Multi-omics Prognostic Analysis of Lysine Acetylation Regulators in Glioma

Zewei Tu
The Second Affiliated Hospital of Nanchang University
Shigang Lv
The Second Affiliated Hospital of Nanchang University
Lei Wu
The Second Affiliated Hospital of Nanchang University
Qing Hu
The Second Affiliated Hospital of Nanchang University
Chuming Tao
The Second Affiliated Hospital of Nanchang University
Kuangxun Li
The Second Affiliated Hospital of Nanchang University
Xingen Zhu (zxg2008vip@163.com)
The Second Affiliated Hospital of Nanchang University https://orcid.org/0000-0002-8556-0941
Kai Huang
The Second Affiliated Hospital of Nanchang University

Research

Keywords: glioma, lysine acetylation regulator, epigenetic, prognostic signature, biomarker

Posted Date: June 25th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-37033/v1

Multi-omics prognostic analysis of lysine acetylation regulators in glioma

Zewei Tu[1†], Shigang Lv[1†], Lei Wu[1,3†], Qing Hu[1,2,3], Chuming Tao[1,2,3], Kuangxun Li[4], Xingen Zhu[1,3*], Kai Huang[1,2,3*]

[1] Department of Neurosurgery, The Second Affiliated Hospital of Nanchang University, Nanchang, Jiangxi 330006, P.R.China;

[2] East China Institute of Digital Medical Engineering, Shangrao, Jiangxi 334000, P.R.China.

[3] Institute of Neuroscience, Nanchang University, Nanchang, Jiangxi 330006, P.R.China

[4]College of Queen Mary, Nanchang University, Nanchang, Jiangxi 330031, P.R.China.

***Correspondence:**

Xingen Zhu

zxg2008vip@163.com

Huang Kai

beige666@163.com

† These authors contribute equally to this work

Abstract

Background: Lysine acetylation is a crucial kind of protein modification and is related to the malignant development of various cancers. But their roles in glioma are still unclear and needed concluded comprehensively.

Methods: In this study, we comprehensively analyzed the expression levels of 33 lysine acetylation regulators (LARs) and prognostic roles by using public data, including the Chinese Glioma Genome Atlas (CGGA) and The Cancer Genome Atlas (TCGA). The prognostic roles of LARs were judged by univariate Cox regression. Consensus clustering was applied to result in three stratified glioma subtypes (LA1, 2, and 3) with different clinical outcomes. We also constructed a risk signature for predicting the overall survival of glioma patients by using least absolute shrinkage and selection operator regression (LASSO regression). Besides, copy number variations (CNVs) and single nucleotide polymorphism (SNP) of LARs were also analyzed in our study.

Results: We found the mRNA expression levels of most of LARs were dysregulated in gliomas and associated with the prognosis of glioma patients. The risk signature constructed by 14 LARs presented an independent prognostic role in both the CGGA (HR:1.96, 95%CI:1.33-2.90) and TCGA (HR:1.48, 95%CI:1.08-2.03) datasets and robust predictive effects in the ROC curves with all of area under curves more than 0.800. Moreover, the copy number variations of LARs were also significantly related to the prognosis of glioma patients, in which *HDAC1*(1p) was one of the oncogenes lost in 1p/19q codeletion events, while *SIRT2*(19q) and *EP300*(22q) may act as tumor suppressors in gliomas with 19q or 22q deletions, respectively.

Conclusion: LARs are potential biomarkers for the malignant progression of gliomas, and our results could be useful for predicting the OS of glioma patients and provide some clues in searching the functions of LARs in glioma progression.

Keywords: glioma, lysine acetylation regulator, epigenetic, prognostic signature, biomarker.

Background

Glioma, the most common and fatal intracranial primary tumor in adults, is known for its rapid progression, high infiltration rate, and relative resistance to chemoradiotherapy [1, 2]. Although comprehensive integrated treatment programs are currently available, the clinical outcomes for glioma patients remain poor [1-3]. According to the Chinese Glioma Genome Atlas (CGGA), patients with malignant glioma have a dissatisfactory prognosis with median overall survival (OS) of 78.1 months for low-grade gliomas (LGGs; WHO grade II), 37.6 months for anaplastic gliomas, and 14.4 months for glioblastomas (GBMs) [1]. In recent years, numerous glioma neuropathological biomarkers and molecular stratification of glioma patients have been identified based on the rapid development of biomedical and bioinformatics technology. However, the identification of new and efficient prognostic and therapeutic biomarkers and targets remains a priority for glioma-tailored treatment.

Epigenetic regulation is essential for cellular homeostasis and its dysregulation is associated with a variety of cancers [4-6]. Post-translational modifications (PTMs) are key elements of epigenetic regulation and function as signaling markers within oncocytes [7, 8]. Lysine acetylation is a dynamic, reversible PTM that has been widely investigated in recent years due to its ubiquity as a mechanism for cellular protein modification that regulates numerous cellular biological processes, including transcription, cell cycle, cell division, DNA damage repair, cellular signaling transduction, protein folding and aggregation, cytoskeleton organization, RNA processing and stability [9, 10]. Both histone and non-histone proteins, such as p53, STAT proteins, NF-κB, FoxO proteins, and tubulins, are targeted by lysine acetylation regulators (LARs), and several are the products of oncogenes or tumor-suppressor genes and are directly involved in tumorigenesis, tumor progression, and metastasis [9, 11].

Lysine acetylation is dynamically regulated by “writers” (acetyltransferases) and “erasers” (deacetylases). The main LARs comprise of acetyltransferase families, including the GCN5 family (KAT2A and KAT2B), p300 family (KAT3A [CREBBP] and KAT3B [EP300]), MYST family (KAT5, KAT6A, KAT6B, KAT7, and KAT8), and others, such as the SLC16A10, KAT1 (HAT1), ESCO1, and ESCO2; and deacetylase families, including the histone deacetylase family (HDAC1–11), Sirtuin deacetylase family (SIRT1–7), and others, such as the TCF1 (HNF1A) and LEF1[9]. Increasing evidence supports that LARs directly or indirectly participate in cancer initiation and progression, which led us to explore the roles of acetylation in glioma in greater detail. Although numerous studies have investigated the acetylation-related molecular regulatory mechanisms in gliomas, the role of lysine acetylation in glioma is still poorly understood, and clarifying the effects of impaired regulation of lysine acetylation could pave the way for new therapeutic approaches to treat patients with these diseases.

In this study, we utilized RNA-seq data for 905 gliomas from the Chinese Glioma Genome Atlas (CGGA) (n = 307) and The Cancer Genome Atlas (TCGA) (n = 598) datasets, and matched copy number variation (CNV; n = 598) and single nucleotide polymorphism (SNP; n = 583) data from the TCGA dataset. Based on bioinformatic and statistical analyses of these open-source datasets, several LARs were found to be involved in malignant progression and prognosis of glioma, and a predictive independent risk signature involving 14 screened LARs was developed to predict the prognosis of

glioma patients. The results showed that several LARs were included in the frequent chromosome alterations observed in gliomas and show prognostic values. Tumor mutation burden (TMB) was also calculated for samples with mutation data in the TCGA dataset and we found TMB showed a positive correlation with our risk score, which may mean that DNA repair system is highly impaired in gliomas with higher risk score and dysregulation of lysine acetylation may lead to malignant progression in glioma.

Methods

Data Acquisition

The RNA-seq data and corresponding clinicopathological information for the CGGA training set were downloaded from the Chinese Glioma Genome Atlas (CGGA) (http://www.cgga.org.cn/). The RNA-seq data, CNV data, and clinicopathological data for the validation set in TCGA were downloaded from the University of California, Santa Cruz Xena browser (UCSC Xena; https://xenabrowser.net/datapages/). The mRNA expression levels in the two databases were detected in clinical tumor samples rather than detached glioma cells, which means the RNA-seq data represent the mRNA levels in glioma microenvironment. SNP data in the TCGA dataset were downloaded from the Genomic Data Commons Data Portal (GDC; https://portal.gdc.cancer.gov/). The RNA-seq transcriptome data for the CGGA and TCGA samples were normalized by log2(n+1) transformation. The GISTIC2 method was applied to generate gene-level copy number estimates. GISTIC2 further thresholded the estimated values to −2, −1, 0, 1, 2, representing homozygous deletion, single copy deletion, diploid normal copy, low-level copy number amplification, or high-level copy number amplification, respectively. Immunohistochemistry images of LARs were obtained from the website of The Human Protein Atlas (HPA: https://www.proteinatlas.org/) The clinicopathological information for the CGGA and TCGA datasets is summarized in Supplementary Table S1. Copy number variation information of the 33 LARs is summarized in Supplementary Table S2.

Selection of LARs

A list of LARs was compiled from the published literature and subsequently restricted to genes for which RNA expression data was available in both the CGGA and TCGA datasets. We obtained a final list of 33 LARs consisting of 13 lysine acetyltransferases and 20 lysine deacetylases. The extracted mRNA expression matrix of these 33 genes was used for the subsequent bioinformatics analysis.

Bioinformatic Analysis

Consensus clustering and screening of molecular subtypes based on the expression profiles of the LARs were performed using the R package "ConsensusClusterPlus" [12]. The Euclidean distance was utilized to compute the similarity distance between samples, and the k-means method was used for clustering based on 50 iterations, with each iteration containing 80% of samples. Then principal component analysis (PCA) was performed to evaluate different expression patterns among glioma subgroups using the R programming language (https://www.r-project.org/). Differential gene expression analysis between the LA3 and LA1/2 subgroups was performed using the R package "limma" [13], and the differentially expressed genes were input into the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) for GO and KEGG pathway enrichment analyses. GSEA software (http://software.broadinstitute.org/gsea/index.jsp) was used to investigate the enriched tumor hallmarks in the LA3 subgroup compared with those in the LA1/2 subgroups. Interactions among LARs were evaluated using the STRING database (https://string-db.org/), and the "Cytoscape" software was used to perform the protein-protein interaction network analysis[14]. The

tumor mutation burden (TMB) was calculated using Perl scripts (https://www.perl.org/), and the algorithm to calculate the TMB included nonsynonymous mutation counts per tumor, with germline mutations filtered out.

Based on the expression of 33 LARs in the CGGA dataset, univariate Cox regression analyses were first used to judge their prognostic power. We screened 23 genes associated with OS ($p < 0.05$) and used the LASSO Cox regression algorithm to develop a potential risk signature. Finally, 14 genes with their coefficients were determined according to minimum criteria, which involved selecting the best penalty parameter λ associated with the smallest 10-fold cross-validation within the training dataset. The risk score for the signature was calculated using the following formula:

$$Risk\ score = \sum_{i=1}^{n} Coef_i * x_i$$

in which $Coef_i$ is the coefficient, and x_i is the *log2(n+1)*-transformed relative expression value for each screened gene. The formula was used to compute a risk score for each patient in both the CGGA and TCGA datasets.

Statistical Analyses

Differential LAR expression levels among WHO grades and between different 1p/19q codeletion status and different IDH mutation status were compared by the Wilcox test. Chi-square tests were used to compare the distribution of gender, age, WHO grade, IDH mutation status, and 1p/19q codeletion status among the three subgroups (clustered by consensus expression of LARs) and between low- and high-risk subgroups (partitioned by the median of risk scores) in gliomas.

The prognostic abilities of the risk score and other clinicopathological characteristics were evaluated by univariate and multivariate Cox regression analyses. The prediction efficiency of our risk signature, age, and WHO grade for 1/3/5-year survival was assessed by receiver operating characteristic (ROC) curves. Kaplan–Meier curves used to compare the OS for patients in different groups were tested by the log-rank test. Spearman correlation test was performed to analyze the correlation between TMB and risk score. All statistical analyses were performed using R v.3.6.1 (https://www.r-project.org/) and SPSS Statistics 25 (https://www.ibm.com/products/software).

Results

Correlation between mRNA Expressions and Clinicopathological Features

Given the crucial biological roles of each LAR, we systematically analyzed the correlation between LAR mRNA expression levels and clinicopathological characteristics (including WHO grades, IDH mutation status, and 1p/19q codeletion status) in gliomas. The heatmaps (Fig. 1A, B) show the expression levels of each LAR in diverse WHO grades, and indicate that most of the LARs were aberrantly expressed in different WHO grades in the CGGA dataset; these differential expression levels were validated in the TCGA dataset (Fig. S1A, B). We found that the mRNA expression of 11 lysine acetyltransferases and 15 lysine deacetylases were significantly correlated with WHO grades in the CGGA dataset (Fig. S2A, B). For acetyltransferases, the mRNA expression levels of most KATs (except *KAT5* and *KAT8*) decreased significantly with increasing WHO grade. In contrast, the mRNA expression of the other four acetyltransferases (*SLC16A10*, *KAT1*, *ESCO1*, and *ESCO2*) showed marked increases. For deacetylases, the mRNA expression of *SIRT1*/*2*/*3*/*5*, and that of *HDAC4*/*5*/*11* decreased with increasing WHO grade, while the mRNA expressions of *HDAC1*/*2*/*3*/*7*/*8*, and that of

SIRT6/7 and *LEF1*, showed an increase. Among these LARs that showed increased expression with increasing WHO grade in both the CGGA and TCGA datasets, *HDAC1* is the best-studied LAR in glioma, whereas the potential functions of *ESCO2*, *KAT1*, *LEF1*, and *SLC16A10* are unreported in this cancer (Fig. 1C, D).

The differential expression levels of LARs according to IDH mutation status were investigated in LGGs and GBMs (Fig. 1E, F). Our results showed that HDACs (except *HDAC8* and *10*), KATs (except *KAT1*), *SIRT1*, *SIRT2*, and *LEF1* were all significantly associated with IDH mutation status in LGGs. The expression levels of *HDAC1*/2/3/4/5/6/7/*10*/*11*, as well as those of KATs (except *KAT1* and *KAT8*), *SIRT1*/2/3/*4*, *ESCO2*, *SLC16A10*, and *LEF1*, were significantly correlated with IDH mutation status in GBMs. We also evaluated the mRNA expression of the 33 LARs according to 1p/19q codeletion status in LGGs with mutated IDH. We found that *HDAC1*/2/3/4/5/*6*/*10*/*11*, as well as *KAT1*/*2A*/*2B*/*7*, *SIRT2*/*3*/4/*5*/*7*, *LEF1*, *TCF1*, *ESCO1*/*2* were closely associated with 1p/19q codeletion status in LGGs with mutated IDH (Fig. 1G).

For further investigating the functional status of LARs in gliomas, we downloaded immunohistochemistry images of several LARs from the Human Protein Atlas database (Fig. S3). Most of investigated LARs (including HDAC1/2/3/5/8, SIRT5/7, KAT2A/2B and LEF1) were differential expressed between LGG and GBM, and have similar expression tendency with mRNA expression.

These results suggested that aberrant LAR expression was highly correlated with clinicopathological characteristics of glioma and may be involved in its malignant progression. Therefore, we concluded that it would be worthwhile investigating the correlation between LARs and glioma malignancy in greater detail.

Identification of Subgroups by Consensus Clustering

The mRNA expression of the 33 LARs was analyzed to determine the glioma subtypes in the CGGA dataset. A total of 307 samples were divided into k (k = 2 to 9) subtypes using the R package "Consensus Cluster Plus". We elected $k = 3$ (Fig. 2A-C) as our subtype-dividing value for further study due to the similar number of samples in each cluster and distinct clinical prognoses among the subgroups when we divided gliomas into three subgroups. To investigate the differences among the three subgroups in more detail, we performed PCA to compare the mRNA expression profiles among the three subgroups and the analysis showed that significant differences existed among the three subgroups (Fig. 2D). Furthermore, survival analysis was conducted and results showed that the LA3 subgroup had the poorest OS time and rate while the LA1 subgroup showed the longest OS time among the three groups (Fig. 2E). Subsequently, we evaluated the differences in clinicopathological features and expression levels among the three clusters (LA1, LA2, and LA3) (Fig. 2F), and found that, compared with the other two groups, LA3 was significantly related to increased age at diagnosis ($p < 0.001$), higher WHO grade ($p < 0.001$), fewer IDH mutations ($p < 0.001$), and fewer 1p/19q codeletions ($p < 0.001$) (Table S3). In contrast, the other two subgroups correlated with younger age at diagnosis, lower WHO grade, more IDH mutations, and more 1p/19q codeletions.

Gene Ontology and Gene Set Enrichment Analysis

The above findings implied that clustering was closely related to glioma malignancy. As the LA3 subgroup had the poorest prognosis, we identified genes that were significantly upregulated (log(fold change) ≥ 1 and $p < 0.05$) in the LA3 subgroup compared with the LA1 and LA2 subgroups, and annotated their functions by gene ontology (GO) pathway analysis for biological processes (BPs) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. The GO–BP results showed that the

upregulated genes were enriched in malignancy-related biological processes, such as positive regulation of the ERK1 and ERK2 cascade, angiogenesis, cell proliferation, tumor necrosis factor-mediated signaling pathway, immune response, and negative regulation of apoptotic process (Fig. 3A). Similar results, such as JAK–STAT signaling pathway, cell adhesion molecules (CAMs), and extracellular matrix (ECM)-receptor interaction, were also significantly enriched in the LA3 subgroup based on KEGG pathway analysis (Fig. 3B). We also performed gene set enrichment analysis (GSEA) between the LA3 and LA1/2 subgroups. The results revealed that malignant hallmarks of tumors, including IL2/STAT5 signaling, epithelial–mesenchymal transition, apoptosis, p53 pathway, IL6/JAK/STAT3 signaling, angiogenesis, TNF signaling via NF-κB, and KRAS signaling, were enriched in the LA3 subgroup (Fig. 3C). These results indicated that the three categories were highly associated with glioma malignancy, and that LA3 was the most malignant of the subgroups.

Correlations and Interactions among LARs

To better understand the correlations among the LARs, we performed correlative and protein-protein interaction (PPI) network analyses. We found that genes within the same functional class showed significantly correlated expression patterns and that a high correlation existed between acetyltransferases and deacetylases (Fig. 4A). In the correlation analysis, five acetyltransferases (*KAT6A*, *KAT6B*, *KAT7*, *EP300*, and *CREBBP*) presented a strong co-expression relationship, they were also positively associated with the expression of *HDAC4*, *HDAC5*, *HDAC6*, and *SIRT1* and negatively associated with *SLC16A10*, *KAT1*, *HDAC1*, *HDAC3*, and *LEF1* expression. The HDAC family seems to be the hub family in lysine regulation, as it showed strong co-expression not only among the family members, but also with that of KATs and SIRTs. In contrast, few SIRTs (except *SIRT1*) showed strong correlations with the other LARs.

Analysis of PPI networks also showed that these LARs frequently interacted (Fig. 4B), and that HDAC1, HDAC2, and CREBBP presented the greatest number of links to other LARs. In the PPI networks, we concluded that HDACs had an especially high number of interactions with other LARs, indicating that members of the histone deacetylase family have crucial functions in the regulation of lysine acetylation. Taken together, these findings revealed that several co-expression patterns existed among the LARs, and HDACs are the hub family. The results further indicate that cross-talk among LARs might play critical roles in the malignant progression of glioma.

Building a Risk Signature by LASSO Cox Regression

To investigate the prognostic value of LARs, univariate Cox regression analysis was performed on the mRNA expression of the 33 LARs in the CGGA training dataset. We found that 26 of the 33 genes were correlated with OS ($p < 0.05$) of glioma patients (Fig. 4C). Among the 26 genes, *SLC16A10*, *ESCO1*, *ESCO2*, *KAT1*, *HDAC1*, *HDAC2*, *HDAC3*, *HDAC7*, *HDAC8*, *SIRT6*, *SIRT7*, and *LEF1* were found to be risk factors in glioma, with hazard ratios >1, whereas *KAT2A*, *KAT2B*, *KAT5*, *KAT6A*, *KAT6B*, *KAT7*, *CREBBP*, *EP300*, *HDAC11*, *HDAC4*, *HDAC5*, *SIRT1*, *SIRT3*, and *SIRT5* were protective factors, with hazard ratios <1. The 26 LARs identified as having prognostic value were selected for use with the least absolute shrinkage and selection operator (LASSO) Cox regression algorithm in the CGGA training dataset. Based on the minimum criteria, we determined a 14-gene risk signature (Fig. 4D), and the coefficients (Fig. 4E) obtained by the LASSO algorithm were used to compute the risk score for each sample in the CGGA and TCGA datasets for further study. Besides, prognostic value of each LARs in LGG and GBM with different IDH mutant status and 1p/19q codeletion status were concluded in Supplementary Table S5 and Table S6.

Testing and Validating the Risk Signature

We performed chi-square test and generated heatmaps to evaluate whether the risk score reflected the different distributions of clinicopathological features among gliomas in the CGGA dataset (Fig. 5A) and the TCGA dataset (Fig. S4A). Significant differences in clinicopathological features were observed between the low- and high-risk subgroups (Table S7). The high-risk subgroup was highly associated with older age ($p = 0.014$), higher WHO grade ($p < 0.001$), wild-type IDH ($p < 0.001$), and non-codeletion of 1p/19q ($p < 0.001$), and overall survival time decreased with the risk score increasing. Based on the ROC curves, we concluded that the risk score could perfectly predict 1/3/5-year survival rates in glioma patients with AUC = 0.812/0.866/0.881, respectively (Fig. 5B–D), and was more efficient than WHO grade and age in predicting 1/3/5-year survival rates. ROC curves of the TCGA validation set proved that the risk signature had a stable and robust predictive ability (Fig. S4B-D).

Univariate and multivariate Cox regression analyses were then performed to determine whether the risk signature was an independent prognostic indicator. We included age, risk score, 1p/19q codeletion status, WHO grade, IDH mutation status, gender, chemotherapy and radiotherapy in the univariate Cox regression analysis and the results of the univariate and multivariate Cox regression analysis showed that risk score, WHO grade, age and 1p/19q status were independent predictors for glioma patients (Fig. 5E,F).

Here, we confirmed that the risk score had prognostic value for different WHO grades. The Kaplan–Meier survival curves indicated that low-risk patients had longer OS time and a higher OS rate than high-risk patients in each and all WHO grades in the CGGA dataset (Fig. 6A), and the prognostic ability of the risk score was further validated in the TCGA dataset (Fig. 6B).

Mutation Analysis of LARs

To study the mutation status of the 33 LARs and the relationship between the risk signature and gene mutations, 583 samples with matched SNP data were divided into low- ($n = 303$) and high-risk ($n = 280$) groups. A waterfall plot was generated depicting the mutation frequency of the 33 LARs and the percentage at which they occurred in gliomas (Fig. 6C). We found that 36 (6.17%) of the 583 samples contained mutations in genes coding for the LARs, in which *KAT6B* (8/583) and *CREBBP* (5/583) were the most frequently mutated genes. Within the eight mutations found in *KAT6B*, six were present in oligodendrogliomas and six of the samples were in the low-risk subgroup. All the mutations in *CREBBP* were in samples from high-risk patients, and comprised two glioblastomas and three astrocytomas.

In the waterfall plot depicting the 30 most frequently mutated genes in gliomas (Fig. S5A), we noticed that glioma patients with a high-risk score often carried a higher frequency of gene mutations. This indicated that DNA repair system is highly damaged in patients with higher risk score. Therefore, we calculated the TMB for each sample with SNP data in the TCGA dataset. We found that high-risk patients had higher TMB values (Fig. 6D), and Spearman's correlation analysis confirmed the positive correlation between our risk signature and TMB ($R = 0.52$, $p < 0.0001$) (Fig. 6E). This result implied that impaired regulation of lysine acetylation may affect glioma malignancy through the modulation of factors involved in DNA replication or repair.

CNV Analysis of LARs

In the heatmap depicting the CNV of the 33 LARs, the high-risk section of the heatmap showed more CNV envents than the low-risk section (Fig. S5B). We selected 7 LARs for which the CNV was highly

associated with the risk score for further analyses (Fig. 7A). These genes are located on chromosome arms 1p, 7p, 10q, 19q, and 22q which are characteristically altered in gliomas [15].

HDAC1 and *SIRT2* are located on chromosome arms 1p and 19q, respectively. Codeletion of these genes is frequently observed in oligodendrogliomas, and is highly associated with improved responses to radiochemotherapy and longer survival than diffuse gliomas without these alterations [2]. Although Kaplan–Meier curves revealed that copy number deletions of *HDAC1* and *SIRT2* are related to a better prognosis, we could not determine whether 1p/19q codeletions resulted in differential OS. Therefore, to exclude the potential influences of 1p/19q codeletion, we compared the mRNA expression levels of *HDAC1* and *SIRT2* according to CNV status, as well as the OS rates between low and high levels of *HDAC1* and *SIRT2* mRNA expression in gliomas without *HDAC1* or *SIRT2* CNVs (Fig. 7B, C). We found that, for both genes, copy number deletions were associated with lower mRNA expression, and in gliomas without *HDAC1* or *SIRT2* copy number variations, patients with lower *HDAC1* expression or higher *SIRT2* expression showed better clinical prognosis. These results indicated that *HDAC1* may be one of the oncogenes lost in gliomas with 1p deletion, while *SIRT2*, as a protective factor, is lost with 19q codeletions in glioma patients.

The *EP300* gene is located on chromosome 22q, and deletion of this gene is also common in gliomas. Although *EP300* copy number deletion was associated with a worse prognosis when compared with the diploid state, we could not exclude that loss of other genes located in 22q may also influence prognosis. Therefore, we compared *EP300* mRNA expression levels according to CNV status, as well as the OS rates between high and low levels of *EP300* expression in gliomas without *EP300* CNVs. We found that copy number deletions of *EP300* were associated with lower *EP300* mRNA expression levels, and reduced expression of *EP300* in gliomas without *EP300* CNVs was related to a worse clinical prognosis (Fig. 7D). This indicates that *EP300* may play a tumor suppressor role in glioma and *EP300*(22q) may be one of the tumor suppressor genes lost in the 22q− event [16].

The CNVs for the other 4 LARs – *KAT6B*(10q), *SIRT1*(10q), *HDAC10*(22q), and *HDAC9*(7q) – were highly associated with the risk signature and may be affected in chromosomal alterations such as 10q−, 22q− and 7+ in gliomas. However, we did not find significant differences in OS between low and high levels of expression of these four genes in patients without CNVs (Fig. S6A–D). Therefore, we regarded the differences in OS rates between patients with or without copy number loss of these four genes as passive changes resulting from chromosomal variations, indicating that they may have little effect in related clinical outcomes.

Discussion

In this report, we have shown that the mRNA expression levels of most of the evaluated LARs are closely associated with clinicopathological features of glioma. We further identified three subgroups, LA1/2/3, by consensus clustering of 26 OS-related LARs, and confirmed that LA3 was the most malignant subtype with the poorest prognosis. Moreover, the LA3 subgroup was tightly associated with malignancy-related biological processes, key signaling pathways, and tumor hallmarks. In addition, we also constructed a prognostic signature and divided glioma patients into low- and high-risk categories by the median of risk scores. We noticed a close relationship between the risk signature and clinicopathological features of glioma, and ROC curves, univariate and multivariate analyses, and Kaplan–Meier curves were used to determine the potential prognosis value of the risk signature in glioma. We also included SNP and CNV data of LARs to identify potential therapeutic targets that may play a prognostic role in gliomas.

Our 14-gene signature included seven members of the HDAC family (*HDAC1*, *HDAC2*, *HDAC3*, *HDAC4*, *HDAC5*, *HDAC7*, and *HDAC8*), two of the sirtuin family of deacetylases (*SIRT5* and *SIRT7*), two of the GCN5 family of lysine acetyltransferase (*KAT2A* and *KAT2B*), establishment of sister chromatid cohesion N-acetyltransferase 2 (*ESCO2*), solute carrier family 16 member 10 (*SLC16A10*), and lymphoid enhancer binding factor 1 (*LEF1*).

Based on univariate analysis, the increased expression of five genes from the risk-signature group was associated with longer OS. The expression of *HDAC4* and *HDAC5* was downregulated in high-grade gliomas when compared with low-grade ones, and was associated with a favorable clinical outcome [17, 18]. Interestingly, *HDAC4* and *HDAC5* have been reported to act as oncogenes by promoting the proliferation of glioma cells, as well as their invasive ability [19, 20]. *KAT2A*, also known as *GCN5*, functions primarily as a transcriptional activator, although it also represses NF-κB signaling by promoting the ubiquitination of the NF-κB subunit RELA [21]. Moreover, KAT2A can enhance glioma proliferation and invasion via the STAT3 and AKT signaling pathways [22]. *KAT2B*, also known as *PCAF* (p300/CBP-associated factor), can upregulate the transcriptional activity of p53 through acetylation of lys320 in the C-terminal portion of p53 [23, 24], as well as mediate the acetylation of AKT1, thereby enhancing the proliferative capacity of glioblastoma cells [25]. SIRT5 functions both as a lysine acetyltransferase, and as a desuccinylase, demalonylase, or deglutarylase [26-28]. Although several reports have indicated that *SIRT5* can act as an oncogene in different cancers [29-32], *SIRT5* downregulation has also been associated with poor prognosis in glioblastoma patients [33].

Compared with our result, the action of the five genes above could be extremely contradictory. We thought that this discrepancy could be due to three reasons: Firstly, their reduced expression with higher WHO grades may have obscured their carcinogenic effects in the univariate analysis, higher grade with lower expression may mislead us to believe they are protective factors. Moreover, most of the LARs evaluated have comprehensive and widespread regulatory targets and that could provide them with dual functions in oncogenic regulation.

The expression levels of the other nine genes were negatively related to OS time in the univariate analysis. HDAC1, one of the most studied histone deacetylases, functions as a powerful tumor promotor and enhances chemoresistance in a variety of cancers [34-39]. Our results showed that it was the main regulatory factor in the LAR interaction network and loss of *HDAC1* copy number with 1p deletion contributed to a better prognosis in gliomas. And it was a hub and potential driver gene in gliomas, its reported that it could activate PI3K/AKT and MEK/ERK signaling pathways to promote glioma cell proliferation and invasion and knockdown of it could induce glioma cell apoptosis and invasion suppression[34, 40]. *HDAC2* and *HDAC3*, both of which were upregulated with higher WHO grades, have been reported to be involved in glioma malignancy and chemoresistance [41-43]. The expression of *HDAC7* is positively associated with a mesenchymal subtype of glioblastoma [44], and can enhance the malignant phenotype of glioma, while its inhibition may suppress STAT3 tumorigenic activity [45, 46]. *HDAC8* has been found to be correlated with several types of cancer, although seldom in gliomas[47, 48]. SIRT7 can activate the ERK/STAT3 signaling pathway, thereby promoting glioma proliferation and invasion [49]. *LEF1*, a recent addition to the deacetylase family, is involved in the Wnt signaling pathway and can promote glioma malignancy [50, 51]. *SLC16A10*, coding for a member of a family of plasma membrane amino acid transporters, was shown to be significantly upregulated in higher-grade gliomas [52]; however, its role in glioma and other cancer types remains unclear. *ESCO2*, required for the establishment of sister chromatid cohesion during the S phase of mitosis [53], can enhance the malignancy of gastric cancer while inhibiting colorectal tumor metastasis [54, 55], and its role in glioma has not been previously reported.

Based on the above results, whether these 14 genes are responsible for the differential OS in a biological meaning or reflect other inner associations remains to be elucidated. Our results showing the presence of distinct biological processes and signaling pathways between the LA1/2 and LA3 subgroups might provide some clues. We clearly identified a correlation between LAR levels and immune characteristics as GO biological process and KEGG pathway analyses both revealed that genes enriched in the LA3 subgroup were significantly involved in immune-related activities such as immune response, complement activation (classical complement pathway), B cell activation and T cell receptor signaling pathway. This indicates that LARs may have a prominent role in the immune response of glioma cells. However, although numerous studies have investigated the relationship between acetylation regulation and immune response, its effect in glioma remains unclear [56-58]. Therefore, further explorations are needed to seek the inner relationship between LARs and immune activities and it may lead to an unexpected breakthrough in glioma therapy.

Mutations of LARs were explored in our study, in which KAT6B and CREBBP were found with highest frequency of mutations in glioma samples. KAT6B is a tumor suppressor histone H3 lysine 23 acetyltransferase, loss or mutation of KAT6B may impair acetylation of histone H3 at lysine 23[59, 60]. CREBBP is one of the most frequent mutant genes in small cell lung cancer (SCLC), and also common in some kinds of paediatric cancers like high-grade gliomas, medulloblastoma and T-lineage acute lymphoblastic leukaemia[61, 62]. CREBBP inactivation mutations could result reduced histone acetylation in genes associated with cell adhesion, and SCLC patients with inactivation CREBBP mutation have a higher sensitivity to HADC inhibitors[61].Though study have revealed that mutations of KAT6B and CREBBP could result the loss of functions, the potential influence of KAT6B and CREBBP mutations remain to be investigated in glioma by deep-going experiments.

Our risk signature was also shown to be pertinent to the TMB of glioma patients. The TMB is associated with neoantigen abundance and increased immunogenicity [63], and is used to quantitatively assess mutations carried by tumor cells. It is defined as the total number of somatic gene coding errors, base substitutions, gene insertions, or deletion errors that are detected per million bases. In recent years, several studies have demonstrated that dysregulation of lysine acetylation may result in errors during DNA damage repair. For instance, PCAF/GCN5-mediated K163 acetylation of RPA1 (replication protein A) is crucial for nucleotide excision repair (NER) [64], SIRT7 is recruited in a PARP1-dependent manner to sites of DNA damage, where it modulates H3K18Ac levels [65], and TET1 (ten-eleven translocation-1) forms a complex with KAT8 to modulate its function and the level of H4K16Ac, which ultimately affects gene expression and DNA repair [66]. Based on these observations, we speculate that dysregulation of lysine acetylation of both histone and non-histone proteins may play a pivotal role in impairing the DNA damage repair response, which would then lead to hypermutations and an increased neoantigen load, leading to malignant progression of tumors.

We have systematically revealed the mRNA expression, underlying functions, and prognostic values of LARs in glioma, and shown that acetylation regulators may have an immune-related effect on the malignant progression of glioma. Moreover, we identified that several underexplored LARs, such as *ESCO2*, *HAT1(KAT1)*, and *LEF1*, may have prognostic value in lower grade glioma patients (Table S5) and may be potential glioma biomarkers. We further found that specific chromosomal alterations in gliomas were highly related to the CNVs of several LARs. *HDAC1* was shown to be one of the oncogenes deleted in the 1p deletion event, and *SIRT2* and *EP300* were two cancer suppressors lost in 19q deletion and 22q deletion events, respectively. Our results also revealed that dysfunction of LARs may partially explain the hypermutation state of gliomas, which is associated with unfavorable prognosis. Taken together, we believe that substantial research is still required to illuminate the detailed

mechanisms involved in lysine acetylation-mediated regulation of glioma malignancy that may ultimately lead to new and effective targeted therapies for glioma patients.

Conclusions

LARs play a crucial role in the prognosis and malignant progression of gliomas. The identified prognostic LARs and CNVs can promote our understanding of the underlying mechanisms in glioma pathological processes and provide new diagnostic and therapeutic strategies.

Abbreviations

LAR: lysine acetylation regulators; LGG: lower-grade glioma; GBM: glioblastoma; CNV: Copy Number Variation; SNP: Single Nucleotide Polymorphism; TCGA: The Cancer Genome Atlas; CGGA: Chinese Glioma Genome Atlas; OS: Overall survival; LASSO: Least absolute shrinkage and selection operator; ROC: Receiver operating characteristic; AUC: Area under curve.

Declarations

Consent for publication

All authors proofread and approved the final manuscript.

Acknowledgments

The authors gratefully acknowledge contributions from the CGGA network and the TCGA network.

Authors' contributions

Xingen Zhu and Kai Huang designed the research. Zewei Tu, Shigang Lv, Lei Wu performed the bioinformatic and statistical analysis and created the figures and tables. Qing Hu, Chuming Tao and Kuangxun Li performed the literature search, and were involved in manuscript writing and proofreading. Xingen Zhu and Lei Wu supervised the research and critically read the draft manuscript.

Funding

The current study was funding by the National Natural Science Foundation (grant nos. 81660420, 81860448 and 81960456), the Natural Science Foundation of Jiangxi Province (grant no. 20171ACB20035,20192BAB205077) and Jiangxi Province Department of Education Science and Technology Research project, China (grant no. GJJ190018).

Availability of data and materials

The datasets of RNA-seq data, clinical information, SNV and SNP data for this study can be found in the Chinese Glioma Genome Atlas (CGGA) (http://www.cgga.org.cn/), the University of California, Santa Cruz Xena browser (UCSC Xena; https://xenabrowser.net/datapages/) and the Genomic Data Commons Data Portal (GDC; https://portal.gdc.cancer.gov/). The Immunohistochemistry images of LARs could be acquired from the The Human Protein Atlas (HPA: https://www.proteinatlas.org/) More details could be found in "Data Acquisition" part of materials and methods.

Competing interests

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

Ethics approval and consent to participate

Not applicable.

References

1. Jiang T, Mao Y, Ma W, Mao Q, You Y, Yang X, Jiang C, Kang C, Li X, Chen L *et al*: **CGCG clinical practice guidelines for the management of adult diffuse gliomas**. *Cancer Lett* 2016, **375**(2):263-273.
2. Cancer Genome Atlas Research N, Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, Rheinbay E, Miller CR, Vitucci M *et al*: **Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas**. *N Engl J Med* 2015, **372**(26):2481-2498.
3. Cancer Genome Atlas Research N: **Comprehensive genomic characterization defines human glioblastoma genes and core pathways**. *Nature* 2008, **455**(7216):1061-1068.
4. Henikoff S, Greally JM: **Epigenetics, cellular memory and gene regulation**. *Curr Biol* 2016, **26**(14):R644-648.
5. Chen Y, Hong T, Wang S, Mo J, Tian T, Zhou X: **Epigenetic modification of nucleic acids: from basic studies to medical applications**. *Chem Soc Rev* 2017, **46**(10):2844-2872.
6. Dawson MA, Kouzarides T: **Cancer epigenetics: from mechanism to therapy**. *Cell* 2012, **150**(1):12-27.
7. Serrano-Gomez SJ, Maziveyi M, Alahari SK: **Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications**. *Mol Cancer* 2016, **15**:18.
8. Toh TB, Lim JJ, Chow EK: **Epigenetics in cancer stem cells**. *Mol Cancer* 2017, **16**(1):29.
9. Narita T, Weinert BT, Choudhary C: **Functions and mechanisms of non-histone protein acetylation**. *Nat Rev Mol Cell Biol* 2019, **20**(3):156-174.
10. Gil J, Ramirez-Torres A, Encarnacion-Guevara S: **Lysine acetylation and cancer: A proteomics perspective**. *J Proteomics* 2017, **150**:297-309.
11. Glozak MA, Sengupta N, Zhang X, Seto E: **Acetylation and deacetylation of non-histone proteins**. *Gene* 2005, **363**:15-23.
12. Wilkerson MD, Hayes DN: **ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking**. *Bioinformatics* 2010, **26**(12):1572-1573.
13. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies**. *Nucleic Acids Res* 2015, **43**(7):e47.
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome Res* 2003, **13**(11):2498-2504.
15. Bredel M, Scholtens DM, Harsh GR, Bredel C, Chandler JP, Renfrow JJ, Yadav AK, Vogel H, Scheck AC, Tibshirani R *et al*: **A network model of a cooperative genetic landscape in brain tumors**. *JAMA* 2009, **302**(3):261-275.
16. Hartmann C, Numann A, Mueller W, Holtkamp N, Simon M, von Deimling A: **Fine mapping of chromosome 22q tumor suppressor gene candidate regions in astrocytoma**. *Int J Cancer* 2004, **108**(6):839-844.
17. Cheng W, Li M, Cai J, Wang K, Zhang C, Bao Z, Liu Y, Wu A: **HDAC4, a prognostic and chromosomal instability marker, refines the predictive value of MGMT promoter**

methylation. *J Neurooncol* 2015, **122**(2):303-312.
18. Dali-Youcef N, Froelich S, Moussallieh FM, Chibbaro S, Noel G, Namer IJ, Heikkinen S, Auwerx J: **Gene expression mapping of histone deacetylases and co-factors, and correlation with survival time and 1H-HRMAS metabolomic profile in human gliomas**. *Sci Rep* 2015, **5**:9087.
19. Cai JY, Xu TT, Wang Y, Chang JJ, Li J, Chen XY, Chen X, Yin YF, Ni XJ: **Histone deacetylase HDAC4 promotes the proliferation and invasion of glioma cells**. *Int J Oncol* 2018, **53**(6):2758-2768.
20. Liu Q, Zheng JM, Chen JK, Yan XL, Chen HM, Nong WX, Huang HQ: **Histone deacetylase 5 promotes the proliferation of glioma cells by upregulation of Notch 1**. *Mol Med Rep* 2014, **10**(4):2045-2050.
21. Mao X, Gluck N, Li D, Maine GN, Li H, Zaidi IW, Repaka A, Mayo MW, Burstein E: **GCN5 is a required cofactor for a ubiquitin ligase that targets NF-kappaB/RelA**. *Genes Dev* 2009, **23**(7):849-861.
22. Liu K, Zhang Q, Lan H, Wang L, Mou P, Shao W, Liu D, Yang W, Lin Z, Lin Q *et al*: **GCN5 Potentiates Glioma Proliferation and Invasion via STAT3 and AKT Signaling Pathways**. *Int J Mol Sci* 2015, **16**(9):21897-21910.
23. Liu L, Scolnick DM, Trievel RC, Zhang HB, Marmorstein R, Halazonetis TD, Berger SL: **p53 sites acetylated in vitro by PCAF and p300 are acetylated in vivo in response to DNA damage**. *Mol Cell Biol* 1999, **19**(2):1202-1209.
24. Sakaguchi K, Herrera JE, Saito S, Miki T, Bustin M, Vassilev A, Anderson CW, Appella E: **DNA damage activates p53 through a phosphorylation-acetylation cascade**. *Genes Dev* 1998, **12**(18):2831-2841.
25. Zhang S, Sun G, Wang Z, Wan Y, Guo J, Shi L: **PCAF-mediated Akt1 acetylation enhances the proliferation of human glioblastoma cells**. *Tumour Biol* 2015, **36**(3):1455-1462.
26. Tan M, Peng C, Anderson KA, Chhoy P, Xie Z, Dai L, Park J, Chen Y, Huang H, Zhang Y *et al*: **Lysine glutarylation is a protein posttranslational modification regulated by SIRT5**. *Cell Metab* 2014, **19**(4):605-617.
27. Peng C, Lu Z, Xie Z, Cheng Z, Chen Y, Tan M, Luo H, Zhang Y, He W, Yang K *et al*: **The first identification of lysine malonylation substrates and its regulatory enzyme**. *Mol Cell Proteomics* 2011, **10**(12):M111 012658.
28. Du J, Zhou Y, Su X, Yu JJ, Khan S, Jiang H, Kim J, Woo J, Kim JH, Choi BH *et al*: **Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase**. *Science* 2011, **334**(6057):806-809.
29. Lu W, Zuo Y, Feng Y, Zhang M: **SIRT5 facilitates cancer cell growth and drug resistance in non-small cell lung cancer**. *Tumour Biol* 2014, **35**(11):10699-10705.
30. Shi L, Yan H, An S, Shen M, Jia W, Zhang R, Zhao L, Huang G, Liu J: **SIRT5-mediated deacetylation of LDHB promotes autophagy and tumorigenesis in colorectal cancer**. *Mol Oncol* 2019, **13**(2):358-375.
31. Chang L, Xi L, Liu Y, Liu R, Wu Z, Jian Z: **SIRT5 promotes cell proliferation and invasion in hepatocellular carcinoma by targeting E2F1**. *Mol Med Rep* 2018, **17**(1):342-349.
32. Ma Y, Qi Y, Wang L, Zheng Z, Zhang Y, Zheng J: **SIRT5-mediated SDHA desuccinylation promotes clear cell renal cell carcinoma tumorigenesis**. *Free Radic Biol Med* 2019, **134**:458-467.
33. Chen X, Xu Z, Zeng S, Wang X, Liu W, Qian L, Wei J, Yang X, Shen Q, Gong Z *et al*: **SIRT5 downregulation is associated with poor prognosis in glioblastoma**. *Cancer Biomark* 2019, **24**(4):449-459.
34. Li S, Chen X, Mao L, Zahid KR, Wen J, Zhang L, Zhang M, Duan J, Duan J, Yin X *et al*: **Histone deacetylase 1 promotes glioblastoma cell proliferation and invasion via activation**

of PI3K/AKT and MEK/ERK signaling pathways. *Brain Res* 2018, **1692**:154-162.
35. Staberg M, Michaelsen SR, Rasmussen RD, Villingshoj M, Poulsen HS, Hamerlik P: **Inhibition of histone deacetylases sensitizes glioblastoma cells to lomustine**. *Cell Oncol (Dordr)* 2017, **40**(1):21-32.
36. Tang Z, Ding S, Huang H, Luo P, Qing B, Zhang S, Tang R: **HDAC1 triggers the proliferation and migration of breast cancer cells via upregulation of interleukin-8**. *Biol Chem* 2017, **398**(12):1347-1356.
37. Huang Y, Chen J, Lu C, Han J, Wang G, Song C, Zhu S, Wang C, Li G, Kang J *et al*: **HDAC1 and Klf4 interplay critically regulates human myeloid leukemia cell proliferation**. *Cell Death Dis* 2014, **5**:e1491.
38. Cacan E, Ali MW, Boyd NH, Hooks SB, Greer SF: **Inhibition of HDAC1 and DNMT1 modulate RGS10 expression and decrease ovarian cancer chemoresistance**. *PLoS One* 2014, **9**(1):e87455.
39. Liu X, Yu Y, Zhang J, Lu C, Wang L, Liu P, Song H: **HDAC1 Silencing in Ovarian Cancer Enhances the Chemotherapy Response**. *Cell Physiol Biochem* 2018, **48**(4):1505-1518.
40. Wang XQ, Bai HM, Li ST, Sun H, Min LZ, Tao BB, Zhong J, Li B: **Knockdown of HDAC1 expression suppresses invasion and induces apoptosis in glioma cells**. *Oncotarget* 2017, **8**(29):48027-48040.
41. Zhang Z, Wang Y, Chen J, Tan Q, Xie C, Li C, Zhan W, Wang M: **Silencing of histone deacetylase 2 suppresses malignancy for proliferation, migration, and invasion of glioblastoma cells and enhances temozolomide sensitivity**. *Cancer Chemother Pharmacol* 2016, **78**(6):1289-1296.
42. Sun XY, Qu Y, Ni AR, Wang GX, Huang WB, Chen ZP, Lv ZF, Zhang S, Lindsay H, Zhao S *et al*: **Novel histone deacetylase inhibitor N25 exerts anti-tumor effects and induces autophagy in human glioma cells by inhibiting HDAC3**. *Oncotarget* 2017, **8**(43):75232-75242.
43. Li ZY, Li QZ, Chen L, Chen BD, Wang B, Zhang XJ, Li WP: **Histone Deacetylase Inhibitor RGFP109 Overcomes Temozolomide Resistance by Blocking NF-kappaB-Dependent Transcription in Glioblastoma Cell Lines**. *Neurochem Res* 2016, **41**(12):3192-3205.
44. Murata H, Yoshimoto K, Hatae R, Akagi Y, Mizoguchi M, Hata N, Kuga D, Nakamizo A, Amano T, Sayama T *et al*: **Detection of proneural/mesenchymal marker expression in glioblastoma: temporospatial dynamics and association with chromatin-modifying gene expression**. *J Neurooncol* 2015, **125**(1):33-41.
45. Yu X, Wang M, Wu J, Han Q, Zhang X: **ZNF326 promotes malignant phenotype of glioma by up-regulating HDAC7 expression and activating Wnt pathway**. *J Exp Clin Cancer Res* 2019, **38**(1):40.
46. Peixoto P, Blomme A, Costanza B, Ronca R, Rezzola S, Palacios AP, Schoysman L, Boutry S, Goffart N, Peulen O *et al*: **HDAC7 inhibition resets STAT3 tumorigenic activity in human glioblastoma independently of EGFR and PTEN: new opportunities for selected targeted therapies**. *Oncogene* 2016, **35**(34):4481-4494.
47. Adhikari N, Amin SA, Jha T: **Selective and nonselective HDAC8 inhibitors: a therapeutic patent review**. *Pharm Pat Anal* 2018, **7**(6):259-276.
48. Banerjee S, Adhikari N, Amin SA, Jha T: **Histone deacetylase 8 (HDAC8) and its inhibitors with selectivity to other isoforms: An overview**. *Eur J Med Chem* 2019, **164**:214-240.
49. Mu P, Liu K, Lin Q, Yang W, Liu D, Lin Z, Shao W, Ji T: **Sirtuin 7 promotes glioma proliferation and invasion through activation of the ERK/STAT3 signaling pathway**. *Oncol Lett* 2019, **17**(2):1445-1452.
50. Gao X, Mi Y, Ma Y, Jin W: **LEF1 regulates glioblastoma cell proliferation, migration, invasion, and cancer stem-like cell self-renewal**. *Tumour Biol* 2014, **35**(11):11505-11511.

51. Rosmaninho P, Mukusch S, Piscopo V, Teixeira V, Raposo AA, Warta R, Bennewitz R, Tang Y, Herold-Mende C, Stifani S *et al*: **Zeb1 potentiates genome-wide gene transcription with Lef1 to promote glioblastoma cell invasion**. *EMBO J* 2018, **37**(15).
52. Kim DK, Kanai Y, Matsuo H, Kim JY, Chairoungdua A, Kobayashi Y, Enomoto A, Cha SH, Goya T, Endou H: **The human T-type amino acid transporter-1: characterization, gene organization, and chromosomal location**. *Genomics* 2002, **79**(1):95-103.
53. Alomer RM, da Silva EML, Chen J, Piekarz KM, McDonald K, Sansam CG, Sansam CL, Rankin S: **Esco1 and Esco2 regulate distinct cohesin functions during cell cycle progression**. *Proc Natl Acad Sci U S A* 2017, **114**(37):9906-9911.
54. Chen H, Zhang L, He W, Liu T, Zhao Y, Chen H, Li Y: **ESCO2 knockdown inhibits cell proliferation and induces apoptosis in human gastric cancer cells**. *Biochem Biophys Res Commun* 2018, **496**(2):475-481.
55. Guo XB, Huang B, Pan YH, Su SG, Li Y: **ESCO2 inhibits tumor metastasis via transcriptionally repressing MMP2 in colorectal cancer**. *Cancer Manag Res* 2018, **10**:6157-6166.
56. Jeong Y, Du R, Zhu X, Yin S, Wang J, Cui H, Cao W, Lowenstein CJ: **Histone deacetylase isoforms regulate innate immune responses by deacetylating mitogen-activated protein kinase phosphatase-1**. *J Leukoc Biol* 2014, **95**(4):651-659.
57. Liu J, Qian C, Cao X: **Post-Translational Modification Control of Innate Immunity**. *Immunity* 2016, **45**(1):15-30.
58. Gogoi M, Chandra K, Sarikhani M, Ramani R, Sundaresan NR, Chakravortty D: **Salmonella escapes adaptive immune response via SIRT2 mediated modulation of innate immune response in dendritic cells**. *PLoS Pathog* 2018, **14**(11):e1007437.
59. Yan K, Rousseau J, Littlejohn RO, Kiss C, Lehman A, Rosenfeld JA, Stumpel CTR, Stegmann APA, Robak L, Scaglia F *et al*: **Mutations in the Chromatin Regulator Gene BRPF1 Cause Syndromic Intellectual Disability and Deficient Histone Acetylation**. *Am J Hum Genet* 2017, **100**(1):91-104.
60. Huang F, Abmayr SM, Workman JL: **Regulation of KAT6 Acetyltransferases and Their Roles in Cell Cycle Progression, Stem Cell Maintenance, and Human Disease**. *Mol Cell Biol* 2016, **36**(14):1900-1907.
61. Jia D, Augert A, Kim DW, Eastwood E, Wu N, Ibrahim AH, Kim KB, Dunn CT, Pillai SPS, Gazdar AF *et al*: **Crebbp Loss Drives Small Cell Lung Cancer and Increases Sensitivity to HDAC Inhibition**. *Cancer Discov* 2018, **8**(11):1422-1437.
62. Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, Ma J, Edmonson MN, Hedlund EK, Rusch MC *et al*: **The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes**. *Nat Commun* 2014, **5**:3630.
63. Hodges TR, Ott M, Xiu J, Gatalica Z, Swensen J, Zhou S, Huse JT, de Groot J, Li S, Overwijk WW *et al*: **Mutational burden, immune checkpoint expression, and mismatch repair in glioma: implications for immune checkpoint immunotherapy**. *Neuro Oncol* 2017, **19**(8):1047-1057.
64. Zhao M, Geng R, Guo X, Yuan R, Zhou X, Zhong Y, Huo Y, Zhou M, Shen Q, Li Y *et al*: **PCAF/GCN5-Mediated Acetylation of RPA1 Promotes Nucleotide Excision Repair**. *Cell Rep* 2017, **20**(9):1997-2009.
65. Vazquez BN, Thackray JK, Simonet NG, Kane-Goldsmith N, Martinez-Redondo P, Nguyen T, Bunting S, Vaquero A, Tischfield JA, Serrano L: **SIRT7 promotes genome integrity and modulates non-homologous end joining DNA repair**. *EMBO J* 2016, **35**(14):1488-1503.
66. Zhong J, Li X, Cai W, Wang Y, Dong S, Yang J, Zhang J, Wu N, Li Y, Mao F *et al*: **TET1 modulates H4K16 acetylation by controlling auto-acetylation of hMOF to affect gene regulation and DNA repair function**. *Nucleic Acids Res* 2017, **45**(2):672-684.

648

Figures

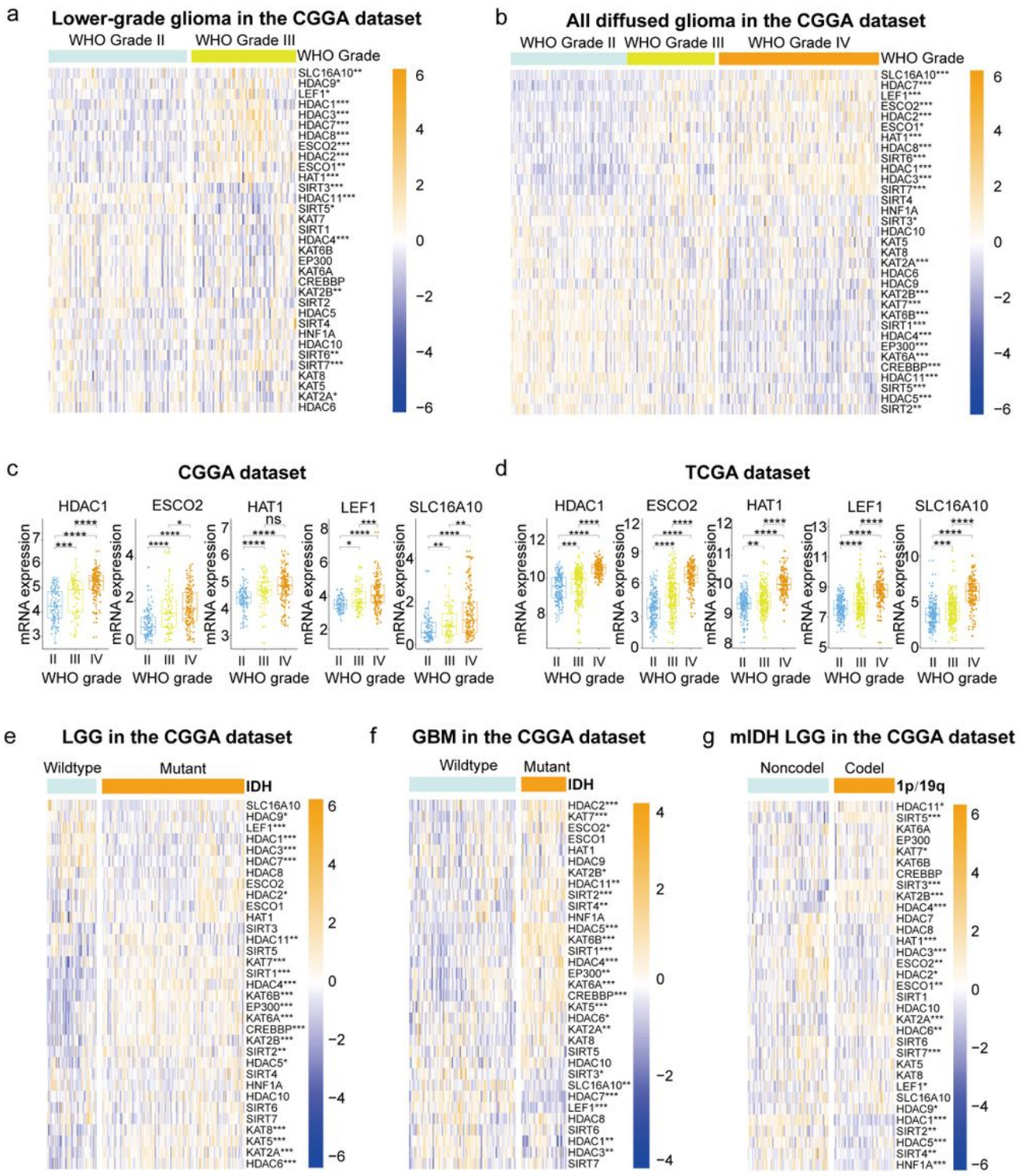

Figure 1

(A, B) The expression levels of 33 lysine acetylation regulators (LARs) in gliomas with different WHO grades. (C, D) HDAC1,ESCO2, KAT1, LEF1, and SCL16A10 expression levels increased with increasing WHO grade. (E, F) The expression levels of LARs in low-grade gliomas (LGGs) and glioblastomas (GBMs)

with differentIDHmutation status. (G) The expression levels of LARs in LGGs with IDH mutations (mIDH) with different 1p/19q codeletion status. *p< 0.05, **p< 0.01, ***p< 0.001, and ****p< 0.0001.

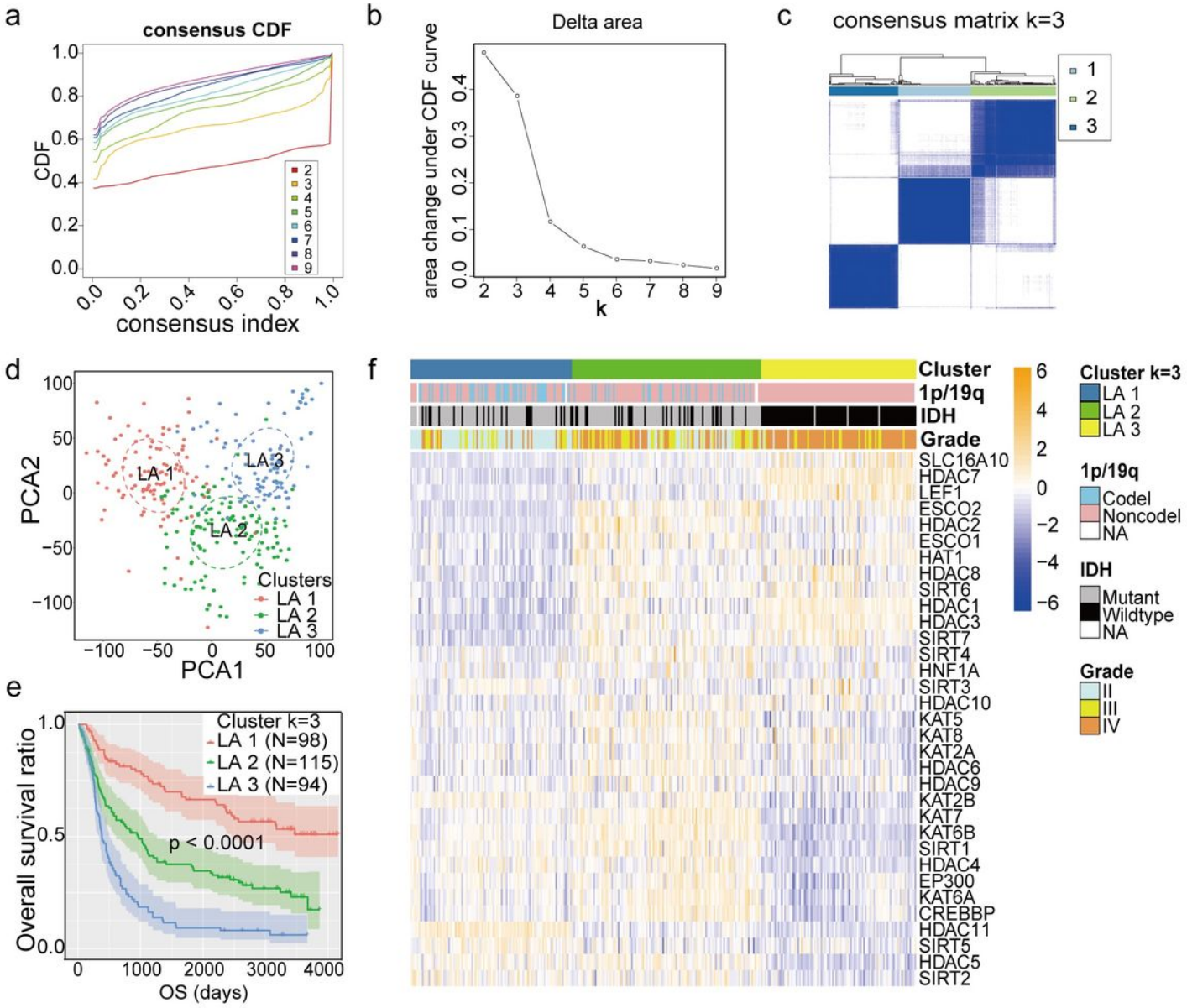

Figure 2

(A) Consensus clustering cumulative distribution function (CDF) for k = 2 to 9. (B) Relative change in area under the CDF curve for k = 2 to 9. (C) Consensus clustering matrix for k = 3. (D) Principal component analysis of the total RNA expression profile in the CGGA dataset. (E) Kaplan–Meier overall survival curves for 307 glioma patients in the LA1, 2, and 3 subgroups.(F) The different expression levels of lysine acetylation regulators (LARs) and clinicopathological feature contributions of the three subgroups defined by the consensus expression of 33 LARs.

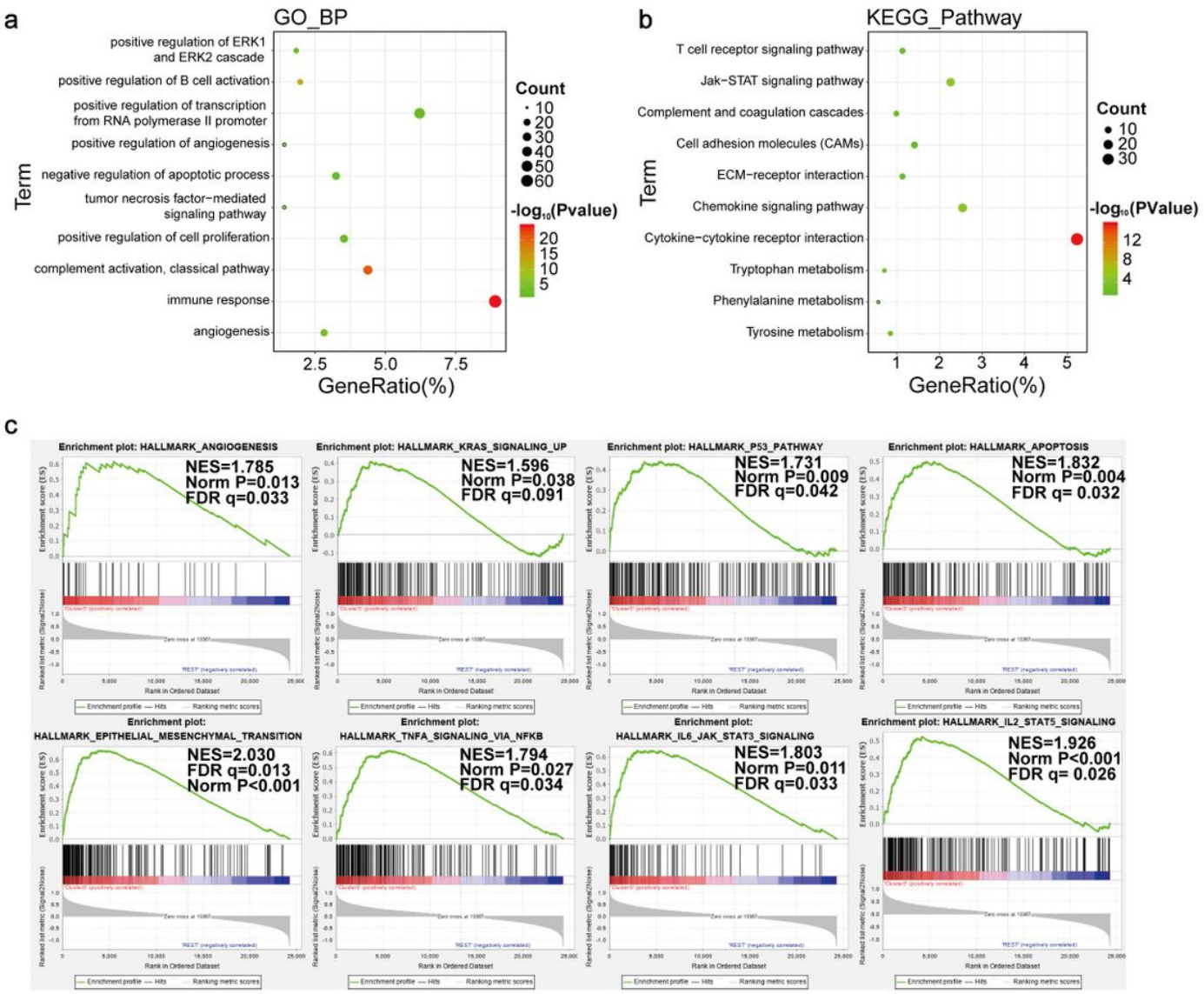

Figure 3

(A, B) Functional annotation of the genes with higher expression in the LA3 subgroup using gene ontology (GO) terms of biological processes (A) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway (B) analyses. (C) Gene set enrichment analysis indicated that genes with higher expression in the LA3 subgroup were enriched for hallmarks of malignant tumors.

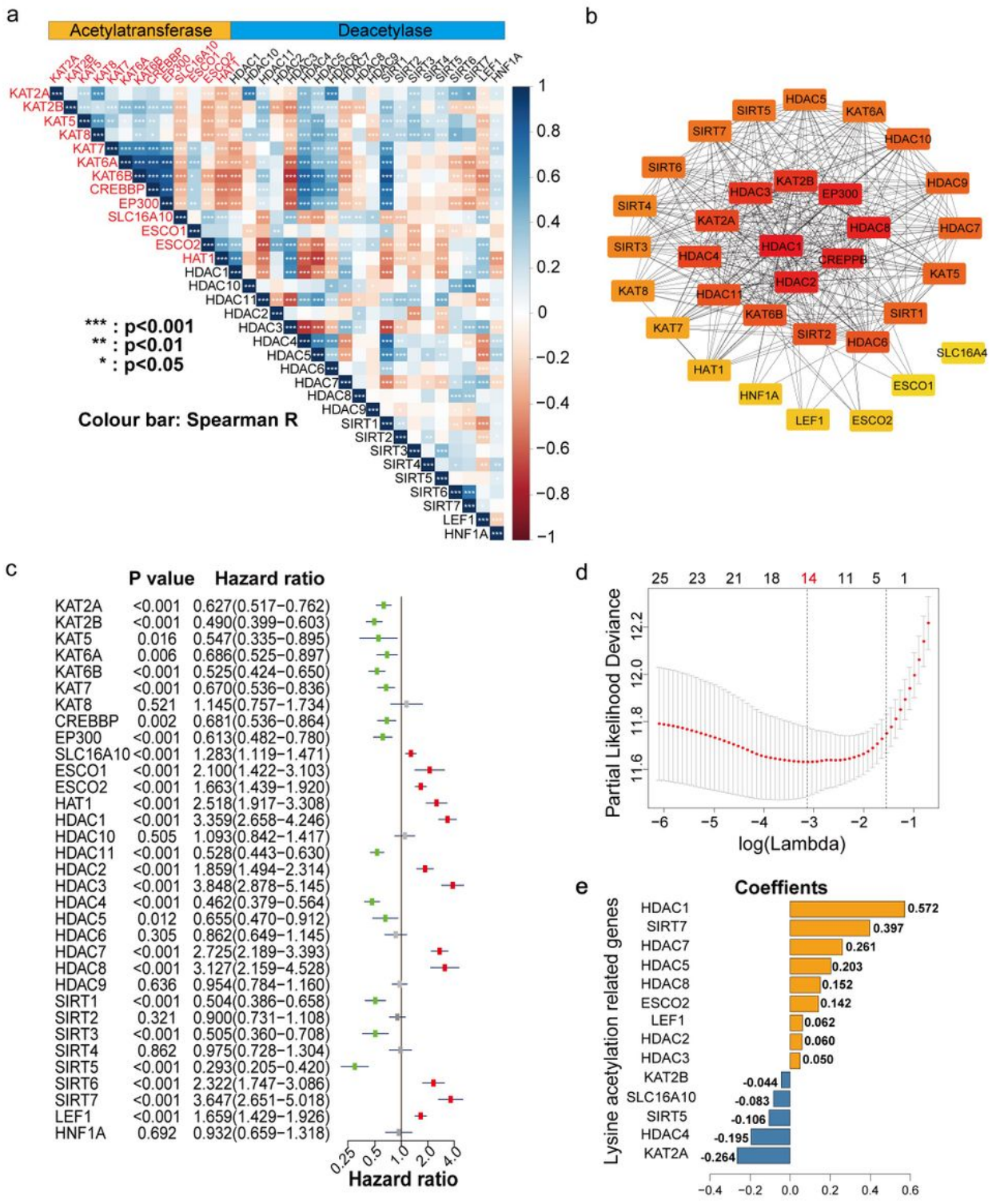

Figure 4

(A) Spearman's correlation analysis among the 33 lysine acetylation regulators (LARs). (B) Protein interaction network of the 33 LARs. (C) Univariate Cox regression analysis was used to calculate the hazard ratios (HRs), 95% confidence intervals, and p-values for screening the prognostic LARs. (D, E) Least absolute shrinkage and selection operator (LASSO) regression was performed to calculate the minimum criteria(D) and coefficients(E).

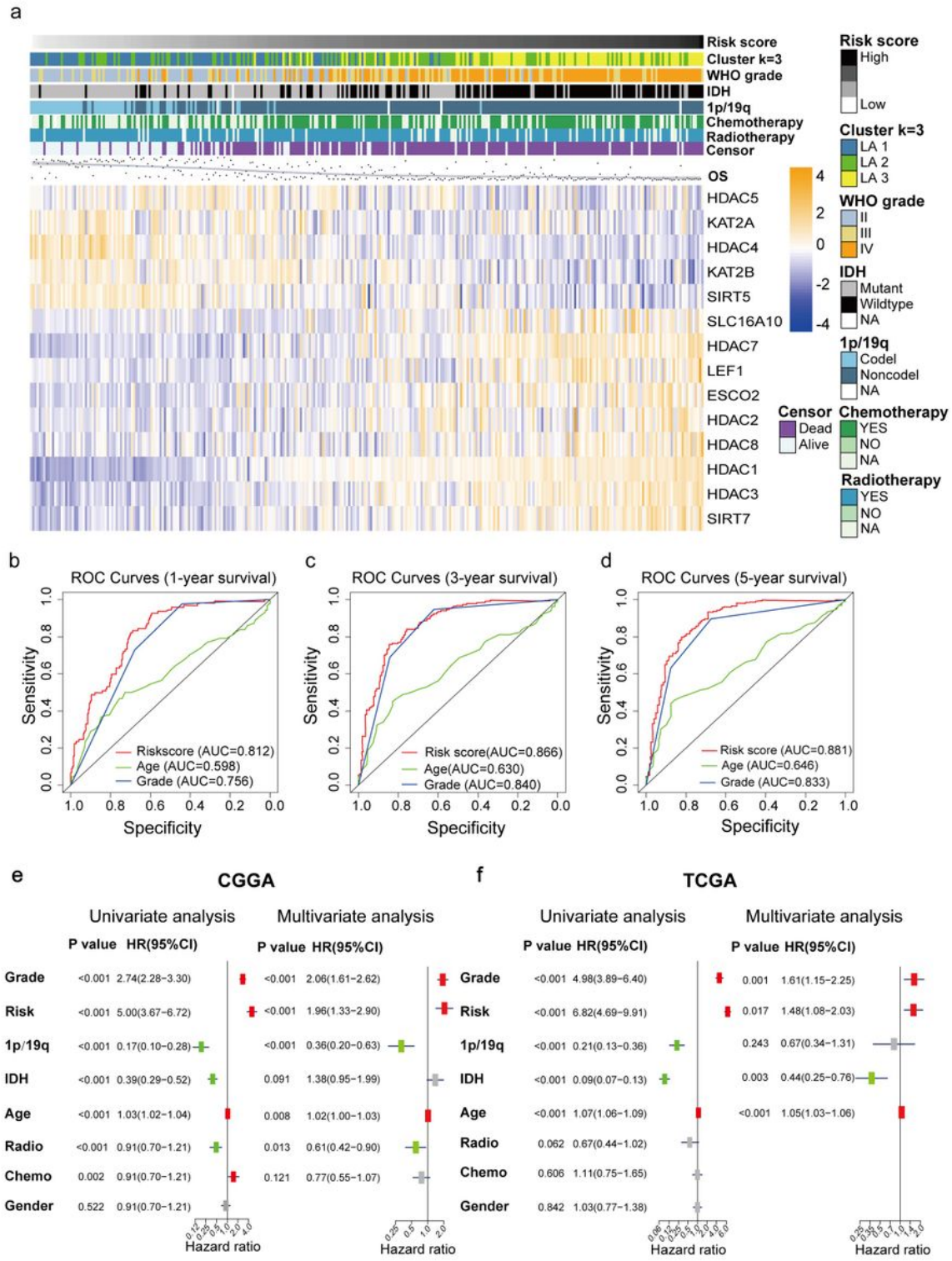

Figure 5

(A) The differential expression levels of the included 14 lysine acetylation regulators (LARs) and the distributions of clinicopathological characteristics were compared between low- and high- risk subgroups. (B–D) Receiver operating characteristic (ROC) curves showed the predictive efficiency of the risk signature, WHO grade, and age on 1/3/5-year survival rate. (E, F) Univariate and multivariate Cox

regression analyses of the overall survival and clinicopathological features of patients fromthe Chinese Glioma Genome Atlas (CGGA) (E) and The Cancer Genome Atlas (TCGA) (F) datasets.

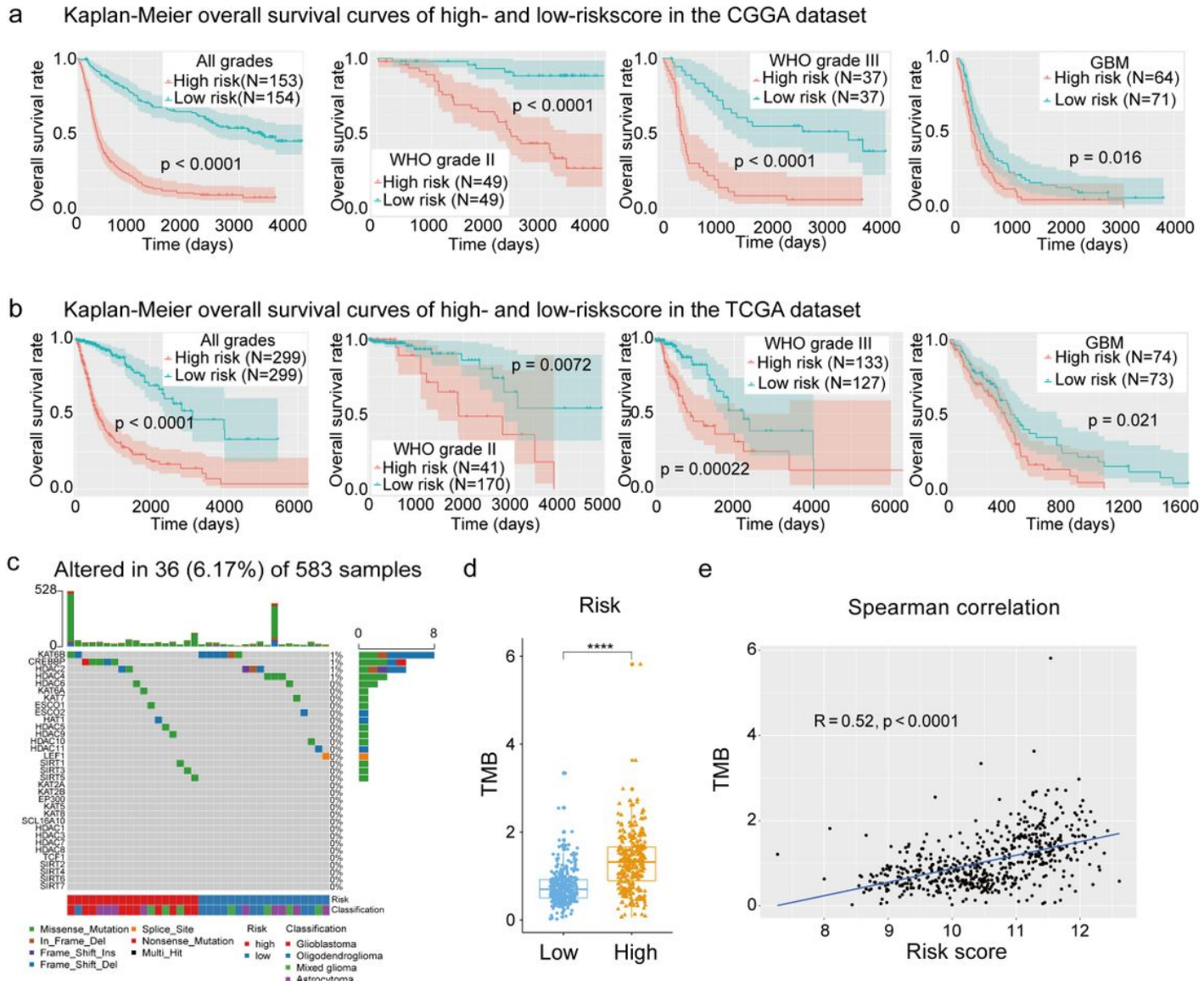

Figure 6

(A) Kaplan–Meier curves of low- and high-risk subgroups of all grades and each grade in the Chinese Glioma Genome Atlas (CGGA) training dataset. (B) Kaplan–Meier curves of low- and high-risk subgroups of each grade in The Cancer Genome Atlas (TGGA) training dataset. (C) Waterfall plot depicting the mutant status of lysine acetylation regulators (LARs). (D, E) Box plot showing that risk subgroups were significantly associated with the tumor mutation burden (TMB) and Spearman correlation analysis showed that the risk score was strongly correlated with TMB.

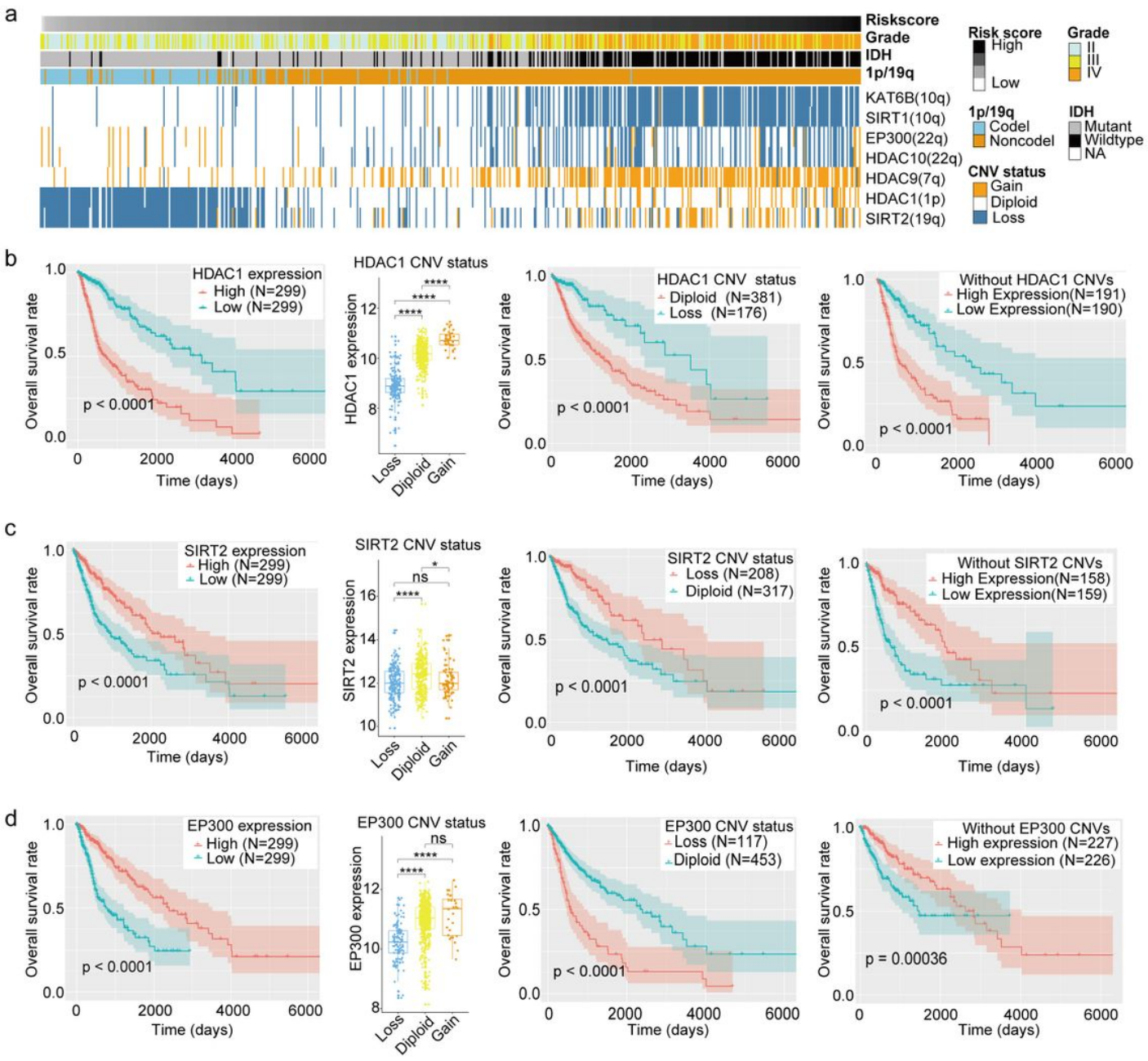

Figure 7

(A) Copy number variation (CNV) of seven lysine acetylation regulators (LARs) were significantly correlated with the risk score and other clinicopathological features. (B–D) Box plots indicated that different CNV status of HDAC1, SIRT2, and EP300 have different mRNA expression levels and the Kaplan–Meier curves revealed that HDAC1, SIRT2, and EP300 expression levels were associated with OS rates both in all samples and samples without their CNVs and their CNV status were also related to OS rates of glioma patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS6.tif
- FigureS5.tif
- FigureS4.tif
- FigureS3.tif
- FigureS2.tif
- SupplementaryTables.docx
- FigureS1.tif