Title Page

Title: Comprehensive analysis of Histone deacetylases genes in the prognosis and immune infiltration of glioma patients

Running head: Histone deacetylases and glioma

Authors and Affiliations

Aibin Liu[1,2], Yanyan Li[3], Lin Shen[4], Na Li[4], Liangfang Shen[4], Zhanzhan Li[4*]

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

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

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

4. Department of Oncology, 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, China, 410008. **E-mail:** lizhanzhan@csu.edu.cn

Abstract

Background: The occurrence and development of tumors are closely related to histone deacetylases (HDACs). However, the overall biology and prognosis are still unknown in glioma. In the present study, we comprehensively explored the biology function and prognosis of eleven HDAC genes in glioma, which may contribute the more understanding of molecular mechanisms and potential therapeutic targets for glioma patients.

Methods: We systematically described the expression files, molecular subtypes, prognostic value, immune filtration and tumor microenvironment and gene alteration, function and pathways enrichment, and drug sensitivity using TCGA and CGGA datasets. We developed and validated the prognostic model based on HDACs genes in glioma using LASSO, univariate, and multivariate cox regression. Receiver operating characteristic analyses were used for model evaluating. We also validated the expressions of HDACs genes included in the model in non-tumor and glioma tissues samples.

Results: Glioma patients can be divided into two subclasses based on eleven HDAC

genes, and patients from two subclasses had markedly different survival outcomes. Then, using six HDAC genes (HDAC1, HDAC3, HDAC4, HDAC5, HDAC7, and HDAC9), we established a prognostic model in glioma patients, and this prognostic model was well validated in an independent cohort population. Furthermore, the calculated risk score from six HDACA genes expression was suggested to be an independent prognostic factor, which can predict the five-year overall survival of glioma patients well. High-risk patients can be attributed to multiple complex function and molecular signaling pathways, and the genes alterations of high- and low-risk patients were significantly different. We also found that different survival outcomes of high- and low- risk patients could be involved in the differences of immune filtration level and tumor microenvironment. Subsequently, we identified several small molecular compounds that could be favorable for glioma patients' treatment. And finally, the expression levels of HDAC genes from prognostic model were validated in glioma and non-tumor tissues samples.

Conclusion: Our results revealed the clinical utility and potential molecular mechanisms of HDAC genes in glioma. Model based on six HDAC genes can predict the overall survival of glioma patients well, which can be served as potential therapeutic targets.

Keywords: Glioma, Histone deacetylases, bioinformatics, prognostic factor, immune infiltration

Introduction

Glioma was deemed to be the most aggressive tumor in central nervous system. The annual incidence of glioma was reported to be about 30-80/1 million in the world, and increased by 1-2% annually, the 5-year survival rate was only 10%-20%(Weller et al.,2015). According to the WHO standard, glioma was divided into four grades by its pathological characteristics(Chen et al.,2017): grade I, pilocytic astrocytoma, which manifested as benign tumor, patients may have full clinical recovery after total tumor resection. Grade II, which had a poor prognosis compared with grade I, but was still considered to be low-grade glioma. Grade III, such as anaplastic astrocytoma as well as Grade IV, GBM were types of advanced grade glioma lined to high degree of malignancy, strong invasive ability, poor prognosis and multiple differentiation potentials, the median survival time was only about one year(Kan et al.,2020). Despite various cancer therapies have been applied over the past decades, the prognosis of glioma patients remains dismal. After combining the isocitrate dehydrogenase (IDH1/2) mutation and whether the 1p19q code is missing, the HWO classification of central nervous system tumors is more refined(Bai et al.,2020). However, the clinical

outcomes and side effects of patients with same grade and classification of tumors are not the same after being comprehensively treated(Snape et al.,2015). This suggests that we still need to explore more instructive molecular targets in the research of glioma with the consideration of unclear and complex molecular mechanisms.

The underlying cause of malignant tumors is the disorder of gene expression system, including oncogenes, tumor suppressor genes and genes related to DNA repair(Hanahan et al.,2011). With the continuous development of epigenetics, it is gradually recognized that almost all malignant tumors have epigenetic abnormalities, which together with gene changes cause tumorigenesis(Dawson et al.,2012). The epigenetic phenomena involved in the occurrence of malignant tumors mainly include abnormal DNA methylation, histone modification and their interaction caused by abnormal expression of non-coding RNA and chromosomal remodeling(Nebbioso et al.,2018). These epigenetic changes lead to abnormal activation of certain genes and silencing, thereby allowing cell growth to enter an uncontrolled state. The occurrence and development of tumors are closely related to histone deacetylases (HDACs)(Hai et al.,2021). Studies have shown that genome-wide histone acetylation levels are generally reduced in tumor cells, among which HDAC1, 5 and 7 are regarded as tumor markers(Mahmud et al.,2015). Second, studies have shown that gene knockout of HDAC1/2 in breast cancer cells or HDAC1/2/3 in colon cancer cells can induce tumor cell apoptosis, suggesting that the activity of HDACs is related to tumor cell survival(Li et al.,2013;Thangaraju et al.,2009). Similarly, abnormal binding of HDACs to oncogene fusion proteins to certain gene loci is also regarded as an

important mechanism of tumorigenesis. The HDAC consists of four classes: class Ⅰ (HDAC1, HDAC2, HDAC3, HDAC8), class Ⅱ (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, HDAC10), class Ⅲ (SIRT1-SIRT7), and class Ⅳ (HDAC11)(Ropero et al.,2007). Previous study had explored the roles of HDACs in the prognosis of clear cell renal cell carcinoma(Cheng et al.,2021). However, the overall biology and prognosis are still unknown in glioma. In the present study, we comprehensively explored the biology function and prognosis of eleven HDAC genes in glioma, which may contribute the more understanding of molecular mechanisms and potential therapeutic targets for glioma patients.

Materials and methods

Data source

The mRNA expression data of glioma patients from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and Chinese Glioma Genome Atlas (CGGA, http://www.cgga.org.cn/). The gene mutation data were also from the TCGA database. Drug response data were available from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/downloads). The immune filtration data from TCIA (https://tcia.at/home). The histone deacetylases genes (HDAC1, HDAC2, HDAC2, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9, HDAC10, HDAC11) were from the molecular signatures database (MSigDB).

Clustering analysis

Using R “Consensus Cluster Plus” package, we performed the clustering analysis. The Consensus Cluster Plus allowed data clustering with negative value. Using

K-means method, we achieved the most approximate number of clusters by extracting 1000 times from 80% of sample size. The results were presented using consensus matrix heatmap. We also used the principal component analysis (PCA) and the t-distributed stochastic neighbor embedding (tSNE) to further validate the clustering analysis.

Development and validation of HDAC genes prognostic model

Using 11 HDAC genes, we developed an overall survival (OS) prognosis model in TCGA dataset. The Least Absolute Shrinkage and Selection Operator (LASSO) regression was used to select the number of entering the model, then a multivariate cox regression was performed to get the regression coefficient of each HDAC gene. We calculated the risk score of each sample using the following formula: risk score $=coef_1*gene1$ expression$+coef_2*$ $gene_2$ expression$+\ldots coef_n*gene_n$ expression. Using the established prognostic model, we performed the validation in CGGA dataset. Glioma patients were separated into high- and low-risk group according the median of risk score. Kaplan-Meier survival cure to compare the difference of OS between high-and low-risk group. The receiver operating characteristic curve (ROC) was used to evaluate the predictive ability at 1-year, 2-year, and 3-year of patients' OS of TCGA and CGGA. The PCA analysis was used to identify the risk type.

Clinical correlation and independent and analysis

To investigate the association between risk score and prognosis in glioma patients, we first performed the stratified analysis in different clinical parameters, then we compared the difference of HDAC genes between two clustering groups, and the risk

score distributions were also observed in different clinical parameters. To validate the independence of risk score, we performed univariate and multiple variate cox regression analyses by adjusting the clinic parameters (TCGA: age, gender, grade; CGGA: age, gender, grade, history, TNM stage, radiotherapy, chemotherapy, occurrence type, IDH and 1p19q status). We evaluated the diagnostic ability of risk score and other parameters using ROC. We built the nomograph to evaluate the clinical application of HDAC genes prognostic model, and the nomogram-predicted probability of 1-year OS, 3-year OS and 5-year OS were used to assess the model fitting ability.

Functional, pathway enrichment and mutations analysis

To explore the function and pathway enrichment of different high- and low-risk groups, we performed the GO functional enrichment and KEEG pathway analysis using the “clusterProfiler” package. Using the masked copy number segmentation data, we investigate the gene mutations frequency of different risk group using the “maftool” package (gene alteration, variant classification, variant type, co-occurrence and mutually exclusive).

Immune filtration, tumor microenvironment, and drug sensitivity analysis

We assessed the difference of 16 immune cells related infiltrating score and 13 immune-related pathways between high- and low-risk groups. Using TCGA dataset, we calculated the immune and stromal estimate scores of high- and low-risk groups using R “estimate” package. To explore the correlation between small molecular drugs and the identified prognostic signature genes, Pearson correlation coefficients

were calculated. |R|>0.25 and P<0.05 were considered significantly correlated.

Validation of HDAC genes in glioma and non-tumor tissue

To validate the expression of six HDAC genes included the prognostic model (HDAC1, HDAC3, HDAC4, HDAC5, HDAC7, and HDAC9), we detected the expression of these HDAC genes in glioma and non-tumor tissues. mRNA expression data were collected from 23 samples of epilepsy patients and 157 tumor samples. The tissue collection was approved by the NCI IRB committee with informed consent obtained from all subjects(Sun et al.,2006). Quantitative polymerase chain reaction was performed to detect mRNA expression level of HDAC genes. These methods have been previously described. Briefly, Total RNA was isolated using TRIzol (Invitrogen). 1,000ng RNA was reverse-transcribed into cDNA in a total reaction volume of 20□l with PrimeScriptTM RT Reagent Kit according to manufacturer's instructions (Takara). The qPCR primer sequences were designed by Santa-cruz Biotechnology Co, Inc. Real-time PCR was performed using SYBR premix Taq via CFX96 Real-Time PCR Detection System (Bio-Rad, Richmond, CA, USA). The 2-ΔΔCT method was used to calculate the relative expression of mRNA.

Statistical analysis

Differentially expressed genes analysis were performed using "limma" package. Differences for category variables were performed using Chi-square test. The comparisons of OS curve were achieved using log-rank test. One-way ANOVA was used for comparing the differences of HDAC genes expression among non-tumor and different grade glioma. SNK methods was used for multiple comparisons. All statical

analysis were finished using R software 4.0.1, and $P<0.05$ was considered significant.

Results

Identification of two subclasses in glioma

The flow chart was plotted to comprehensively describe our study (Figure 1A). For 11 HDAC genes, the correlations among HDAC members were different. HDAC6 and HDAC8 showed positive association with other HDAC genes, while HDAC3 and HDAC4 showed negative correlation with other HDAC genes (Figure 1B). We performed clustering analysis using 11 HDAC genes (HDAC1-HDAC11). The consistency coefficient was calculated to achieve the optimal clustering number (K value), and k=2 was finally selected as optimal clustering number. The sharp and clear boundaries showed stable and robust clustering for glioma patients (Figure 1C). To validate the two subclasses, we further performed individual PCA and t-SNE with decreased dimensions of features. We found that glioma patients were well distributed at two components (Figure 1D and Supplementary material Table S1, Cluster 1 and Cluster 2), and t-SNE also suggested that the samples presented two-dimensional distribution model (Figure 1E). The Kaplan-Meier survival curve indicated that the cluster 2 had worse OS than cluster 1 (Figure 1F). The clustering group was also associated with some clinical parameters (age, gender and grade and survival outcomes). HDAC1, HDAC4, HDAC5, HDAC6 and HDAC10 were highly expressed in cluster2, and HDAC1, HDAC2, HDAC3, and HDAC7, and HDAC9 were significantly highly expressed in cluster 1 (Figure 1G).

Development and validation of HDAC genes prognostic model in glioma

We first developed a prognostic model of OS in TCGA training dataset. The univariate cox regression indicated that high expressions of HDAC1, HDAC3, HDAC7 and HDAC9 were associated with poor OS in glioma, and the elevated expression of HDAC4, HDAC5, HDAC6 and HDAC11 were associated with favorable prognosis in glioma (Figure 2A). The HDAC2, HDAC2 and HDAC10 seemed not to be related to prognosis. Furthermore, the LASSO regression identified six HDAC genes (HDAC1, HDAC3, HDAC4, HDAC5, HDAC7 and HDAC9) that entered the final model (Figure 2B and 2C). Risk score of each sample was calculated according to the following formula: risk score=$0.179*HDAC1_{expression} + 0.502*HDAC3_{expression} - 0.671*HDAC4_{expression} - 0.567*HDAC5_{expression} + 0.488*HDAC7_{expression} + 0.216*HDAC9_{expression}$ (Supplementary material Table S2). The glioma patients were categorized into high-risk group and low-risk group using the median of risk score (-0.360). The Kaplan-Meier curve showed that high-risk group had poorer OS than the low-risk group ($P<0.001$, Figure 2D). The risk score and survival time were also separately distributed (Figure 2E). PCA showed two obvious risk distribution patterns (Figure 2F).

Patients from CGGA dataset were used to validate the calculated risk score, and they were also separated into high- and low-risk groups according to the median calculated using the formula established in TCGA training set. Similarly, survival analysis suggested that high-risk groups had a poorer OS than low-risk group ($P<0.001$, Figure 2G), and risk score and survival time were also visually scattered (Figure 2H). Likewise, PCA showed two-dimensional distribution patterns (Figure 2I). The are

under the curves (AUC) of 1-year, 2-year and 3-year were 0.873, 0.884 and .904 in the TCGA training set, respectively (Figure 2J). The AUCs of 1-year, 2-year and 3-year were 0.705, 0.765, and 0.761 in the CGGA validation set (Figure 2K).

Stratified analysis

To further validate the prognostic model in different sub glioma patients, we performed stratified analysis in different sub-population. We found low-risk patients based on risk score had prolonged OS compared with high-risk group, which was not affected by different age, gender, histology, occurrence type, IDH codeletion status, 1p19q mutation, previous history of radiotherapy and chemotherapy (Figure 3A-3F, Figure 3H-3P). However, the OS showed insignificant differences for glioma patients with WHO II (Figure 3G), which means that the developed risk score may be inappropriate in such sub-population.

Clinical correlation and independent and analysis

The analysis of expression difference indicated that all eleven HDAC genes showed significant differences between two subclasses (Figure 4A). The chi-square indicated that the high-risk patients tend to be GBM, WHO III/IV, older, occurrence or secondary, and have IDH mutation and 1p19q codeletion ($P<0.05$). No significant differences between high-and low-risk groups were observed for radiotherapy ratio and gender ratio ($P>0.05$, Figure 4B). We further compared the expression levels of six HDAC genes included in the prognostic model, and found HDAC1 ($P<0.001$), HDAC3 ($P<0.001$), HDAC7($P<0.001$) and HDAC9 were highly expressed in high-risk group, while HDAC4 ($P<0.001$) and HDAC5 ($P<0.001$) were lowly

expressed in high-risk group (*P*<0.001). Then we compared the risk score differences among different clinical parameters. Our results indicated that patients with >41 years old, advanced WHO stage and grade had higher risk scores (*P*<0.005, Figure 4C-4F). Patients with IDH mutation and 1p19q codeletion had lower risk score (*P*<0.05, Figure 4H and 4I). However, the risk score showed no significant differences among different gender (Figure 4D), recurrent or secondary (Figure 4G), or radiotherapy status (Figure 4J). Patients who received chemotherapy also have higher risk score compared those without chemotherapy (Figure 4K).

To investigate whether risk score was an independent prognostic factor for glioma patients or not, we performed univariate and multiple variates cox regression in the TCGA training set and CGGA validation set. In TCGA dataset, the univariate and multiple variate cox regression indicated that risk score was associated with OS in glioma patients (univariate: HR=2.084, 95%CI: 1.890-2.297, P<0.001, Figure 5A; multiple: HR=1.425, 95%CI: 1.247-1.629, P<0.001, Figure 5B). The ROC result showed the risk score have optimal predictive ability (AUC=0.828) for 5-year OS (Figure 5C). Similarly, the elevated risk score also increased the risk of poor OS (univariate: HR=7.801 95%CI: 5.887-10.338, P<0.001, Figure 5D; multiple: HR=2.184, 95%CI: 1.484-3.213, P<0.001, Figure 5E). The AUC was 0.808 that was higher than any of other clinical parameters (Figure 5F). Besides, recurrence, advanced grade and age, were also risk factor for poor OS, while patients who received chemotherapy and have IDH mutation and 1p19q codeletion had favorable OS (Figure 5E).

Functional, pathway enrichment and mutations analysis

To explore the functional and pathway enrichment of high-and low-risk groups, we performed GO and KEGG analysis. We first identified potential differently expressed gens between high- and low-risk groups (Log fold change >1, $P<0.05$). We finally identified 2598 differently expressed genes, including 1723 upregulated genes and 875 downregulated genes in the high-risk groups (Supplementary material Table S3). GO enrichment analysis indicated that high risk group was mainly enriched in immune-related function in biological process, collagen and lumen in cellular component, , and antigen binding, extracellular matrix structural constituent, some regulator, receptors and inhibitors binding in molecular function (Figure 6A). The KEGG pathways analysis showed that the high-risk group was involved PI3K-Akt signaling pathways, AGE-RAGE signaling, HIF-1 signaling, Relaxin signaling, and p53 signaling. Focal adhesion, ECM-receptor and cytokine-cytokine receptor interaction, cell cycle and pyrimidine metabolism were also significantly enriched (Figure 6B). The occurrence of glioma was involved in the integrations of multiple molecular functions and signaling pathways.

We also explored the gene mutation differences between high-risk and low-risk groups. Our results indicated that there are significant differences in gene mutation between high- and low-risk groups. The high-risk group showed high gene alteration in EGFR, PTEN, FLG, PKHD1(Figure 7A), while the gene alteration rate of IDH, ATRX, CIC, was higher in the low-risk group than in the high-risk group (Figure 7B). High- and low-risk groups showed similar results in variant classification, variant type,

SNV class (Figure 7C and 7D). Furthermore, HYDIN-PI3CA, AHNAK2-SPTA1, COL6A3-PTEN, and IDH1-TP53 showed highly co-occurrence, and PTEN-TTN, IDH1-EGFR showed mutually exclusive in high-risk groups. The SSPO-HMCN1, LRP2, NIPBL, MYH1-CIC, TTN, MUC16, APOB, RYR2, DNMT3A, NIPBL-IDH2, APOB, NOTCH1, LRP2 showed highly co-occurrence (Figure 7E). The IDH2-TP53, PI3CA-TP53, CIC-TP53 showed mutually exclusive in low-risk group (Figure 7F).

Immune filtration, tumor microenvironment, and drug sensitivity analysis

To explore the immune response difference between high- and low-risk group, we compared the immune filtration cells and immune-related pathways between high- and low-risk groups. Our results indicated that the aDCs, B cells, CD8+ Tcells, iDCs, macrophages, NK cells, pDCS, T helper cells, Tfh, Th1 cell, Th2 cells, TIL, and Treg were highly expressed in high-risk group, and no significant differences were observed in DCs, Mast cells and neutrophils between different risk groups (Figure 8A). All immune-related pathways were highly enriched in high-risk group (Figure 8B). we also found that the ESTIMATE score, immune score, and stromal score were higher in the high-risk group than in the low-risk group (Figure 8C-8E). The Pearson correlation analysis indicated that macrophages M0, M1, and M2 showed positive associations with risk score (Figure 8F, 8I, and 8J), while monocytes, NK cells activated, mast cells activated showed negative associations with risk score (Figure 8G, 8H, 8K).

To explore the potential small molecular drug related with HDAC genes, we performed Pearson correlation analysis between HDAC genes expression and some

small molecular drug (Figure 9). Our results indicated that HDAC7 showed significant drug resistance with Selumetinib, Cobimetinib, Trametinib, PD-98059, Dabrafenib, and Dolastatin 10, while Everolimus, Rapamycin, and Temsirolimus showed positive associations with HDAC7. PX-316, Chelerythrine, Selumetinib were positively associated with HDAC4 expression. HDAC9 showed a negative correlation with By-product o. Chelerythrine and Acrichine were also positively associated with HDAC1. These results provided some potential clues for HDAC-targeted treatment in glioma.

Validation of HDAC genes in glioma and non-tumor tissue

The qPCR was adopted methods to detect the expression of HDAC1, HDAC3, HDAC4, HDAC5, HDAC7 and HDCA9 of non-tumor and glioma tissue. The results were presented in Figure 10. The expression levels of HDAC1 (Figure 10A), HDAC3 (Figure 10B), HDAC7 (Figure 10E), and HDAC9 (Figure 10F) was significantly elevated in the glioma patients compared with non-tumor group. However, the expression levels of HDAC4 and HDAC5 were lower in the glioma patients than in the non-tumor control groups (Figure 10 C and 10D). This result was consistent with the role of these HDAC genes in prognosis.

Discussion

HDACs, as key enzymes that catalyze the acetylation of histones, are involved in many processes such as the growth and proliferation of malignant tumor cells, expression regulation, etc(Rroji et al.,2021). In the epigenetic research of tumor occurrence and development, it has gradually attracted academic wide attention in the

world. Most of the current research focuses on the chemical modification and structural modification of the existing antitumor drugs with HDACs inhibitory activity to enhance the therapeutic effect of the drug and alleviate the toxic and side effects(Chen et al.,2020). There are currently few antitumor drugs designed to act on specific targets and specific pathways. At the same time, it is of great of need to explore the molecular signatures for more understanding of biological relationship between tumor genotype and phenotypes.

In the present study, we found that glioma patients can be divided into two subclasses based on eleven HDAC genes, and patients from two subclasses had markedly different survival outcomes. Then, using six HDAC genes (HDAC1, HDAC3, HDAC4, HDAC5, HDAC7, and HDAC9), we established a prognostic model in glioma patients, and this prognostic model was well validated in an independent cohort population. Furthermore, the calculated risk score from six HDACA genes expression was suggested to be an independent prognostic factor, which can predict the five-year overall survival of glioma patients well. High-risk patients can be attributed to multiple complex function and molecular signaling pathways, and the genes alterations of high- and low-risk patients were significantly different. We also found that different survival outcomes of high- and low- risk patients could be involved in the differences of immune filtration level and tumor microenvironment. Subsequently, we identified several small molecular compounds that could be favorable for glioma patients' treatment. And finally, we validated the expression levels of HDAC genes from prognostic model using glioma and non-tumor tissues

samples. Our study provided new and simple molecular subtypes and prognosis prediction methods, and added more understanding for biology and molecular mechanisms of glioma.

We identified six HDAC genes in established prognostic model. HDAC1 and HDAC3 belongs to Class I of HDAC. Previous studies had indicated that HDAC1 was over-expressed in diverse human malignancies, such as prostate cancer, breast cancer, liver cancer, and lung cancer(Burdelski et al.,2015;Dong et al.,2018;Rivas et al.,2021;Zhang et al.,2021). HDAC1 was also highly expressed in glioma tissue, and highly expression of glioma is associated with glioma cell proliferation, migration, invasion, angiogenesis, and poor prognosis(Fan et al.,2021). Besides, it was suggested that increased activation of HDAC1/2/6 and Sp1 underlies therapeutic resistance and tumor growth in glioblastoma(Yang et al.,2020). We also found expression of HDAC1 was elevated in glioma tissues and be associated with poor prognosis. HDAC3 has become a focus in recent research, many scholars worldwide have found that it plays a role of carcinogenesis in human tumors. After the expression of HDAC3 is reduced by the inhibitors, the growth and invasive ability of human glioma cell appears significantly weakened, which provides a new way for the cancer treatment(Zhong et al.,2018). HDAC4, HDAC5, HDAC7, and HDAC9 belong to Class II of HDACs(Hou et al.,2018). HDAC4 is frequently dysregulated in human malignancies, and we also confirmed the downregulated expression in glioma tissues. However, previous studies reported that HDAC4 was significantly upregulated in glioma tissues. The proliferation, adenosine triphosphate (ATP) levels and invasion

ability were substantially enhanced in U251 cells with HDAC4 overexpression, and suppressed in U251 cells with a knockdown of HDAC4 compared with that in U251 cells transfected with the negative control(Cai et al.,2018;Wang et al.,2019). This may be associated with glioma grade, stage and histology, and further research was required. Just like HDAC4, HDAC5 was also found to be lowly expressed in glioma tissue. HDAC7 plays an oncogene role in glioma. It was reported that ZNF326 could activate HDAC7 transcription by binding to a specific promoter region via its transcriptional activation domain and zinc-finger structures in glioma cell(Yu et al.,2019). Furthermore, ZNF326 was not only highly expressed in glioma but was also positively correlated with the expression of HDAC7, which identified the oncogene role of HDAC7(Wang et al.,2020). HDAC9, like most class II HDACs, has a conserved histone deacetylase domain, catalyzes the removal of acetyl moieties in the N-terminal tail of histones, and possesses a long regulatory N-terminal domain to interact with tissue-specific transcription factors and co-repressors, the amino-terminal domain contains highly conserved serine residues that are subjected to phosphorylation(Fernandez-Ruiz,2020). Signal-dependent phosphorylation of HDAC9 is a critical event that determines whether it is localized in the cytoplasm or nucleus. High expression of HDAC9 has been reported in many cancers(Liang et al.,2020;Linares et al.,2019;Xu et al.,2019). In glioma, the high expression of HDAC9 can promote the proliferation and tumor formation, and accelerated cell cycle in part by potentiating the EGFR signaling pathways(Zhang et al.,2016). With the emerging and rapid development of disciplines such as structural biomechanics and

computer-aided drug design, the development of new HDACs inhibitors with antitumor activity targeting HDACs is bound to have a very broad development space and development prospects.

Sum it up, our results reveled the clinical utility and potential molecular mechanisms of HDAC genes in glioma. Model based on six HDAC genes (HDAC1, HDAC3, HDAC4, HDAC5, HDAC7, and HDAC9) can predict the overall survival of glioma patients well, which can be served as potential therapeutic targets. The future research should validate this model in a large cohort population, and experiments in vivo and vitro will favor the understanding of molecular mechanisms of glioma.

Authors' contributions

ZZL designed this study and directed the research group in all aspects, including planning, execution, and analysis of the study. LAB drafted the manuscript. LYY, NL, SL, LAB collected the data. LZZ provided the statistical software, performed the data analysis, SL arranged the Figures and Tables. SLF revised the manuscript. All authors have read and approved the final version of the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (LZZ: 82003239), and Science Foundation of Xiangya Hospital for Young Scholar (LZZ: 2018Q012), and National Natural Science Foundation of China (SLF:81974466).

Conflict of interest

The authors declare that they have no conflict of interest.

Data availability Statements

The TCGA data can be obtained from the https://portal.gdc.cancer.gov/, and CGGA data can be available from the Chinese Glioma Genome Atlas (http://www.cgga.org.cn/). The expression levels of HDAC genes from non-tumor and glioma patients can be available from the GEO (GSE4290).

References

Bai, J., Varghese, J., Jain, R. (2020). Adult glioma WHO classification update, genomics, and imaging: What the radiologists need to know. *Top Magn Reson Imaging* 29, 71-82. doi: 10.1097/RMR.0000000000000234

Burdelski, C., Ruge, O. M., Melling, N., Koop, C., Simon, R., and Steurer, S., et al. (2015). HDAC1 overexpression independently predicts biochemical recurrence and is associated with rapid tumor cell proliferation and genomic instability in prostate cancer. *Exp. Mol. Pathol.* 98, 419-426. doi: 10.1016/j.yexmp.2015.03.024

Cai, J. Y., Xu, T. T., Wang, Y., Chang, J. J., Li, J., and Chen, X. Y., et al. (2018). Histone deacetylase HDAC4 promotes the proliferation and invasion of glioma cells. *Int. J. Oncol.* 53, 2758-2768. doi: 10.3892/ijo.2018.4564

Chen, R., Smith-Cohn, M., Cohen, A. L., Colman, H. (2017). Glioma subclassifications and their clinical significance. *Neurotherapeutics* 14, 284-297. doi: 10.1007/s13311-017-0519-x

Chen, X., He, Y., Fu, W., Sahebkar, A., Tan, Y., and Xu, S., et al. (2020). Histone deacetylases (HDACs) and atherosclerosis: A mechanistic and pharmacological review. *Front Cell Dev Biol* 8, 581015. doi: 10.3389/fcell.2020.581015

Cheng, F., Zheng, B., Wang, J., Zhao, G., Yao, Z., and Niu, Z., et al. (2021). Comprehensive analysis of a new prognosis signature based on histone deacetylases in clear cell renal cell carcinoma. *Cancer Med* 10, 6503-6514. doi: 10.1002/cam4.4156

Dawson, M. A., Kouzarides, T. (2012). Cancer epigenetics: From mechanism to therapy. *Cell* 150, 12-27. doi: 10.1016/j.cell.2012.06.013

Dong, Z. Y., Zhou, Y. R., Wang, L. X. (2018). HDAC1 is indirectly involved in the epigenetic regulation of p38 MAPK that drive the lung cancer progression. *Eur Rev Med Pharmacol Sci* 22, 5980-5986. doi: 10.26355/eurrev_201809_15932

Fan, Y., Peng, X., Wang, Y., Li, B., Zhao, G. (2021). Comprehensive analysis of HDAC family identifies HDAC1 as a prognostic and immune infiltration indicator and HDAC1-Related signature for prognosis in glioma. *Front Mol Biosci* 8, 720020. doi: 10.3389/fmolb.2021.720020

Fernandez-Ruiz, I. (2020). HDAC9 linked to aortic calcification. *Nat. Rev. Cardiol.* 17, 6-7. doi: 10.1038/s41569-019-0308-9

Hai, R., He, L., Shu, G., Yin, G. (2021). Characterization of histone deacetylase mechanisms in cancer development. *Front Oncol* 11, 700947. doi: 10.3389/fonc.2021.700947

Hanahan, D., Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* 144, 646-674. doi: 10.1016/j.cell.2011.02.013

Hou, F., Li, D., Yu, H., Kong, Q. (2018). The mechanism and potential targets of class II HDACs in angiogenesis. *J. Cell. Biochem.* 119, 2999-3006. doi:

10.1002/jcb.26476

Kan, L. K., Drummond, K., Hunn, M., Williams, D., O'Brien, T. J., and Monif, M. (2020). Potential biomarkers and challenges in glioma diagnosis, therapy and prognosis. *BMJ Neurol Open* 2, e69. doi: 10.1136/bmjno-2020-000069

Li, S., Paterno, G. D., Gillespie, L. L. (2013). Nuclear localization of the transcriptional regulator MIER1alpha requires interaction with HDAC1/2 in breast cancer cells. *PLoS One* 8, e84046. doi: 10.1371/journal.pone.0084046

Liang, Z., Feng, A., Shim, H. (2020). MicroRNA-30c-regulated HDAC9 mediates chemoresistance of breast cancer. *Cancer Chemother Pharmacol* 85, 413-423. doi: 10.1007/s00280-019-04024-9

Linares, A., Assou, S., Lapierre, M., Thouennon, E., Duraffourd, C., and Fromaget, C., et al. (2019). Increased expression of the HDAC9 gene is associated with antiestrogen resistance of breast cancers. *Mol Oncol* 13, 1534-1547. doi: 10.1002/1878-0261.12505

Mahmud, I., Liao, D. (2015). Microarray gene expression profiling reveals potential mechanisms of tumor suppression by the class I HDAC-selective benzoylhydrazide inhibitors. *Genom Data* 5, 257-259. doi: 10.1016/j.gdata.2015.06.019

Nebbioso, A., Tambaro, F. P., Dell'Aversana, C., Altucci, L. (2018). Cancer epigenetics: Moving forward. *PLoS Genet.* 14, e1007362. doi: 10.1371/journal.pgen.1007362

Rivas, M., Johnston, M. N., Gulati, R., Kumbaji, M., Margues, A. T., and Timchenko, L., et al. (2021). HDAC1-Dependent repression of markers of hepatocytes and p21 is

involved in development of pediatric liver cancer. *Cell Mol Gastroenterol Hepatol* 12, 1669-1682. doi: 10.1016/j.jcmgh.2021.06.026

Ropero, S., Esteller, M. (2007). The role of histone deacetylases (HDACs) in human cancer. *Mol Oncol* 1, 19-25. doi: 10.1016/j.molonc.2007.01.001

Rroji, O., Kumar, A., Karuppagounder, S. S., Ratan, R. R. (2021). Epigenetic regulators of neuronal ferroptosis identify novel therapeutics for neurological diseases: HDACs, transglutaminases, and HIF prolyl hydroxylases. *Neurobiol. Dis.* 147, 105145. doi: 10.1016/j.nbd.2020.105145

Snape, T. J., Warr, T. (2015). Approaches toward improving the prognosis of pediatric patients with glioma: Pursuing mutant drug targets with emerging small molecules. *Semin. Pediatr. Neurol.* 22, 28-34. doi: 10.1016/j.spen.2014.12.003

Sun, L., Hui, A. M., Su, Q., Vortmeyer, A., Kotliarov, Y., and Pastorino, S., et al. (2006). Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. *Cancer Cell* 9, 287-300. doi: 10.1016/j.ccr.2006.03.003

Thangaraju, M., Carswell, K. N., Prasad, P. D., Ganapathy, V. (2009). Colon cancer cells maintain low levels of pyruvate to avoid cell death caused by inhibition of HDAC1/HDAC3. *Biochem. J.* 417, 379-389. doi: 10.1042/BJ20081132

Wang, M., Han, Q., Su, Z., Yu, X. (2020). Transcription factor ZNF326 upregulates the expression of ERCC1 and HDAC7 and its clinicopathologic significance in glioma. *Labmedicine* 51, 377-384. doi: 10.1093/labmed/lmz075

Wang, Y., Xia, Y., Hu, K., Zeng, M., Zhi, C., and Lai, M., et al. (2019). MKK7 transcription positively or negatively regulated by SP1 and KLF5 depends on

HDAC4 activity in glioma. *Int. J. Cancer* 145, 2496-2508. doi: 10.1002/ijc.32321

Weller, M., Wick, W., Aldape, K., Brada, M., Berger, M., and Pfister, S. M., et al. (2015). Glioma. *Nat Rev Dis Primers* 1, 15017. doi: 10.1038/nrdp.2015.17

Xu, G., Li, N., Zhang, Y., Zhang, J., Xu, R., and Wu, Y. (2019). MicroRNA-383-5p inhibits the progression of gastric carcinoma via targeting HDAC9 expression. *Braz. J. Med. Biol. Res.* 52, e8341. doi: 10.1590/1414-431X20198341

Yang, W. B., Hsu, C. C., Hsu, T. I., Liou, J. P., Chang, K. Y., and Chen, P. Y., et al. (2020). Increased activation of HDAC1/2/6 and Sp1 underlies therapeutic resistance and tumor growth in glioblastoma. *Neuro Oncol* 22, 1439-1451. doi: 10.1093/neuonc/noaa103

Yu, X., Wang, M., Wu, J., Han, Q., Zhang, X. (2019). ZNF326 promotes malignant phenotype of glioma by up-regulating HDAC7 expression and activating Wnt pathway. *J Exp Clin Cancer Res* 38, 40. doi: 10.1186/s13046-019-1031-4

Zhang, Y., Nalawansha, D. A., Herath, K. E., Andrade, R., Pflum, M. (2021). Differential profiles of HDAC1 substrates and associated proteins in breast cancer cells revealed by trapping. *Mol Omics* 17, 544-553. doi: 10.1039/d0mo00047g

Zhang, Y., Wu, D., Xia, F., Xian, H., Zhu, X., and Cui, H., et al. (2016). Downregulation of HDAC9 inhibits cell proliferation and tumor formation by inducing cell cycle arrest in retinoblastoma. *Biochem Biophys Res Commun* 473, 600-606. doi: 10.1016/j.bbrc.2016.03.129

Zhong, S., Fan, Y., Wu, B., Wang, Y., Jiang, S., and Ge, J., et al. (2018). HDAC3 expression correlates with the prognosis and grade of patients with glioma: A

diversification analysis based on transcriptome and clinical evidence. *World Neurosurg*. 119, e145-e158. doi: 10.1016/j.wneu.2018.07.076

Figure legends

Figure 1 Glioma patients can be separated into two subclasses using HDAC genes. **A:** The flow chart of data analysis. **B:** The correlation circle plot among eleven HDAC genes. **C:** The consensus matrix plot identified the best grouping (k=2). **C.** Principal component analysis of glioma subclasses in the TCGA dataset. **D:** The corrected t-SNE2 analysis for two subclasses. **E:** The Kaplan-Meier survival curve for two subclasses in TCGA dataset.

Figure 2 Development and validation of prognostic model based on HDAC genes. **A:** Forest plot of univariate cox regression for HDAC genes in glioma patients. **B:** LASSO regression of the 11 OS-related HDAC genes. **C:** Cross-validation for turning parameters selection in the LASSO regression. **D:** Kaplan-Meier survival curve of high- and low-risk groups from developed prognostic model based on 6 HDAC genes in TCGA. **E:** Distributions of risk scores and survival time of glioma patients in TCGA. **F:** PCA plot for high- and low-risk group in TCGA. **G:** Kaplan-Meier survival curve of high- and low-risk groups from validated prognostic model based on 6 HDAC genes in CGGA. **H:** Distributions of risk scores and survival time of glioma patients in CGGA. **I:** PCA plot for high- and low-risk group in CGGA. **J and K:** The receiver operating characteristic curve for predicting 1-year, 2-year, and 3-year survival rate of glioma patients in TCGA and CGGA.

Figure 3 Stratified analyses of established HDAC-related genes prognostic model in TCGA. **A and B:** Age (>41 vs <=41). **C and D:** Gender (male vs female). **E and F:** Histology (LGG vs GBM). **G, H and I:** WHO stage (II, III and IV). **J-L:** Type of tumors (Primary, secondary vs recurrent). **M and N:** 1p19q (Non-codel and codel). **O and P:** mutant and wildtype. **Q and R:** Radiotherapy (Yes vs No). **S and T:** Chemotherapy (Yes vs No).

Figure 4 Association between HDAC genes and clinical characteristics in glioma patients. **A:** Heatmap indicated the expression of HDAC genes between two subclasses. **B:** Heatmap of associations among risk stratifications and clinical parameters and six HDAC gens expression. Comparisons of risk score among different clinical parameters: **C:** age (>41 vs <=41), **D:** gender (male vs female), **E:** WHO stage (II, III, IV). **F:** histology (LGG vs GBM). **G:** PRS type (primary, recurrent, and secondary). **H:** IDH mutation status (mutant vs wild type). **I:** 1p19q codeletion status (codel vs non-codel). **J:** radiotherapy status (No vs Yes). **K:** chemotherapy (No vs Yes).

Figure 5 Independent prognosis analyses of HDAC-related genes model. **A and B:** univariate and multivariate cox regression of risk score based on HDAC genes in TCGA. **C:** The receiver operating characteristic curve of risk score for predicting 5-year survival rate in TCGA. **D and E:** univariate and multivariate cox regression of risk score based on HDAC genes in CGGA. **F:** The receiver operating characteristic curve of risk score for predicting 5-year survival rate in CGGA. **G-I:** Calibration curves of 1-eyar, 3-year, and 5-year OS in TCGA. **K:** Nomograph model established

in CGGA cohort.

Figure 6 GO enrichment (**A**) and KEGG pathways analysis (**B**) based on differently expressed genes between high- and low risk groups.

Figure 7 Landscape of mutation profiles between high- and low-risk groups. **A and B:** Waterfall plots of mutation information in each sample. **C and D:** Bar graph of variant classification. **E and F:** somatic interactions plot (co-occurrence and exclusive)

Figure 8 Immune status analysis between high- and low-risk group. **A:** The ssGSEA scores of immune cells. **B:** The ssGSEA scores of immune-related functions. **C-E:** Comparisons of Estimated, immune and stromal score between high-and low-risk group. **F-K:** Correlation between risk score and immune markers (Macrophages M0, Monocytes, NK cells activated, Macrophages M1, M2, and Mast cells activated) in glioma patients.

Figure 9 Top 16 kinds of drug associated with HDAC member.

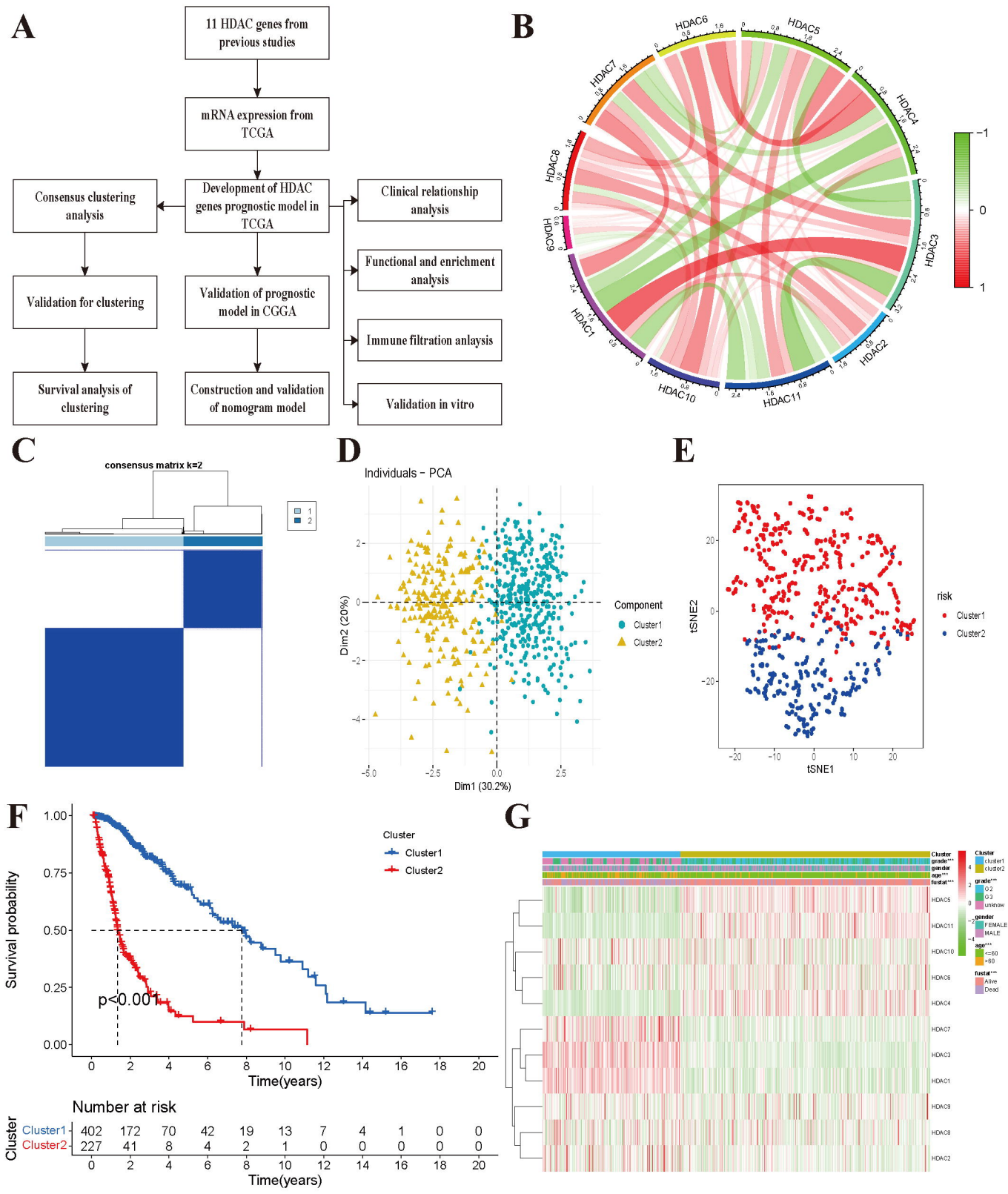
Figure 10 Expression of HDAC genes in glioma and non-tumor tissue. **A:** HDAC1, **B:** HDAC3, **C:** HDAC4, **D:** HDAC5, **E:** HDAC7, **F:** HDAC9

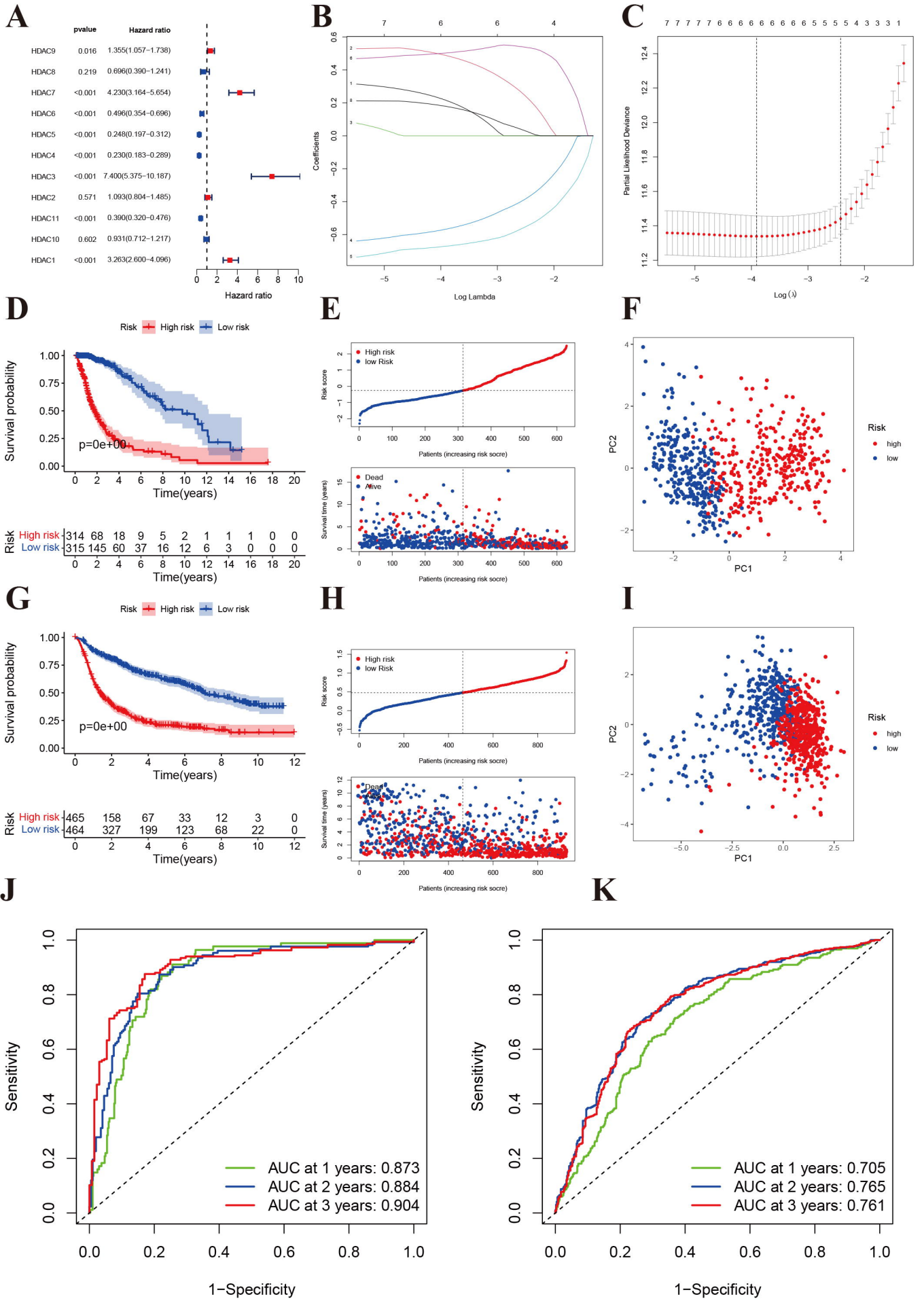
Supplementary materials

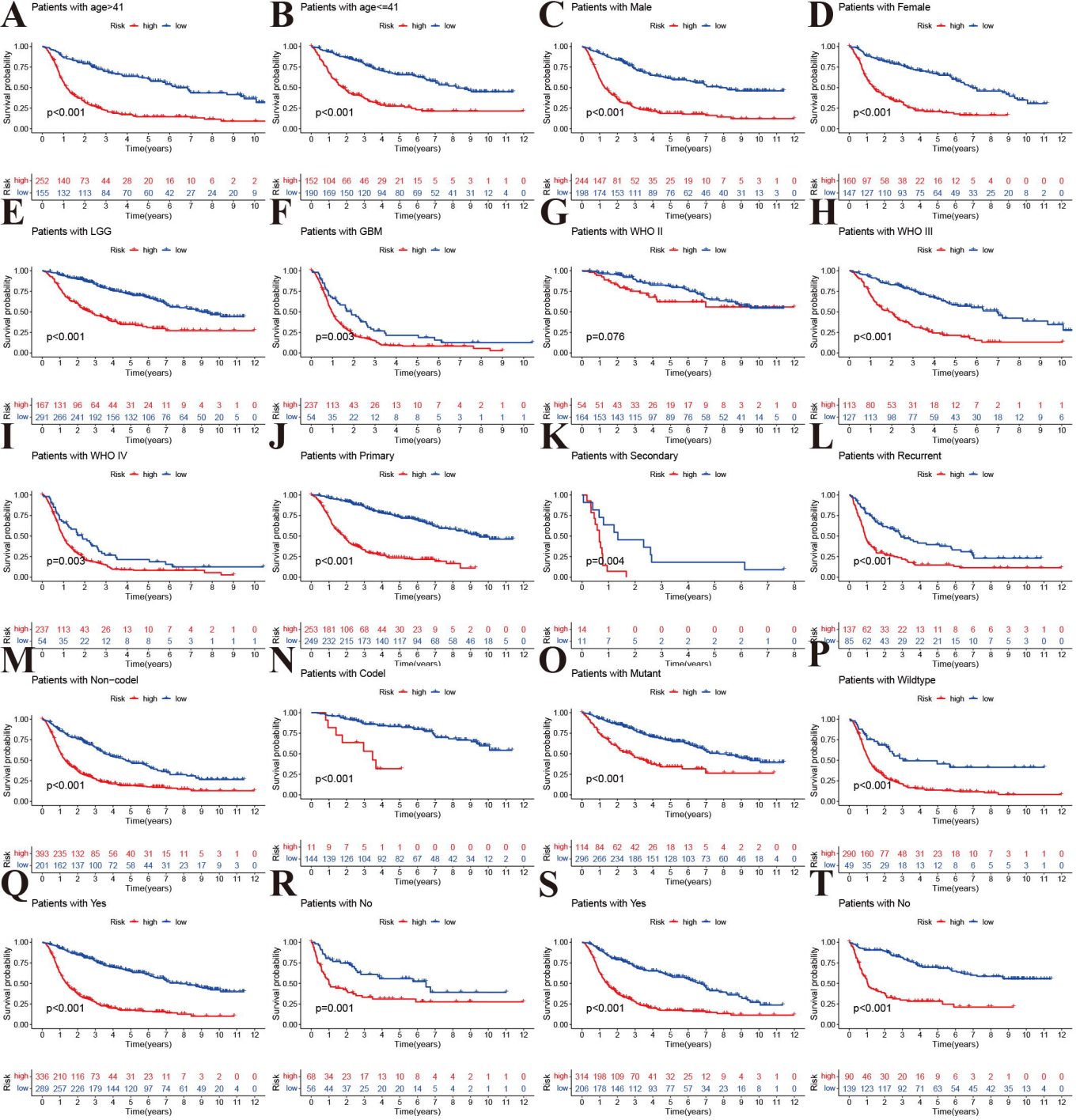
Table S1 Clustering of glioma based on 11 HDAC genes in TCGA

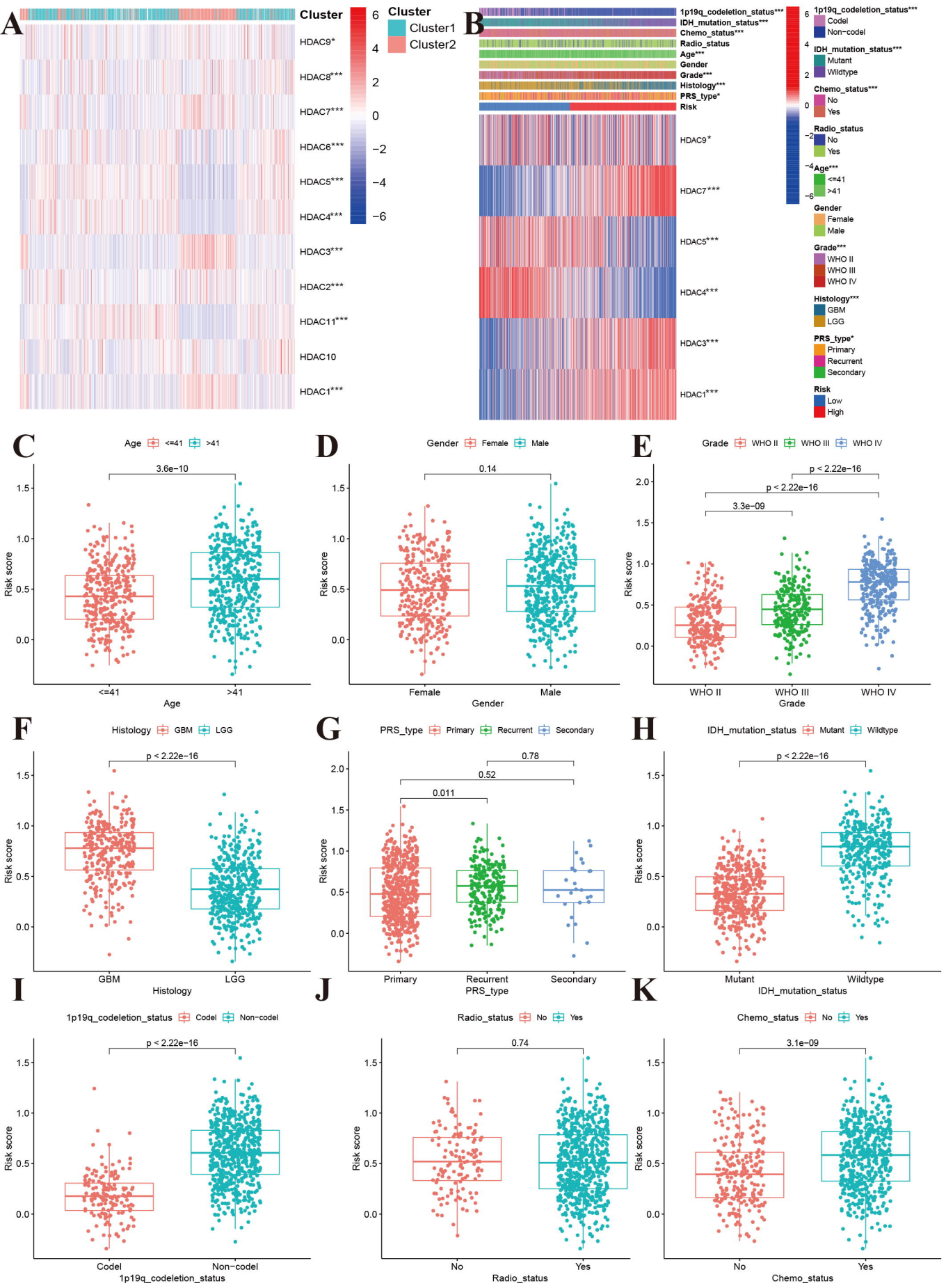
Table S2 Coefficient of HDAC in the included model

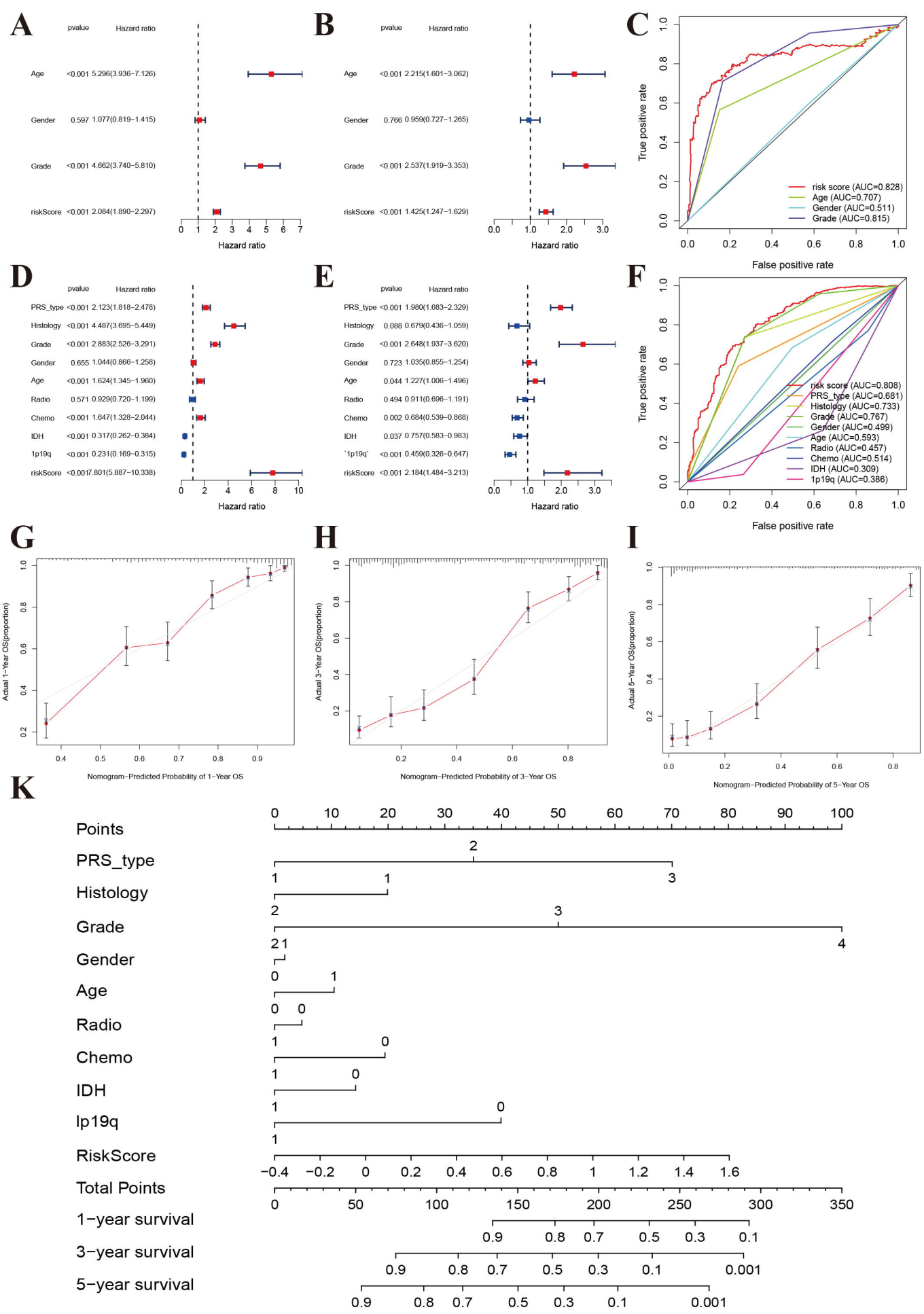
Table S3 Differential expressed genes between high-and low-risk group

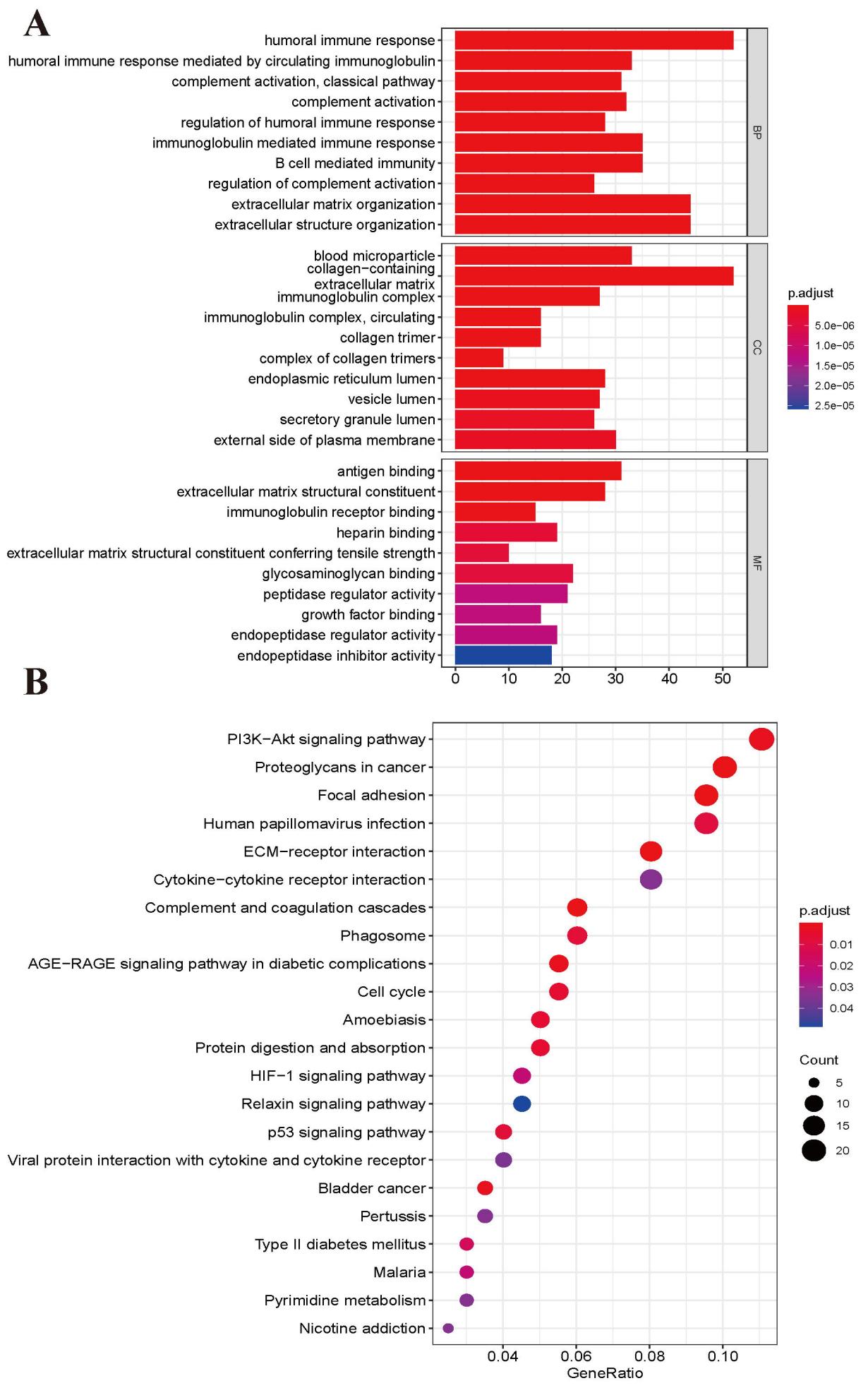

A
11 HDAC genes from previous studies
mRNA expression from TCGA
Development of HDAC genes prognostic model in TCGA
Consensus clustering analysis
Validation for clustering
Survival analysis of clustering
Validation of prognostic model in CGGA
Construction and validation of nomogram model
Clinical relationship analysis
Functional and enrichment analysis
Immune filtration anlaysis
Validation in vitro
B
HDAC6
HDAC5
HDAC7
HDAC4
HDAC8
HDAC9
HDAC3
HDAC1
HDAC2
HDAC10
HDAC11
C
consensus matrix k=2
D
Individuals - PCA
Dim1 (30.2%)
Dim2 (20%)
Component
Cluster1
Cluster2
E
tSNE1
tSNE2
risk
Cluster1
Cluster2
F
Survival probability
Time(years)
Cluster
Cluster1
Cluster2
p<0.001
Number at risk
Cluster1 402 172 70 42 19 13 7 4 1 0 0
Cluster2 227 41 8 4 2 1 0 0 0 0 0
G
Cluster
grade***
gender
age***
fustat***
HDAC5
HDAC11
HDAC10
HDAC6
HDAC4
HDAC7
HDAC3
HDAC1
HDAC9
HDAC8
HDAC2

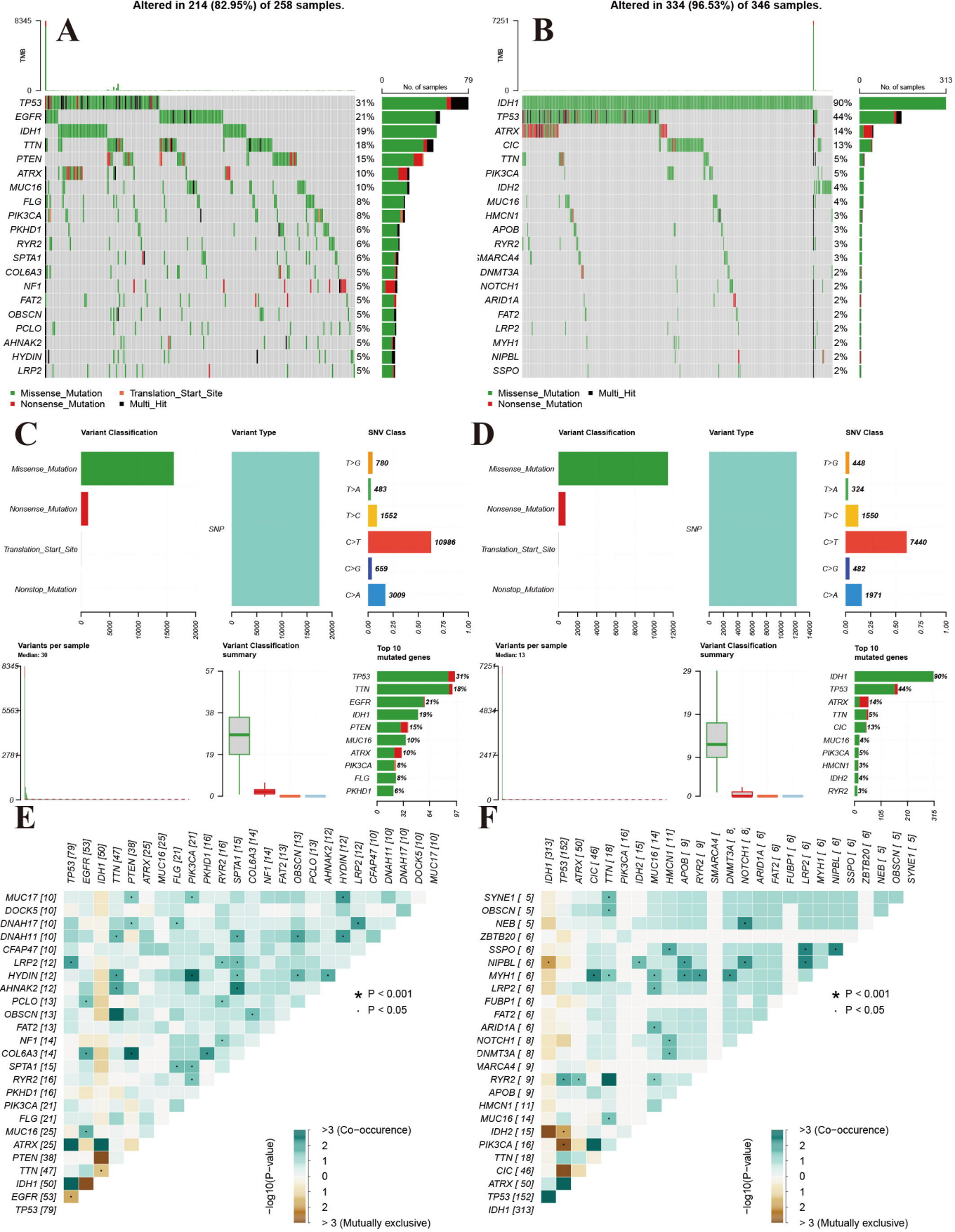

A
pvalue
Hazard ratio
HDAC9 0.016 1.355(1.057−1.738)
HDAC8 0.219 0.696(0.390−1.241)
HDAC7 <0.001 4.230(3.164−5.654)
HDAC6 <0.001 0.496(0.354−0.696)
HDAC5 <0.001 0.248(0.197−0.312)
HDAC4 <0.001 0.230(0.183−0.289)
HDAC3 <0.001 7.400(5.375−10.187)
HDAC2 0.571 1.093(0.804−1.485)
HDAC11 <0.001 0.390(0.320−0.476)
HDAC10 0.602 0.931(0.712−1.217)
HDAC1 <0.001 3.263(2.600−4.096)
Hazard ratio
B
Coefficients
Log Lambda
C
Partial Likelihood Deviance
Log (λ)
D
Risk High risk Low risk
Survival probability
p=0e+00
Time(years)
Risk High risk 314 68 18 9 5 2 1 1 1 0 0
Low risk 315 145 60 37 16 12 6 3 0 0 0
Time(years)
E
High risk
low Risk
Risk score
Patients (increasing risk socre)
Dead
Alive
Survival time (years)
Patients (increasing risk socre)
F
PC2
PC1
Risk
high
low
G
Risk High risk Low risk
Survival probability
p=0e+00
Time(years)
Risk High risk 465 158 67 33 12 3 0
Low risk 464 327 199 123 68 22 0
Time(years)
H
High risk
low Risk
Risk score
Patients (increasing risk socre)
Survival time (years)
Patients (increasing risk socre)
I
PC2
PC1
Risk
high
low
J
Sensitivity
1−Specificity
AUC at 1 years: 0.873
AUC at 2 years: 0.884
AUC at 3 years: 0.904
K
Sensitivity
1−Specificity
AUC at 1 years: 0.705
AUC at 2 years: 0.765
AUC at 3 years: 0.761

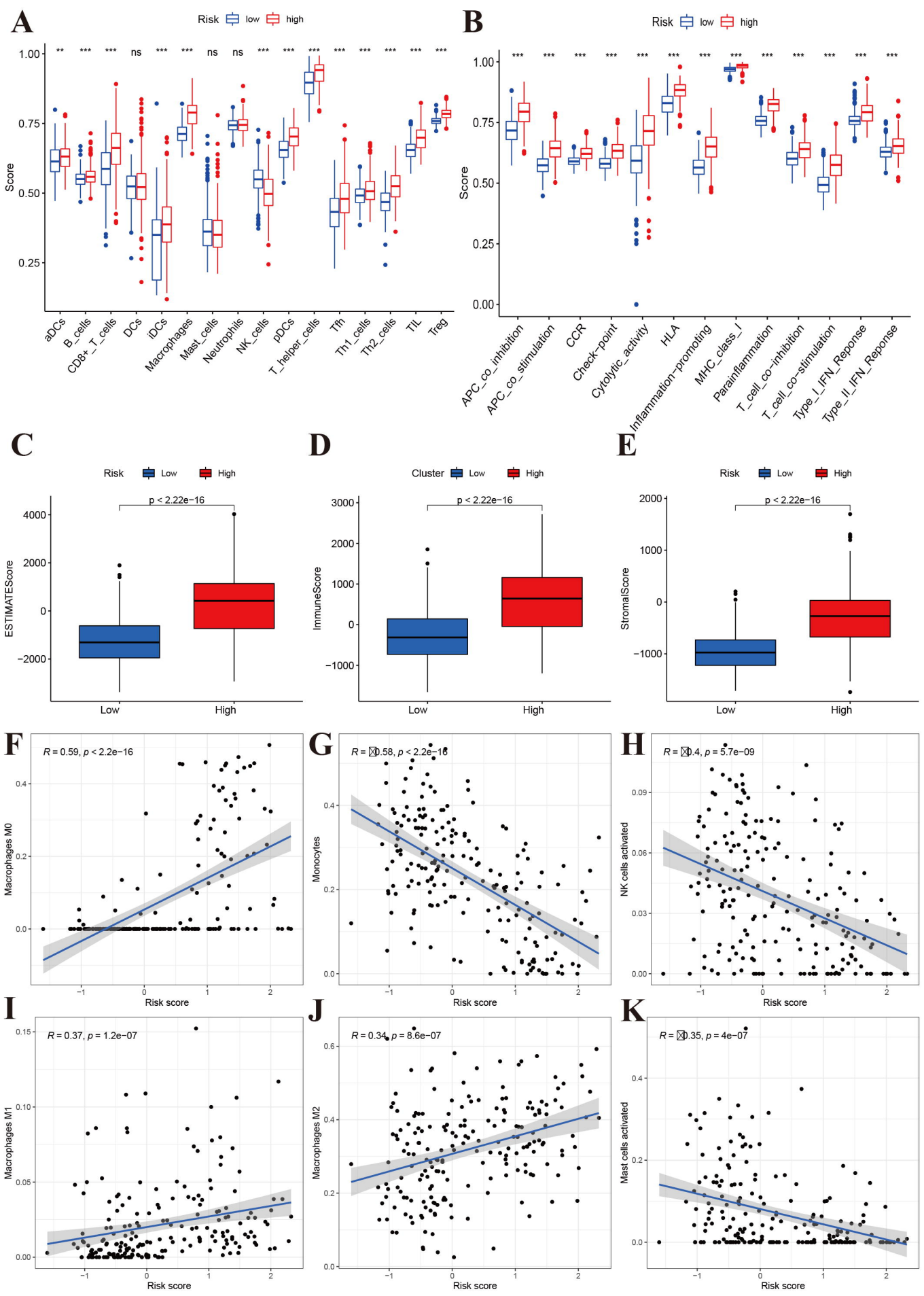

A
Patients with age>41
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 252 140 73 44 28 20 16 10 6 2 2
Risk low 155 132 113 84 70 60 42 27 24 20 9
B
Patients with age<=41
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 152 104 66 46 29 21 15 5 5 3 1 1 0
Risk low 190 169 150 120 94 80 69 52 41 31 12 4 0
C
Patients with Male
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 244 147 81 52 35 25 19 10 7 5 3 1 0
Risk low 198 174 153 111 89 76 62 46 40 31 13 3 0
D
Patients with Female
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 160 97 58 38 22 16 12 5 4 0 0 0 0
Risk low 147 127 110 93 75 64 49 33 25 20 8 2 0
E
Patients with LGG
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 167 131 96 64 44 31 24 11 9 4 3 1 0
Risk low 291 266 241 192 156 132 106 76 64 50 20 5 0
F
Patients with GBM
Risk high low
Survival probability
p=0.003
Time(years)
Risk high 237 113 43 26 13 10 7 4 2 1 0
Risk low 54 35 22 12 8 8 5 3 1 1 1
G
Patients with WHO II
Risk high low
Survival probability
p=0.076
Time(years)
Risk high 54 51 43 33 26 19 17 9 8 3 2 1 0
Risk low 164 153 143 115 97 89 76 58 52 41 14 5 0
H
Patients with WHO III
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 113 80 53 31 18 12 7 2 1 1 1
Risk low 127 113 98 77 59 43 30 18 12 9 6
I
Patients with WHO IV
Risk high low
Survival probability
p=0.003
Time(years)
Risk high 237 113 43 26 13 10 7 4 2 1 0
Risk low 54 35 22 12 8 8 5 3 1 1 1
J
Patients with Primary
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 253 181 106 68 44 30 23 9 5 2 0 0 0
Risk low 249 232 215 173 140 117 94 68 58 46 18 5 0
K
Patients with Secondary
Risk high low
Survival probability
p=0.004
Time(years)
Risk high 14 1 0 0 0 0 0 0 0 0
Risk low 11 7 5 2 2 2 2 1 0
L
Patients with Recurrent
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 137 62 33 22 13 11 8 6 6 3 3 1 0
Risk low 85 62 43 29 22 21 15 10 7 5 3 0 0
M
Patients with Non−codel
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 393 235 132 85 56 40 31 15 11 5 3 1 0
Risk low 201 162 137 100 72 58 44 31 23 17 9 3 0
N
Patients with Codel
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 11 9 7 5 1 1 0 0 0 0 0 0 0
Risk low 144 139 126 104 92 82 67 48 42 34 12 2 0
O
Patients with Mutant
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 114 84 62 42 26 18 13 5 4 2 2 0 0
Risk low 296 266 234 186 151 128 103 73 60 46 18 4 0
P
Patients with Wildtype
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 290 160 77 48 31 23 18 10 7 3 1 1 0
Risk low 49 35 29 18 13 12 8 6 5 5 3 1 0
Q
Patients with Yes
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 336 210 116 73 44 31 23 11 7 3 2 0 0
Risk low 289 257 226 179 144 120 97 74 61 49 20 4 0
R
Patients with No
Risk high low
Survival probability
p=0.001
Time(years)
Risk high 68 34 23 17 13 10 8 4 4 2 1 1 0
Risk low 56 44 37 25 20 20 14 5 4 2 1 1 0
S
Patients with Yes
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 314 198 109 70 41 32 25 12 9 4 3 1 0
Risk low 206 178 146 112 93 77 57 34 23 16 8 1 0
T
Patients with No
Risk high low
Survival probability
p<0.001
Time(years)
Risk high 90 46 30 20 16 9 6 3 2 1 0 0 0
Risk low 139 123 117 92 71 63 54 45 42 35 13 4 0

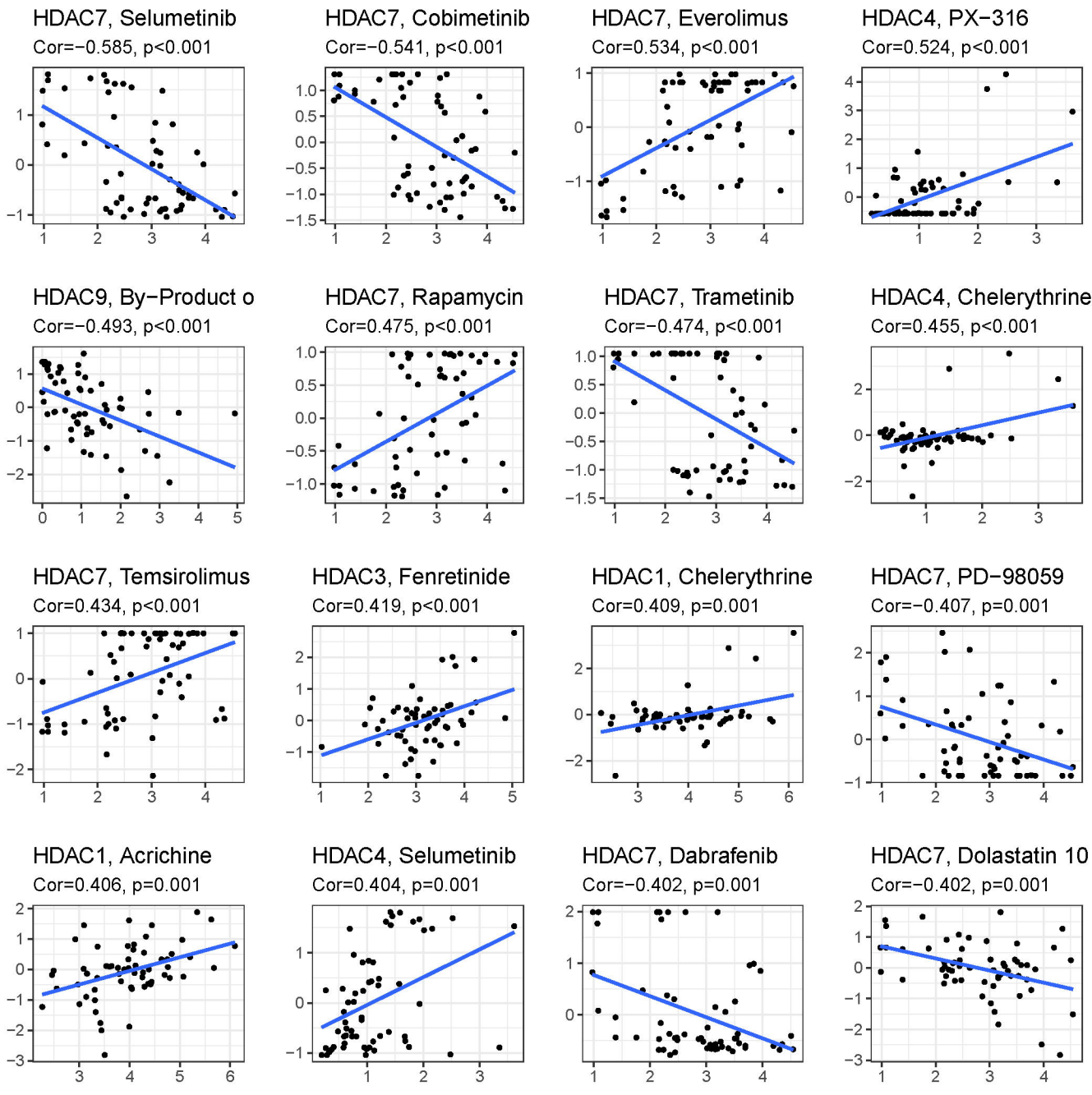

A
Cluster
Cluster1
Cluster2
HDAC9*
HDAC8***
HDAC7***
HDAC6***
HDAC5***
HDAC4***
HDAC3***
HDAC2***
HDAC11***
HDAC10
HDAC1***
B
1p19q_codeletion_status***
IDH_mutation_status***
Chemo_status***
Radio_status
Age***
Gender
Grade***
Histology***
PRS_type*
Risk
Codel
Non−codel
Mutant
Wildtype
No
Yes
<=41
>41
Female
Male
WHO II
WHO III
WHO IV
GBM
LGG
Primary
Recurrent
Secondary
Low
High
C
Age
3.6e−10
Risk score
D
Gender
0.14
E
Grade
p < 2.22e−16
3.3e−09
F
Histology
G
PRS_type
0.78
0.52
0.011
H
IDH_mutation_status
I
1p19q_codeletion_status
J
Radio_status
0.74
K
Chemo_status
3.1e−09

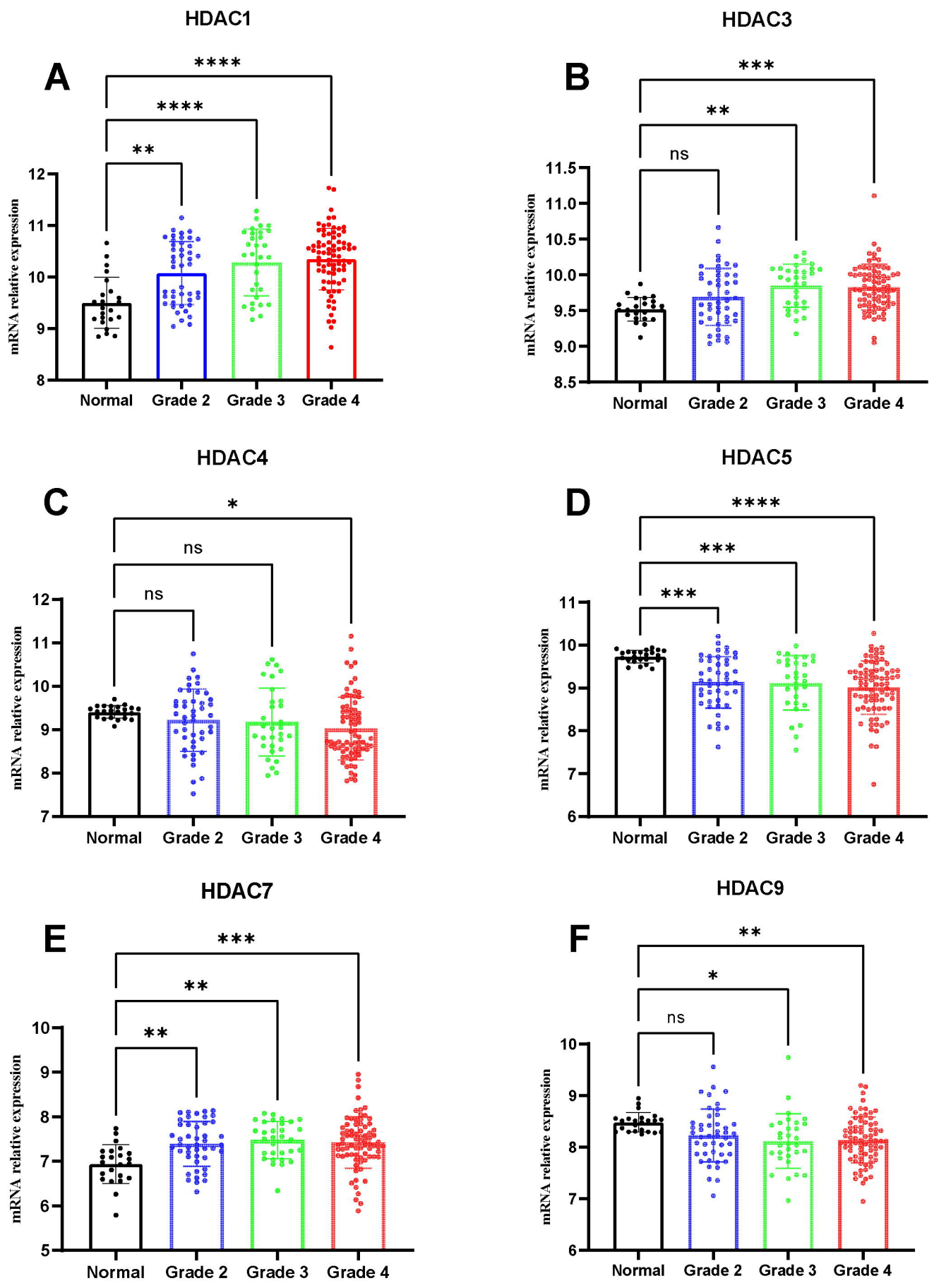

A
pvalue
Hazard ratio
Age <0.001 5.296(3.936−7.126)
Gender 0.597 1.077(0.819−1.415)
Grade <0.001 4.662(3.740−5.810)
riskScore <0.001 2.084(1.890−2.297)
Hazard ratio
B
pvalue
Hazard ratio
Age <0.001 2.215(1.601−3.062)
Gender 0.766 0.959(0.727−1.265)
Grade <0.001 2.537(1.919−3.353)
riskScore <0.001 1.425(1.247−1.629)
Hazard ratio
C
True positive rate
False positive rate
risk score (AUC=0.828)
Age (AUC=0.707)
Gender (AUC=0.511)
Grade (AUC=0.815)
D
pvalue
Hazard ratio
PRS_type <0.001 2.123(1.818−2.478)
Histology <0.001 4.487(3.695−5.449)
Grade <0.001 2.883(2.526−3.291)
Gender 0.655 1.044(0.866−1.258)
Age <0.001 1.624(1.345−1.960)
Radio 0.571 0.929(0.720−1.199)
Chemo <0.001 1.647(1.328−2.044)
IDH <0.001 0.317(0.262−0.384)
1p19q <0.001 0.231(0.169−0.315)
riskScore <0.001 7.801(5.887−10.338)
Hazard ratio
E
pvalue
Hazard ratio
PRS_type <0.001 1.980(1.683−2.329)
Histology 0.088 0.679(0.436−1.059)
Grade <0.001 2.648(1.937−3.620)
Gender 0.723 1.035(0.855−1.254)
Age 0.044 1.227(1.006−1.496)
Radio 0.494 0.911(0.696−1.191)
Chemo 0.002 0.684(0.539−0.868)
IDH 0.037 0.757(0.583−0.983)
1p19q <0.001 0.459(0.326−0.647)
riskScore <0.001 2.184(1.484−3.213)
Hazard ratio
F
True positive rate
False positive rate
risk score (AUC=0.808)
PRS_type (AUC=0.681)
Histology (AUC=0.733)
Grade (AUC=0.767)
Gender (AUC=0.499)
Age (AUC=0.593)
Radio (AUC=0.457)
Chemo (AUC=0.514)
IDH (AUC=0.309)
1p19q (AUC=0.386)
G
Actual 1−Year OS(proportion)
Nomogram−Predicted Probability of 1−Year OS
H
Actual 3−Year OS(proportion)
Nomogram−Predicted Probability of 3−Year OS
I
Actual 5−Year OS(proportion)
Nomogram−Predicted Probability of 5−Year OS
K
Points
PRS_type
Histology
Grade
Gender
Age
Radio
Chemo
IDH
lp19q
RiskScore
Total Points
1−year survival
3−year survival
5−year survival


A
humoral immune response
humoral immune response mediated by circulating immunoglobulin
complement activation, classical pathway
complement activation
regulation of humoral immune response
immunoglobulin mediated immune response
B cell mediated immunity
regulation of complement activation
extracellular matrix organization
extracellular structure organization
BP
blood microparticle
collagen-containing extracellular matrix
immunoglobulin complex
immunoglobulin complex, circulating
collagen trimer
complex of collagen trimers
endoplasmic reticulum lumen
vesicle lumen
secretory granule lumen
external side of plasma membrane
CC
antigen binding
extracellular matrix structural constituent
immunoglobulin receptor binding
heparin binding
extracellular matrix structural constituent conferring tensile strength
glycosaminoglycan binding
peptidase regulator activity
growth factor binding
endopeptidase regulator activity
endopeptidase inhibitor activity
MF
0
10
20
30
40
50
p.adjust
5.0e-06
1.0e-05
1.5e-05
2.0e-05
2.5e-05
B
PI3K-Akt signaling pathway
Proteoglycans in cancer
Focal adhesion
Human papillomavirus infection
ECM-receptor interaction
Cytokine-cytokine receptor interaction
Complement and coagulation cascades
Phagosome
AGE-RAGE signaling pathway in diabetic complications
Cell cycle
Amoebiasis
Protein digestion and absorption
HIF-1 signaling pathway
Relaxin signaling pathway
p53 signaling pathway
Viral protein interaction with cytokine and cytokine receptor
Bladder cancer
Pertussis
Type II diabetes mellitus
Malaria
Pyrimidine metabolism
Nicotine addiction
0.04
0.06
0.08
0.10
GeneRatio
p.adjust
0.01
0.02
0.03
0.04
Count
5
10
15
20


A
Altered in 214 (82.95%) of 258 samples.
TMB
No. of samples
TP53 31%
EGFR 21%
IDH1 19%
TTN 18%
PTEN 15%
ATRX 10%
MUC16 10%
FLG 8%
PIK3CA 8%
PKHD1 6%
RYR2 6%
SPTA1 6%
COL6A3 5%
NF1 5%
FAT2 5%
OBSCN 5%
PCLO 5%
AHNAK2 5%
HYDIN 5%
LRP2 5%
Missense_Mutation
Translation_Start_Site
Nonsense_Mutation
Multi_Hit
B
Altered in 334 (96.53%) of 346 samples.
IDH1 90%
TP53 44%
ATRX 14%
CIC 13%
TTN 5%
PIK3CA 5%
IDH2 4%
MUC16 4%
HMCN1 3%
APOB 3%
RYR2 3%
SMARCA4 3%
DNMT3A 2%
NOTCH1 2%
ARID1A 2%
FAT2 2%
LRP2 2%
MYH1 2%
NIPBL 2%
SSPO 2%
Missense_Mutation
Multi_Hit
Nonsense_Mutation
C
Variant Classification
Variant Type
SNV Class
Missense_Mutation
Nonsense_Mutation
Translation_Start_Site
Nonstop_Mutation
SNP
T>G 780
T>A 483
T>C 1552
C>T 10986
C>G 659
C>A 3009
Variants per sample
Median: 30
Variant Classification summary
Top 10 mutated genes
TP53 31%
TTN 18%
EGFR 21%
IDH1 19%
PTEN 15%
MUC16 10%
ATRX 10%
PIK3CA 8%
FLG 8%
PKHD1 6%
D
Variant Classification
Variant Type
SNV Class
Missense_Mutation
Nonsense_Mutation
Translation_Start_Site
Nonstop_Mutation
SNP
T>G 448
T>A 324
T>C 1550
C>T 7440
C>G 482
C>A 1971
Variants per sample
Median: 13
Variant Classification summary
Top 10 mutated genes
IDH1 90%
TP53 44%
ATRX 14%
TTN 5%
CIC 13%
MUC16 4%
PIK3CA 5%
HMCN1 3%
IDH2 4%
RYR2 3%
E
TP53 [79]
EGFR [53]
IDH1 [50]
TTN [47]
PTEN [38]
ATRX [25]
MUC16 [25]
FLG [21]
PIK3CA [21]
PKHD1 [16]
RYR2 [16]
SPTA1 [15]
COL6A3 [14]
NF1 [14]
FAT2 [13]
OBSCN [13]
PCLO [13]
AHNAK2 [12]
HYDIN [12]
LRP2 [12]
CFAP47 [10]
DNAH11 [10]
DNAH17 [10]
DOCK5 [10]
MUC17 [10]
* P < 0.001
· P < 0.05
−log10(P−value)
>3 (Co−occurence)
> 3 (Mutually exclusive)
F
IDH1 [313]
TP53 [152]
ATRX [50]
CIC [46]
TTN [18]
PIK3CA [16]
IDH2 [15]
MUC16 [14]
HMCN1 [11]
APOB [9]
RYR2 [9]
SMARCA4 [
DNMT3A [8]
NOTCH1 [8]
ARID1A [6]
FAT2 [6]
FUBP1 [6]
LRP2 [6]
MYH1 [6]
NIPBL [6]
SSPO [6]
ZBTB20 [6]
NEB [5]
OBSCN [5]
SYNE1 [5]
* P < 0.001
· P < 0.05
−log10(P−value)
>3 (Co−occurence)
> 3 (Mutually exclusive)


A
Risk low high
Score
aDCs
B_cells
CD8+_T_cells
DCs
iDCs
Macrophages
Mast_cells
Neutrophils
NK_cells
pDCs
T_helper_cells
Tfh
Th1_cells
Th2_cells
TIL
Treg
B
Risk low high
Score
APC_co_inhibition
APC_co_stimulation
CCR
Check-point
Cytolytic_activity
HLA
Inflammation-promoting
MHC_class_I
Parainflammation
T_cell_co-inhibition
T_cell_co-stimulation
Type_I_IFN_Reponse
Type_II_IFN_Reponse
C
Risk Low High
p < 2.22e-16
ESTIMATEScore
Low
High
D
Cluster Low High
p < 2.22e-16
ImmuneScore
Low
High
E
Risk Low High
p < 2.22e-16
StromalScore
Low
High
F
R = 0.59, p < 2.2e-16
Macrophages M0
Risk score
G
R = 0.58, p < 2.2e-16
Monocytes
Risk score
H
R = 0.4, p = 5.7e-09
NK cells activated
Risk score
I
R = 0.37, p = 1.2e-07
Macrophages M1
Risk score
J
R = 0.34, p = 8.6e-07
Macrophages M2
Risk score
K
R = 0.35, p = 4e-07
Mast cells activated
Risk score


HDAC7, Selumetinib
Cor=−0.585, p<0.001
HDAC7, Cobimetinib
Cor=−0.541, p<0.001
HDAC7, Everolimus
Cor=0.534, p<0.001
HDAC4, PX−316
Cor=0.524, p<0.001
HDAC9, By−Product o
Cor=−0.493, p<0.001
HDAC7, Rapamycin
Cor=0.475, p<0.001
HDAC7, Trametinib
Cor=−0.474, p<0.001
HDAC4, Chelerythrine
Cor=0.455, p<0.001
HDAC7, Temsirolimus
Cor=0.434, p<0.001
HDAC3, Fenretinide
Cor=0.419, p<0.001
HDAC1, Chelerythrine
Cor=0.409, p=0.001
HDAC7, PD−98059
Cor=−0.407, p=0.001
HDAC1, Acrichine
Cor=0.406, p=0.001
HDAC4, Selumetinib
Cor=0.404, p=0.001
HDAC7, Dabrafenib
Cor=−0.402, p=0.001
HDAC7, Dolastatin 10
Cor=−0.402, p=0.001


HDAC1
A
****
****
**
mRNA relative expression
12
11
10
9
8
Normal
Grade 2
Grade 3
Grade 4
HDAC3
B
***
**
ns
mRNA relative expression
11.5
11.0
10.5
10.0
9.5
9.0
8.5
Normal
Grade 2
Grade 3
Grade 4
HDAC4
C
*
ns
ns
mRNA relative expression
12
11
10
9
8
7
Normal
Grade 2
Grade 3
Grade 4
HDAC5
D
****
***
***
mRNA relative expression
11
10
9
8
7
6
Normal
Grade 2
Grade 3
Grade 4
HDAC7
E
***
**
**
mRNA relative expression
10
9
8
7
6
5
Normal
Grade 2
Grade 3
Grade 4
HDAC9
F
**
*
ns
mRNA relative expression
10
9
8
7
6
Normal
Grade 2
Grade 3
Grade 4