

## **Title Page**

**Title:** Functional implications of aging-related lncRNAs for predicting prognosis and immune status in glioma patients

**Running Title:** Aging-related lncRNAs and glioma

## **Authors and affiliations**

Guangying Zhang<sup>1</sup>, Yanyan Li<sup>2</sup>, Na Li<sup>1</sup>, Liangfang Shen<sup>1</sup>, Zhanzhan Li<sup>1,3\*</sup>

1. Department of Oncology, Xiangya Hospital, Central South University, Changsha, Hunan Province, P.R. China. 410008

2. Department of Nursing, Xiangya Hospital, Central South University, Changsha, Hunan Province, P.R. China. 410008

3. National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan Province, P.R. China, 410008

**\*Correspondence to:** Zhanzhan Li, No.87, Xiangya Road, Kaifu District, Changsha, Hunan Province P.R. China, 410008

**E-mail:** lizhanzhan@csu.edu.cn

## Abstract

Glioma, is the most prevalent intracranial tumor with high recurrence and mortality rate. Long noncoding RNAs (lncRNAs) play a critical role in the occurrence and progression of tumors as well as in aging regulation. Our study aimed to establish a new glioma prognosis model by integrating aging-related lncRNAs expression profiles and clinical parameters in glioma patients from the Chinese Glioma Genome Atlas (CGGA) and the Cancer Genome Atlas (TCGA) datasets. The Pearson correlation analysis ( $|R| > 0.6$ ,  $P < 0.001$ ) was performed to explore the aging-related lncRNAs, and univariate cox regression and least absolute shrinkage and selection operator (LASSO) regression were used to screening prognostic signature in glioma patients. Based on the fifteen lncRNAs, we can divide glioma patients into three subtypes, and developed a prognostic model. Kaplan-Meier survival curve analysis showed that low-risk patients had longer survival time than high-risk group. Principal component analysis indicated that aging-related lncRNAs signature had a clear distinction between high- and low-risk groups. We also found that fifteen target lncRNAs were closely correlated with 119 genes by establishing a co-expression network. In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis displayed different function and pathways enrichment in high-and low-risk groups. The different missense mutations were observed in two groups, and the most frequent variant types were single nucleotide polymorphism (SNP). This study demonstrated that the novel aging-related lncRNAs signature had an important prognosis prediction and may contribute to individual treatment for glioma.

**Keywords: aging, lncRNA, glioma, prognosis, immune filtration**

## **Introduction**

Glioma is a highly heterogeneous tumor in central nervous system with high recurrence and mortality rate, especially glioblastoma. The median survival time of WHO III grade glioma is approximately 3 years, whereas WHO IV grade glioma has a grave prognosis of less than 15 months[1]. The conventional treatment for gliomas is surgical resection, radiotherapy and temozolomide (TMZ) chemotherapy. The drug treatment of glioma remains challenging due to the unique structure of blood-brain barrier, which prevents numerous antitumor drugs into the brain[2, 3]. Despite various cancer therapies have been applied over the past decades, the prognosis of glioma patients remains dismal. And the prognosis of glioma patients with the same grade is inconsistent. Recently, it is proved that many mutations could assess risk factor and predict prognosis of glioma, such as isocitrate dehydrogenase (IDH) mutation and 1p19q co-deletion, which present relatively favorable survival[4]. But some of these indicators are not very comprehensive to predict the prognosis of all grade glioma patients, for example, IDH (since nearly 80% of patients with low-grade glioma have IDH mutations). Therefore, a robust prognostic model is urgently needed to predict survival of patients with the malignant growth and highly relapse rate of glioma.

Long non-coding RNAs (lncRNAs) are more than 200 nucleotides in length, which do not encode protein and involve in post-transcriptional modulation and gene translation[5]. It is recorded that the diversity of lncRNAs is implicated in many biological functions, such as epigenetic regulation, tumor microenvironment and cell

apoptosis[6]. Some scholars have revealed that overexpression of lncRNA AGAP2-AS1 enhanced breast tumor growth and trastuzumab resistance[7]. The high expression of HOTAIR also promotes proliferation of breast tumor cells and drug resistance to tamoxifen[8]. Luo et al. have found that the depletion of lncRNA AGAP2-AS1 depressed proliferation and invasion, and induced apoptosis in U251 cells[9]. Other study has shown that overexpression of ATB is predicts a poor prognosis in colorectal cancer[10]. With further research on lncRNAs, accumulating researches have been identified lncRNAs as critical roles in regulating cell proliferation, invasion, apoptosis and drug resistance in a variety of tumors[11, 12]. Meanwhile, the vital roles of lncRNAs in degenerative disease of the central nervous system are becoming evident[13]. LncRNAs also promote tumor immune evasion, for example, NKILA escape immunological destruction by sensitizing T cells and inhibit NF- $\kappa$ B activity[14].

As is known to us, neurodegenerative disorders and cancer are age-related diseases. Recent researches have demonstrated that lncRNA expression profiles are influencing aging. An accumulation of studies has suggested that aging is a significant risk factor for the development of cancer[15, 16]. However, the role of aging-related lncRNAs in gliomas has not been fully elucidated. In the present study, we integrated gene matrix and clinical parameters from CGGA and TCGA database. Fifteen lncRNAs related to aging are screened by cox regression analyses. The aging-related lncRNAs prognosis model can predict the prognosis of glioma patients as a potential prognostic indicator, which may provide crucial implication in clinical targeted therapy.

## **Materials and methods**

### **Data source**

The complete RNA-sequencing data and corresponding clinical features of glioma patients were obtained from the Chinese Glioma Genome Atlas (CGGA) and The Cancer Genome Atlas (TCGA; <https://cancergenome.nih.gov/>). The lncRNA and protein-coding genes were classified according to gene annotation in GENCODE project (<https://www.genecodegenes.org/>) from our downloaded raw readings and fragments per kilo-base of transcript per million data. Detailed clinicopathological information, including age, gender, WHO grade, radiotherapy and chemotherapy status, and survival data were obtained for our further analysis. Similarly, the corresponding 1p19q codeletion and IDH mutation status were downloaded from CGGA database. Glioma samples were excluded who died for non-cancer related factors with survival times less than 30 days. In addition, a portion of glioma subjects with imperfect information were also eliminated. No specific ethical approval or patients informed consent was not required necessary because all of these data were publicly available.

### **Identification of aging-related lncRNAs**

The list of aging-related genes were acquired from Human Ageing Genomic Resources (HAGR; Table S1 ). To determine the correlation between aging-related genes and lncRNAs, we performed Pearson correlation coefficients in software R (version 3.6.3). The aging-related lncRNAs were selected out based on the criteria that correlation coefficient  $|R| > 0.6$  and  $P$  value was less than 0.001. According to the

above threshold values, candidate aging-related lncRNAs were identified and further analyzed.

### **Cluster analysis of aging-related lncRNAs**

We performed Principal Component Analysis (PCA) to visualize these aging-related lncRNAs expression patterns in different glioma patients in CGGA database and validated in TCGA database. The ConsensusClusterPlus algorithm was used for clustering samples assessments and then patients were divided into subtypes. The aging-related lncRNAs were identified three subtypes in CGGA.

### **Development and validation of aging-related lncRNAs prognostic signature**

To identify the potential prognostic-related lncRNAs, univariate Cox regression analysis was used to analyze the association between aging-related lncRNAs and overall survival (OS). The least absolute shrinkage and selection operator (LASSO) Cox regression was performed for filtered more meaningful prognosis-related lncRNAs to establish the risk signature. Finally, we developed an aging-related lncRNAs prognostic signature for glioma patients involving fifteen aging-related lncRNAs. Then, we constructed the prognostic signature of aging-lncRNAs. According to risk coefficient and the expression levels of each lncRNA, and a risk score was calculated for each patient. The calculation formula of risk score is as follows: risk score = lncRNA1 $\beta$   $\times$  Expression + lncRNA2 $\beta$   $\times$  Expression + ... lncRNA $n\beta$   $\times$  Expression.

### **Prediction analysis of prognostic signature**

On the basis of the median risk score as the threshold, the glioma patients were divided into high- and low-risk cohorts. We depicted survival curve between two cohorts using Kaplan-Meier curve method with a two-sided log-rank test. Univariate Cox regressions were utilized to evaluate the effect of clinicopathological variables on survival of glioma patients, including age, gender, tumor grade, radiotherapy and chemotherapy status. Furthermore, multivariate Cox regressions was performed to determine whether risk score was an independent prognostic factor. In addition, we performed stratified survival analysis to detect the impact of prognostic value of our risk scores model in different glioma subgroups. To further delve into the effect of single aging-related lncRNAs on glioma patients in the prognostic risk model, the relationship between expression of each screened lncRNA and clinical characteristics was assessed by Student's *t*-test or one-way analysis of variance (ANOVA). In addition, predictive efficiency of the risk score was generated by calculating the area under the curve (AUC) of receiver operating characteristic (ROC) curve.

### **Co-expression network and gene sets enrichment analysis**

Pearson correlation coefficients were calculated among lncRNAs and aging genes in glioma patients using R software (v.3.6.3) ( $R > 0.6$ ,  $p < 0.001$ ). The aging-related lncRNAs and target genes co-expression network were constructed using Cytoscape (v.3.8.2). A Sandkey diagram was used to visualize the co-expression of mRNAs and lncRNAs and showed the risk type. To explore the Functional enrichment of the fifteen lncRNAs, we conducted gene set enrichment analysis (GSEA; v.4.0.3) to determine the biological functions and pathways by this priori defined lncRNAs and

verified whether a showed statistically significant differences between high-risk and low-risk group or not.

### **Immune infiltration analysis**

We obtained immune cells infiltration score for each glioma patient to explore the degree of immune cells infiltration between low- and high-risk groups. And then we assessed the difference in proportion of 22 immune cell subtypes between the low- and high-risk groups. The ESTIMATE algorithm was applied to compare of the estimation, immune and stromal score between the two groups [a]. Furthermore, the scatter plot showed the correlations of risk score with Macrophages M0, Monocytes, and NK cells activated. The correlation between immune cells infiltration and risk scores was calculated by Pearson correlation ( $P < 0.05$  was considered statistically significant).

### **Somatic mutations analysis based on risk score**

Further, we explored the genetic alteration of glioma patients in high- and low- risk groups and represent by waterfall plots. The top 10 somatic mutation were screened in two groups, respectively. In addition, according to different classifications, the mutations were further sorted in detail. The exclusive and co-occurrence of mutated genes in both two groups of glioma patients were visualized.

### **Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)**

We collected 7 glioma and non-neoplastic brain tissues (NBTs) tissues from patients who underwent surgical operation. Fresh tissue samples were frozen and stored at  $-80^{\circ}\text{C}$ . This research was approved by Ethics Committee of Xiangya

Hospital, Central South University. And the informed consent was acquired from involved patients. We extracted total RNA from tissues following manufacturer's instructions. cDNA Synthesis Kit (TaKaRa) was utilized to reverse transcription. Then, qRT-PCR assay was performed following the reaction steps. The related expression of lncRNAs were normalized by GAPDH mRNA and calculated by  $2^{-\Delta\Delta C_t}$  method. Primers sequences were showed as follows: LINC00665 forward 5` -GAGGACTCAGAGGTGGAATT-3`, and reverse 5` - CAGCCAGC TTGTAGGG-3`; LINC00339 forward 5` -ACCATGCTAGAAAGCCTCCC-3`,and reverse 5` - CGTCCAGCAAGGTCCTAGAG-3`; SNHG16 forward 5` - ACAT CGGCATGATGGCAGAA-3`, and reverse 5` - TCACAAAAGGCGGGACCAC-3`; PAXIP1.AS2 forward 5` -GCTTCCAGGGGAGAT-3`, and reverse 5` - ATC AGACTGCCTGAAGA -3`; LINC00092 forward 5` - CCTATGATTTGGCCTCT GGA-3`, and reverse 5` - GAGAGCAGCGTTCAGGAAAC-3`.

### **Statistical analysis**

All statistical analyses were conducted using R software (v.4.0.5). Kaplan-Meier curves and log-rank tests were applied to evaluate survival data between various subgroups. Univariate and multivariate Cox regression analyses were used to assessed the independent prognostic factors for glioma patients. The risk coefficients of prognostic signature were calculated by LASSO regression. Prognostic accuracy of the nomogram and predicted signatures for 1-, 3-, and 5-year OS was estimated by ROC curves. Single nucleotide variation was analyzed with maftools R package. The differences between variables were evaluated by independent t-tests. Chi-square test

was performed to predict the association of variables of clinical parameters between high- and low risk groups. Pearson and Spearman correlation analyses were conducted to determine the correlation. Probability value of less than 0.05 or 0.001 was regarded statistically significant.

## **Results**

### **Identification of aging-related lncRNAs**

We obtained 307 aging related genes from HAGR public database (Table S1). The RNA-sequencing profile data and corresponding clinical parameters of glioma patients were downloaded from CGGA and TCGA. A total of 928 lncRNAs were screened from CGGA and 629 lncRNAs were identified from TCGA (Table S2 and Table S3). Then, based on Pearson correlation analysis ( $|R| > 0.6$  and  $P < 0.001$ ), 226 and 152 aging-related lncRNAs were selected from CGGA and TCGA, respectively. We further performed univariate cox regression analysis to filtrate potential prognostic-related lncRNAs from aging-related lncRNAs, and found that 33 lncRNAs were significantly associated with OS of glioma patients. Detail data analyses were provided in Figure 1A. The correlation among 33 aging-related lncRNAs were presented by circle plot (Figure 1B).

### **Clustering analysis of aging-related lncRNAs associated with prognosis**

To identify aging-related glioma patterns, we performed ConsensusClusterPlus analysis to classify glioma patients into three subgroups using 33 lncRNAs (Figure 1C). The heat map showed sample clustering results with optimal clustering stability (the value of  $k=3$ ). Subsequently, a PCA analysis suggested that glioma samples

could be completely distinguished. There were three subtypes obtained in CGGA (Figure 1D) and validated in TCGA (Figure 1E). We further compared the prognosis of the three clusters by Kaplan-Meier curves of overall survival. In addition, the OS of the cluster1 was found to be shorter than these two clusters and the prognoses of the cluster2 were the best (Figure 1F and 1G). The three clusters were observably separated in CGGA database ( $P < 0.001$ ). While there is no different distribution between cluster2 and cluster3 about survival time in TCGA.

### **Establishment and validation of aging-lncRNAs prognostic model**

Based on the survival information of glioma samples, we applied univariate cox regression and screened 33 aging-related lncRNAs were highly related to OS. To better explore the prognostic role of those aging-related lncRNAs in glioma patients, we applied LASSO cox regression algorithm to validate that of 15 lncRNAs were most correlated with prognostics (Figure 2A). The optimal value of penalty parameter was determined for tuning the parameter selection in LASSO regression (Figure 2B). A risk score was calculated according to coefficients of each lncRNA (Figure 2C). The risk scores of glioma samples were calculated as follows: risk score = (  $0.220 \times \text{Exp}_{\text{LINC00665}}$  ) + (  $0.177 \times \text{Exp}_{\text{LINC00339}}$  ) + (  $0.126 \times \text{Exp}_{\text{SNHG16}}$  ) + (  $-0.026 \times \text{Exp}_{\text{PAXIP1.AS2}}$  ) + (  $0.035 \times \text{Exp}_{\text{LINC00092}}$  ) + (  $0.030 \times \text{Exp}_{\text{LINC00265}}$  ) + (  $-0.164 \times \text{Exp}_{\text{SOC2.AS1}}$  ) + (  $0.002 \times \text{Exp}_{\text{SNHG9}}$  ) + (  $-0.010 \times \text{Exp}_{\text{LINC00237}}$  ) + (  $-0.026 \times \text{Exp}_{\text{SLC25A21.AS1}}$  ) + (  $-0.078 \times \text{Exp}_{\text{PEPB41L4A.AS1}}$  ) + (  $-0.126 \times \text{Exp}_{\text{HAR1A}}$  ) + (  $-0.164 \times \text{Exp}_{\text{GDNF.AS1}}$  ) + (  $-0.262 \times \text{Exp}_{\text{SNA13.AS1}}$  ) + (  $-0.283 \times \text{Exp}_{\text{WDFY3.AS2}}$  ). Taking the median risk score, glioma patients from CGGA and TCGA datasets were divided into high- and low-risk group,

respectively. We found that there was a marked difference in prognosis between two groups ( $P < 0.001$ ). Patients in high-risk group had a worse OS than those in the low-risk group (Figure 2D, 2G). Further, the risk score and survival status distributions of glioma patients were showed in Figure 2E and 2H. Survival rates of glioma patients were correlated with risk scores, the mortality rate increased with a higher risk score. In addition, tSNE2 method were then performed to classify the samples into obvious two components in CGGA and TCGA (Figure 2F, 2I). Besides, ROC analysis indicated that risk score could accurately predict prognosis of glioma patients (1-year AUC = 0.760, 2-year AUC =0.832 , 3-year OS =0.827 in CGGA and 1-year AUC = 0.858, 2-year AUC =0.889 , 3-year OS = 0.9014 in TCGA; Figure 2J, 2K). These results demonstrated that the aging-lncRNAs prognostic model had a robust and stable prognostic value ability for glioma patients.

### **Correlations of prognostic signature lncRNAs with clinical features**

Finally, fifteen aging-related lncRNAs were involving in the prognosis signature. We performed univariate cox regression analysis to evaluate their prognostic roles. The forest plot illustrated that EPB41L4A.AS1, GDNF.AS1, HAR1A, LINC00237, SLC25A21.AS1, SNA13.AS1, and WDFY3.AS2 were protective factors with  $HR < 1$ , while LINC00092, LINC00265, LINC00339, LINC00665, PAXIP1.AS2, SNHG16, SNHG9 and SOCS2.AS1 were risk factors with  $HR > 1$  in glioma patients (Figure 3A). Subsequently, we investigated the distribution of clinic pathological features and expression of the fifteen aging-related lncRNAs between high- and low-risk groups, which for testing whether the aging-related lncRNAs signature could predict glioma

clinical parameters. The heat map indicated that there were significant differences in age, grade, PRS type, histology, clusters, chemotherapy status, 1p19q codeletion and IDH mutation status ( $P < 0.01$ ) between high- and low-risk groups (Figure 3B). Then, we further explored the relevance of risk score and each clinicopathological features. As expected, we found that the score of clusters 3, older age, advanced grade tumors, 1p19q non-codeletion, IDH wild type, recurrent tumors and chemotherapy status were significantly increased (Figure 3). The above results elucidate the aging-related lncRNAs signature may serve a pivotal role in tumor progression of glioma.

In addition, to prove the applicability of our prognostic model, we performed stratification analysis to identify whether it had ability to evaluate prognosis in each subgroup. As shown in Figure 4, in contrast with higher risk group, lower risk patients had longer survival time by Kaplan-Meier survival curve analysis. However, the OS rate was statistically similar for patients with secondary tumor, which because of the smaller sample size probably. These results indicated that our prognosis risk signature may be a potential predictor for glioma patients.

### **Prediction ability and independent analysis of prognostic signature**

Univariate and multivariate Cox regression analysis indicated that age (HR = 5.296, 95% CI: 3.936–7.1263,  $p < 0.001$ ), grade (HR = 4.662, 95% CI: 3.740–5.810,  $p < 0.001$ ) and risk score (HR = 6.556, 95% CI: 5.119–8.396,  $p < 0.001$ ) were remarkably associated with OS in CGGA training cohort. Meanwhile, similar conclusions were observed in TCGA validating cohort (Figure 5A, B). In order to develop a clinically applicable tool to predict OS in glioma patients by taking risk score of aging-related

lncRNAs prognostic signature and other clinicopathological parameters. We built a nomogram to evaluate the probability of survival at 1, 3 and 5 years. The C-index value of the nomogram was 0.801(95%CI:0.783-0.820). Calibration curves demonstrated that concordances between actual and predicted of survival rates of glioma patients after bias corrections of nomogram in CGGA cohort (Figure 5C-E). These results indicated that our aging-related lncRNAs prognostic signature was reliable and precise.

### **lncRNA-mRNA co-expression and pathways enrichment**

Considering that miRNA and lncRNA can affect progression of cancers through mutual regulation, next we explored the potential functions of the fifteen aging-related lncRNAs in glioma by establishing a co-expression network. We found that fifteen target lncRNAs were closely correlated with 119 genes, which constructed a complex co-expression network. The detailed showed in Figure 7A). A Sankey diagram was depicted to visualize the relationship among lncRNAs, mRNAs and outcomes (risk/protective) (Figure 7B). In addition, KEGG analysis was performed to study the potential biological function and pathway in high- and low- risk groups. We observed that high-risk group was enriched in lysosome, N glycan biosynthesis, pathogenic *Escherichia coli* infection, primary immunodeficiency, pyrimidine metabolism and regulation of actin cytoskeleton, and low-risk group was enriched in long term depression, long term potentiation, neuroactive ligand receptor, phosphatidylinositol signaling, and WNT signaling pathway, respectively. The top six significant gene sets in the two groups were presented in Figure 7C, 7D. The above data provided valuable

insights for future finding potential individualized treatments in different risk score group of glioma patients.

### **Immune status based on risk score**

In order to further investigate the various infiltration of immune cells between high- and low- groups, we compared the differences of 22 immune cells infiltration in the two groups. The proportions of B cells naive ( $p=0.030$ ), plasma cells ( $P=0.046$ ), T cells CD8 ( $P < 0.001$ ), T cells CD4 memory activated ( $P=0.002$ ), T cells regulatory ( $P=0.001$ ), T cells gamma delta ( $P=0.001=3$ ), NK cells resting ( $P=0.021$ ), Macrophages M1 ( $P < 0.001$ ), Macrophages M2 ( $P < 0.001$ ), Mast cells resting ( $P=0.005$ ) and Neutrophils ( $P=0.004$ ) were increased in high-risk group (Figure 8A). Besides, we calculated ESTIMATE score, immune score and stromal score by using ESTIMATE algorithm in two groups. Our result indicated that the scores of high risk group were higher than those of low risk group ( $P < 2.22e-16$ ) (Figure 8B-D). As shown in Figure 7F and 7G, risk scores were negatively correlated with the activation of Monocytes ( $R=0.58$ ,  $P=2.2e-16$ ) and NK cells ( $R=0.43$ ,  $P=4.3e-10$ ). However, the Macrophages M0 activation was upregulated with increased risk score (Figure 8E). These findings suggested that different immune infiltration status was emerged between two groups of glioma patients.

### **Somatic mutations analysis based on risk score**

In addition, we performed somatic mutation profiles to analyze the gene mutation in high- and low-risk groups of 604 glioma patients. The waterfall plots exhibited that top 10 mutated genes were TP53, IDH1, EGFR, TTN, PTEN, ATRX, MUC16, FLG,

PIK3CA and RYR2 in high-risk group, and were IDH1, TP53, ATRX, CIC, IDH2, PIK3CA, TTN, MUC16, 6MARCA4 and DNMT3A in low-risk group (Figure 9A,9B). As we can see, some mutated genes were high mutation frequency in two groups. Based on different classifications, mutations were further sorted. The missense mutations were accounted for the majority in two groups. And the most frequent variant types were single nucleotide polymorphism (SNP), of which occurred as C > T and T > C in low-risk group and C > T and C > A in high group (Fig, 9C,9D). Recently, co-occurrence and mutual exclusivity of genetics were often observed in cancer. In high-risk group, the co-occurrence mutations were much than those of in low-risk group (Details were showed in Figure 9E, 9F). However, they display significant exclusivity of mutations, such as IDH1 were mutually exclusive with PTEN and EGFR in high-risk group, while were mutually exclusive with IDH2 in low-risk group. This interrelated mutation could suggest functional interactions, which may provide new insights into clinical treatment.

### **Validation of aging-related lncRNAs expression levels**

To further verify our study, we detected five aging-relate lncRNAs expression in 7 gliomas (4 WHO II gliomas and 3 WHO III gliomas) and NBT samples by qRT-PCR assay. Our results showed that mean expression levels of LINC00665, LINC00339, SNHG16, PAXIPI.AS2 and LINC00092 in glioma tissues were higher than those in non-tumor brain tissues ((Supplementary Figure S1A-E). Moreoer, the higher-grade patients were detected to obtain the higher lncRNAs expression. The above results confirmed the reliability of our analysis.

## **Discussion**

Glioma has a high recur and lead to a mortal outcome. The therapeutic effect of glioma is remains not satisfactory, especially glioblastoma. It is necessary for highly heterogeneous disease of glioma to identify potential prognostic indicators. Accumulating evidences provided that lncRNAs show a pivotal role in the process of tumor occurrence, development, metastasis, drug resistance and display the potential as a novel biomarker[17]. Moreover, with wide application of high-throughput technologies and the increasing maturity of data-sharing, tumor data about lncRNAs are accumulating in the public databases[18]. LncRNAs modulate diverse biological processes and the role in aging have recently attracted attention. It can regulate cell senescence, telomere length, and stem cell differentiation in aging process[19].

Aging is an inevitable process and is considered one of the predominant risk factors for most chronic diseases, including cancer[20]. Aging and cancer are interrelated. Aging-related genes can regulate cell senescence and tumor malignancy. The current view is that cell aging may promote occurrence and development of gliomas, because gliomas are more common in the elderly patients, who's the number of senescent cells is increased dramatically in the brain[21]. Some scholars have also found that aging-associated genes linked to progression and prognosis of gliomas[22, 23]. Because of DNA damage from radiochemotherapy therapies have been shown to induce cells aging, which may be associated with glioma recurrence after treatment. Moreover, aging brain cells secrete excessive levels of factors to promote cell survival and invasion, such as MMP-2 and MMP-9[24, 25]. At the present, studies on

indicators of glioma based on lncRNA signature is mounting. However, the potential role of aging-related lncRNA prognosis signature in glioma is not adequately researched. This study was performed to assess the role of aging-related lncRNAs in glioma using CGGA dataset as training cohort and TCGA dataset as validation cohort.

In this study, we identified a fifteen aging-related lncRNAs risk signature in gliomas through uni-cox regression and LASSO analysis. Finally, eight lncRNAs (LINC00092, LINC00265, LINC00339, LINC00665, PAXIP1.AS2, SNHG16, SNHG9 and SOCS2.AS1) were associated with high risk, and patients with high expression of these lncRNAs had an unfavorable prognosis. The remaining seven lncRNAs (EPB41L4A.AS1, GDNF.AS1, HAR1A, LINC00237, SLC25A21.AS1, SNA13.AS1, and WDFY3.AS2) were related to low risk. The AUC values to predict the OS of glioma in training cohort (1-year AUC = 0.760, 2-year AUC =0.832, 3-year OS =0.827 in CGGA and 1-year AUC = 0.858, 2-year AUC =0.889, 3-year OS = 0.9014), which indicate that the prognostic risk model is reliable and stable. Recently, a large number of studies have revealed the important role of lncRNA as oncogenes or cancer suppressor genes in various tumors. Studies have presented that lncRNAs play a complex regulatory role in the progress of tumor. Zhao et al. found that LINC00092-silenced cells presented obviously compromised metastatic potential and reduced invasive capacity in ovarian cancer, and which involved in glycolysis process[26]. Some scholars have revealed that overexpression of LINC00665 could reversed the ability of invasion and migration through encoding micropeptide in

triple-negative breast tumor cells[27]. Meanwhile, Lu et al. showed that LINC00665 regulate stemness and EMT properties to promote gemcitabine resistance in gemcitabine resistance[28]. And in glioma cells, LINC00665 could also inhibit tumor progress via STAU1-mediated mRNA degradation[29]. LncRNA WDFY3-AS2 as a ceRNA to inhibit invasion properties, which is correlated with lymph node metastasis and TNM stage in oesophageal squamous cell carcinoma[30]. Wu et al. indicated that WDFY3-AS2 upregulated SDC4 expression to promotes cisplatin resistance in ovarian cancer. And si-WDFY3-AS2 should reduce invasion and migration of tumor cells[31].

A PCA based on aging-related lncRNAs showed three significantly different clustering patterns in glioma samples, and there was a difference OS among three clusters. Continuous inflammatory response led to cancer. In addition, it should be noted that the relationships between aging-related lncRNAs and immune cell infiltration. Our result revealed that different immune infiltration status was emerged between two groups of glioma patients. The risk scores were negatively correlated with the activation of Monocytes and NK cells, while the Macrophages M0 activation was upregulated with increased risk score. Further, we systemically analyzed the corresponding key signaling pathways of those lncNRNAs by KEGG. The majority of enriched pathways manifested the immunomodulatory functions, and the top five significantly enriched pathways involved in phosphatidylinositol signaling, and WNT signaling pathway. Somatic mutations are generally considered as tumor-initiating events[32]. Then, we also used waterfall plots to show the common mutate in glioma,

that TP53 and IDH1 genes were both high mutation frequency in two groups. This fifteen lncRNAs risk signature not only effectively predicted the prognosis of gliomas but also reflected clinicopathological factors, such as grade, chemotherapy status, 1p19q codeletion and IDH mutation status, which have the potential to be a precise indicator for prognosis of gliomas.

In conclusion, we constructed an independent and robust prognostic signature using fifteen aging-related lncRNAs. The present risk signature in our study was based on transcriptome databases, so validation with some fundamental experiments is required. Next, further to clarify investigation of underlying mechanisms and to reveal how it regulates infiltrating immune cells in glioma. This study may be beneficial for clinicians to more accurately identify patients with high risk factors and improve prognosis of glioma.

### **Data availability**

Original data can be obtained from Supplementary materials.

### **Author contributions**

ZZZ designed this study and directed the research group in all aspects, including planning, execution, and analysis of the study. ZGY drafted the manuscript. YYL, and NL, collected the data. LZZ provided the statistical software, performed the data analysis, LZZ arranged the Figures and Tables. ZGY and SLF revised the manuscript.

All authors have read and approved the final version of the manuscript.

### **Funding**

This work was partly supported by the Science Foundation of Xiangya Hospital for

Young Scholar (LZZ: NO. 2018Q012), National Natural Science Foundation of China (LZZ: No. 82003239 and SLF: 81974466).

### **Conflict of interest**

The authors declare that they have no conflicts of interest.

### **References**

1. Wang FY, Kang CS, Wang-Gou SY, Huang CH, Feng CY, Li XJ. EGFL7 is an intercellular EGFR signal messenger that plays an oncogenic role in glioma. *Cancer Lett.* 2017; 384: 9-18. <https://10.1016/j.canlet.2016.10.009>
2. Chua J, Nafziger E, Leung D. Evidence-Based Practice: Temozolomide Beyond Glioblastoma. *Curr Oncol Rep.* 2019; 21: 30. <https://10.1007/s11912-019-0783-5>
3. Oberoi RK, Parrish KE, Sio TT, Mittapalli RK, Elmquist WF, Sarkaria JN. Strategies to improve delivery of anticancer drugs across the blood-brain barrier to treat glioblastoma. *Neuro Oncol.* 2016; 18: 27-36. <https://10.1093/neuonc/nov164>
4. Labreche K, Kinnersley B, Berzero G, Di Stefano AL, Rahimian A, Detrait I, Marie Y, Grenier-Boley B, Hoang-Xuan K, Delattre JY, Idbaih A, Houlston RS, Sanson M. Diffuse gliomas classified by 1p/19q co-deletion, TERT promoter and IDH mutation status are associated with specific genetic risk loci. *Acta Neuropathol.* 2018; 135: 743-755. <https://10.1007/s00401-018-1825-z>
5. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet.* 2014; 15: 7-21. <https://10.1038/nrg3606>
6. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. *Nature.* 2012; 482: 339-346. <https://10.1038/nature10887>

7. Dong H, Wang W, Mo S, Chen R, Zou K, Han J, Zhang F, Hu J. SP1-induced lncRNA AGAP2-AS1 expression promotes chemoresistance of breast cancer by epigenetic regulation of MyD88. *J Exp Clin Cancer Res.* 2018; 37: 202. <https://10.1186/s13046-018-0875-3>
8. Xue X, Yang YA, Zhang A, Fong KW, Kim J, Song B, Li S, Zhao JC, Yu J. LncRNA HOTAIR enhances ER signaling and confers tamoxifen resistance in breast cancer. *Oncogene.* 2016; 35: 2746-2755. <https://10.1038/onc.2015.340>
9. Luo W, Li X, Song Z, Zhu X, Zhao S. Long non-coding RNA AGAP2-AS1 exerts oncogenic properties in glioblastoma by epigenetically silencing TFPI2 through EZH2 and LSD1. *Aging (Albany NY).* 2019; 11: 3811-3823. <https://10.18632/aging.102018>
10. Abedini P, Fattahi A, Agah S, Talebi A, Beygi AH, Amini SM, Mirzaei A, Akbari A. Expression analysis of circulating plasma long noncoding RNAs in colorectal cancer: The relevance of lncRNAs ATB and CCAT1 as potential clinical hallmarks. *J Cell Physiol.* 2019; 234: 22028-22033. <https://10.1002/jcp.28765>
11. Xi J, Sun Q, Ma L, Kang J. Long non-coding RNAs in glioma progression. *Cancer Lett.* 2018; 419: 203-209. <https://10.1016/j.canlet.2018.01.041>
12. Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, Jin X, Yin J, Chen L, Zhang Y, Xu J, Li X. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat Commun.* 2020; 11: 1000. <https://10.1038/s41467-020-14802-2>
13. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE,

Finch CE, St LGR, Kenny PJ, Wahlestedt C. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med.* 2008; 14: 723-730. <https://10.1038/nm1784>

14. Huang D, Chen J, Yang L, Ouyang Q, Li J, Lao L, Zhao J, Liu J, Lu Y, Xing Y, Chen F, Su F, Yao H, Liu Q, Su S, Song E. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *Nat Immunol.* 2018; 19: 1112-1125. <https://10.1038/s41590-018-0207-y>

15. Childs BG, Gluscevic M, Baker DJ, Laberge RM, Marquess D, Dananberg J, van Deursen JM. Senescent cells: an emerging target for diseases of ageing. *Nat Rev Drug Discov.* 2017; 16: 718-735. <https://10.1038/nrd.2017.116>

16. Carreno G, Guiho R, Martinez-Barbera JP. Cell senescence in neuropathology: A focus on neurodegeneration and tumours. *Neuropathol Appl Neurobiol.* 2021; 47: 359-378. <https://10.1111/nan.12689>

17. Huarte M. The emerging role of lncRNAs in cancer. *Nat Med.* 2015; 21: 1253-1261. <https://10.1038/nm.3981>

18. Li H, Gao C, Liu L, Zhuang J, Yang J, Liu C, Zhou C, Feng F, Sun C. 7-lncRNA Assessment Model for Monitoring and Prognosis of Breast Cancer Patients: Based on Cox Regression and Co-expression Analysis. *Front Oncol.* 2019; 9: 1348. <https://10.3389/fonc.2019.01348>

19. Grammatikakis I, Panda AC, Abdelmohsen K, Gorospe M. Long noncoding RNAs(lncRNAs) and the molecular hallmarks of aging. *Aging (Albany NY).* 2014; 6: 992-1009. <https://10.18632/aging.100710>

20. Maleki F, Sadigh ZA, Sadeghi F, Muhammadnejad A, Farahmand M, Parvin M, Shirkoohi R. Human cytomegalovirus infection in Iranian glioma patients correlates with aging and tumor aggressiveness. *J Med Virol.* 2020; 92: 1266-1276. <https://10.1002/jmv.25673>
21. Latorre E, Harries LW. Splicing regulatory factors, ageing and age-related disease. *Ageing Res Rev.* 2017; 36: 165-170. <https://10.1016/j.arr.2017.04.004>
22. Carreno G, Guiho R, Martinez-Barbera JP. Cell senescence in neuropathology: A focus on neurodegeneration and tumours. *Neuropathol Appl Neurobiol.* 2021; 47: 359-378. <https://10.1111/nan.12689>
23. Lasry A, Ben-Neriah Y. Senescence-associated inflammatory responses: aging and cancer perspectives. *Trends Immunol.* 2015; 36: 217-228. <https://10.1016/j.it.2015.02.009>
24. Ladomersky E, Zhai L, Lauing KL, Bell A, Xu J, Kocherginsky M, Zhang B, Wu JD, Podojil JR, Plataniias LC, Mochizuki AY, Prins RM, Kumthekar P, Raizer JJ, Dixit K, Lukas RV, Horbinski C, Wei M, Zhou C, Pawelec G, Campisi J, Grohmann U, Prendergast GC, Munn DH, Wainwright DA. Advanced Age Increases Immunosuppression in the Brain and Decreases Immunotherapeutic Efficacy in Subjects with Glioblastoma. *Clin Cancer Res.* 2020; 26: 5232-5245. <https://10.1158/1078-0432.CCR-19-3874>
25. Boisvert MM, Erikson GA, Shokhirev MN, Allen NJ. The Aging Astrocyte Transcriptome from Multiple Regions of the Mouse Brain. *Cell Rep.* 2018; 22: 269-285. <https://10.1016/j.celrep.2017.12.039>

26. Zhao L, Ji G, Le X, Wang C, Xu L, Feng M, Zhang Y, Yang H, Xuan Y, Yang Y, Lei L, Yang Q, Lau WB, Lau B, Chen Y, Deng X, Yao S, Yi T, Zhao X, Wei Y, Zhou S. Long Noncoding RNA LINC00092 Acts in Cancer-Associated Fibroblasts to Drive Glycolysis and Progression of Ovarian Cancer. *Cancer Res.* 2017; 77: 1369-1382. <https://10.1158/0008-5472.CAN-16-1615>
27. Guo B, Wu S, Zhu X, Zhang L, Deng J, Li F, Wang Y, Zhang S, Wu R, Lu J, Zhou Y. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *Embo J.* 2020; 39: e102190. <https://10.15252/emj.2019102190>
28. Lu M, Qin X, Zhou Y, Li G, Liu Z, Geng X, Yue H. Long non-coding RNA LINC00665 promotes gemcitabine resistance of Cholangiocarcinoma cells via regulating EMT and stemness properties through miR-424-5p/BCL9L axis. *Cell Death Dis.* 2021; 12: 72. <https://10.1038/s41419-020-03346-4>
29. Ruan X, Zheng J, Liu X, Liu Y, Liu L, Ma J, He Q, Yang C, Wang D, Cai H, Li Z, Liu J, Xue Y. lncRNA LINC00665 Stabilized by TAF15 Impeded the Malignant Biological Behaviors of Glioma Cells via STAU1-Mediated mRNA Degradation. *Mol Ther Nucleic Acids.* 2020; 20: 823-840. <https://10.1016/j.omtn.2020.05.003>
30. Zhang Q, Guan F, Fan T, Li S, Ma S, Zhang Y, Guo W, Liu H. LncRNA WDFY3-AS2 suppresses proliferation and invasion in oesophageal squamous cell carcinoma by regulating miR-2355-5p/SOCS2 axis. *J Cell Mol Med.* 2020; 24: 8206-8220. <https://10.1111/jcmm.15488>
31. Wu Y, Wang T, Xia L, Zhang M. LncRNA WDFY3-AS2 promotes cisplatin

resistance and the cancer stem cell in ovarian cancer by regulating hsa-miR-139-5p/SDC4 axis. *Cancer Cell Int.* 2021; 21: 284. <https://10.1186/s12935-021-01993-x>

32. Balmain A. The critical roles of somatic mutations and environmental tumor-promoting agents in cancer risk. *Nat Genet.* 2020; 52: 1139-1143. <https://10.1038/s41588-020-00727-5>

### **Figure legends**

**Figure 1** Molecular classification based on aging-related lncRNAs. **A:** The flow chart of data analyses in the study. **B:** The circle plot showed the correlation among 33 aging-related lncRNAs. **C:** Glioma patients were divided into three clusters in CGGA. **D:** PCA indicated that three subclasses were obtained in CGGA. **E:** Three subclasses were validated in TCGA. **F:** Kaplan-Meier curves of overall survival for three clusters in CGGA. **G:** Kaplan-Meier curves of overall survival for three clusters in TCGA.

**Figure 2** Development and validation of aging-related lncRNAs prognosis signature. **A:** LASSO regression of 15 aging-related lncRNAs. **B:** Cross-validation for tuning the parameter selection in the LASSO regression. **C:** Coefficient of prognosis model regression. **D:** Kaplan-Meier curves of high-risk group and low-risk group in CGGA. **E:** Distribution of risk score and patients based on the risk score in CGGA. **F:** The tSNE2 method showed obvious two components in CGGA. **G:** Kaplan-Meier curves of high-risk group and low-risk group in TCGA. **H:** Distribution of risk score and patients based on the risk score in TCGA. **I:** The tSNE2 method showed obvious two components in TCGA. **J and K:** ROC curves of prognostic signature based on risk

score in CGGA and TCGA.

**Figure 3** Correlations of clinical characteristic with identified aging-related lncRNAs signature. **A:** Forest of univariate COX regression for 15 signature lncRNAs. **B:** Heatmap showed that correlation of clinical parameters with risk scores and expression of 15 lncRNAs in high- and low-risk group. Boxplot showed the comparisons of risk score in different subgroups: **C:** Cluster1 vs Cluster2 vs Cluster3. **D:** age $\leq$ 41 vs  $>$ 41, **E:** WHO II vs WHO III vs WHOIV. **F:** GBM vs LGG, **G:** 1p19q\_codeletion vs non-codel. **H:** IDH mutation vs wildtype. **I:** Primary vs Recurrent vs Secondary. **J:** Chemotherapy (Yes vs No). **K:** Radiotherapy (Yes vs No)

**Figure 4** Stratified analyses of high- and low-risk group. **A and B:** age. **C and D:** Gender. **E and F:** Histology. **G-I:** WHO stage. **J-L:** Pathology type. **M and N:** 1p19q codeletion status. **O and P:** IDH mutation status. **Q and R:** Chemotherapy. **S and T:** Radiotherapy.

**Figure 5** Independent prognosis analysis of risk score. **A and B:** Univariate and multivariate cox forest plot of risk score in CGGA and TCGA. **C-E:** Calibration plots of the nomogram for predicting the probability of OS at 1, 3, and 5 years in the CGGA.

**Figure 6** Nomograph of 1-, 3-, and 5-year overall survival probabilities predicted based on aging-related lncRNA signature.

**Figure 7** Functional and enrichment pathways analysis. **A:** lncRNAs-mRNA co-expression regulatory network based-on fifteen aging-related lncRNAs. **B:** A Sankey diagram was depicted to visualize the co-occurrences of lncRNAs, mRNAs

and outcomes. **C:** KEGG pathway enrichment analysis in high-risk group. **D:** KEGG pathway enrichment analysis in low-risk group.

**Figure 8** Immune filtration analysis between high- and low-risk groups. **A:** Differential analysis of immune-related cells based on risk score. **B-D:** Boxplot showed the comparisons of Estimation, immune and stromal score between high- and low-risk groups. **E-G:** Scatter plot showed that the correlations of risk score with Macrophages M0, Monocytes, and NK cells activated

**Figure 9** Landscape of mutation profiles between high- and low-risk glioma patients. **A and B:** Waterfall plots showed the mutations information in each sample of high- and low-risk group glioma patients. **C and D:** The variant classification in high-and the low-risk group glioma patients. **E and F:** The exclusive and co-occurrence in high-and the low-risk group glioma patients.

### **Supplementary materials**

#### **Supplementary material 1**

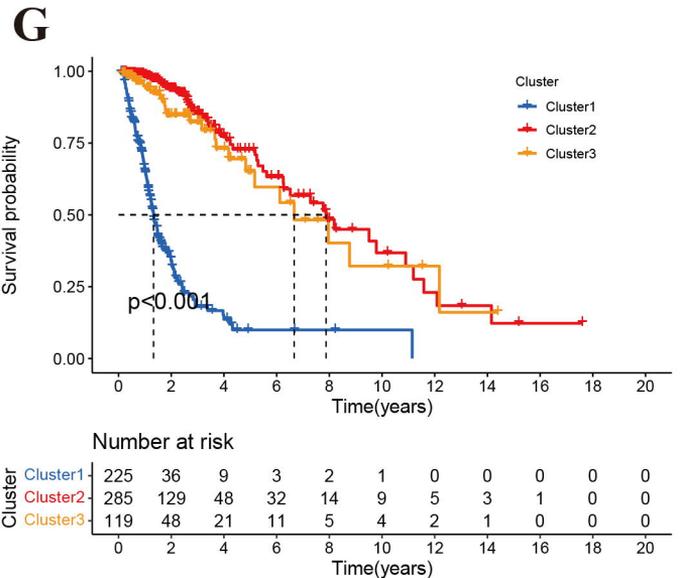
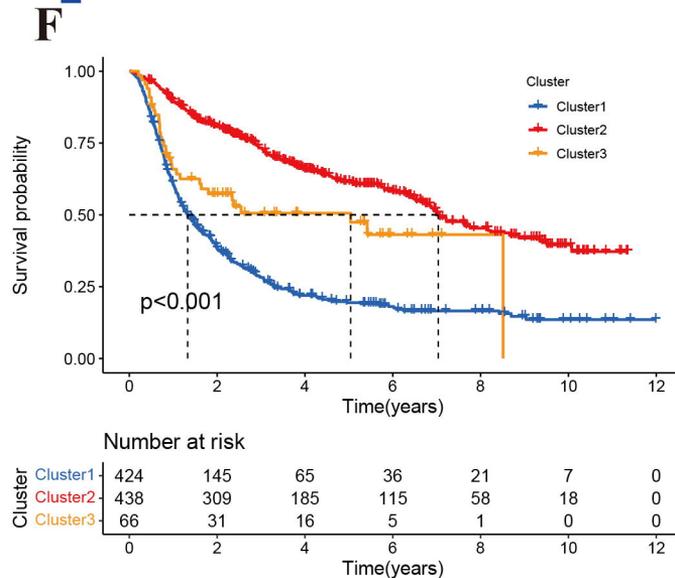
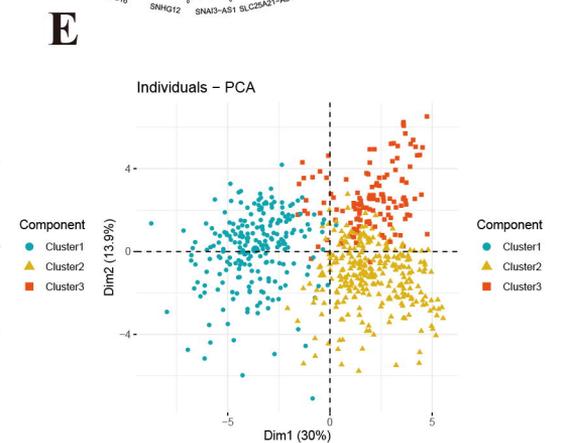
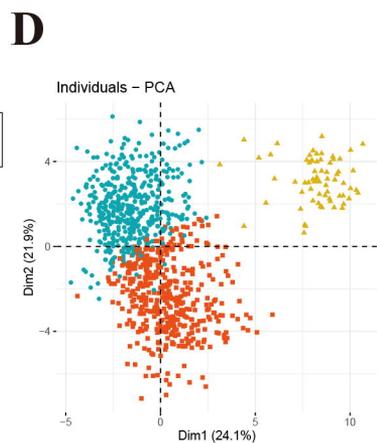
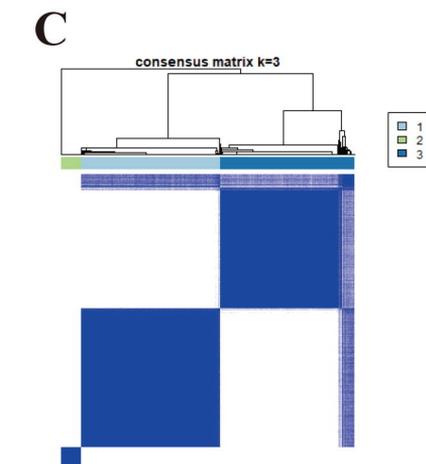
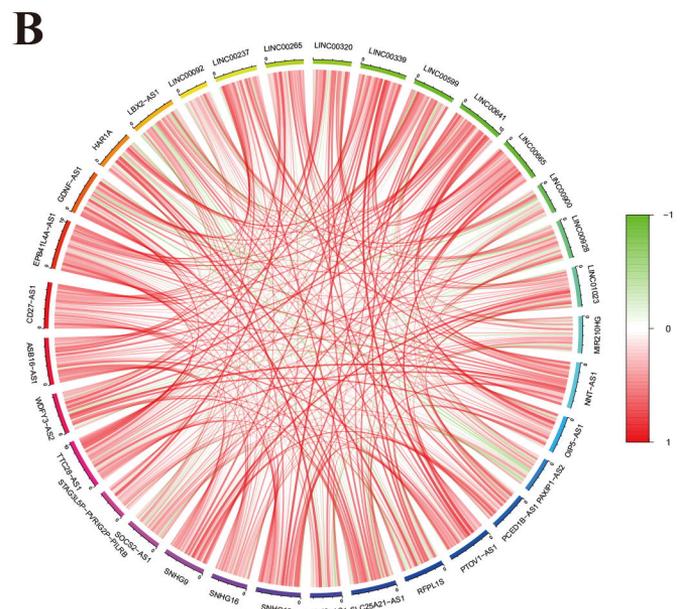
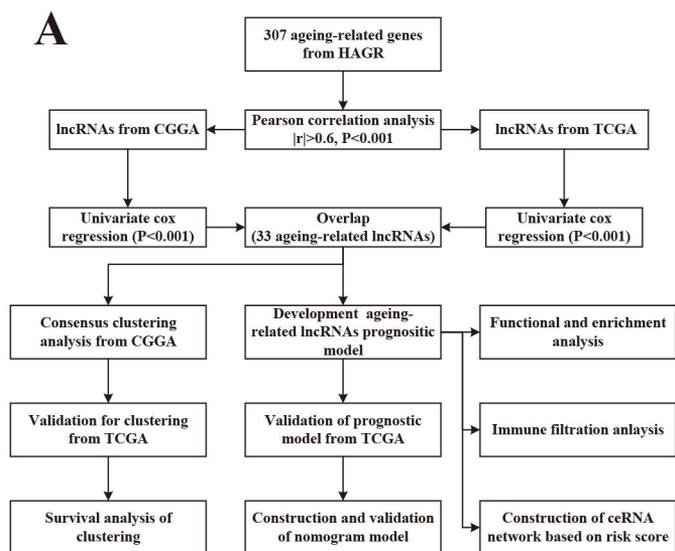
**Table S1** Aging-related genes

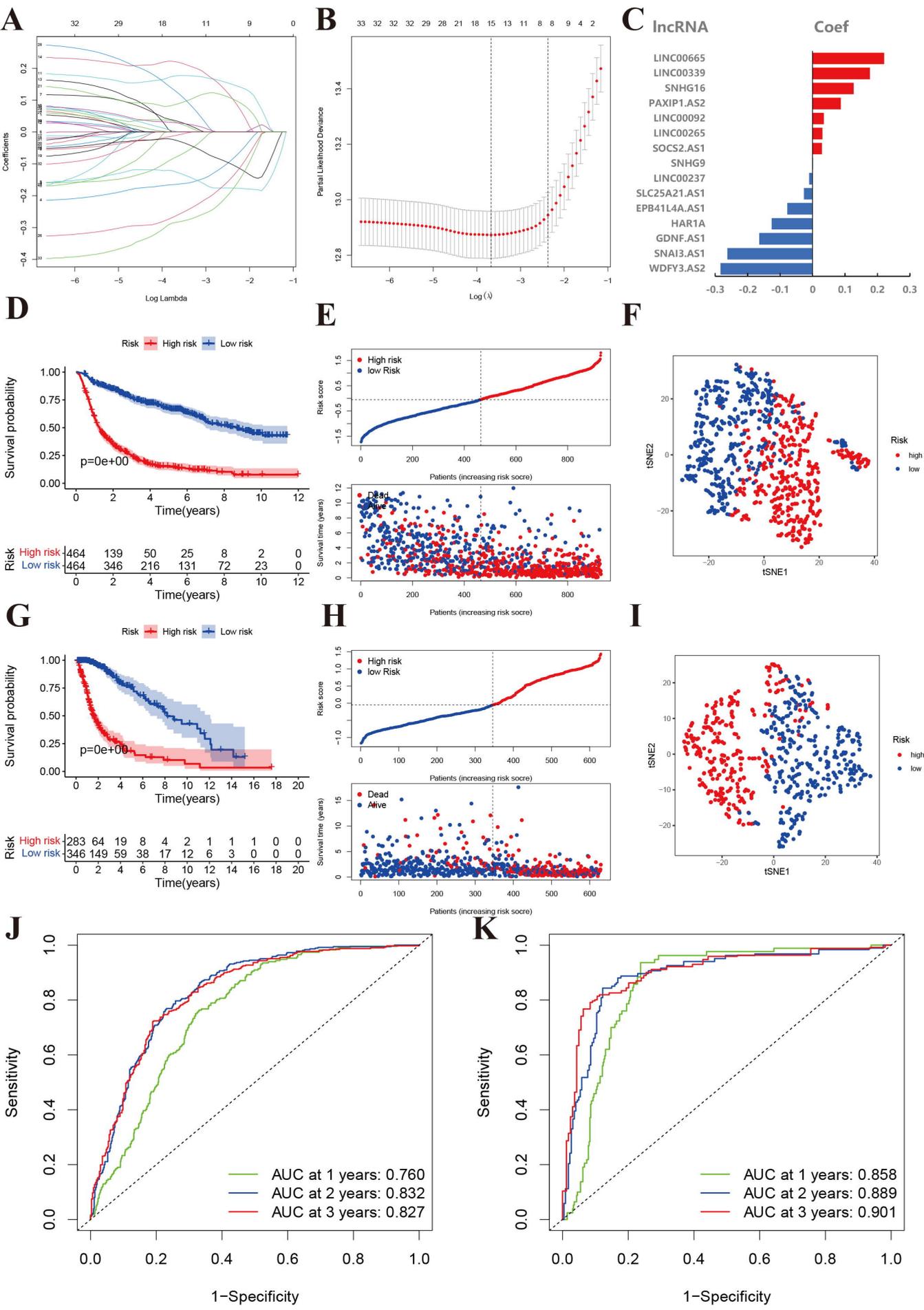
**Table S2** 33 aging-related lncRNAs expression in TCGA

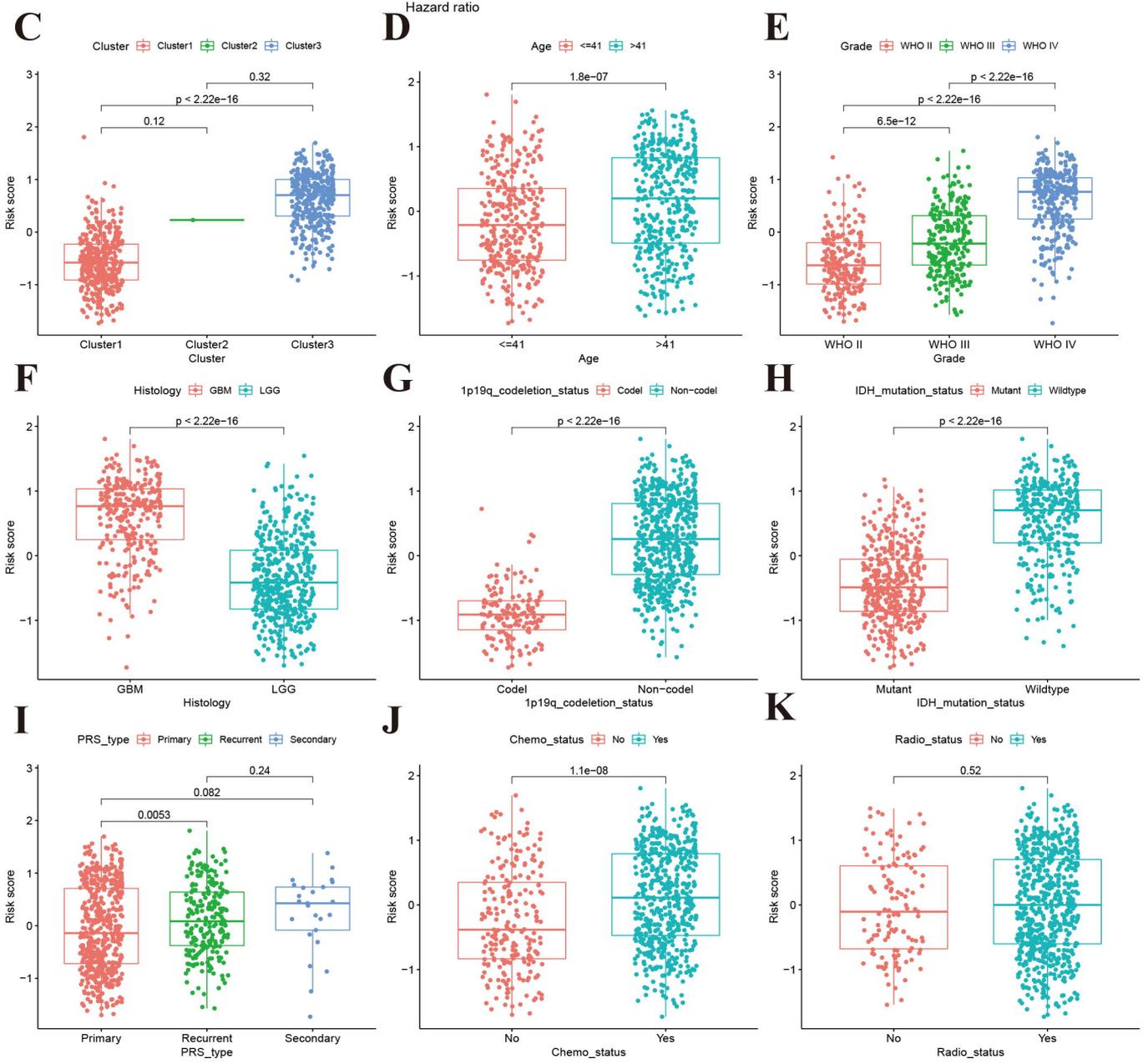
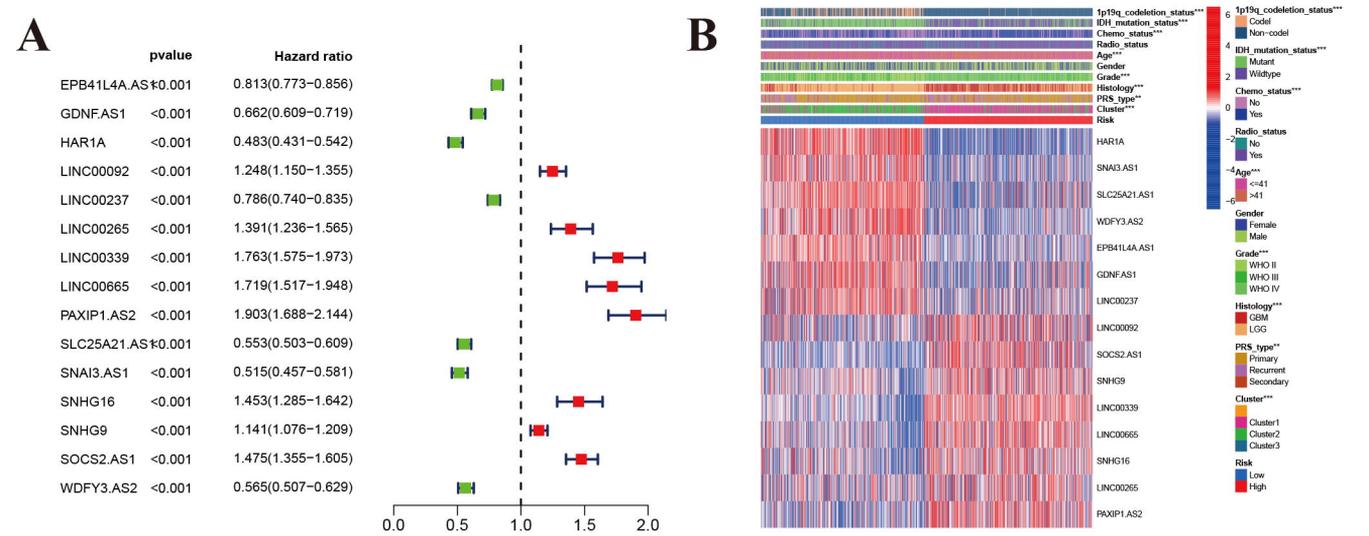
**Table S3** 33 aging-related lncRNAs expression in CGGA

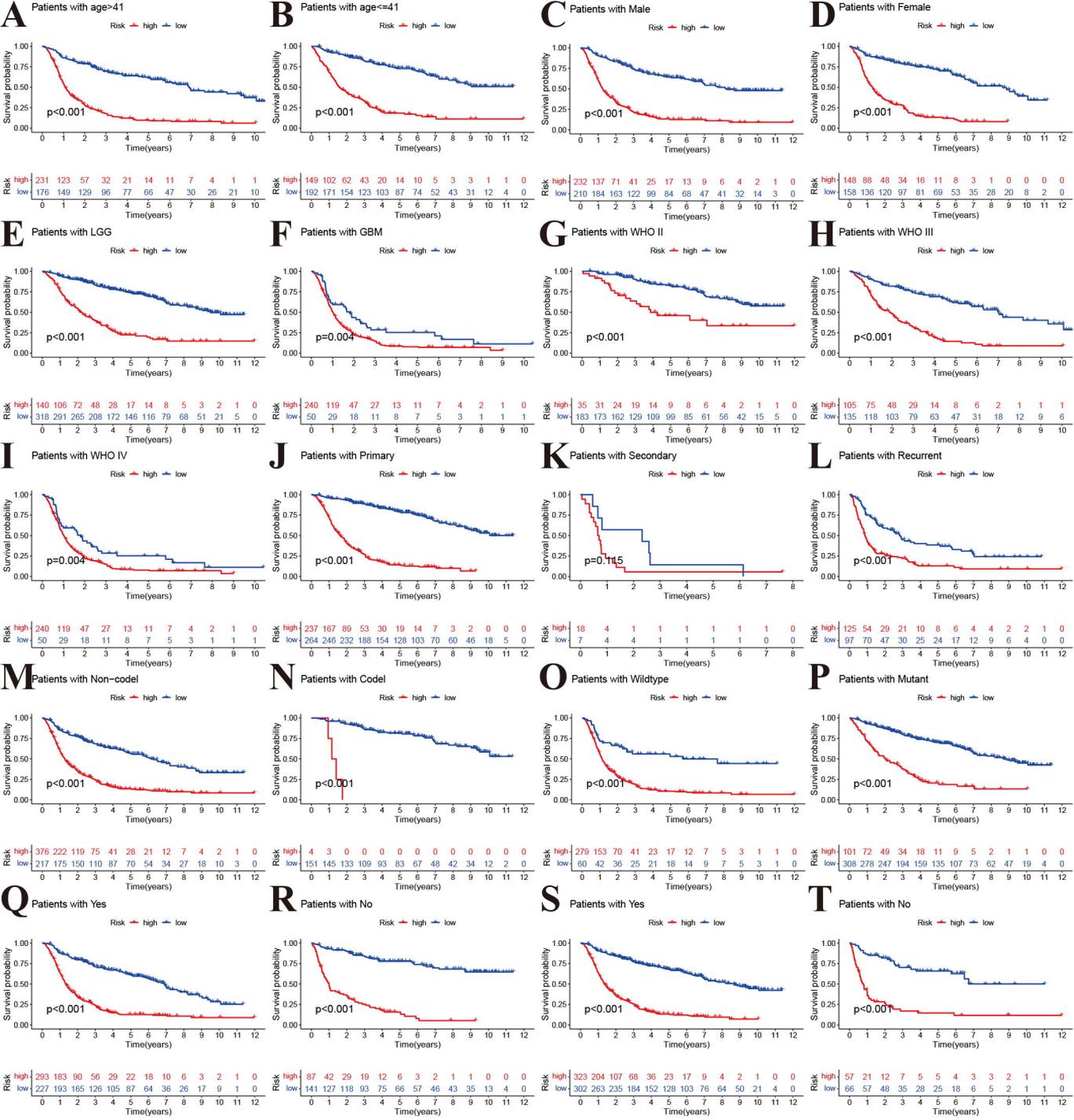
#### **Supplementary material 2**

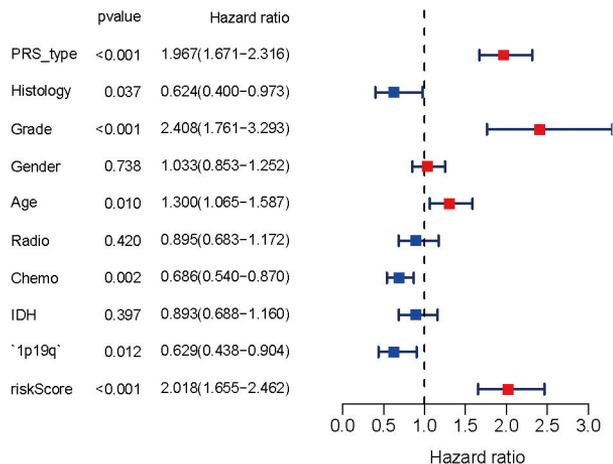
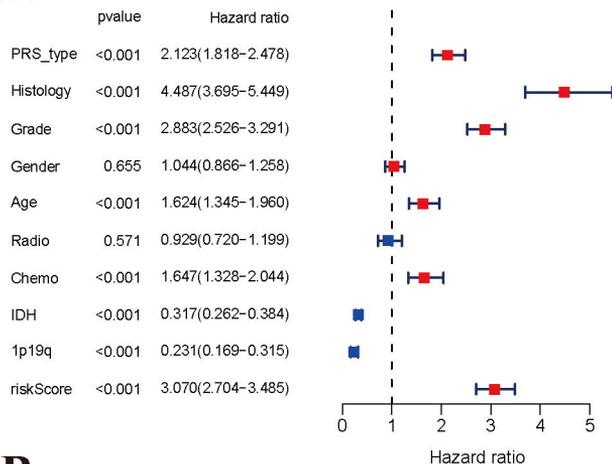
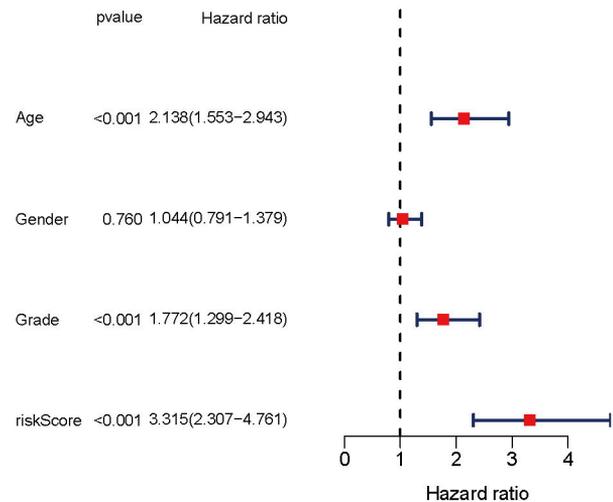
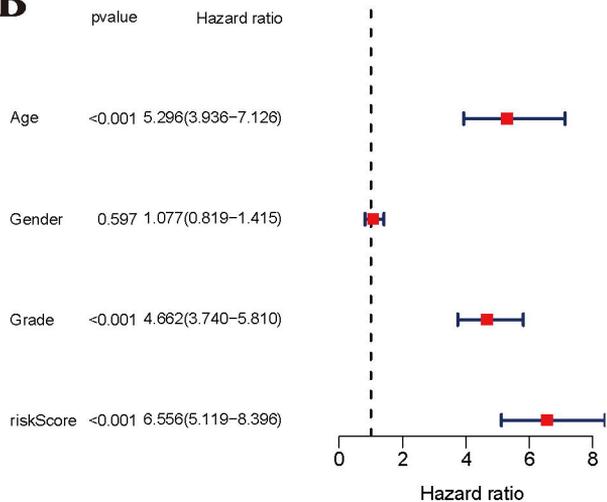
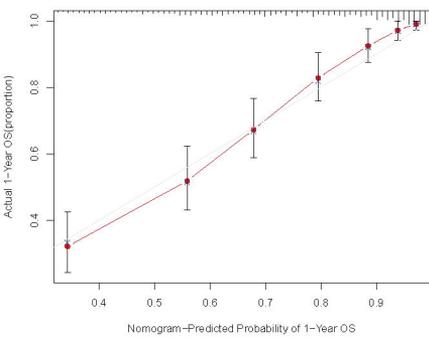
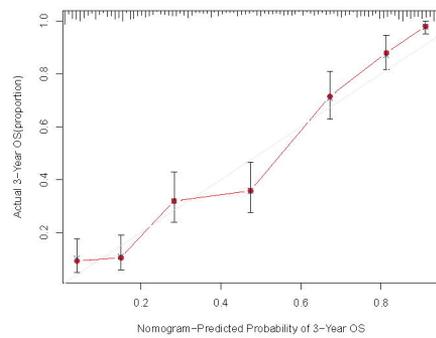
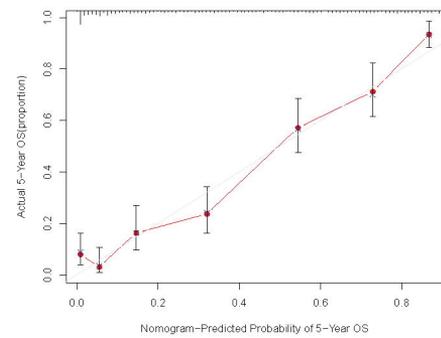
**Figure S1** Validations of several identified lncRNAs in glioma tissue. **A-E:** LINC00665, LINC00339, SNHG16, PAXIPI.AS2, LINC00092

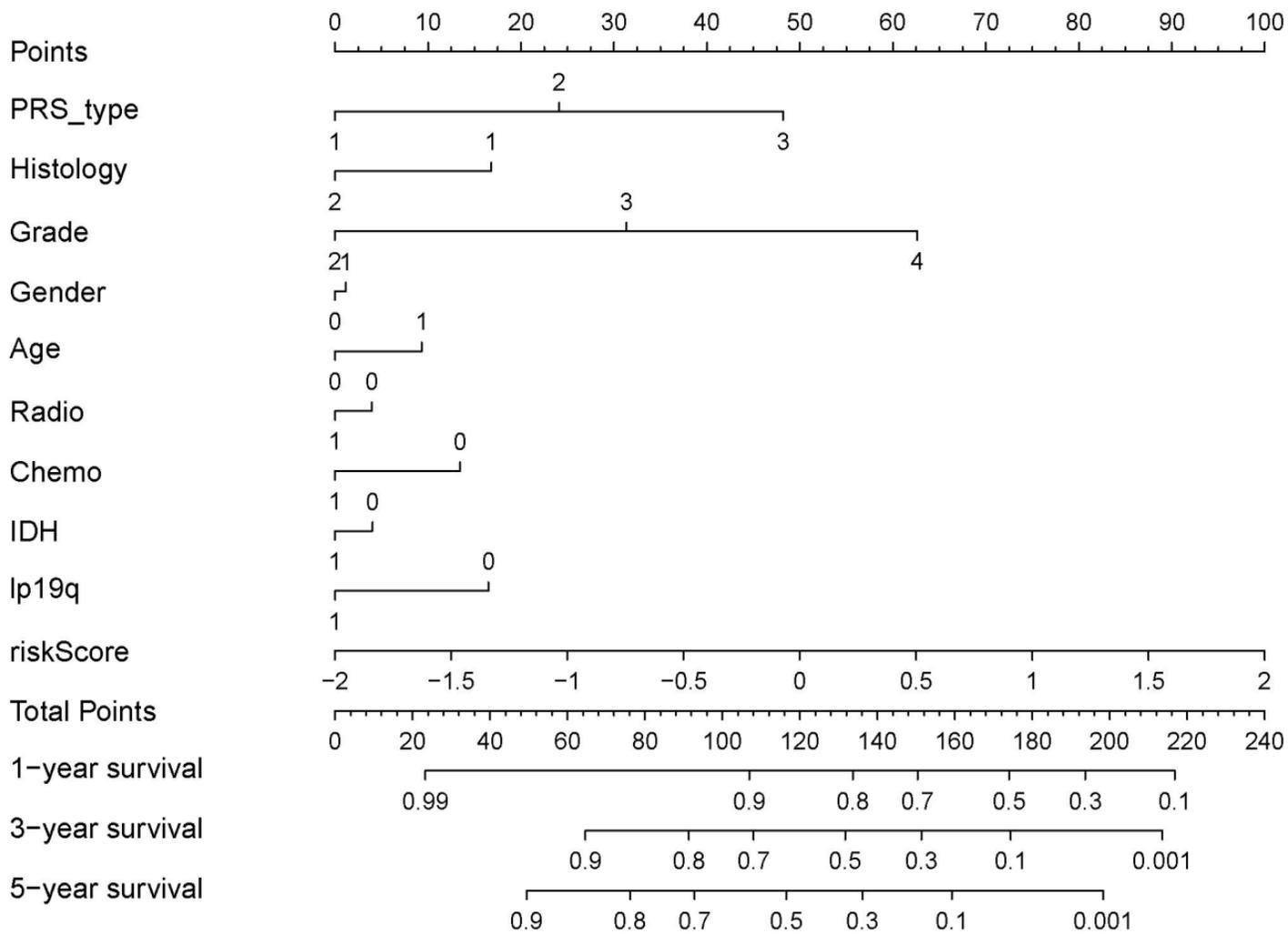




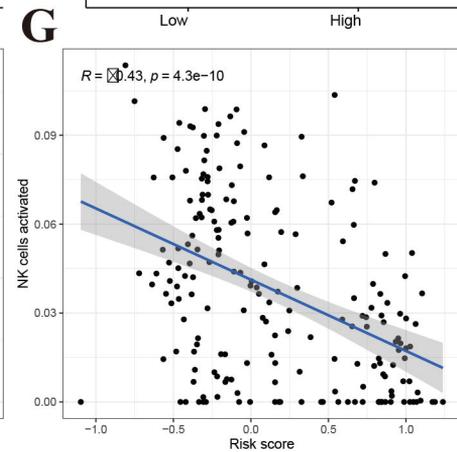
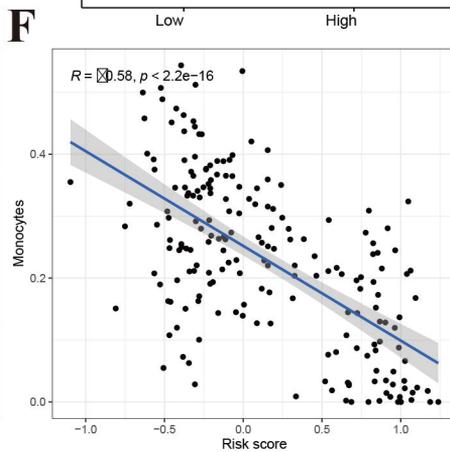
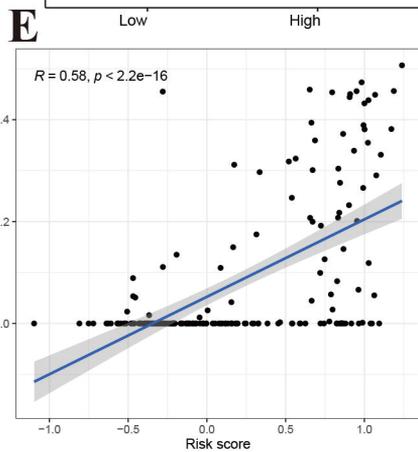
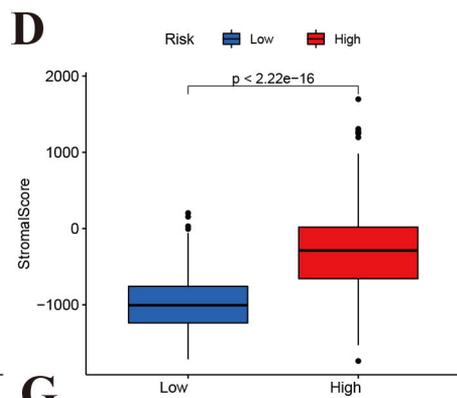
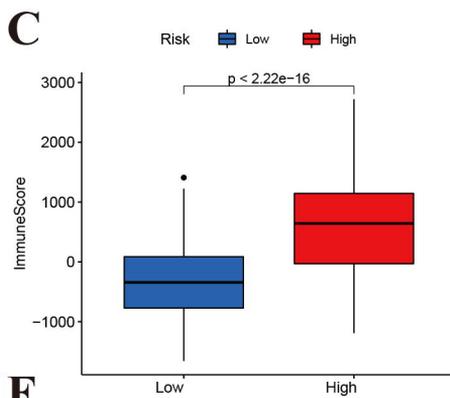
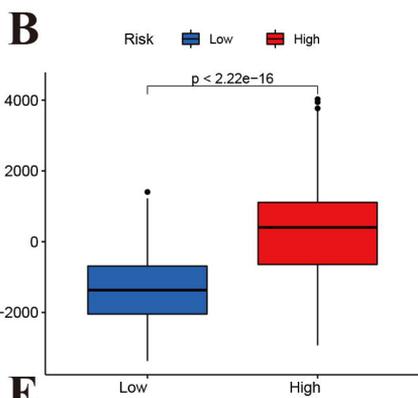
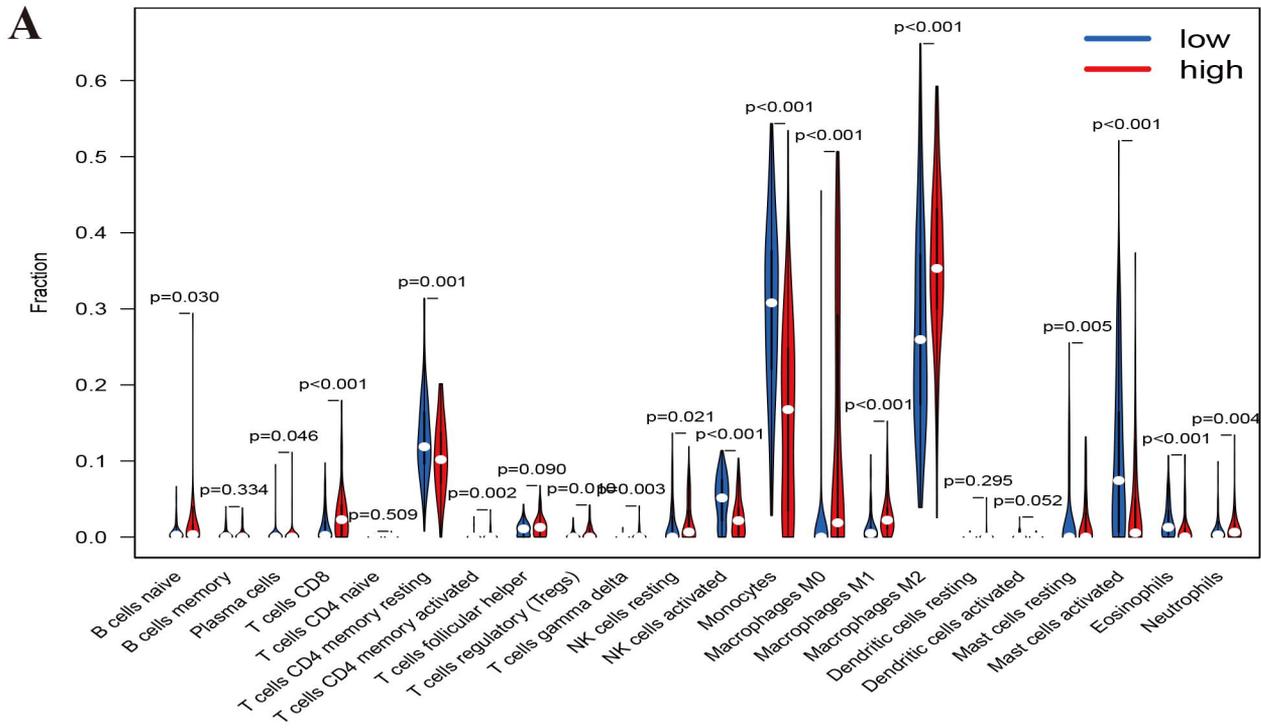


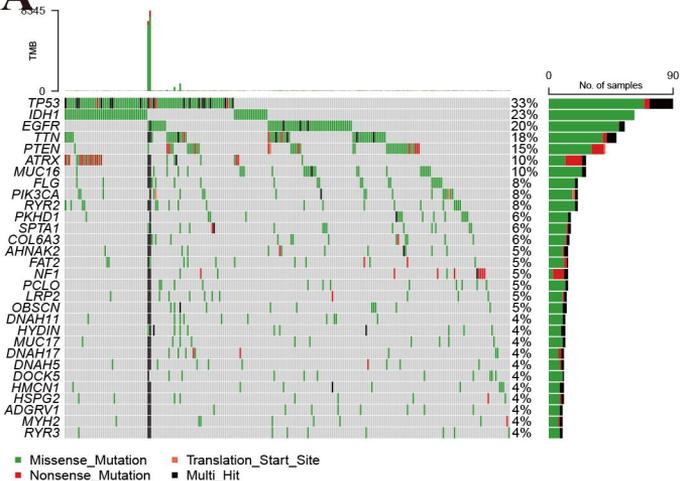


**A****B****C****D****E**







**A** Altered in 237 (87.45%) of 271 samples.**B** Altered in 322 (96.7%) of 333 samples.