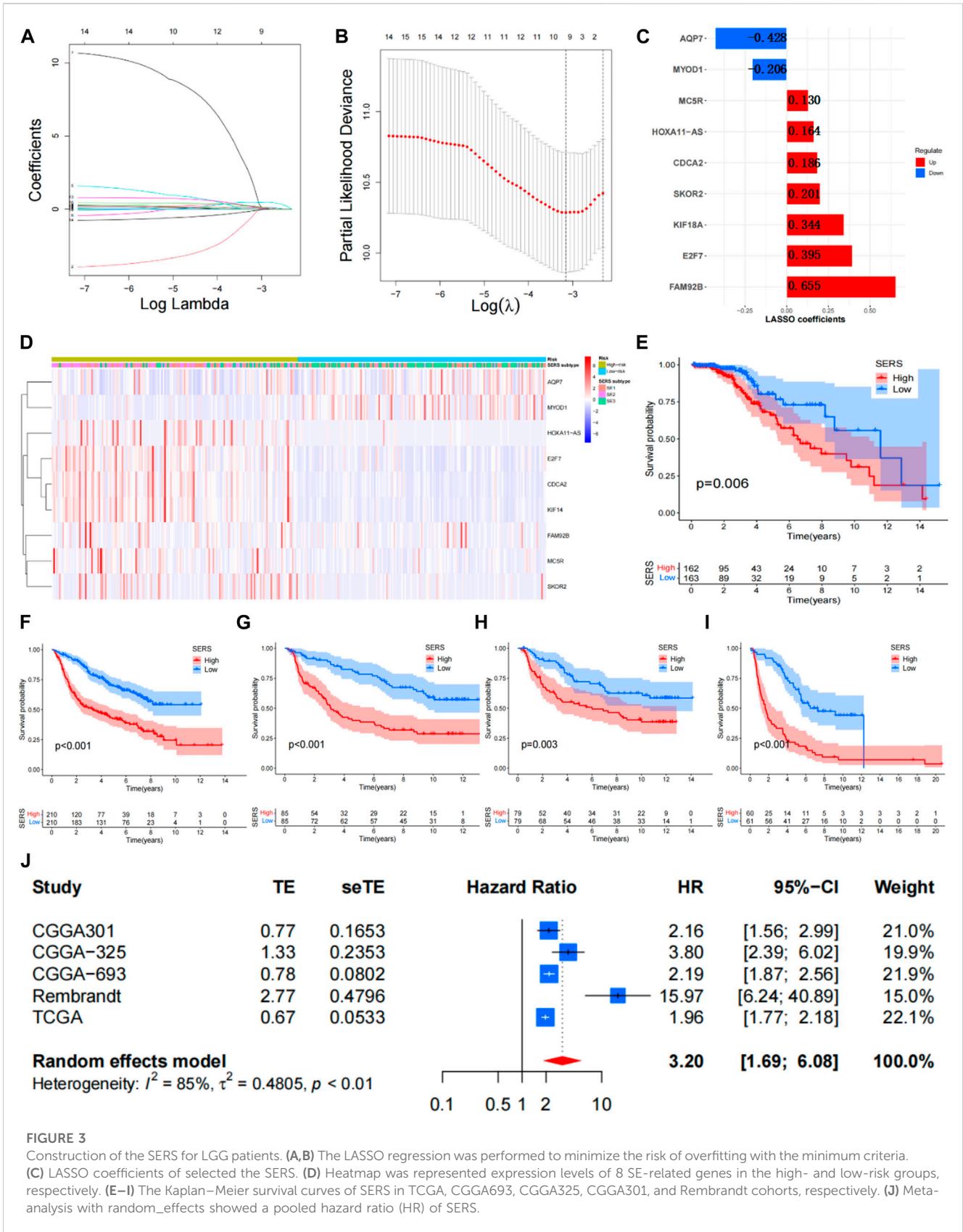


FIGURE 3

Construction of the SERS for LGG patients. (A,B) The LASSO regression was performed to minimize the risk of overfitting with the minimum criteria. (C) LASSO coefficients of selected the SERS. (D) Heatmap was represented expression levels of 8 SE-related genes in the high- and low-risk groups, respectively. (E–I) The Kaplan–Meier survival curves of SERS in TCGA, CGGA693, CGGA325, CGGA301, and Rembrandt cohorts, respectively. (J) Meta-analysis with random\_effects showed a pooled hazard ratio (HR) of SERS.



SERS and clinicopathological characteristics of LGG patients in the CGGA693, CGGA325, CGGA301 and Rembrandt cohorts was also identified the similar results to the TCGA cohort (Supplementary Figures S4–S7).

To better address the prognostic features associated with SERS, the correlation between common cancer markers and SERS were also explored. The correlation heatmap showed that SERS was significantly positively correlated with many well-known cancer hallmarks including DNA repair, cell cycle, hypoxia, and metabolism (Figure 5A). The correlation between SERS and 29 immune signatures was illustrated by a correlation heatmap in TCGA cohort (Figure 5B). Subsequently, further analysis showed that SERS was significantly positively associated with TMB, mutation count, copy number gain and loss burden at the focal level, and copy number gain burden at the arm level (Figures 5C–H). The distribution of TMB, mutation counts, copy number burdens at focal and arm levels between high and low-risk groups were also compared in TCGA cohort (Supplementary Figures S10A–F). Based on the above data, it is indicated that high SERS may represent a higher frequency of genomic alterations to some extent.

### 3.5 Establishment and evaluation of a nomogram

The univariate Cox regression and multivariate Cox regression analyses were performed to identify independent prognostic factors in the TCGA (Figures 6A, B), CGGA693, CGGA325, CGGA301, and Rembrandt cohorts (Supplementary Figures S8A–H). As we expected, SERS including nine selected genes was confirmed as an independent prognostic factor in all cohorts. The nomogram was established to predict 1-, 3-, and 5-year survival time in LGG patients based on the independent prognostic factors (age, WHO grade, and SERS) identified in the TCGA cohort (Figure 6C). The nomogram was firstly internally assessed, and the C-index was 0.862 (95% CI: 0.811–0.896), 0.833 (95% CI: 0.786–0.896), 0.812 (95% CI: 0.761–0.856) at 1, 3, and 5 years, respectively. The 1-year, 3-year and 5-year ROC curves showed that compared with SERS or age, the nomogram had the highest AUC values with the 1-year, 3-year and 5-year AUC values were 0.911, 0.913, and 0.812, respectively, which indicating that the nomogram had the optimal prediction effect (Figures 6D–F). The calibration curves showed a good fit between the actual and nomogram-predicted results for 1-, 3-, and 5-year OS (Figure 6G). In the same way, external validation of this nomogram was performed in the CGGA693, CGGA325, CGGA301 and Rembrandt cohorts. The accuracy in predicting 1-, 3-, and 5-year survival was good, and calibration curve analysis showed that the predicted and actual outcomes were basically conformity in all 4 cohorts (Supplementary Figures S9A–P). Therefore, this nomogram has potential as a quantitative predictor of prognosis in LGG patients.

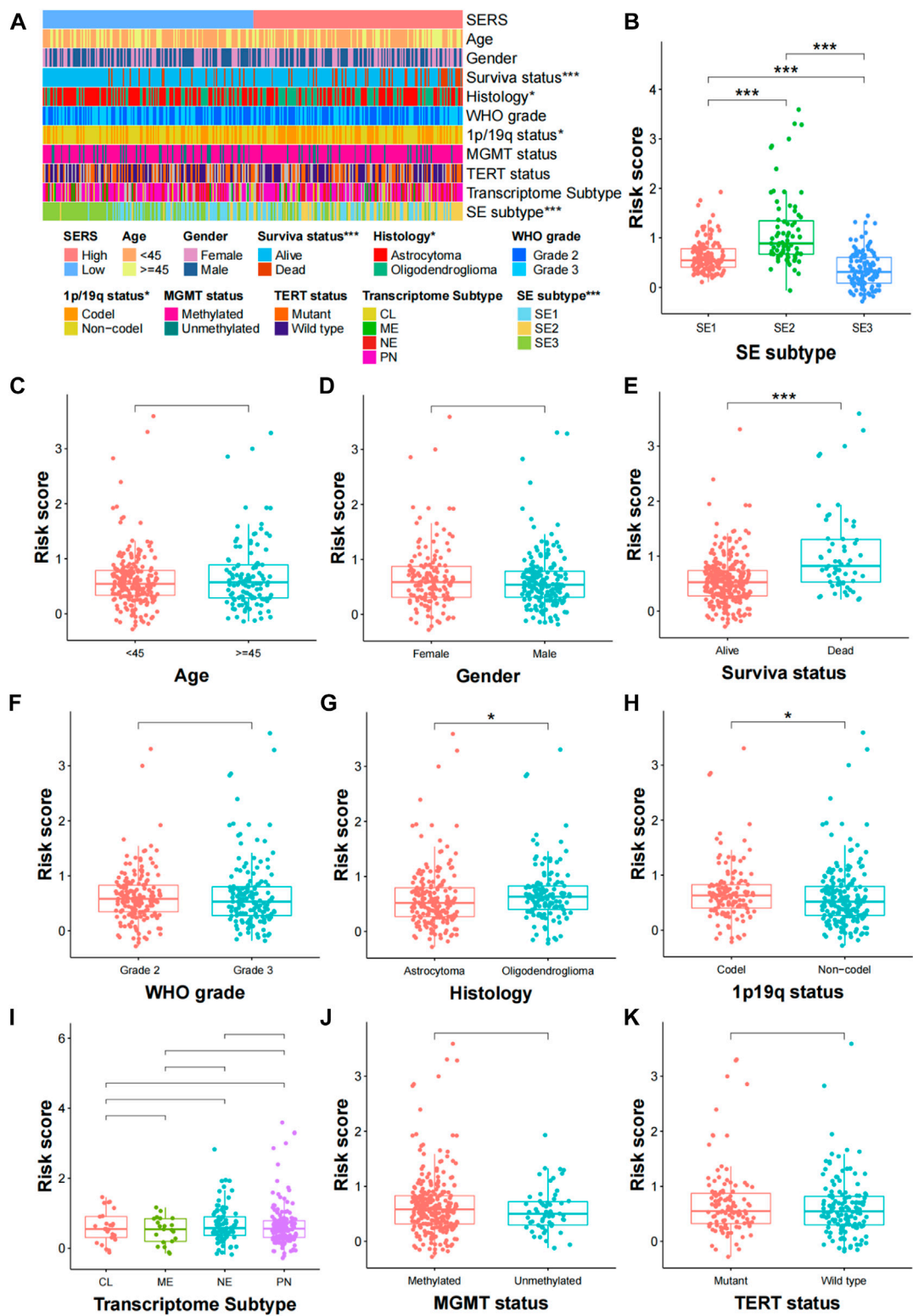
### 3.6 Correlation of SERS with the LGG immune microenvironment and immunotherapy

Based on the differential expression analysis of high and low risk groups in the TCGA cohort, there were 462 DEGs ( $|\log_2FC| > 2$  and the BH method adjusted  $p < 0.05$ .) We then further performed

functional enrichment analysis to characterize the biological functions of DEGs between the two risk subgroups. The results of GO analysis revealed that DEGs are enriched in several immune-related biological processes, such as regulation of T cell activation, positive regulation of T cell activation, negative regulation of immune system process, and positive regulation of lymphocyte activation (Figure 7A). Following, KEGG pathway analysis also showed significant enrichment of immune-related pathways, including cytokine-cytokine receptor interactions and chemokine signaling pathways (Figure 7B).

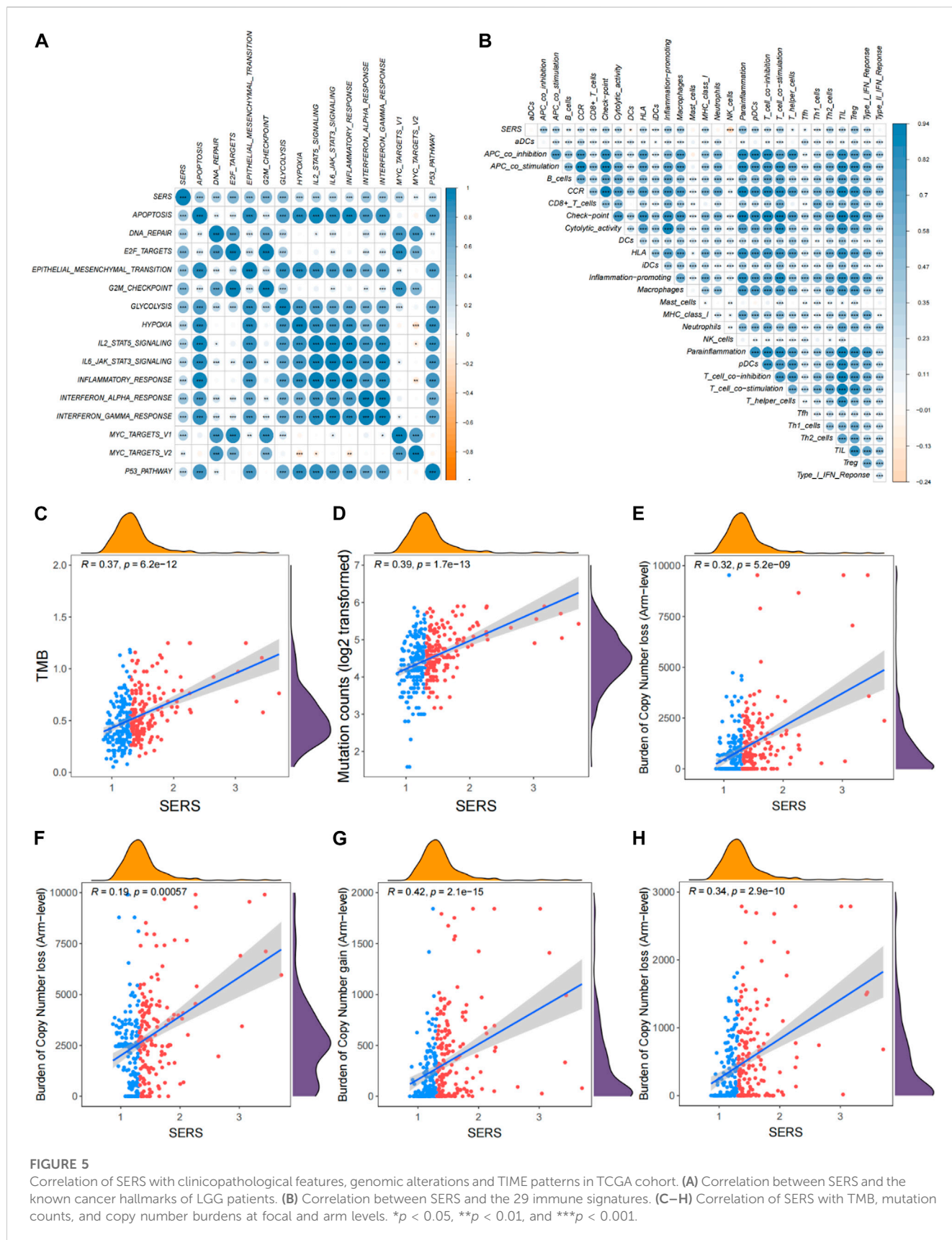
Given the findings that DEGs are enriched in immune-related functions, we further investigated the correlation of SERS with the immune microenvironment of LGG in the TCGA cohort. It turned out that SERS was significantly positively correlated with immune, stromal, and ESTIMATE scores, but negatively correlated with tumor purity, suggesting that the infiltration levels of immune cells and stromal cells increased with SERS (Figures 7C–F). The distribution of ESTIMATE scores, immune scores, stromal scores and tumor purity were no significant between high and low-risk groups in TCGA cohort (Supplementary Figures S10G–J). Further correlation analysis was performed between SERS and the infiltration levels of 22 immune cells quantified by the CIBERSORT algorithm. The results showed that the expressions of cells CD4 memory resting, T cells CD4 memory activated, NK cells activated, Monocytes, Macrophages M1, Mast cells activated and Neutrophils were significantly different in high and low risk groups. Among of them, the abundance of T cells CD4 memory resting, Macrophages M1, and Neutrophils was lower in the high-risk group, but the abundance of T cells CD4 memory activated, NK cells activated, Mast cells activated and Monocytes was higher in the high-risk group (Figure 7G).

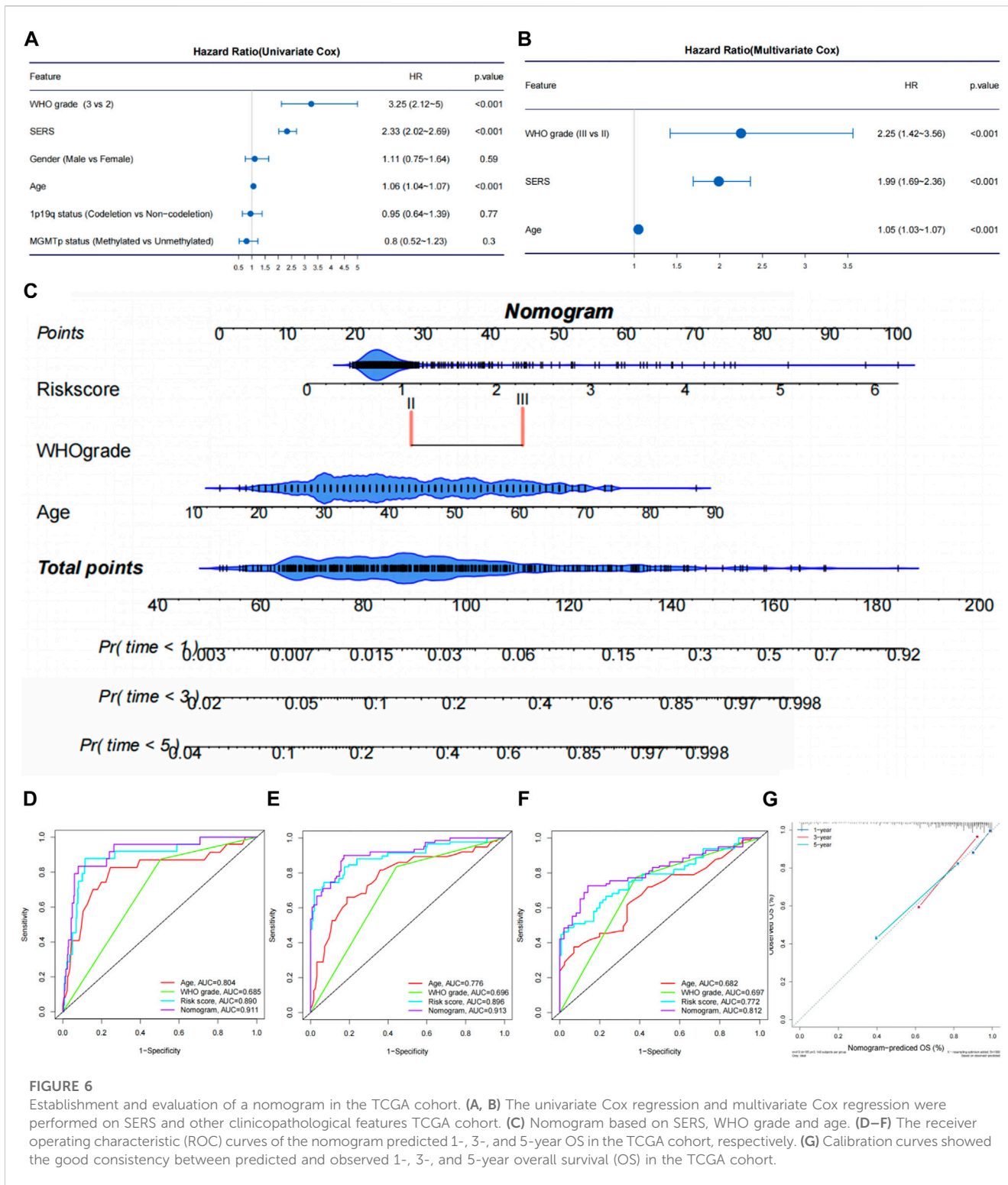
In addition, we evaluated the correlation of SERS with immune checkpoints (PD-1, PD-L1, LAG-3, and B7-H3) and macrophage-associated molecules (CCL2, CCR2, CXCR4, and CSF1). The results showed that all immune checkpoints and macrophage-associated molecules were upregulated in the high-risk group except for LAG-3 (Figure 8A). We next determined whether there is a correlation between immune checkpoints and prognostic SERGs. The heat map showed that immune checkpoint proteins were significantly positively correlated with CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2, and significantly negatively correlated with AQP7 and MYOD1 (Figure 8B). In the TCGA cohort, SERS was positively correlated with TIDE and T-cell exclusion score, and negatively correlated with MSI score and T-cell dysfunction score (Figure 8C). The distribution difference can also be clearly observed in the high and low risk groups (Supplementary Figures S10K–N). In view of the TIDE algorithm, the distribution of SERS for the non-responder and responder groups to ICI indicated that the non-responder group had a significantly higher SERS, which just happened to explain the poorer prognosis of LGG patients who did not respond to ICI (Figure 8D). The high SERS subgroup had a lower proportion of responders to ICI treatment compared with the low SERS subgroup ( $p < 0.05$ , Figure 8E). Figure 8F showed that SERS had a satisfactory prediction in immunotherapy effect, which can provide a reference for whether patients should undergo immunotherapy.



**FIGURE 4**

Correlation analysis between the prognostic SERS and clinicopathological characteristics in the TCGA cohort. (A) A heatmap was represented expression levels of eight selected SERS and the distribution of clinicopathological characteristics in the high- and low-risk groups, respectively. (B–K) Different levels of risk scores in glioma patients stratified by age, gender, Survival status, WHO grade, Histology, 1p19q codeletion, MGMT status, SERS subtype, TERT status and Transcriptome subtype. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.



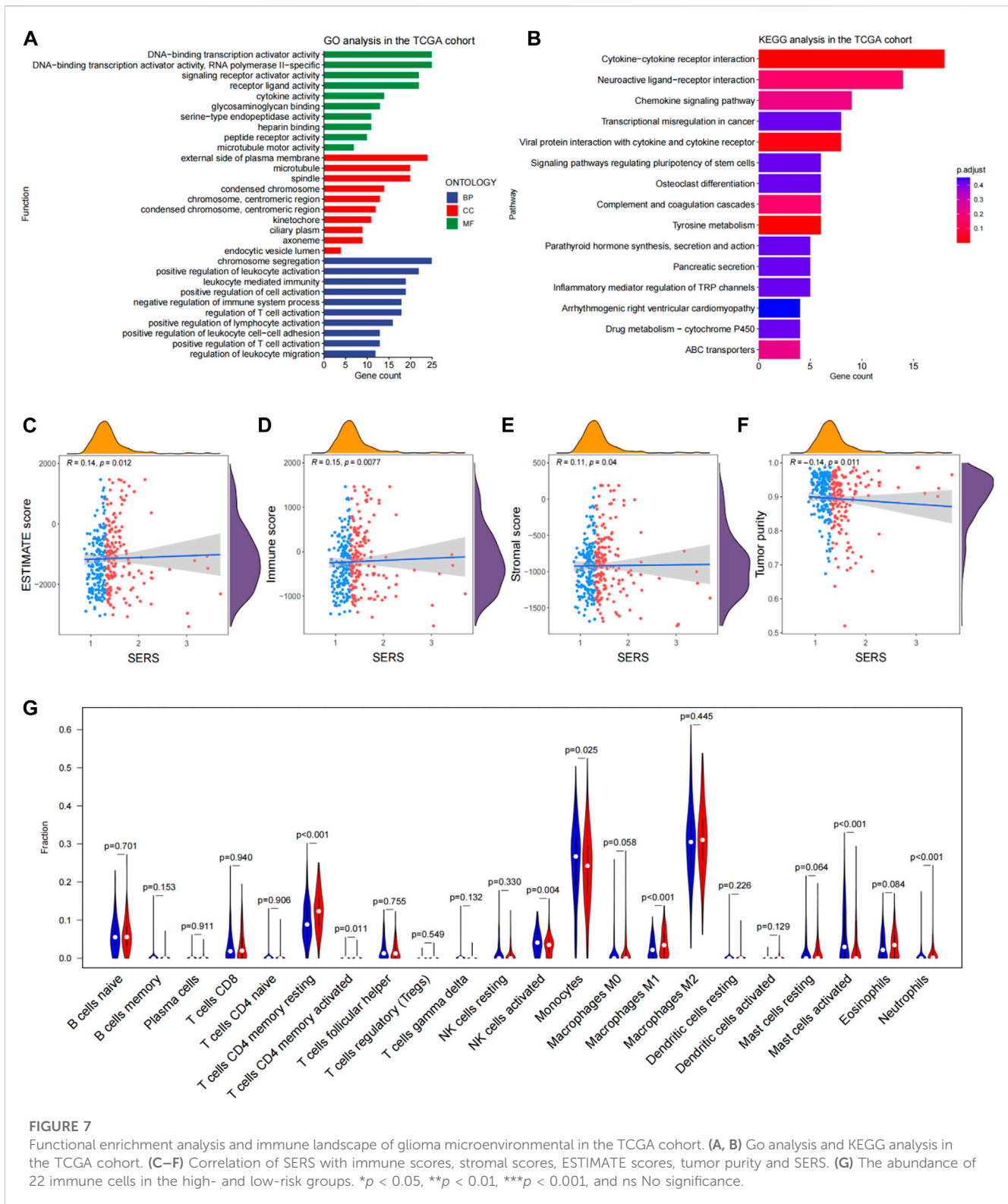


**FIGURE 6** Establishment and evaluation of a nomogram in the TCGA cohort. (A, B) The univariate Cox regression and multivariate Cox regression were performed on SERS and other clinicopathological features TCGA cohort. (C) Nomogram based on SERS, WHO grade and age. (D–F) The receiver operating characteristic (ROC) curves of the nomogram predicted 1-, 3-, and 5-year OS in the TCGA cohort, respectively. (G) Calibration curves showed the good consistency between predicted and observed 1-, 3-, and 5-year overall survival (OS) in the TCGA cohort.

### 3.7 The expression levels of selected SE-related genes

Two SE-related genes (AQP7, and E2F7) were selected to detect their transcriptional levels in cell lines, LGG tissues and normal brain tissues. The qRT-PCR results showed that compared with HA1800, the mRNA expression levels of AQP7 in human glioma cell

lines were generally decreased, while the mRNA expression levels of E2F7 were generally increased (Figure 9A). Subsequently, we also detected their expression levels in 10 normal brain tissues and 10 glioma tissues. The qRT-PCR results of the tissue samples were consistent with those of the cell lines (Figure 9B). The representative IHC staining images of AQP7 and E2F7 were shown in Figure 9C.



**FIGURE 7**

Functional enrichment analysis and immune landscape of glioma microenvironmental in the TCGA cohort. (A, B) Go analysis and KEGG analysis in the TCGA cohort. (C–F) Correlation of SERS with immune scores, stromal scores, ESTIMATE scores, tumor purity and SERS. (G) The abundance of 22 immune cells in the high- and low-risk groups. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.

In addition, the relations between the selected gene (AQP7 and E2F7) and tumor immune features also analyzed. It turned out that AQP7 was significantly positively correlated with immune, stromal, and ESTIMATE scores, but negatively correlated with tumor purity (Figure 10A). While E2F7 was not significantly related with immune, stromal, and ESTIMATE

scores, and tumor purity (Figure 10C). Further correlation analysis was also performed between the selected gene and the infiltration levels of 22 immune cells. The results showed that the expressions of 22 immune cells were significantly different whatever in high and low expression of AQP7 or E2F7 groups (Figures 10B, D).

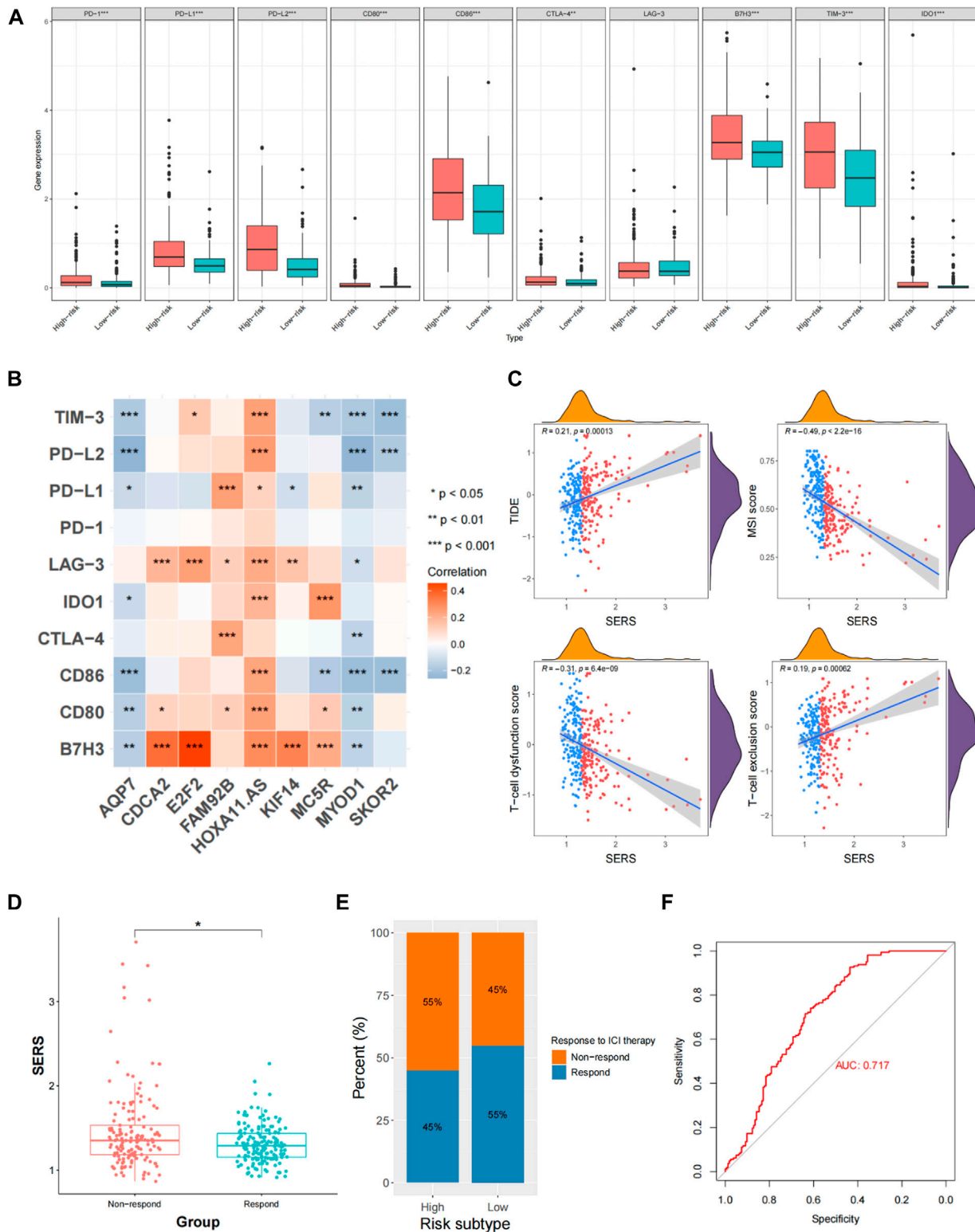


FIGURE 8

Evaluation of immune checkpoints and immunotherapy responsiveness in the TCGA cohort. (A) The expression levels of immune checkpoints and macrophage associated molecules in the high- and low-risk groups. (B) Correlation analysis between the prognostic SERS and immune checkpoints. (C) Correlation of SERS with T-cell dysfunction score, TIDE, MSI score and T-cell exclusion score. (D) The distributions of risk scores between non-respond and respond groups. (E) Comparative analysis of the response rates to ICI treatment in the high- and low-risk groups. (F) The ROC curve of predicting immunotherapeutic benefit.

## 4 Discussion

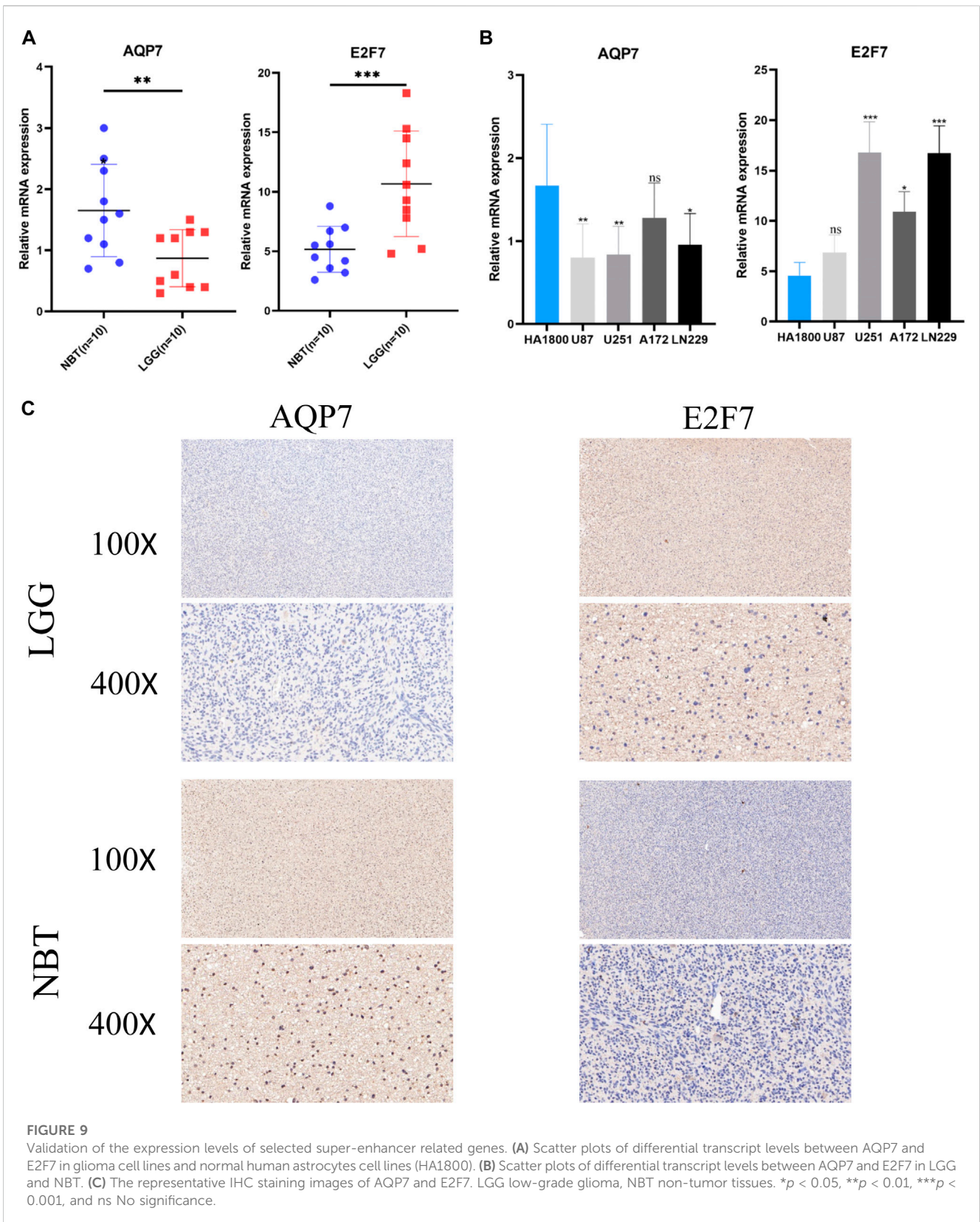
LGG patients, with the better prognosis than GBM patients, account for about half of all glioma patients. But their survival time varies widely, ranging from 1 year to more than 10 years (Chen et al., 2022), notwithstanding LGG patients with the same WHO grade had same standardized sequential therapy including surgery, radiotherapy, and chemotherapy. The high heterogeneity of LGG, which results in inconsistent treatment effects and prognosis, is a clinical conundrum faced by most neurosurgeons. And so, there is an urgent need to develop accurate and robust prognostic prediction models for data-assisted clinical decision-making. With the rapid development of bioinformatics and sequencing technologies, some studies have reported gene markers as prognostic indicators to predict the prognosis of LGG, such as hypoxia-related genes (Dao et al., 2018), ferroptosis-related genes (Wan et al., 2021), immune-related genes (Zhou et al., 2018) and the corresponding lncRNAs. Compared with other biomarkers, SE, as important distal regulatory DNA elements, are direct drivers of carcinogenesis and are highly tissue-specific. Therefore, SE are good candidates for predicting prognosis in various cancers. Ropri et al. (2021) and Huang et al. (2022) performed mechanistic exploration and prognostic prediction in breast cancer and hepatocellular carcinoma, respectively. However, whether super-enhancer related genes can serve as prognostic markers for LGG needs further discussion.

In this study, the NMF algorithm was used to identify three LGG subtypes in 325 LGG patients based on the expression profiles of DEGs between LGG and NBT. Then, significant differences in prognostic, clinicopathological features of the three LGG subtypes were observed with the naked eyes. A prognostic signature, called SERS, was constructed by univariate Cox regression and LASSO Cox regression for an individualized comprehensive assessment. The results showed that SERS was significantly associated with the prognosis, clinicopathological features, genomic alterations and TIME pattern of LGG patients, and the predictive ability of SERS for ICI treatment was also outstanding. In addition, a clinically accessible nomogram was constructed based on SERS, age, and WHO classification, which maintained excellent predictive accuracy in both the internal cohort and 4 external cohorts (CGGA693, CGGA325, CGGA301, and Rembrandt). So, it can provide a good net clinical benefit for screening LGG patients at high risk of death.

The SERS was constructed on 9 SE-related genes in our study, incorporating AQP7, MYOD1, CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2. Among these genes, CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2 were risky genes, which are associated with poor prognosis for LGG patients. Whereas the remaining two genes with good prognosis. Conversely, AQP7 and MYOD1 are related to good prognosis. AQP7, named Aquaporin 7, is a water and glycerol channel. Chen et al. demonstrated that low expression of AQP7 correlates with tumor grade and aggressive features of hepatocellular carcinoma (Chen et al., 2016). In a mouse model of breast cancer, lower AQP7 expression resulted in a reduction in primary tumor burden and lung metastases, thus suggesting that AQP7 is a prognostic indicator of overall survival in breast cancer patients (Dai et al., 2020). Myogenic differentiation 1 (MYOD1), as a transcription factor, promoted expression of muscle-specific genes. Wu et al. (2020) found that the expression of MYOD1 is positively correlated with the migration and invasion of gastric cancer cells.

The cell division cycle associated 2 (CDCA2) has been proved to play an important role in the tumorigenesis of some cancers. The study showed that the high expression of CDCA2 was significantly correlated with the expression of related components of cell cycle phase transition and G2/M phase transition pathway, and suggested that CDCA2 could be a potential target for regulating tumor growth and radiation resistance in patients with oesophageal square cell carcinoma (Xu et al., 2021b). FAM92B, HOXA11-AS, E2F7, MC5R, and SKOR2 are important epigenetic regulators that can be targeted for cancer therapy. Specifically, E2F7 is an atypical E2F transcription factor family member with two independent DNA-binding domains. Some studies have found that E2F7 is upregulated in endometrial cancer, skin squamous cell carcinoma and other malignant tumors, promote tumor progression and metastasis in these cancers (Endo-Munoz et al., 2009; Li et al., 2015). KIF18A, a member of the kinesin-8 subfamily, has low expression in most human normal tissues and abnormally high expression in a variety of malignant tumor tissues (Marquis et al., 2021), which is associated with malignant pathological features and poor prognosis of cancer patients, and it promotes the proliferation, invasion and metastasis of tumor cells (Sepaniac et al., 2021). KIF18A may be a novel molecular targeted therapy for cancers. PTCRA (pre-T cell antigen receptor) is a protein-coding gene, together with the TCRB and CD3 complexes, encodes a protein that forms the T-cell pre-receptor complex, which regulates early T cell development (Liu et al., 2010).

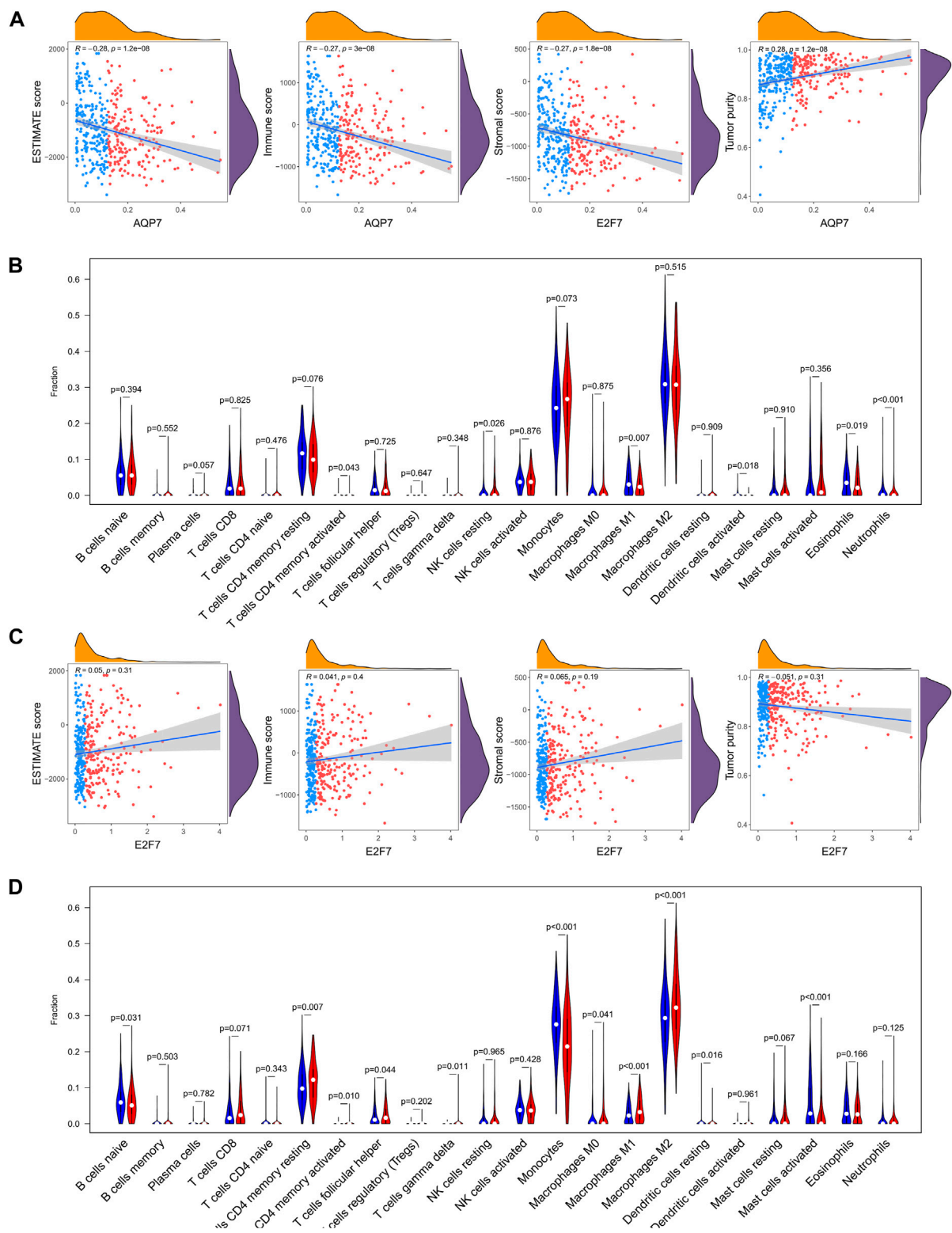
The SERS is very effective in predicting prognosis of LGG patients in this study. To further explore the specific mechanism, we identified DEGs in high- and low-risk groups without hesitation. Then, the GO and KEGG analysis were performed to explore the detailed biological processes and pathways of these genes affecting the prognosis of LGG patients. Functional enrichment analysis revealed that DEGs between different risk subgroups were enriched in many immune-related biological processes and pathways. Therefore, we subsequently analyzed immune scores and immune cell infiltration between the two risk subgroups. Further analysis found that high risk was positively correlated with immune score, the abundance of T cells CD4 memory resting, and T cells CD4 memory activated. On the contrary, activated NK cells (tumor killer cells) showed higher abundance in the low-risk group. These results suggest that super-enhancer-related genes are related to the LGG immune microenvironment to a certain extent. From the above results, it can be concluded that the anti-tumor immunity of LGG patients in the high-risk group is significantly weakened, so we speculated that this may be one of the important reasons for their poor prognosis. The research of cancer immunotherapy has been very hot in recent years, especially immune checkpoint inhibitors have been quite mature as the first generation of immunotherapy, they play a therapeutic role in various cancers through mainly blocking PD-1/PD-L1 pathway and molecular receptors and/or ligands such as CTLA-4 (Topalian et al., 2015). Several previous studies have described therapeutic effect for immune checkpoints in some cancers, with findings consistent with favorable clinical outcomes in patients with many cancers, such as glioma (Puigdelloses et al., 2021), hepatocellular carcinoma (Sangro et al., 2020), lung cancer (Kartolo et al., 2021), and more. Therefore, we also evaluated the relationship between SERS and the expression levels of immune



checkpoints, macrophage-related molecules, and immunotherapy response. It found that SERS was positively correlated with the expression levels of immune checkpoints and macrophage-related

molecules. The response rate to ICI was significantly lower than that of the low-risk group. Therefore, we surmised that this may be another reason for the poor prognosis of LGG patients in the high-





**FIGURE 10**

The relations between the selected gene and tumor immune features. **(A)** The abundance of 22 immune cells in the high-expression and low-expression of AQP7. **(B)** Correlation of the expression of AQP7 with immune scores, stromal scores, ESTIMATE scores, and tumor purity. **(C)** The abundance of 22 immune cells in the high-expression and low-expression of E2F7. **(D)** Correlation of the expression of E2F7 with immune scores, stromal scores, ESTIMATE scores, and tumor purity. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.

risk group. Taken together, the SERS proposed in our study can be used to screen clinically high-risk LGG patients, and then to prescribe professionally-informed treatment.

Without doubt, there are some inevitable shortcomings in this study. Firstly, this is a retrospective study based on public databases, thus yielding more reliable results in a prospective study. Secondly, the five cohorts of LGG patients have varying degrees of lack of clinical information, which may lead to varying degrees of selection bias. Thirdly, GO and KEGG enrichment analysis and subsequent immune microenvironment and immune checkpoint analysis were not validated in the other 4 cohorts. Fourthly, we only analyzed transcriptome information and did not perform multi-omics analysis including methylation and gene copy number. Finally, further experiments are needed to explore the specific molecular mechanism of super-enhancer related genes for further elucidation.

## 5 Conclusion

In conclusion, three novel LGG subtypes were established based on SE-related genes. Subsequently, an accurate and independently validated model were proposed for predicting overall survival in LGG. In addition, we also found that SERS was associated with prognosis, clinicopathological features, tumor immune microenvironment, cancer hallmarks, and genomic alterations and the effect of immunotherapy in patients with LGG. The findings can be as the novel biomarkers for predicting prognosis and potential therapeutic targets for LGG, which will help physicians and patients to evaluate prognosis, determine follow-up period, and make immunotherapy decisions.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving human participants were reviewed and approved by Wuhan University of Science and Technology Affiliated Xiaogan Central Hospital. The ethics committee waived the requirement of written informed consent for participation.

## Author contributions

Conception and design: SF. Data collection and analysis and interpretation: YH, QY, and WW. Language editing and grammar correction: SC. Manuscript writing: YH and QY. Final approval of the manuscript: All authors.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1085584/full#supplementary-material>

### SUPPLEMENTARY FIGURE S1

Comparisons of age, gender, KPS, survival status, histology, WHO grade, MGMT status, Transcriptome subtype and TERT status among SE subtypes in TCGA cohort.

### SUPPLEMENTARY FIGURE S2

The Kaplan–Meier curves of TCGA cohort showed that there is different overall survival for patients with different expression levels of nine selected super-enhancer related prognostic genes.

### SUPPLEMENTARY FIGURE S3

Risk scores, survival status in high and low-risk group and Kaplan–Meier survival curves of SERS in TCGA, CGGA693, CGGA325, CGGA301 and Rembrandt cohorts, respectively.

### SUPPLEMENTARY FIGURE S4

Correlation analysis between the prognostic SERS and clinicopathological characteristics in the CGGA693 cohort. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.

### SUPPLEMENTARY FIGURE S5

Correlation analysis between the prognostic SERS and clinicopathological characteristics in the CGGA325 cohort. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.

### SUPPLEMENTARY FIGURE S6

Correlation analysis between the prognostic SERS and clinicopathological characteristics in the CGGA301 cohort. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.

### SUPPLEMENTARY FIGURE S7

Correlation analysis between the prognostic SERS and clinicopathological characteristics in the Rembrandt cohort. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and ns No significance.

### SUPPLEMENTARY FIGURE S8

The univariate Cox regression and multivariate Cox regression were performed on SERS and other clinicopathological features CGGA693, CGGA325, CGGA301, Rembrandt cohorts, respectively.

### SUPPLEMENTARY FIGURE S9

The ROC curves of the nomogram predicted 1-, 3-, and 5-year OS and corresponding calibration curves in CGGA693, CGGA325, CGGA301, Rembrandt cohorts, respectively.

## SUPPLEMENTARY FIGURE S10

The distribution of TMB, mutation counts, copy number burdens at focal and arm levels, immune scores, stromal scores, ESTIMATE

## References

- Bi, F., Chen, Y., and Yang, Q. (2020). Significance of tumor mutation burden combined with immune infiltrates in the progression and prognosis of ovarian cancer. *Cancer Cell. Int.* 20, 373. doi:10.1186/s12935-020-01472-9
- Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., Cooper, L. A., et al. (2015). Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* 372, 2481–2498. doi:10.1056/NEJMoa1402121
- Chen, X. F., Li, C. F., Lu, L., and Mei, Z. C. (2016). Expression and clinical significance of aquaglyceroporins in human hepatocellular carcinoma. *Mol. Med. Rep.* 13, 5283–5289. doi:10.3892/mmr.2016.5184
- Chen, S., Pu, J., Bai, J., Yin, Y., Wu, K., Wang, J., et al. (2018). Ezh2 promotes hepatocellular carcinoma progression through modulating mir-22/galectin-9 axis. *J. Exp. Clin. Cancer Res.* 37, 3. doi:10.1186/s13046-017-0670-6
- Chen, J., Shen, S., Li, Y., Fan, J., Xiong, S., Xu, J., et al. (2022). Apollo: An accurate and independently validated prediction model of lower-grade gliomas overall survival and a comparative study of model performance. *Ebiomedicine* 79, 104007. doi:10.1016/j.ebiom.2022.104007
- Dai, C., Charlestin, V., Wang, M., Walker, Z. T., Miranda-Vergara, M. C., Facchine, B. A., et al. (2020). Aquaporin-7 regulates the response to cellular stress in breast cancer. *Cancer Res.* 80, 4071–4086. doi:10.1158/0008-5472.CAN-19-2269
- Dao, T. P., Rosch, S., Mairbaurl, H., Pusch, S., Unterberg, A., Herold-Mende, C., et al. (2018). Identification of a prognostic hypoxia-associated gene set in idh-mutant glioma. *Int. J. Mol. Sci.* 19, 2903. doi:10.3390/ijms19102903
- Endo-Munoz, L., Dahler, A., Teakle, N., Rickwood, D., Hazar-Rethinam, M., Abdul-Jabbar, I., et al. (2009). E2f7 can regulate proliferation, differentiation, and apoptotic responses in human keratinocytes: Implications for cutaneous squamous cell carcinoma formation. *Cancer Res.* 69, 1800–1808. doi:10.1158/0008-5472.CAN-08-2725
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22. doi:10.18637/jss.v033.i01
- Gittleman, H., Sloan, A. E., and Barnholtz-Sloan, J. S. (2020). An independently validated survival nomogram for lower-grade glioma. *Neuro Oncol.* 22, 665–674. doi:10.1093/neuonc/noz191
- Hillman, R. T., Chisholm, G. B., Lu, K. H., and Futreal, P. A. (2018). Genomic rearrangement signatures and clinical outcomes in high-grade serous ovarian cancer. *J. Natl. Cancer Inst.* 110, 265–272. doi:10.1093/jnci/djx176
- Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-Andre, V., Sigova, A. A., et al. (2013). Super-enhancers in the control of cell identity and disease. *Cell.* 155, 934–947. doi:10.1016/j.cell.2013.09.053
- Hu, J., Lai, Y., Huang, H., Ramakrishnan, S., Pan, Y., Ma, V., et al. (2022). Tcof1 upregulation in triple-negative breast cancer promotes stemness and tumour growth and correlates with poor prognosis. *Br. J. Cancer* 126, 57–71. doi:10.1038/s41416-021-01596-3
- Hu, X., Ni, S., Zhao, K., Qian, J., and Duan, Y. (2022). Bioinformatics-led discovery of osteoarthritis biomarkers and inflammatory infiltrates. *Front. Immunol.* 13, 871008. doi:10.3389/fimmu.2022.871008
- Huang, P., Zhang, B., Zhao, J., and Li, M. D. (2022). Integrating the epigenome and transcriptome of hepatocellular carcinoma to identify systematic enhancer aberrations and establish an aberrant enhancer-related prognostic signature. *Front. Cell. Dev. Biol.* 10, 827657. doi:10.3389/fcell.2022.827657
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of t cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1
- Jiang, T., Nam, D. H., Ram, Z., Poon, W. S., Wang, J., Boldbaatar, D., et al. (2021). Clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* 499, 60–72. doi:10.1016/j.canlet.2020.10.050
- Jiao, W., Chen, Y., Song, H., Li, D., Mei, H., Yang, F., et al. (2018). Hpsc enhancer rna promotes cancer progression through driving chromatin looping and regulating hnrnpu/p300/egr1/hpsc axis. *Oncogene* 37, 2728–2745. doi:10.1038/s41388-018-0128-0
- Kartolo, A., Feilotter, H., Hopman, W., Fung, A. S., and Robinson, A. (2021). A single institution study evaluating outcomes of pd-1 high kras-mutant advanced non-small cell lung cancer (nscl) patients treated with first line immune checkpoint inhibitors. *Cancer Treat. Res. Commun.* 27, 100330. doi:10.1016/j.ctarc.2021.100330
- Kenny, P. A., Nelson, C. M., and Bissell, M. J. (2006). The ecology of tumors: By perturbing the microenvironment, wounds and infection may be key to tumor development. *Scientist* 20, 30.
- Lee, J. H., Xiong, F., and Li, W. (2020). Enhancer rnas in cancer: Regulation, mechanisms and therapeutic potential. *Rna Biol.* 17, 1550–1559. doi:10.1080/15476286.2020.1712895
- Li, Q., Qiu, X. M., Li, Q. H., Wang, X. Y., Li, L., Xu, M., et al. (2015). MicroRNA-424 may function as a tumor suppressor in endometrial carcinoma cells by targeting e2f7. *Oncol. Rep.* 33, 2354–2360. doi:10.3892/or.2015.3812
- Liu, H., Chi, A. W., Arnett, K. L., Chiang, M. Y., Xu, L., Shestova, O., et al. (2010). Notch dimerization is required for leukemogenesis and t-cell development. *Genes. Dev.* 24, 2395–2407. doi:10.1101/gad.1975210
- Liu, X., Li, Y., Qian, Z., Sun, Z., Xu, K., Wang, K., et al. (2018). A radiomic signature as a non-invasive predictor of progression-free survival in patients with lower-grade gliomas. *Neuroimage Clin.* 20, 1070–1077. doi:10.1016/j.nicl.2018.10.014
- Lou, J., Hao, Y., Lin, K., Lyu, Y., Chen, M., Wang, H., et al. (2020). Circular rna cdr1as disrupts the p53/mdm2 complex to inhibit gliomagenesis. *Mol. Cancer* 19, 138. doi:10.1186/s12943-020-01253-y
- Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., et al. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* 346, 1373–1377. doi:10.1126/science.1259037
- Marquis, C., Fonseca, C. L., Queen, K. A., Wood, L., Vandal, S. E., Malaby, H., et al. (2021). Chromosomally unstable tumor cells specifically require kif18a for proliferation. *Nat. Commun.* 12, 1213. doi:10.1038/s41467-021-21447-2
- Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. doi:10.1101/gr.239244.118
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi:10.1038/nmeth.3337
- Ostrom, Q. T., Price, M., Ryan, K., Edelson, J., Neff, C., Cioffi, G., et al. (2022). Cbtrus statistical report: Pediatric brain tumor foundation childhood and adolescent primary brain and other central nervous system tumors diagnosed in the United States in 2014–2018. *Neuro Oncol.* 24, i1–i38. doi:10.1093/neuonc/noac161
- Peng, L., Jiang, B., Yuan, X., Qiu, Y., Peng, J., Huang, Y., et al. (2019). Super-enhancer-associated long noncoding rna hcl5 is activated by zeb1 and promotes the malignancy of hepatocellular carcinoma. *Cancer Res.* 79, 572–584. doi:10.1158/0008-5472.CAN-18-0367
- Puigdellos, M., Garcia-Moure, M., Labiano, S., Laspidea, V., Gonzalez-Huarriz, M., Zalacain, M., et al. (2021). Cd137 and pd-1 targeting with immunovirotherapy induces a potent and durable antitumor immune response in glioblastoma models. *J. Immunother. Cancer* 9, e002644. doi:10.1136/jitc-2021-002644
- Ropri, A. S., DeVaux, R. S., Eng, J., Chittur, S. V., and Herschkowitz, J. I. (2021). Cis-acting super-enhancer lncrnas as biomarkers to early-stage breast cancer. *Breast Cancer Res.* 23, 101. doi:10.1186/s13058-021-01479-8
- Sangro, B., Chan, S. L., Meyer, T., Reig, M., El-Khoueiry, A., and Galle, P. R. (2020). Diagnosis and management of toxicities of immune checkpoint inhibitors in hepatocellular carcinoma. *J. Hepatol.* 72, 320–341. doi:10.1016/j.jhep.2019.10.021
- Sepaniac, L. A., Martin, W., Dionne, L. A., Stearns, T. M., Reinholdt, L. G., and Stumpff, J. (2021). Micronuclei in kif18a mutant mice form stable micronuclear envelopes and do not promote tumorigenesis. *J. Cell. Biol.* 220, e202101165. doi:10.1083/jcb.202101165
- Shen, R., Li, P., Li, B., Zhang, B., Feng, L., and Cheng, S. (2019). Identification of distinct immune subtypes in colorectal cancer based on the stromal compartment. *Front. Oncol.* 9, 1497. doi:10.3389/fonc.2019.01497
- Topalian, S. L., Drake, C. G., and Pardoll, D. M. (2015). Immune checkpoint blockade: A common denominator approach to cancer therapy. *Cancer Cell.* 27, 450–461. doi:10.1016/j.ccell.2015.03.001
- Wan, R. J., Peng, W., Xia, Q. X., Zhou, H. H., and Mao, X. Y. (2021). Ferroptosis-related gene signature predicts prognosis and immunotherapy in glioma. *Cns Neurosci. Ther.* 27, 973–986. doi:10.1111/cns.13654
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., et al. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* 153, 307–319. doi:10.1016/j.cell.2013.03.035
- Wu, F., Qin, Y., Jiang, Q., Zhang, J., Li, F., Li, Q., et al. (2020). MyoD1 suppresses cell migration and invasion by inhibiting fut4 transcription in human gastric cancer cells. *Cancer Gene Ther.* 27, 773–784. doi:10.1038/s41417-019-0153-3

Xu, S., Tang, L., Liu, Z., Luo, C., and Cheng, Q. (2021). Hypoxia-related lncrna correlates with prognosis and immune microenvironment in lower-grade glioma. *Front. Immunol.* 12, 731048. doi:10.3389/fimmu.2021.731048

Xu, B., Chen, H., Xu, Z., Yao, X., Sun, X., and Cheng, H. (2021). Cdc2 promotes tumorigenesis and induces radioresistance in oesophageal squamous cell carcinoma cells. *Mol. Med. Rep.* 24, 530. doi:10.3892/mmr.2021.12169

Ye, L., Wang, L., Yang, J., Hu, P., Zhang, C., Tong, S., et al. (2021). Identification of tumor antigens and immune subtypes in lower grade gliomas for mrna vaccine development. *J. Transl. Med.* 19, 352. doi:10.1186/s12967-021-03014-x

Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Zakharova, G., Efimov, V., Raevskiy, M., Rumiantsev, P., Gudkov, A., Belogurova-Ovchinnikova, O., et al. (2022). Reclassification of tcga diffuse glioma profiles linked to transcriptomic, epigenetic, genomic and clinical data, according to the 2021 who cns tumor classification. *Int. J. Mol. Sci.* 24, 157. doi:10.3390/ijms24010157

Zhou, M., Zhang, Z., Zhao, H., Bao, S., Cheng, L., and Sun, J. (2018). An immune-related six-lncrna signature to improve prognosis prediction of glioblastoma multiforme. *Mol. Neurobiol.* 55, 3684–3697. doi:10.1007/s12035-017-0572-9