Translational Oncology 14 (2021) 100979

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Hub gene identification and prognostic model construction for isocitrate dehydrogenase mutation in glioma

Yanfei Jia[#], Wenzhen Yang[#], Bo Tang, Qian Feng, Zhiqiang Dong[*]

Department of Neurosurgery, Lanzhou University Second Hospital, No. 82 Cuiyingmen, Chengguan District, Lanzhou, Gansu Province 730000, China

ARTICLE INFO

Keywords:
Glioma
IDH mutation
WGCNA
Machine learning
Prognosis

ABSTRACT

Our study attempted to identify hub genes related to isocitrate dehydrogenase (IDH) mutation in glioma and develop a prognostic model for IDH-mutant glioma patients. In a first step, ten hub genes significantly associated with the IDH status were identified by weighted gene coexpression analysis (WGCNA). The functional enrichment analysis demonstrated that the most enriched terms of these hub genes were cadherin binding and glutathione metabolism. Three of these hub genes were significantly linked with the survival of glioma patients. 328 samples of IDH-mutant glioma were separated into two datasets: a training set ($N = 228$) and a test set ($N = 100$). Based on the training set, we identified two IDH-mutant subtypes with significantly different pathological features by using consensus clustering. A 31 gene-signature was identified by the least absolute shrinkage and selection operator (LASSO) algorithm and used for establishing a differential prognostic model for IDH-mutant patients. In addition, the test set was employed for validating the prognostic model, and the model was proven to be of high value in classifying prognostic information of samples. The functional annotation revealed that the genes related to the model were mainly enriched in nuclear division, DNA replication, and cell cycle. Collectively, this study provided novel insights into the molecular mechanism of IDH mutation in glioma, and constructed a prognostic model which can be effective for predicting prognosis of glioma patients with IDH-mutation, which might promote the development of IDH target agents in glioma therapies and contribute to accurate prognostication and management in IDH-mutant glioma patients.

Introduction

As the most common primary intracranial tumor, glioma represents more than 80% of malignant brain tumors [40]. Glioma originates from glial cells and is a relatively rare tumor with significant mortality and recurrence rate [64]. Chemotherapy, radiotherapy and surgical excision are the common treatments for glioma [55]. The invasive growth of glioma blurs the boundaries of peripheral nerve tissue, which complicates the surgical removal of glioma [8]. Since glioma is one of central nervous system diseases, the lack of targeted treatment would easily cause serious side-effects, thus hampering the therapeutic efficacy. In recent years, despite substantial progress in the understanding of the molecular pathogenesis and the development of multimodal treatments of glioma, the prognosis of patients with gliomas remains dismal [39,52].

Isocitrate dehydrogenase (IDH) mutation is widely used as a molecular biomarker in glioma, which has been paid a strong clinical attention [10]. Several studies found that IDH mutations are also linked to other cancers, including acute myeloid leukemia and cholangiocarcinoma [42,47]. However, significant diagnostic and prognostic relevance of IDH mutation has been only proven in glioma [46]. Recently, the World Health Organization (WHO) has included the IDH mutational status into the classification of gliomas [34]. IDH mutation mostly occurs in lower grade glioma (LGG) and about 5% of primary glioblastomas (GBM), but absent in the other primary brain tumors. Therefore, we speculated that the genes associated with the IDH mutation might be used as potential biomarkers for glioma diagnosis. In recent years, more and more studies on the core genes and mechanisms that regulate IDH mutations in gliomas have emerged (Y. Q. [33,65]). Li et al. indicated that, in glioma, MEGF10 is markedly correlated with IDH mutation [30]. Several recent studies have revealed therapeutic targets of IDH-mutant LGG by bioinformatics analysis [29,65]. Nevertheless, most of the studies on IDH mutation in glioma were concentrated on LGG; the expression pattern and predictive value of IDH-related genes in subjects with glioma remain to be further elucidated.

As one of the widely used R libraries, weighted gene co-expression network analysis (WGCNA) is important because it has a big merit in the discovery of gene modules containing genes with analogous expression blueprints, called co-expressed genes, that are associated with a given

* Corresponding author.
E-mail address: dongzqiang443@aliyun.com (Z. Dong).
# These two authors contributed equally to this work.

https://doi.org/10.1016/j.tranon.2020.100979
Received 21 September 2020; Received in revised form 9 November 2020; Accepted 25 November 2020

feature [28]. After the modules are identified, the correlation between the modules and clinical traits can be easily evaluated by using eigengene network methodology. This powerful tool has been commonly acknowledged and successfully applied to various diseases such as prostate cancer [6], ischemic stroke [62] and Type 2 diabetes mellitus [15].

Novel technologies have empowered and accelerated the accumulation of patient data in recent years. However, with a character of high-throughput, these datasets are often cross-sectional, noisy, sparse, and lack of statistical power, rendering conventional data analytical approaches unable to mine useful biological insights from these data. Machine learning, an integral part of computational biology, is being increasingly adopted for analyzing health-care and biomedical data to address some of these challenges. In the last decade, machine learning algorithms have gained popularity in medicine and scientific research [35,58,75]. A study conducted by Zhang and co-workers successfully discriminated hepatocellular carcinoma tissues from cirrhosis tissues by establishing a computational model based on machine learning method [75]. Hammond et al. revealed different roles of Aβ, tau, and neurodegeneration in Alzheimer's disease development by using the random forest machine learning method [19]. In the selected datasets, machine learning methods are capable of identifying statistically significant patterns that could provide predictions with good accuracy.

The current study sought to identify the hub genes correlated with IDH mutation and establish a prognostic model for glioma patients with IDH-mutation. A total of ten hub genes were discovered by using WGCNA, which showed strong correlation with survival of glioma patients. By using bioinformatics analysis and machine learning methods, we identified a 31 gene-signature which was used to establish a prognostic model based on the profile of IDH-mutant glioma patients from TCGA database. The results of this study indicated two novel subtypes of IDH mutation with distinct molecular and clinical characteristics and provided a robust model for predicting the prognosis of IDH-mutant glioma patients.

Materials and methods

Data collection and processing

In brief, gene expression RNA-seq of glioma samples were downloaded from TCGA database along with clinical and subtype information of corresponding patients. A total of 669 samples, containing 515 samples of LGG patients and 154 samples of GBM patients were extracted for identification of hub genes associated with IDH status. The flowchart for this part of analysis was depicted in **Figure S1**.

For identifying gene signature of IDH-mutant glioma and establishing prognostic model, gene expression and subtype data of 328 IDH-mutant glioma samples were downloaded from TCGA database and split into two sets: a training set ($N = 228$) and a test set ($N = 100$).

WGCNA construction and key module identification

Before constructing WGCNA, outlier samples were detected by using the standardized connectivity (Z.K) method as proposed before, and the samples with a Z.K score < -2 were screened as outliers [38]. In a network, the Z.K score represents the overall strength of connections between a given node and all of the other nodes. After eliminating 27 outlier samples, the expression data profile was used for network generation by the R library WGCNA [28]. Initially, the soft-thresholding for network construction was determined: first, the similarity amid two co-expressed genes (m, n) was set as $S_{mn} = |\text{cor (m, n)}|$. Meanwhile, correlation of gene adjacency was conducted using a power function (pickSoftThreshold function of WGCNA): $a_{mn} = \text{power}(S_{mn}, \beta) = |S_{mn}|^{\beta}$. Then, after setting the power value in the range of 1 to 30, average connectivity and scale freedom were detected by a gradient method. The modules were generated by hierarchical average linkage clustering approach. The modules were assigned to different colors for visualization and a cut line of 0.25 was set to combine the similar expression modules. The genes assembled into the gray module were removed in the following analysis. The module with first rank absolute MS was defined as the key module that was significantly associated with the clinical trait we studied.

Identification of transcription factors

In order to further understand the regulatory information of the key module, transcription factors (TFs) were identified by the comprehensive web-based tool Enrichr (http://amp.pharm.mssm.edu/Enrichr/). The TFs-target genes interactome was generated after importing key module genes into Enrichr, and the visualization of their regulatory network was performed by using Cytoscape 3.4.0 software.

Analysis of hub genes associated with IDH status

Hub genes were identified by the module connectivity. First, full weighted network of the key module was exported for extracting the obtained gene network. Then, the subnetworks were generated with MCODE in Cytoscape. Finally, subnetwork genes displaying high module membership (MM), high gene significance (GS) and high weight in the correlation network were screened as the hub genes. The correlation of the hub genes and IDH status was validated by the Wilcoxon test (p-value cutoff =0.05). Moreover, further analyses (overall survival (OS) analysis and disease-free survival (RFS) analysis) for the hub genes were performed in Gene Expression Profiling Interactive Analysis (GEPIA) (http://gepia.cancer-pku.cn/).

Consensus clustering

Based on the comparison of gene expression profile, the R package "CancerSubtypes" [69] was applied to conduct consensus clustering for determining subgroups of IDH mutant glioma. The cumulative distribution function (CDF) and consensus matrices were used to assess the optimal number (k) of subgroups. Then the survival analysis of subgroups was performed by the "survival" R package and visualized by the "survminer" package.

Screening of differential expressed genes (DEGs)

The IDH-mutant glioma samples were assigned to cluster1 or cluster2 according to the result of consensus clustering analysis. Next, the R "limma" package was run for screening of DEGs, with P value < 0.05 and $|\log$ (fold change) (FC) $| > 2$ as cutoff criteria. The "ComplexHeatmap" package was used to visualize the expression heatmap of DEGs in R.

Gene signature identification and prognostic model establishment

With the aim of computing the predictive significance of the identified DEGs, the univariate Cox regression analysis was performed. Afterward, the DEGs with high predictive significance (p-value cutoff = 0.05) were used for feature selection by using the least absolute shrinkage and selection operator (LASSO). Additionally, the candidate prognostic model was established via different machine learning models and the accuracy of each model was gaged by performing the receiver characteristic operator curve (ROC) analysis. Then the model which showed the highest value of area under curve (AUC) among all candidate models was selected as the prognostic model for further analysis. The risk score was calculated by LASSO based on the gene expression level and regression coefficients.

Evaluation of the prognostic model

After calculating the risk scores, the median risk score was set as a threshold to divide patients of test set ($N = 100$) into two groups: low-

risk patients and high-risk patients. The overall survival (OS) was evaluated by means of the Kaplan–Meier survival analysis accompanied with a two-sided log-rank test. Pathological and molecular differences among both risk groups were assessed using the Wilcoxon and Kruskal-Wallis tests. The visualization of this process was implemented in the R package "ggpubr". $P < 0.05$ was considered statistically significant.

Functional enrichment analysis

Using ClusterProfiler (R package), we performed functional gene enrichment analysis to reveal the functional role of the identified genes. Briefly, using the gene lists as input, the results of the enriched terms of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were generated after running clusterProfiler. The adjusted P-value <0.05 was chosen as the threshold for the identification of significant GO terms and pathways.

Results

Identification of the hub genes related to IDH status by WGCNA

To find the key module associated with IDH status in glioma, WGCNA was performed based on 669 tumor samples from TCGA database. By using Z.K method, a total of 27 outlier samples were detected and removed. After eliminating the outlier samples, the sample dendrogram and trait heatmap were constructed (Fig. 1**A**). As shown in Fig. 1**B**, the network had a relatively high-average connectivity when the scale independence was 0.85. In this case, $\beta = 24$ was retained as the power threshold and 20 modules were generated after hierarchical clustering analysis (Fig. 1**C**). Then, the modules whose eigengenes showed correlation above 0.75 (cut line for merging of modules = 0.25) were merged (Y. [5]), and ten modules were finally obtained. The eigengene adjacency heatmap was shown in Fig. 1**D** and indicated that modules were independent from each other. As shown in Fig. 1**E** and **Figure S2**, the cyan module was most credibly associated with IDH status ($r = -0.87$, $p = 4e\text{-}198$). For this reason, for a more in-depth analysis, we chose the cyan module as the model module to explore the genes most associated with IDH status.

The scatter plots of GS for IDH status vs. MM in the cyan module were depicted in Fig. 2**A**. In the cyan module, a total of six TFs (KLF4, TP63, SOX2, TP53, FOSL2 and SMC3) were identified, and the TF-target interactome was displayed in Fig. 2**B**. There were 572 genes in the cyan module. Setting cut-off value as MM ≥ median and GS ≥ median, 81 genes were obtained. Afterward, full weighted network of the cyan module was exported to extract subnetworks using MCODE in Cytoscape. Ten subnetwork genes displaying high MM, high GS and high weight, which included GPX8, CCDC109B, IGFBP2, LINC00152, LOC541471, METTL7B, S100A4, EMP3, CLIC1, and TAGLN2, were screened as the hub genes (Fig. 2**C**). Boxplots of hub genes based on IDH status were as indicated in Fig. 2**D**, and we observed that all the hub genes were markedly influenced by the IDH status in glioma ($P<0.0001$). Meanwhile, we found that the expressions of these hub genes were also markedly different between cases with and without IDH mutation in both LGG and GBM (**Figure S3 and S4**). The most significant functional terms were as shown in **Table S1**. Functional annotation revealed that these ten hub genes were mainly involved in cadherin binding, insulin-like growth factor binding, glutathione metabolism, and arachidonic acid metabolism.

Validation of the hub genes in the GEPIA

LOC541471 is a long non-coding gene. There was no survival data about it in GEPIA. As shown in **Figure S5A**, LGG patients with up-regulated hub genes had a remarkably shorter survival time. In the meanwhile, except for IGFBP2, LINC00152 and S100A4, the expression of the other six hub genes showed no significant association with the survival time of GBM patients (**Figure S5B**). The disease-free survival (RFS) analysis exhibited a result similar to the OS analysis (**Figure S6**).

Classification of IDH-mutant subtypes based on consensus clustering

Based on the gene profile and clinical phenotype information of 669 glioma patients (including IDH mutant and IDH wild type patients), the above analyses uncovered a set of hub genes which were closely related to IDH status, providing a novel insight of the mechanism of IDH mutation in glioma. To further identify molecular subtypes or new glioma markers for monitoring IDH-mutant glioma patients accurately, we classified IDH-mutant subtypes in glioma by consensus clustering. The flowchart of gene signature identification and prognostic model construction was displayed in **Figure *S7***. A total of 228 samples in the training set were employed for the consensus clustering to obtain the IDH-mutant subtypes. As shown in Fig. 3**A-C**, to guarantee high correlation within the group, consensus clustering divided the patients into two distinctive clusters based on a slow growth rate of CDF value. The heatmap of the two clusters (Fig. 3**E**) revealed an important difference in the prognosis between samples of these two clusters (Fig. 3**D**), suggesting that the patients of cluster1 had significantly poorer prognosis in comparison with those of cluster2. The average silhouette of the two clusters were 0.93 and 0.88 respectively, indicating the high degree of separation of the two clusters (Fig. 3**F**). The clinical characteristics of the two clusters were summarized in Table 1. As we can see, the patients in cluster1 were significantly associated with older age (82.7%, $P = 7.14e\text{-}07$), astrocytomas (75%, $P = 3.96e\text{-}04$), and high grade (100%, $P = 3.43e\text{-}03$). Consistent with the result of survival analysis, cluster1 was strongly correlated with poor clinical outcome ($P = 1.18e\text{-}03$). These results suggested that the two subtypes based on the IDH mutation may provide guidance for clinicians on personalized treatments and diagnoses by identifying differences in prognosis for each epigenetic subtype.

Identification of DEGs between IDH mutant subtypes

Samples divided according to the two clusters were employed for identifying DEGs by using the “limma” R package. The volcano plots for all genes and the heatmap of the genes with the most significant differential expression were shown in Fig. 4**A and 4B**. A total of 283 DEGs between the samples in cluster1 and cluster2 were identified under the criteria of P value < 0.05 and |log FC| > 2. Among these DEGs, 155 were highly-expressed in cluster1 (downregulated in cluster2) while 128 DEGs were lowly-expressed (upregulated in cluster2). After intersecting the key module genes identified by WGCNA with the DEGs, a total of 21 genes were obtained (Fig. 4**C**), which suggested that these 21 genes were not only associated with IDH status but also related to IDH mutant subtypes. However, no hub gene was found in the intersection result. Then we performed the functional enrichment annotation for the DEGs. As shown in Fig. 4**D**, the genes up-regulated in cluster1 were mainly involved in the process of cell division including organelle fission, nuclear division and chromosome segregation; we also found that these genes were equally implicated in cellular processes such as cell cycle and cellular senescence. The p53 signaling pathway was also associated with these DEGs. For the genes downregulated in cluster1, enrichment in transport-related processes including substrate-specific channel activity were the most prevalently recorded (Fig. 4**F**). The function enrichment of the 21 intersection genes was also conducted. The results were listed in Table S2, indicating that these genes were mainly enriched in positive regulation of receptor-mediated endocytosis, regulation of protein processing and regulation of protein maturation.

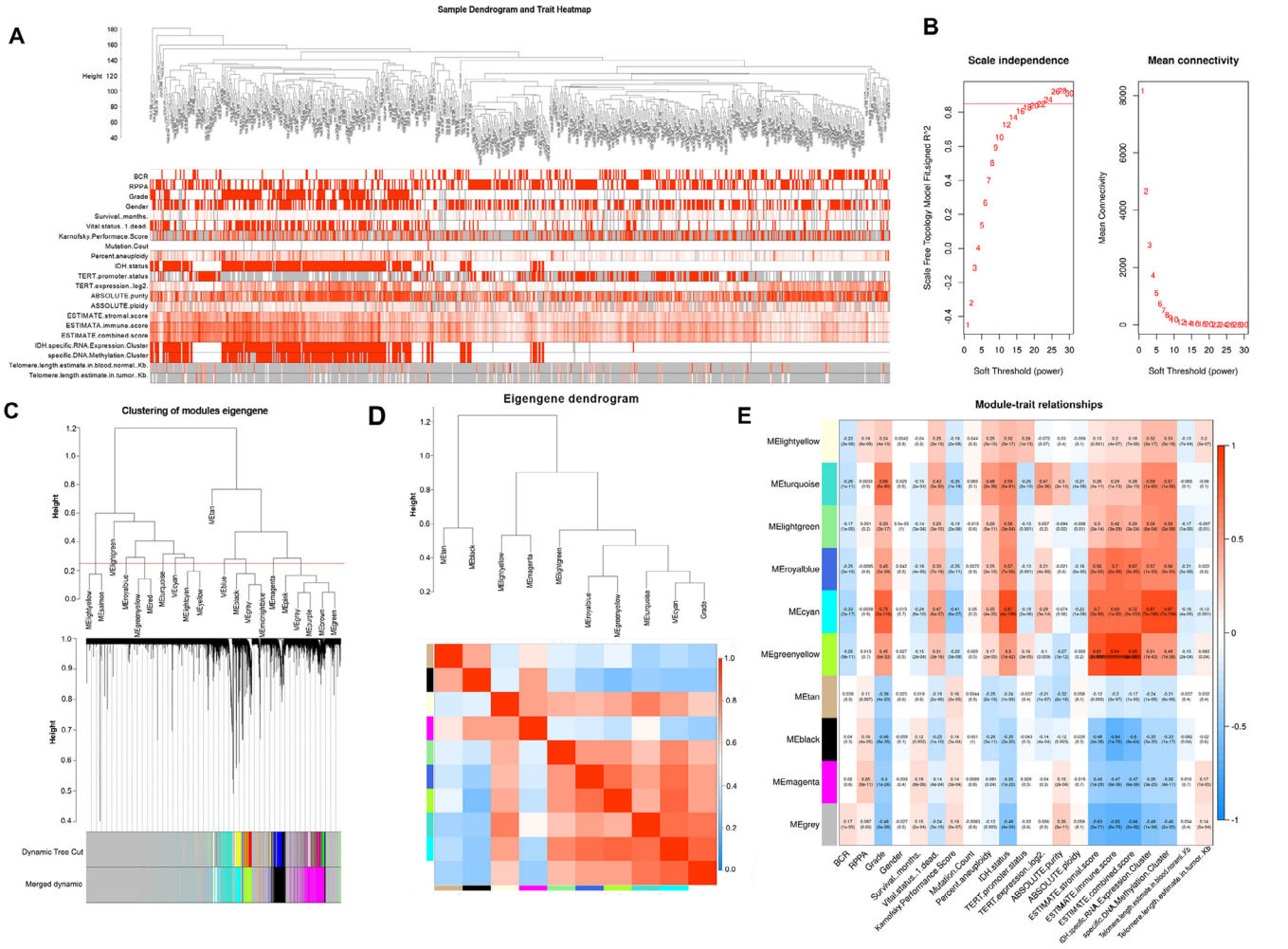

Fig. 1. Identification of the key module significantly associated with IDH status (mutant and wild type) by weighted gene co-expression network analysis (WGCNA). (A) Sample dendrogram and trait heatmap after excluding 27 outlier samples, based on the expression data of glioma patients in TCGA, which contained 515 LGG and 154 GBM samples. (B) Determination of soft-thresholding power (β) by analyzing (left) scale-free fit index and (right) mean connectivity; 24 was the most fit soft-thresholding power value in this study. (C) Dendrogram of consensus module eigengenes. The eigengenes groups (20) were merged according to merging threshold (the red line) due to their similarity; bottom: gene dendrogram was acquired after clustering the dissimilarity. A total of ten modules were finally generated. (D) Heatmap plot of the adjacencies of modules, suggesting modules were independent from each other; red represents positive correlation while blue represents negative correlation. (E) Heatmap of the correlation between different traits of glioma and module eignegenes. Each row and column correspond to a consensus module and trait, respectively. The correlations (with the p values described below) of each trait and module eigengenes were depicted in the corresponding cell. The cells are colored according to the correlation (red, positively correlated; blue, negative correlated), the intensity of the color is related to the strength of the correlation. The cyan module was the most significantly correlated with IDH status ($r = 0.87$, $p = 4e\text{-}198$).

Table 1
Characteristics of IDH-mutant glioma patients in cluster 1 and 2 from training set.

Features		N	Cluster 1	Cluster 2	P-value
Total		228	96	132	
Age	Young (<50)	176	87	89	7.14e-
	Old (≥50)	52	9	43	07
Gender	Male	130	57	73	0.5418
	Female	98	39	59	
Histology	Glioblastoma	3	3	0	3.96e-
	Astrocytoma	70	53	17	04
	Oligodendroglioma	100	10	90	
	Oligoastrocytoma	55	30	25	
Grade	G2	118	40	78	3.43e-
	G3	107	53	54	03
	G4	3	3	0	
Vital	Alive	178	65	113	1.18e-
Status	Dead	50	31	19	03

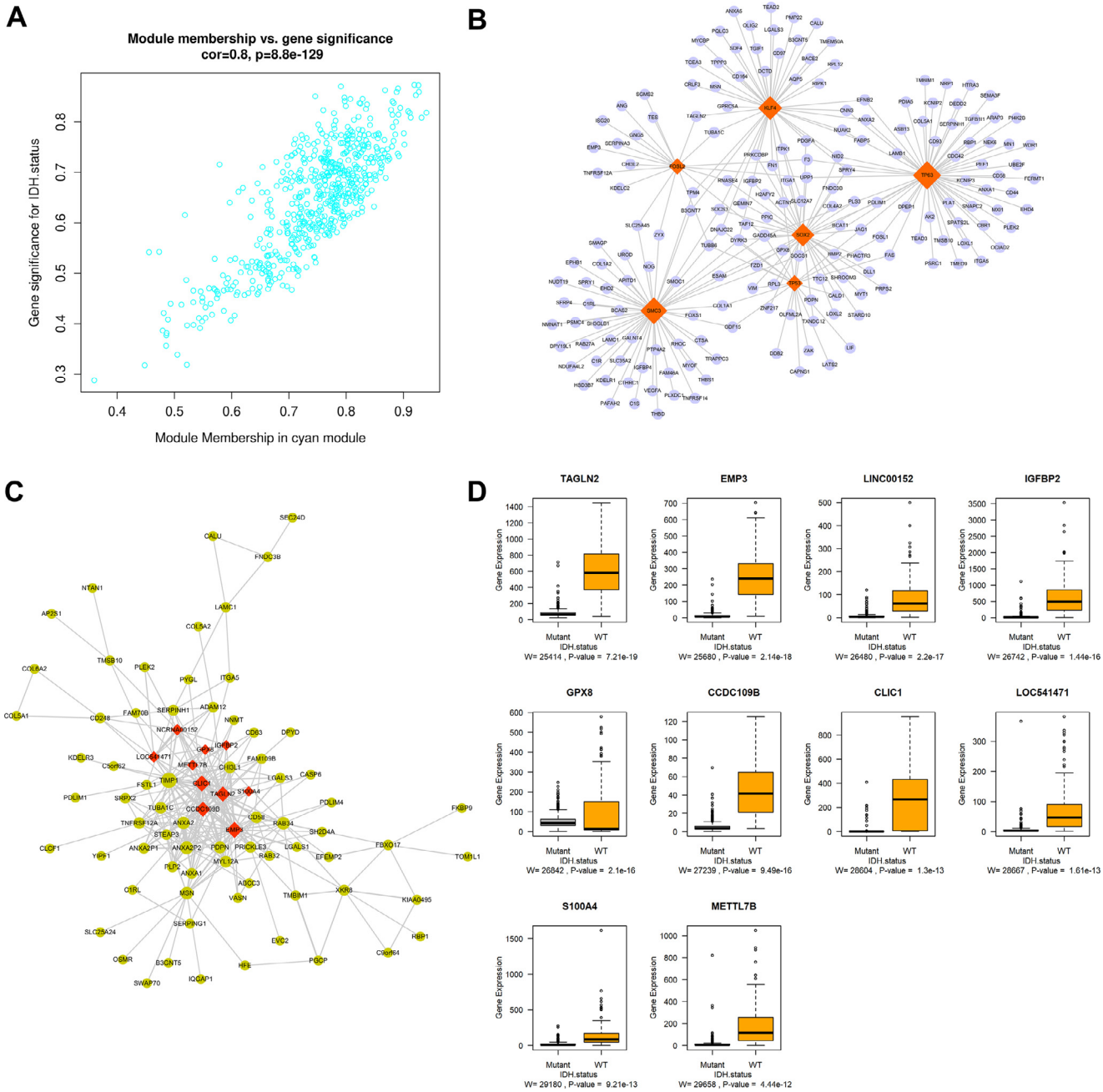

Fig. 2. Analysis of hub genes in the cyan module. (A) Scatter plot of module eigengenes related to IDH status in the cyan module. (B) The regulatory network of TF-target gene in the cyan module. Red diamonds symbolize the TFs, and purple nodes symbolize the genes. (C) The network of hub genes, suggesting the hub genes exhibited high connectivity with other genes. (D) Boxplots of the hub genes showing the expressions of all the hub genes exhibited statistical discrepancy between IDH mutant ($N = 432$) and IDH wild type ($N = 237$) (WT) of glioma patients.
Note: TF, transcription factor.

Construction of prognostic model for IDH-mutant glioma patients

By using a univariate Cox regression, we found that all the DEGs between the cluster1 and cluster2 subtypes were significantly linked with OS of IDH-mutant glioma patients in the training set. Thus, prognosis gene signature were identified based on these 283 DEGs. Through the LASSO regression algorithm, a signature of 31 genes was finally obtained (Fig. 5**A-C**). Among 31 genes in the signature, only CD97 was identified in WGCNA as an IDH status-related gene (Fig. 5**D**). Then, the 31-gene signature was applied to construct prognostic models by using different machine learning approaches including back-propagation neural network (BPNN), support vector machine (SVM), convolutional neural network (CNN), eXtreme Gradient Boosting (Xgboost), random forest (RF), and LASSO. As shown in **Figure S8**, the model constructed by LASSO showed the highest accuracy of prognosis in IDH-mutant glioma patients, with an AUC of 0.8743.

Validation of the prognostic model on the test set

Based on machine learning method (LASSO), the risk scores for IDH-mutant glioma patients were computed. The median risk score was considered as the risk score threshold for grouping IDH-mutant patients in the test set as two patient groups (low-risk group and high-risk group). Before validating the prognostic model established based on the 31-gene

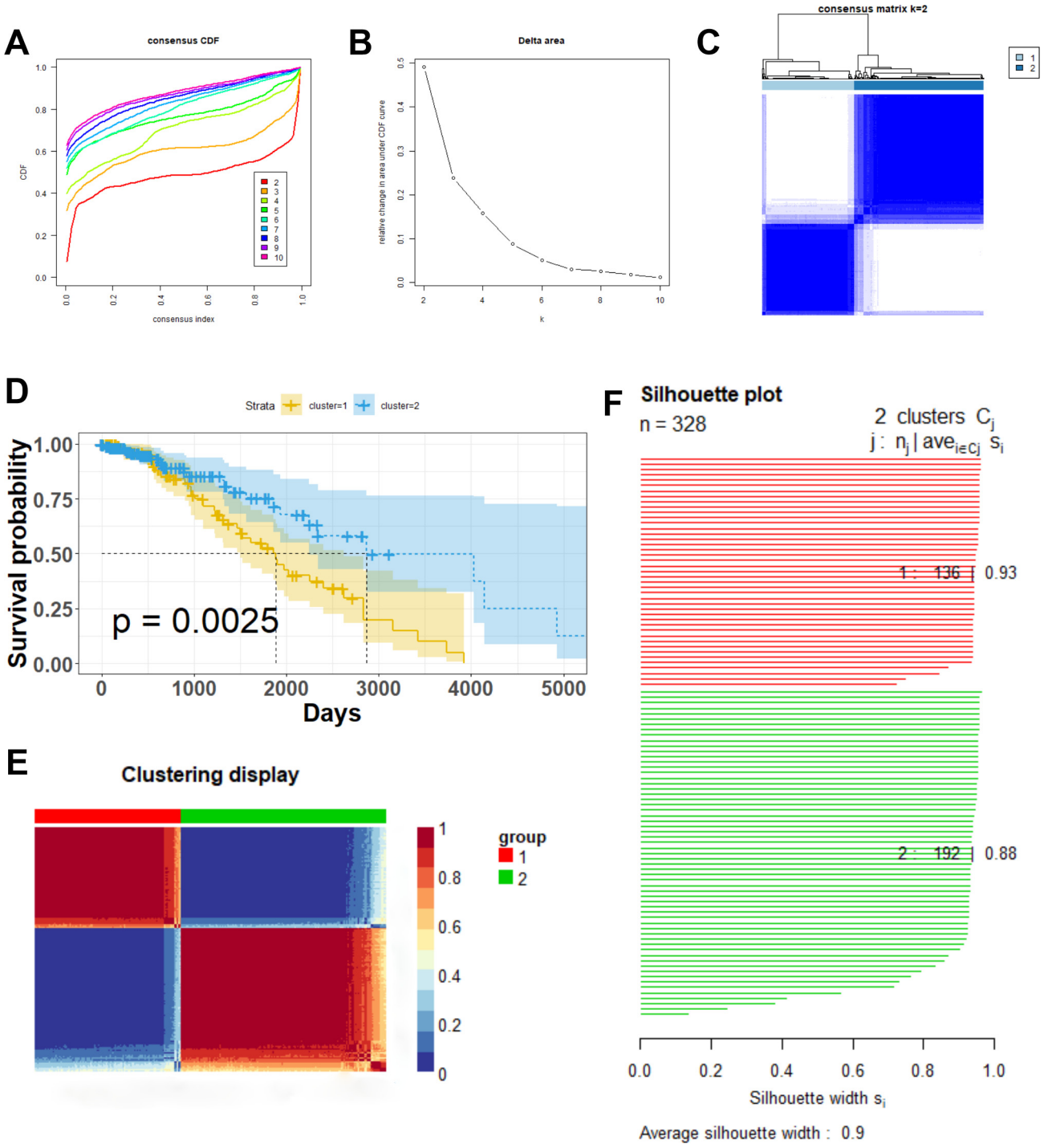

Fig. 3. The consensus clustering of IDH-mutant glioma samples classification in the training set ($N = 228$). (A) The cumulative distribution function (CDF) of the consensus matrix for k with range of 2 to 10. (B) Delta Area Plot showed the change of area under the CDF curve with diverse k values. (C) Color-coded heat map corresponding to the consensus matrix for $k = 2$ obtained by applying consensus clustering. (D) Survival analysis of IDH-mutant glioma patients from Cluster 1 and 2. (E) Heat map of the two clusters. (F) Silhouette plot used to examine the degree of separation of clusters.

signature, the correlation among the risk scores of this signature and the clinicopathological features of patients was evaluated. As shown in Fig. 6**A-D**, a high discrepancy in the risk scores was noted amid patients sorted by age ($P = 0.045$), WHO grade ($P = 0.025$) and histology ($P = 0.038$), while the patient gender did not show a correlation with the risk scores ($P = 0.33$). In ROC curve analysis based on the risk model, age, grade, and histology, we found that only the risk model showed high efficiency in predicting vital status **(**Fig. 6**E)** and distinguishing cluster1 from cluster2 **(**Fig. 6**F)**. The prognostic model was validated based on the test set (Fig. 7). The OS of subjects with high-risk scores was obviously shorter relatively to patients with low-risk scores ($P < 0.0001$, Fig. 7**A**). As shown in Fig. 7**B and** 7**C**, we found that whether the patients were old (age≥50) or young, patients of the high-risk group showed a lower OS compared to those of the low-risk group. Next, we gaged the value of the prognostic model based on the glioma WHO grade II and III, respectively. For the patients with WHO grade II or III, the

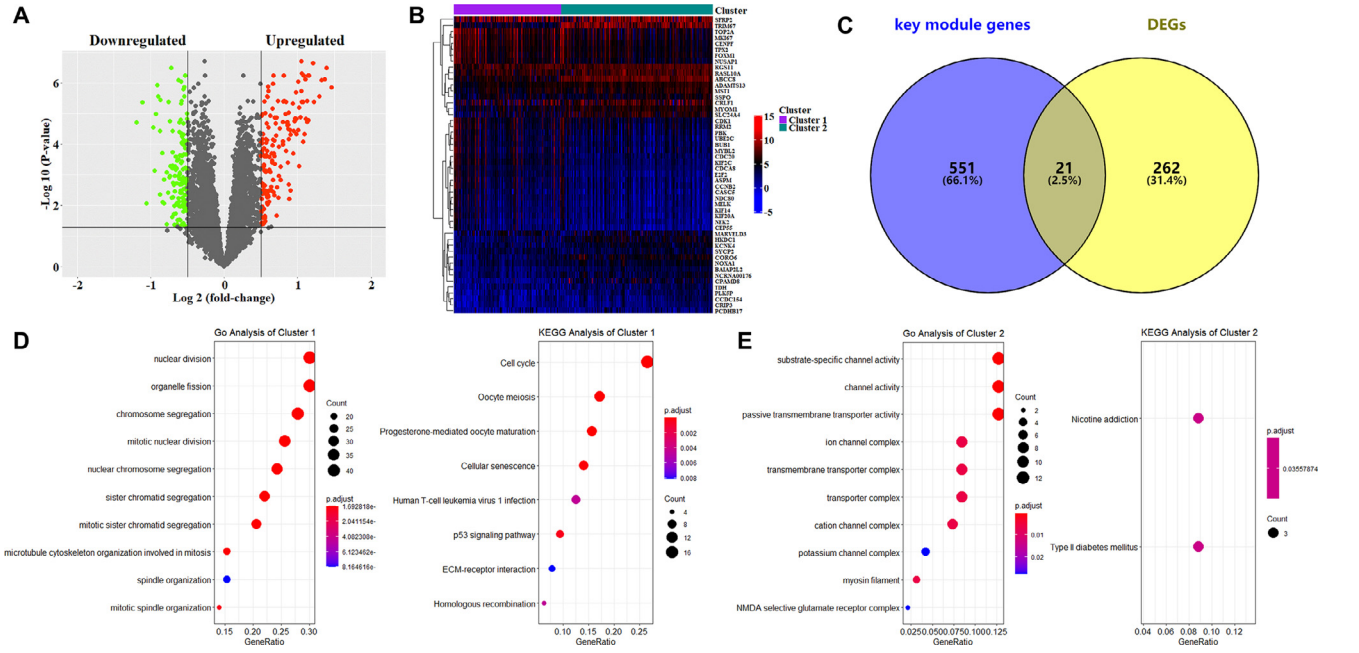

Fig. 4. Differential expressed genes (DEGs) analysis for IDH-mutant glioma patients from Cluster 1 and 2, with a threshold of $\log_2|FC| > 1$ and $P < 0.05$. (A) Volcano plot of DEGs identification. (B) Heatmap of the top 50 DEGs. (C) The Venn diagram showed that a total of 22 common genes between the key module genes and DEGs. (D) Functional enrichment annotation of the up-regulated genes and (E) down-regulated genes. The size of the circle indicates the number of genes and the y-axis represents the terms of GO and KEGG pathway. The color of the circle showed the p.adjusted value of each term. The redder the color the higher the p.adjusted value.

Note: FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

high-risk group patients displayed a strong correlation with shorter OS compared with the low-risk group patients ($P = 0.011$ and $P = 0.0074$, Fig. 7**F and** 7**G**). Additionally, we also investigated the predictive value of the model in IDH-mutant glioma patients which were divided by different types of histological features. As shown in Fig. 7**H-J**, the high-risk group patients had a bad prognosis in both the oligoastrocytoma patients ($P = 0.046$) and oligodendroglioma patients ($P = 0.0032$), but showed no significant difference in the astrocytoma patients ($P = 0.8$).

Functional enrichment of the 31-gene prognostic signature

To further explore the mechanism of the 31-gene prognostic signature for IDH-mutant patients, the GO and KEGG functional enrichment was performed. Initially, by using Pearson correlation analysis, we identified the genes positively ($R > 0.5$ and $P < 0.0001$) and negatively ($R < -0.5$ and $P < 0.0001$) correlated with the risk sore of gene signature. Then, the results were displayed in Fig. 8. The functional analysis of the positively related genes indicated that these genes were those driving the biological processes of nuclear division, chromosome segregation and DNA replication (Fig. 8**A**). In the cellular component ontology, the positively related genes were significantly enriched in chromosomal region and spindle (Fig. 8**B**). The GO terms of ATPase activity and catalytic activity acting on DNA were significantly enriched molecular functions (Fig. 8**C**). Cell cycle was the most mainly enriched pathways resulting from the KEGG pathway analysis (Fig. 8**D**). Notably, the most enriched terms of genes negatively related to risk score were organelle fission and nuclear division in the biological process, nuclear envelope and chromosomal region in cellular component, catalytic activity acting on DNA and helicase activity for molecular function, and cell cycle and Fanconi anemia pathway in the KEGG pathway database (Fig. 8**E-H**).

Discussion

Glioma is the most common primary intracranial tumor with varying malignancy grades I-IV [17]. IDH is identified as a major prognostic marker which is significantly related to tumor grade in glioma. IDH mutation occurs in the majority of LGG (WHO grade II/III), but in only 5% –apparently – primary glioblastomas (WHO grade IV) (GBM) [46]. In this study, we applied WGCNA to screen out ten hub genes associated with IDH status and constructed a prognostic model for IDH-mutant patients via bioinformatics and machine learning tools.

Initially, we applied WGCNA to unearth the main modules and hub genes associated with IDH status in gliomas (LGG and GBM). For similar samples with low connectivity, the power (β) estimation would inflate the output [37]. Thus, the outlier samples with low connectivity were detected and removed by Z.K method before constructing a co-expression network. The Z.K examination of all samples confers both the flexibility and the efficacy required for analyzing large datasets by identifying and removing outliers [38]. IDH status was most significantly correlated with the cyan module, thus the cyan module was selected as the key module. After filtering GS and MM value, a total of ten hub genes (GPX8, CCDC109B, IGFBP2, LINC00152, LOC541471, METTL7B, S100A4, EMP3, CLIC1, and TAGLN2) were obtained and all of them were significant in distinguishing IDH status in glioma. Besides, since IDH mutation occurrence is quite disproportionally distributed in LGG and GBM, we validated whether these genes were true surrogates for IDH status in glioma by comparing their expressions from samples with and without IDH mutation in LGG and GBM, respectively. The results revealed increased expression levels of these hub genes in IDH wild type samples, suggesting that these hub genes could be used as true surrogates for IDH status. Previous studies demonstrated that the IDH wild type patients generally have poorer clinical outcome than the IDH mutant patients [3]. Our results are in line with the published studies, showing that genes correlated with IDH wild type status are expected to be associated with worse prognosis. Then, we found that these hub genes were all cancer-related genes, and many studies reported that some of them had significant correlation with glioma. Xu et al. identified CCDC109B as an oncogene and a prognostic marker in human gliomas; they also revealed that the expression of CCDC109B was regulated by HIF1α [68]. Yao et al. found that HIF1α is associated with

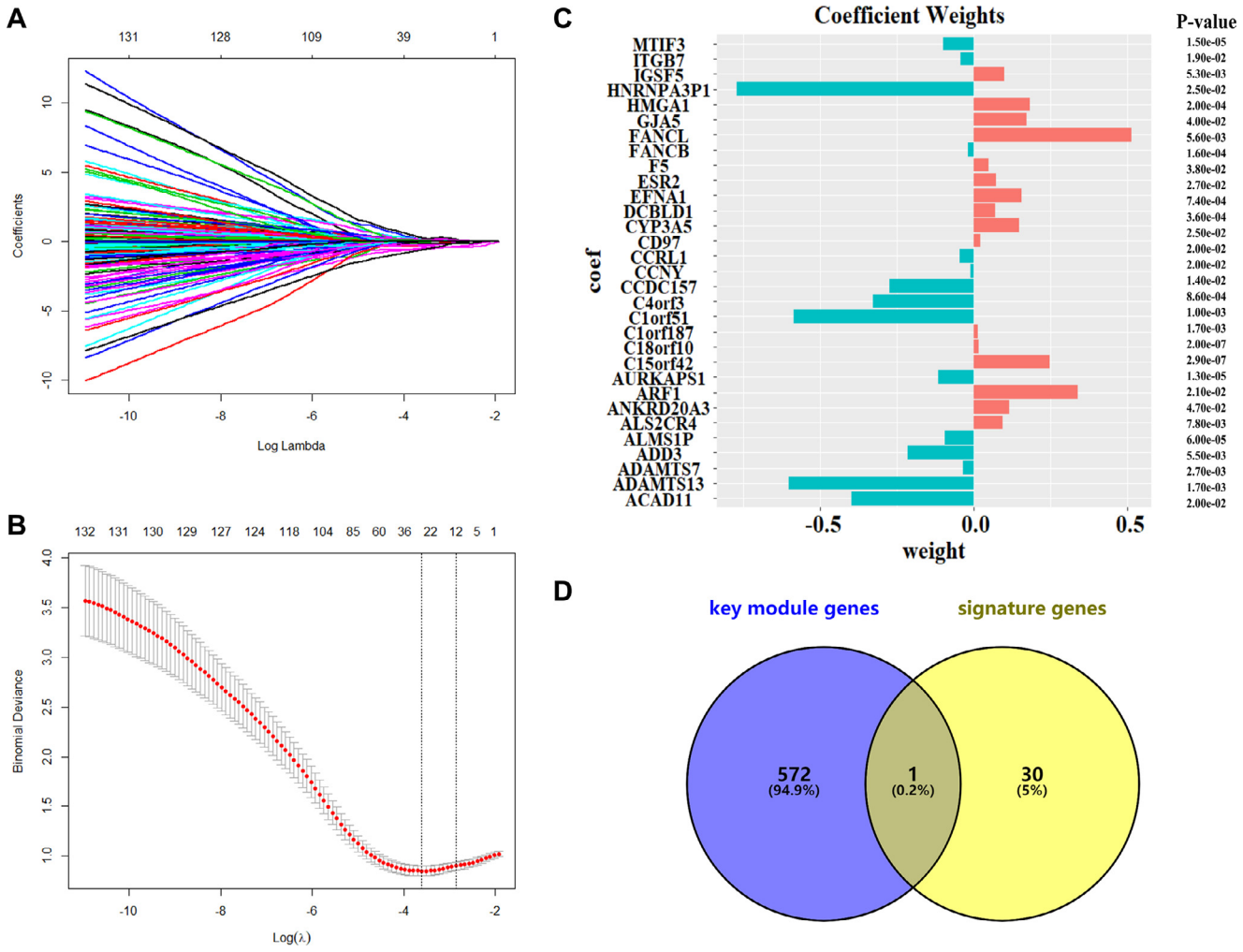

Fig. 5. Identification of gene signatures by LASSO regression approach based on training set from IDH-mutant glioma patients. (A) Cross-validation for choosing tuning parameter of LASSO model; (B) LASSO coefficient profiles of 283 genes. (C) Coefficient values for the 31 genes by LASSO. Note: LASSO, least absolute shrinkage and selection operator. (D) The Venn diagram showed that only one common gene between the key module genes and signature genes.

hypoxia and acidity in wild type but not IDH1-mutant gliomas [72]. As known, IDH1 mutation has been linked with a better prognosis in glioma patients [61,71]. Analogously, our results showed that CCDC109B up-regulated expression was significantly connected with IDH1 wild type and poor prognosis. Thus, we speculated that HIF1α might play a key role in the relation of CCDC109B with IDH mutation in glioma. Several researchers prompted that IGFBP2 is implicated in glioma cell proliferation, migration and invasion, which was also a prognostic biomarker in GBM [2,43,45]. A recent study conducted by Yuan and his co-workers proved that the expression of IGFBP2 was related to the occurrence of IDH1 mutation [74]. LOC541471 was reported to be involved in the oxidative phosphorylation of GBM, which was significantly associated with overall survival (X. [4]). EMP3 is one of the molecular markers associated with clinical outcome in glioma [16]. Interestingly, EMP3 overexpression is most detected in GBM, but the majority of LGG samples showed low or no expression of EMP3, which is the opposite of the IDH mutation [12]. Besides, Mellai et al. indicated that hypermethylation of EMP3 was strongly associated with IDH mutation [36]. CLIC1 is a regulator involved in the aggressive aspects of numerous solid tumors including glioma [44]. Peretti et al. found that the expression level of CLIC1 was correlated with poor prognosis of GBM [44]. Djuric and his colleagues found that CLIC1 was enriched in IDH wild type glioma cells [36]. Han et al. identified TAGLN2 as a possible predictive biomarker driving the tumorigenesis in gliomas [20]. A recent study demonstrated that increased TAGLN2 expression may be related to invasion and poor prognosis in IDH wild type gliomas [1]. Combining these previous literature with our findings, we supposed that such hub genes make a contribution in the correlation of prognosis with IDH status in glioma. Subsequent functional enrichment of such hub genes revealed that the majority of these genes were significantly enriched in cadherin binding, insulin-like growth factor binding, glutathione metabolism, and arachidonic acid metabolism. Increasing studies suggested that IDH mutation is closely related to glutathione metabolism [7,13]. Recently, several studies highlighted the value of glutathione metabolism as the novel therapeutic target of IDH-mutant glioma [57,73]. Based on the results of functional annotation, we anticipated that hub genes identified herein were involved in the glioma IDH mutation via glutathione metabolism related pathways.

In order to clarify the mechanisms beneath IDH mutations in glioma, we built a TF-gene interactome based on genes in the cyan module. All the TFs we identified (KLF4, TP63, SOX2, TP53, FOSL2, SMC3) are significantly related to tumor progression, which suggested a correlation between IDH mutation and glioma progression. Several studies indicated that KLF4 can play a role in suppressing tumors in various human cancer types by regulating cell-cycle genes [26,63,76]. KLF4 also plays a central role in the prevention of Epithelial to mesenchymal transition (EMT) and metastasis [59]. TP63 was reported to promote squamous cancer progression by activating both MEK/ERK1/2 and PI3K/AKT axes

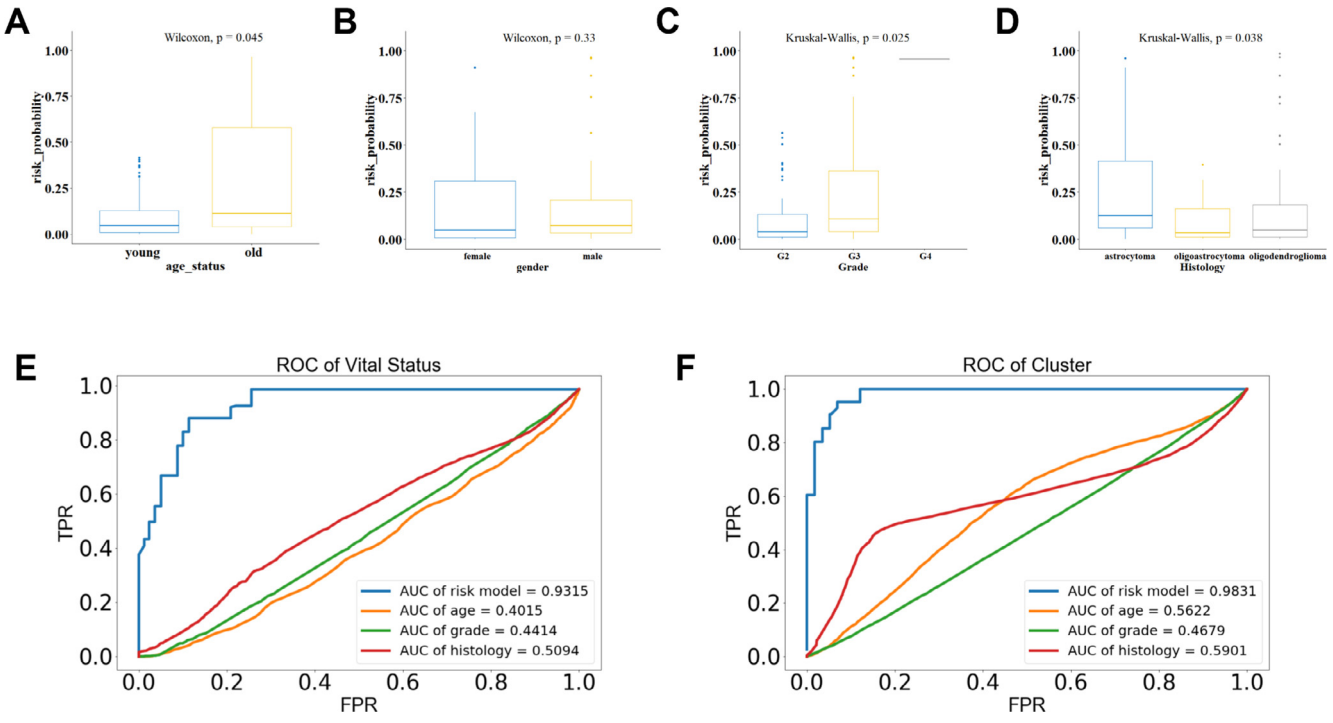

Fig. 6. Association between the gene signature and the pathological characteristics in test set from IDH-mutant glioma patients. The risk scores distribution in glioma patients stratified by (A) age, (B) gender, (C) WHO grade, (D) histology. Receiver characteristic operator curves (ROC) showed the predictive efficiency of the risk signature in (E) vital status and (F) Cluster 1/2.

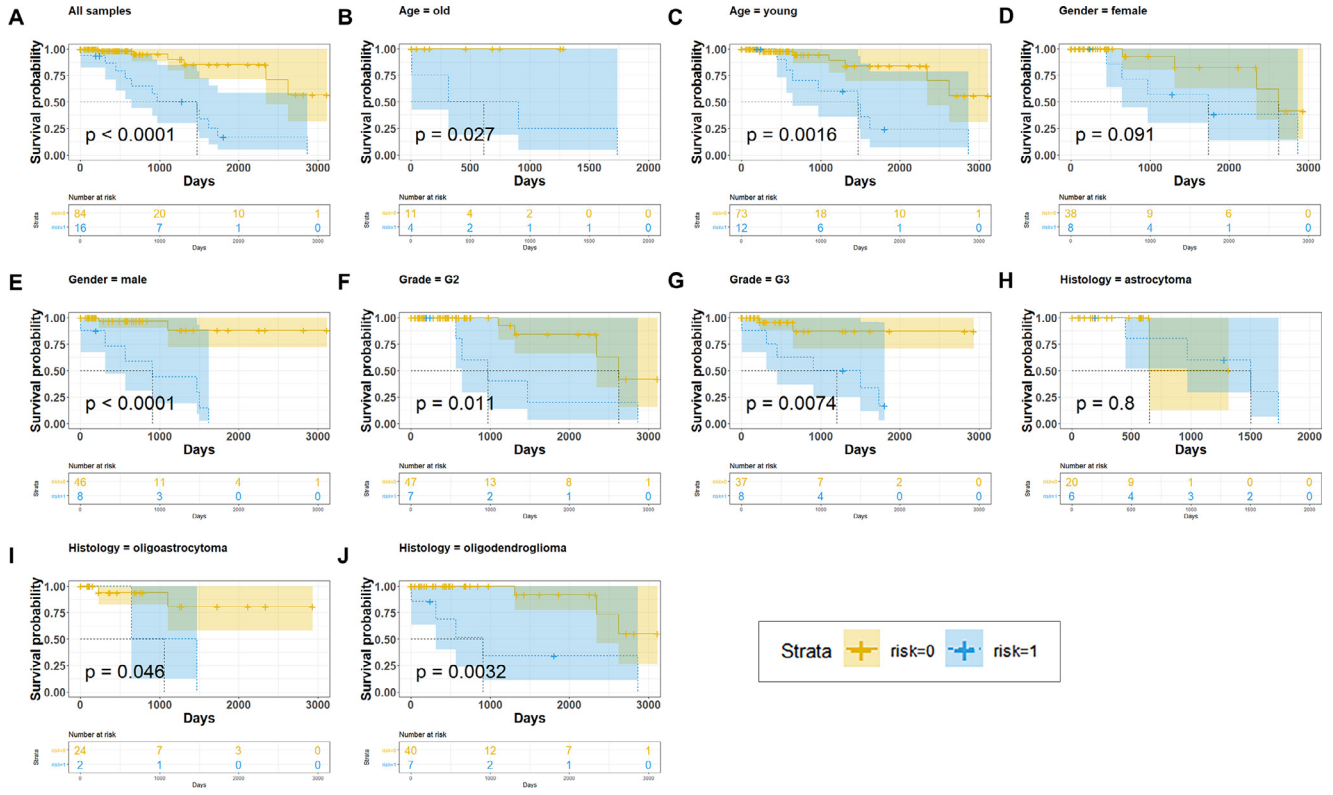

Fig. 7. Verification for significance of the prognostic model derived risk scores in test set from IDH-mutant glioma patients. (A) Kaplan–Meier survival curves for all samples that assigned into high- and low-risk scores groups. Kaplan–Meier survival curves for patients with (B) age ≥ 50 (old), (C) age < 50 (young), (D) female, (E) male, (F) Grade 2, (G) Grade 3, (H) astrocytoma, (I) oligoastrocytoma, and (J) oligodendroglioma that assigned into low- and high-risk scores groups. The yellow line represents high risk scores group, while the blue line represents low risk scores group.

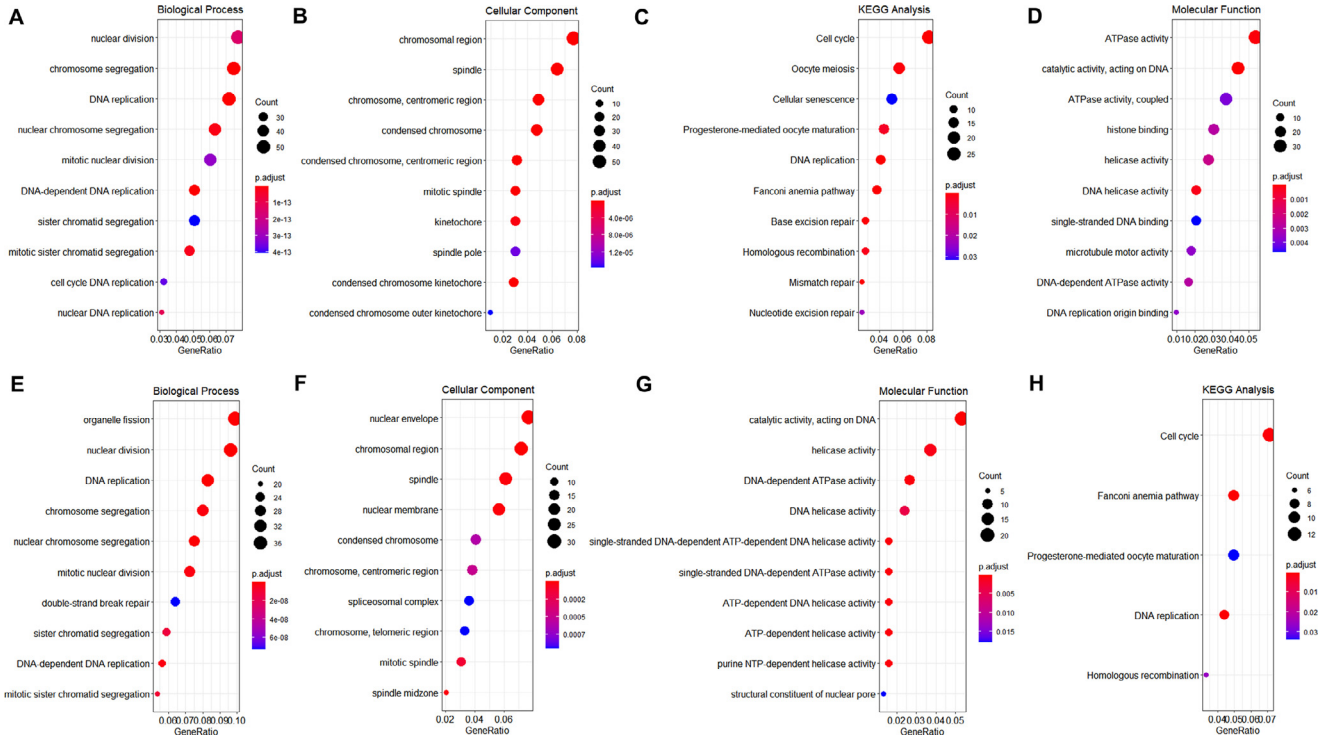

Fig. 8. Functional enrichment analysis for the prognostic model related genes. GO terms and KEGG pathway enrichment analyses for the (A-D) positively related genes and (E-H) negatively related genes. The size of the circle indicates the number of genes and the y-axis represents the terms of GO and KEGG pathway. The color of the circle showed the p.adjusted value of each term. The redder the color the higher the p.adjusted value.
Note: GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

[25]. SOX2 has been implicated in growth, tumorigenicity, drug resistance, and metastasis in the majority of cancers, which included cervical cancer, pancreatic cancer, breast cancer and glioma (C. [22,24]; K. [32,49,53]). Magali et al. indicated that TP53 mutations are indicators of bad prognosis in several cancers [17]. FOSL2 has been reported to induce TGF-β expression which is related to a diversity of biological and cellular processes, including cell proliferation, differentiation, apoptosis, and migration [48]. Reports on SMC3 involvement in cancers are scarce but our study conveyed its probable involvement in gliomagenesis.

As for the Kaplan-Meier survival analysis, the results demonstrated that the expression of IGFBP2, LINC00152 and S100A4 were negatively associated with OS and RFS in glioma patients, suggesting these three genes might be oncogenes in glioma. The expressions of other hub genes were pointedly related to the survival time of LGG but not GBM patients.

According to the above results, we supposed that the hub genes might be involved in the prognosis of glioma. Therefore, we further established a prognostic model based on the expression profile of IDH-mutant glioma patients. To identify the signature associated with IDH mutation for establishing prognostic model, we performed consensus clustering, followed by differentially expressed genes analysis and univariate Cox regression. Consensus clustering, an algorithm for classification of cancer subtype, provides quantitative and visual stability evidence based on repeated subsampling and clustering for calculating the number of unsupervised clusters in a dataset [66]. Two clusters were identified in this study, which were proven to be significantly related to patient survival. The comparison of patients from the two clusters also showed statistical difference in age, WHO grade, histology and prognosis, suggesting these two clusters might represent two novel IDH-mutant subtypes that reflected different features. Our study revealed that the overall survival rate of cluster 1 was remarkably lower than that of cluster 2 ($P = 0.0025$). As reported, patients with oligodendroglioma have a generally better prognosis than those with astrocytoma [14]. Consistent with published data, our results showed that the majority of astrocytoma (75%) are in cluster1 and the majority of oligodendroglioma (90%) are in cluster2. Notably, the distribution of the young patients (with younger age at diagnosis (<50)) in the two clusters was roughly the same; however, 82.7% of old patients were classified into cluster2.

Next, we screened out DEGs between these two clusters, and the result for functional annotation suggested that the up-regulated genes (overexpressed in patients from cluster1) affected patient survival via the regulation of the cell division related process. As known, abnormal cell division is one of the important features in tumors [11,18], and increasing studies emphasized the accumulation of cell divisions is one of the most important biological causes of cancer [60,67]. These genes overexpressed in patients from cluster1 were also mainly enriched in crucial pathways in cancer progression, such as the cell cycle, cellular senescence, and p53 signaling pathways. These results of functional annotation might explain why the prognosis of patients from cluster1 was poorer than those from cluster2. We also intersected DEGs with genes in the cyan module to seek the genes which were related to not only IDH status, but also IDH mutation subtypes. However, among 21 intersected genes, none was hub gene. We speculated that although the expressions of hub genes were closely related to whether glioma patients are accompanied by IDH mutations, they might not relate to the different phenotypes under IDH mutation. Based on this, the expression of hub genes may be mainly associated with wild type IDH status in glioma. Then, by using the univariate Cox regression analysis and LASSO regression method based on the DEGs, a signature of 31 genes was identified. Among such 31 genes, some of them have been incriminated in the progression of various cancers. The overexpression of HMGA1 gene is always found in cancer, and its high levels portend a poor prognosis in diverse tumors [9,56]. Pang et al. reported HMGA1 as a key gene controlling malignancy, proliferation, invasion, and angiogenesis of glioma [41]. Several studies revealed an association between increased CD97 expression and poor survival in glioma patients [27,50]. As we men-

tioned above, the cyan module was significantly related to IDH status. It is worth mentioning that CD97 is a gene in the cyan module. No clear relationship between the expression of CD97 and IDH mutation has been previously described but studies have suggested that both may relate to invasion in glioma ([21]; L. C. [23,51]), yet the mechanism remains obscure. ESR2 is a tumor suppressor gene and its role was reported in the context of GBM and other cancers (J. [31,54]). Additionally, Xu et al. revealed that CCNY regulates cell cycle in glioma [70]. Besides, most of the other signature genes, such as CYP3A5, DCBLD1, ADD3 and ADAMTS, have been proven to be cancer related genes. These published data verified the value of signature gene in glioma prognosis. Although there is no hub gene in the 31 genes that made up the signature, the predictive model composed of the signature was capable of predicting the prognosis of patients with IDH mutations in glioma. Based on the test set, we investigated the efficiency of the prognostic model and revealed that the high-risk group had low OS, implying the predictive model we built could effectively predict prognosis in IDH-mutant glioma patients. We also found that patients at high risk were older patients and WHO grade IV patients, which was consistent with the above findings as the old patients or patients with WHO grade IV commonly mean poor prognosis.

Furthermore, GO and KEGG analysis were performed based on the genes significantly correlated with 31-gene signature for exploring the biological function of this signature. The results showed the 31-gene signatures were closely related to basic cancer related biological processes, including cell cycle, nuclear division, and DNA replication.

There are some limitations in this study. First, our study was entirely focused on data mining and data analysis, and the results lacked verification by experimental data. Further experiments are required to confirm the findings of this study. Second, the data applied to analyze in our study was downloaded from TCGA which is a public database, so some significant patient information might lack in this study. The lack of such data may produce some deviations.

In summary, this study identified a total of ten hub genes which were associated with wild type IDH status in glioma, providing important information for future development and application of IDH target agents in glioma therapies. Moreover, the current study, for the first time, established a prognostic model with high accuracy for IDH-mutant glioma patients, which might have important clinical implications for accurate prognostication and management in IDH-mutant glioma patients with various epigenetic subtypes.

Funding

This work was supported by the Cuiying Scientific and Technological Program of Lanzhou University Second Hospital [Grant No. CY2018-QN16]. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Authors' contributions

YJ, WY, BT and QF performed the study. BT and QF obtained the data from TGCA database. YJ and WY analyzed data and wrote the manuscript. ZD contributed to the design of the study and project administration.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2020.100979.

References

[1] S.J. Beyer, E.H. Bell, J.P. McElroy, J.L. Fleming, T. Cui, A. Becker, A. Chakravarti, Oncogenic transgelin-2 is differentially regulated in isocitrate dehydrogenase wild-type vs. mutant gliomas, Oncotarget 9 (98) (2018) 37097–37111, doi:10.18632/oncotarget.26365.

[2] J. Cai, Q. Chen, Y. Cui, J. Dong, M. Chen, P. Wu, C. Jiang, Immune heterogeneity and clinicopathologic characterization of IGFBP2 in 2447 glioma samples, Oncoimmunology 7 (5) (2018) e1426516, doi:10.1080/2162402x.2018.1426516.

[3] S. Camelo-Piragua, S. Kesari, Further understanding of the pathology of glioma: implications for the clinic, Expert. Rev. Neurother. 16 (9) (2016) 1055–1065, doi:10.1080/14737175.2016.1194755.

[4] X. Chen, C. Pan, C. Xu, Y. Sun, Y. Geng, L. Kong, L. Zhang, Identification of survival-associated key genes and long non-coding RNAs in glioblastoma multiforme by weighted gene co-expression network analysis, Int. J. Mol. Med. 43 (4) (2019) 1709–1722, doi:10.3892/ijmm.2019.4101.

[5] Y. Chen, B. Jiang, W. Wang, D. Su, F. Xia, X. Li, Identifying the transcriptional regulatory network associated with extrathyroidal extension in papillary thyroid carcinoma by comprehensive bioinformatics analysis, Front. Genet. 11 (2020) 453-453, doi:10.3389/fgene.2020.00453.

[6] Y. Cheng, L. Li, Z. Qin, X. Li, F. Qi, Identification of castration-resistant prostate cancer-related hub genes using weighted gene co-expression network analysis, J. Cell. Mol. Med. (2020), doi:10.1111/jcmm.15432.

[7] H. Cho, H. Cho, J. Park, O. Kwon, H. Lee, T. Huh, B. Kang, NADP-dependent cytosolic isocitrate dehydrogenase provides NADPH in the presence of cadmium due to the moderate chelating effect of glutathione, J. Biol. Inorg. Chem.: JBIC: Publ. Soc. Biol. Inorg. Chem. 23 (6) (2018) 849–860, doi:10.1007/s00775-018-1581-5.

[8] A. Claes, A.J. Idema, P. Wesseling, Diffuse glioma growth: a guerilla war, Acta Neuropathol. 114 (5) (2007) 443–458, doi:10.1007/s00401-007-0293-7.

[9] M. Colamaio, N. Tosti, F. Puca, A. Mari, R. Gattordo, Y. Kuzay, A. Fusco, HMGA1 silencing reduces stemness and temozolomide resistance in glioblastoma stem cells, Expert Opin. Ther. Targets 20 (10) (2016) 1169–1179, doi:10.1080/14728222.2016.1220543.

[10] J.E. Eckel-Passow, D.H. Lachance, A.M. Molinaro, K.M. Walsh, P.A. Decker, H. Sicotte, R.B. Jenkins, Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors, N. Engl. J. Med. 372 (26) (2015) 2499–2508, doi:10.1056/NEJMoa1407279.

[11] U. Ehmer, J. Sage, Control of proliferation and cancer growth by the Hippo signaling pathway, Mol. Cancer Res. 14 (2) (2016) 127–140, doi:10.1158/1541-7786.Mcr-15-0305.

[12] A. Ernst, S. Hofmann, R. Ahmadi, N. Becker, A. Korshunov, F. Engel, B. Radlwimmer, Genomic and expression profiling of glioblastoma stem cell-like spheroid cultures identifies novel tumor-relevant genes associated with survival, Clin. Cancer Res. 15 (21) (2009) 6541–6550, doi:10.1158/1078-0432.ccr-09-0695.

[13] F. Fack, S. Tardito, G. Hochart, A. Oudin, L. Zheng, S. Fritah, S. Niclou, Altered metabolic landscape in IDH-mutant gliomas affects phospholipid, energy, and oxidative stress pathways, EMBO Mol. Med. 9 (12) (2017) 1681–1695, doi:10.15252/emmm.201707729.

[14] K.B. Fallon, C.A. Palmer, K.A. Roth, L.B. Nabors, W. Wang, M. Carpenter, A. Perry, Prognostic value of 1p, 19q, 9p, 10q, and EGFR-FISH analyses in recurrent oligodendrogliomas, J. Neuropathol. Exp. Neurol. 63 (4) (2004) 314–322, doi:10.1093/jnen/63.4.314.

[15] T. Feng, K. Li, P. Zheng, Y. Wang, Y. Lv, L. Shen, Y. Yao, Weighted gene coexpression network analysis identified microRNA coexpression modules and related pathways in type 2 diabetes mellitus, Oxidat. Med. Cell. Longev., 2019 (2019) 9567641, doi:10.1155/2019/9567641.

[16] Y. Gao, T. Zhu, C. Mao, Z. Liu, Z. Wang, X. Mao, Z. Liu, PPIC, EMP3 and CHI3L1 are novel prognostic markers for high grade glioma, Int. J. Mol. Sci. 17 (11) (2016), doi:10.3390/ijms17111808.

[17] C.L. Gladson, R.A. Prayson, W.M. Liu, The pathobiology of glioma tumors, Annu. Rev. Pathol. 5 (2010) 33–50, doi:10.1146/annurev-pathol-121808-102109.

[18] S. Gómez-López, R.G. Lerner, C. Petritsch, Asymmetric cell division of stem and progenitor cells during homeostasis and cancer, Cell. Mol. Life Sci. 71 (4) (2014) 575–597, doi:10.1007/s00018-013-1386-1.

[19] T.C. Hammond, X. Xing, C. Wang, D. Ma, K. Nho, P.K. Crane, A.L. Lin, β-amyloid and tau drive early Alzheimer's disease decline while glucose hypometabolism drives late decline, Commun. Biol. 3 (1) (2020) 352, doi:10.1038/s42003-020-1079-x.

[20] M.Z. Han, R. Xu, Y.Y. Xu, X. Zhang, S.L. Ni, B. Huang, J. Wang, TAGLN2 is a candidate prognostic biomarker promoting tumorigenesis in human gliomas, J. Exp. Clin. Cancer Res. 36 (1) (2017) 155, doi:10.1186/s13046-017-0619-9.

[21] H. Hu, Z. Wang, Y. Liu, C. Zhang, M. Li, W. Zhang, T. Jiang, Genome-wide transcriptional analyses of Chinese patients reveal cell migration is attenuated in IDH1-mutant glioblastomas, Cancer Lett. 357 (2) (2015) 566–574, doi:10.1016/j.canlet.2014.12.018.

[22] C. Huang, H. Lu, J. Li, X. Xie, L. Fan, D. Wang, T. Yao, SOX2 regulates radioresistance in cervical cancer via the hedgehog signaling pathway, Gynecol. Oncol. 151 (3) (2018) 533–541, doi:10.1016/j.ygyno.2018.10.005.

[23] L.C. Huang, D.Y. Hueng, CD97 and glioma invasion, J. Neurosurg. 120 (2) (2014) 579–580, doi:10.3171/2012.11.Jns12437.

[24] Y. Jia, D. Gu, J. Wan, B. Yu, X. Zhang, E.G. Chiorean, J. Xie, The role of GLI-SOX2 signaling axis for gemcitabine resistance in pancreatic cancer, Oncogene 38 (10) (2019) 1764–1777, doi:10.1038/s41388-018-0553-0.

[25] Y. Jiang, Y.Y. Jiang, J.J. Xie, A. Mayakonda, M. Hazawa, L. Chen, H.P. Koeffler, Co-activation of super-enhancer-driven CCAT1 by TP63 and SOX2 promotes squamous cancer progression, Nat. Commun. 9 (1) (2018) 3619, doi:10.1038/s41467-018-06081-9.

[26] J.P. Katz, N. Perreault, B.G. Goldstein, L. Actman, S.R. McNally, D.G. Silberg, K.H. Kaestner, Loss of Klf4 in mice causes altered proliferation and differentiation and precancerous changes in the adult stomach, Gastroenterology 128 (4) (2005) 935–945, doi:10.1053/j.gastro.2005.02.022.
[27] J.G. Kuhn, S.M. Chang, P.Y. Wen, T.F. Cloughesy, H. Greenberg, D. Schiff, M.D. Prados, Pharmacokinetic and tumor distribution characteristics of temsirolimus in patients with recurrent malignant glioma, Clin. Cancer Res. 13 (24) (2007) 7401–7406, doi:10.1158/1078-0432.Ccr-07-0781.
[28] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559, doi:10.1186/1471-2105-9-559.
[29] G. Li, Y. Jiang, X. Lyu, Y. Cai, M. Zhang, Z. Wang, Q. Qiao, Deconvolution and network analysis of IDH-mutant lower grade glioma predict recurrence and indicate therapeutic targets, Epigenomics 11 (11) (2019) 1323–1333, doi:10.2217/epi-2019-0137.
[30] G. Li, Z. Wang, C. Zhang, X. Liu, F. Yang, L. Sun, T. Jiang, MEGF10, a glioma survival-associated molecular signature, predicts IDH mutation status, Dis. Mark., 2018 (2018) 5975216, doi:10.1155/2018/5975216.
[31] J. Liu, G. Sareddy, M. Zhou, S. Viswanadhapalli, X. Li, Z. Lai, R. Vadlamudi, Differential effects of estrogen receptor β isoforms on glioblastoma progression, Cancer Res. 78 (12) (2018) 3176–3189, doi:10.1158/0008-5472.Can-17-3470.
[32] K. Liu, F. Xie, A. Gao, R. Zhang, L. Zhang, Z. Xiao, X. Lan, SOX2 regulates multiple malignant processes of breast cancer development through the SOX2/miR-181a-5p, miR-30e-5p/TUSC3 axis, Mol. Cancer 16 (1) (2017) 62, doi:10.1186/s12943-017-0632-9.
[33] Y.Q. Liu, F. Wu, J.J. Li, Y.F. Li, X. Liu, Z. Wang, R.C. Chai, Gene expression profiling stratifies IDH-wildtype glioblastoma with distinct prognoses, Front. Oncol. 9 (2019) 1433, doi:10.3389/fonc.2019.01433.
[34] D.N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W.K. Cavenee, D.W. Ellison, The 2016 world health organization classification of tumors of the central nervous system: a summary, Acta Neuropathol. 131 (6) (2016) 803–820, doi:10.1007/s00401-016-1545-1.
[35] R.J. McDougall, Computer knows best? The need for value-flexibility in medical AI, J. Med. Ethics 45 (3) (2019) 156–160, doi:10.1136/medethics-2018-105118.
[36] M. Mellai, A. Piazzi, V. Caldera, L. Annovazzi, O. Monzeglio, R. Senetta, D. Schiffer, Promoter hypermethylation of the EMP3 gene in a series of 229 human gliomas, BioMed. Res. Int., 2013 (2013) 756302, doi:10.1155/2013/756302.
[37] A. Ohandjo, Z. Liu, E. Dammer, C. Dill, T. Griffen, K. Carey, J. Lillard, Transcriptome network analysis identifies CXCL13-CXCR5 signaling modules in the prostate tumor immune microenvironment, Sci. Rep. 9 (1) (2019) 14963, doi:10.1038/s41598-019-46491-3.
[38] M.C. Oldham, P. Langfelder, S. Horvath, Network methods for describing sample relationships in genomic datasets: application to Huntington's disease, BMC Syst. Biol. 6 (2012) 63, doi:10.1186/1752-0509-6-63.
[39] M. Onishi, T. Ichikawa, K. Kurozumi, I. Date, Angiogenesis and invasion in glioma, Brain Tumor Pathol. 28 (1) (2011) 13–24, doi:10.1007/s10014-010-0007-z.
[40] Q.T. Ostrom, L. Bauchet, F.G. Davis, I. Deltour, J.L. Fisher, C.E. Langer, J.S. Barnholtz-Sloan, Response to "the epidemiology of glioma in adults: a 'state of the science' review", Neuro. Oncol. 17 (4) (2015) 624–626, doi:10.1093/neuonc/nov022.
[41] B. Pang, H. Fan, I.Y. Zhang, B. Liu, B. Feng, L. Meng, Q. Pang, HMGA1 expression in human gliomas and its correlation with tumor proliferation, invasion and angiogenesis, J. Neurooncol. 106 (3) (2012) 543–549, doi:10.1007/s11060-011-0710-6.
[42] P. Paschka, R.F. Schlenk, V.I. Gaidzik, M. Habdank, J. Krönke, L. Bullinger, . . ., K. Döhner, IDH1 and IDH2 mutations are frequent genetic alterations in acute myeloid leukemia and confer adverse prognosis in cytogenetically normal acute myeloid leukemia with NPM1 mutation without FLT3 internal tandem duplication, J. Clin. Oncol. 28 (22) (2010) 3636–3643, doi:10.1200/jco.2010.28.3762.
[43] S.S. Patil, R. Railkar, M. Swain, H.S. Atreya, R.R. Dighe, P. Kondaiah, Novel anti IGFBP2 single chain variable fragment inhibits glioma cell migration and invasion, J. Neurooncol. 123 (2) (2015) 225–235, doi:10.1007/s11060-015-1800-7.
[44] M. Peretti, F.M. Raciti, V. Carlini, I. Verduci, S. Sertic, S. Barozzi, . . ., M. Mazzanti, Mutual influence of ROS, pH, and CLIC1 membrane protein in the regulation of G(1)-S phase progression in human glioblastoma stem cells, Mol. Cancer Ther. 17 (11) (2018) 2451–2461, doi:10.1158/1535-7163.mct-17-1223.
[45] L.M. Phillips, X. Zhou, D.E. Cogdell, C.Y. Chua, A. Huisinga, R.H. K, W. Zhang, Glioma progression is mediated by an addiction to aberrant IGFBP2 expression and can be blocked using anti-IGFBP2 strategies, J. Pathol. 239 (3) (2016) 355–364, doi:10.1002/path.4734.
[46] A. Picca, A.L. Di Stefano, M. Sanson, Current and future tools for determination and monitoring of isocitrate dehydrogenase status in gliomas, Curr. Opin. Neurol. 31 (6) (2018) 727–732, doi:10.1097/wco.0000000000000617.
[47] S. Rizvi, G.J. Gores, Pathogenesis, diagnosis, and management of cholangiocarcinoma, Gastroenterology 145 (6) (2013) 1215–1229, doi:10.1053/j.gastro.2013.10.013.
[48] S. Roy, S. Khanna, A. Azad, R. Schnitt, G. He, C. Weigert, C.K. Sen, Fra-2 mediates oxygen-sensitive induction of transforming growth factor beta in cardiac fibroblasts, Cardiovasc. Res. 87 (4) (2010) 647–655, doi:10.1093/cvr/cvq123.
[49] C.M. Rudin, S. Durinck, E.W. Stawiski, J.T. Poirier, Z. Modrusan, D.S. Shames, S. Seshagiri, Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer, Nat. Genet. 44 (10) (2012) 1111–1116, doi:10.1038/ng.2405.
[50] M. Safaee, A.J. Clark, M.C. Oh, M.E. Ivan, O. Bloch, G. Kaur, A.T. Parsa, Overexpression of CD97 confers an invasive phenotype in glioblastoma cells and is associated with decreased survival of glioblastoma patients, PLoS ONE 8 (4) (2013) e62765, doi:10.1371/journal.pone.0062765.
[51] M. Sanson, Y. Marie, S. Paris, A. Idbaih, J. Laffaire, F. Ducray, J.Y. Delattre, Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas, J. Clin. Oncol. 27 (25) (2009) 4150–4154, doi:10.1200/jco.2009.21.9832.
[52] S. Sathornsumetee, D.A. Reardon, A. Desjardins, J.A. Quinn, J.J. Vredenburgh, J.N. Rich, Molecularly targeted therapy for malignant glioma, Cancer 110 (1) (2007) 13–24, doi:10.1002/cncr.22741.
[53] M. Schmitz, A. Temme, V. Senner, R. Ebner, S. Schwind, S. Stevanovic, B. Weigle, Identification of SOX2 as a novel glioma-associated antigen and potential target for T cell-based immunotherapy, Br. J. Cancer 96 (8) (2007) 1293–1301, doi:10.1038/sj.bjc.6603696.
[54] A. Shergalis, A. Bankhead, U. Luesakul, N. Muangsin, N. Neamati, Current challenges and opportunities in treating glioblastoma, Pharmacol. Rev. 70 (3) (2018) 412–445, doi:10.1124/pr.117.014944.
[55] R. Stupp, M.E. Hegi, M.R. Gilbert, A. Chakravarti, Chemoradiotherapy in malignant glioma: standard of care and future directions, J. Clin. Oncol. 25 (26) (2007) 4127–4136, doi:10.1200/jco.2007.11.8554.
[56] T. Sumter, L. Xian, T. Huso, M. Koo, Y. Chang, T. Almasri, L. Resar, The high mobility group A1 (HMGA1) transcriptome in cancer and development, Curr. Mol. Med. 16 (4) (2016) 353–393, doi:10.2174/1566524016666160316152147.
[57] X. Tang, X. Fu, Y. Liu, D. Yu, S. Cai, C. Yang, IDH1Blockade of glutathione metabolism in -mutated glioma, Mol. Cancer Ther. 19 (1) (2020) 221–230, doi:10.1158/1535-7163.Mct-19-0103.
[58] R. Tian, Z. Cui, D. He, X. Tian, Q. Gao, X. Ma, Z. Hu, Risk stratification of cervical lesions using capture sequencing and machine learning method based on HPV and human integrated genomic profiles, Carcinogenesis 40 (10) (2019) 1220–1228, doi:10.1093/carcin/bgz094.
[59] N. Tiwari, N. Meyer-Schaller, P. Arnold, H. Antoniadis, M. Pachkov, E. van Nimwegen, G. Christofori, Klf4 is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8), PLoS ONE 8 (2) (2013) e57329, doi:10.1371/journal.pone.0057329.
[60] C. Tomasetti, B. Vogelstein, Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions, Science 347 (6217) (2015) 78–81, doi:10.1126/science.1260825.
[61] M. Uno, S.M. Oba-Shinjo, R. Silva, F. Miura, C.A. Clara, J.R. Almeida, S.K. Marie, IDH1 mutations in a Brazilian series of Glioblastoma, Clinics (Sao Paulo) 66 (1) (2011) 163–165, doi:10.1590/s1807-59322011000100028.
[62] M. Wang, L. Wang, L. Pu, K. Li, T. Feng, P. Zheng, L. Jin, LncRNAs related key pathways and genes in ischemic stroke by weighted gene co-expression network analysis (WGCNA), Genomics 112 (3) (2020) 2302–2308, doi:10.1016/j.ygeno.2020.01.001.
[63] D. Wei, W. Gong, M. Kanai, C. Schlunk, L. Wang, J.C. Yao, K. Xie, Drastic down-regulation of Krüppel-like factor 4 expression is critical in human gastric cancer development and progression, Cancer Res. 65 (7) (2005) 2746–2754, doi:10.1158/0008-5472.can-04-3619.
[64] M. Weller, W. Wick, K. Aldape, M. Brada, M. Berger, S.M. Pfister, G. Reifenberger, Glioma, Nat. Rev. Dis. Primers 1 (2015) 15017, doi:10.1038/nrdp.2015.17.
[65] M.M.J. Wijnenga, P.J. French, H.J. Dubbink, W.N.M. Dinjens, P.N. Atmodimedjo, J.M. Kros, M.J. van den Bent, Prognostic relevance of mutations and copy number alterations assessed with targeted next generation sequencing in IDH mutant grade II glioma, J. Neurooncol. 139 (2) (2018) 349–357, doi:10.1007/s11060-018-2867-8.
[66] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (12) (2010) 1572–1573, doi:10.1093/bioinformatics/btq170.
[67] S. Wu, S. Powers, W. Zhu, Y.A. Hannun, Substantial contribution of extrinsic risk factors to cancer development, Nature 529 (7584) (2016) 43–47, doi:10.1038/nature16166.
[68] R. Xu, M. Han, Y. Xu, X. Zhang, C. Zhang, D. Zhang, J. Wang, Coiled-coil domain containing 109B is a HIF1α-regulated gene critical for progression of human gliomas, J. Transl. Med. 15 (1) (2017) 165, doi:10.1186/s12967-017-1266-9.
[69] T. Xu, T.D. Le, L. Liu, N. Su, R. Wang, B. Sun, J. Li, CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization, Bioinformatics 33 (19) (2017) 3131–3133, doi:10.1093/bioinformatics/btx378.
[70] Y. Xu, Z. Wang, J. Wang, J. Li, H. Wang, W. Yue, Lentivirus-mediated knockdown of cyclin Y (CCNY) inhibits glioma cell proliferation, Oncol. Res. 18 (8) (2010) 359–364, doi:10.3727/096504010x12644422320582.
[71] W. Yan, W. Zhang, G. You, Z. Bao, Y. Wang, Y. Liu, T. Jiang, Correlation of IDH1 mutation with clinicopathologic factors and prognosis in primary glioblastoma: a report of 118 patients from China, PLoS ONE 7 (1) (2012) e30339, doi:10.1371/journal.pone.0030339.
[72] J. Yao, A. Chakhoyan, D.A. Nathanson, W.H. Yong, N. Salamon, C. Raymond, B.M. Ellingson, Metabolic characterization of human IDH mutant and wild type gliomas using simultaneous pH- and oxygen-sensitive molecular MRI, Neuro. Oncol. (2019), doi:10.1093/neuonc/noz078.
[73] D. Yu, Y. Liu, Y. Zhou, V. Ruiz-Rodado, M. Larion, G. Xu, C. Yang, Triptolide suppresses IDH1-mutated malignancy via Nrf2-driven glutathione metabolism, Proc. Natl. Acad. Sci. U.S.A. 117 (18) (2020) 9964–9972, doi:10.1073/pnas.1913633117.
[74] Q. Yuan, H.Q. Cai, Y. Zhong, M.J. Zhang, Z.J. Cheng, J.J. Hao, J.H. Wan, Overexpression of IGFBP2 mRNA predicts poor survival in patients with glioblastoma, Biosci. Rep. 39 (6) (2019), doi:10.1042/bsr20190045.
[75] Z.M. Zhang, J.X. Tan, F. Wang, F.Y. Dao, Z.Y. Zhang, H. Lin, Early diagnosis of hepatocellular carcinoma using machine learning method, Front. Bioeng. Biotechnol. 8 (2020) 254, doi:10.3389/fbioe.2020.00254.
[76] W. Zhao, I.M. Hisamuddin, M.O. Nandan, B.A. Babbin, N.E. Lamb, V.W. Yang, Identification of Krüppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer, Oncogene 23 (2) (2004) 395–402, doi:10.1038/sj.onc.1207067.