ORIGINAL RESEARCH

Comprehensive Bioinformatics Analysis of mRNA Expression Profiles and Identification of a miRNA–mRNA Network Associated with the Pathogenesis of Low-Grade Gliomas

Ming Wang
Yan Cui
Yang Cai
Yugang Jiang
Yong Peng [ORCID]

Department of Neurosurgery, The Second Xiangya Hospital of Central South University, Changsha, Hunan, People's Republic of China

Purpose: Low-grade glioma is the most common type of primary intracranial tumour, and the overall survival of patients with low-grade glioma (LGG) has shown no significant improvement over the past few decades. Therefore, it is crucial to understand the precise molecular mechanisms involved in the carcinogenesis of LGG.

Methods: To investigate the regulatory mechanisms of mRNA–miRNA networks related to LGG, in the present study, a comprehensive analysis of the genomic landscape between low-grade gliomas and normal brain tissues from the GEO and TCGA datasets was first conducted to identify differentially expressed genes (DEGs) and differentially expressed miRNAs in LGG. Following a series of analyses, including WGCNA, GO and KEGG analyses, PPI and key model analyses, and survival analysis of the DEGs with clinical phenotypes, the potential key genes were screened and identified, and the related miRNA–mRNA networks were subsequently constructed through miRWalk 3.0. Finally, the potential miRNA–mRNA networks were further validated in CGGA (Chinese Glioma Genome Atlas) datasets and clinical specimens by qRT-PCR.

Results: In our results, six hub genes, MELK, NCAPG, KIF4A, NUSAP1, CEP55, and TOP2A, were ultimately identified. Two regulatory pathways, miR-495-3p-TOP2A and miR-1224-3p-MELK, that regulate the pathogenesis of LGG were ultimately identified. Furthermore, the expression of miR-495-3p-TOP2A and miR-1224-3p-MELK in solid tissues was validated by qRT-PCR.

Conclusion: Our study identified hub genes and related miRNA–mRNA regulatory pathways that contribute to the carcinogenesis of LGG, which may help us reveal the mechanisms underlying the development of LGG.

Keywords: low-grade glioma, miRNA–mRNA network, bioinformatics analysis, weighted gene coexpression

Introduction

Gliomas are the most common and fatal malignancy in the central nervous system (CNS).[1] According to the 2007 World Health Organization (WHO) classification of tumours of the central nervous system, gliomas are classified into low-grade (WHO grades I-II) and high-grade (WHO grades III-IV) gliomas.[2] High-grade gliomas (HGGs) have a high mortality rate, and the median survival time for HGGs is 14 to 17 months.[3] Low-grade gliomas (LGGs) account for 17% to 22% of all primary

Correspondence: Yong Peng
Department of Neurosurgery, The Second Xiangya Hospital of Central South University, Changsha, Hunan, People's Republic of China
Tel +86-18874909698
Fax +86-731 85295110
Email 188201058@csu.edu.cn

brain tumours and tend to be associated with longer survival than HGG.[4] Comprehensive therapy for LGG includes neurosurgical resection, chemotherapy and radiotherapy. Nevertheless, LGG cannot be cured entirely by conventional treatment due to therapeutic resistance, tumour recurrence and progression to HGG. Therefore, it is crucial to understand the precise molecular mechanisms involved in the carcinogenesis of LGG and explore novel therapeutic biomarkers for LGG patients.

In recent years, with the progression of high-throughput sequencing technologies, most researchers have focused on the impact of genetic alterations on tumourigenesis and progression.[5] Like other cancers, LGG is a highly heterogeneous disease with highly distinct molecular features and gene expression signatures and can be classified into different subtypes. Different subtypes of LGG have different molecular characteristics, biological behaviours, and therapeutic responses. The 2016 World Health Organization classification of CNS tumours updated the definition of gliomas,[6] and the new classification defined CNS tumours by adding molecular subtypes to conventional histological features. In these molecular types, isocitrate dehydrogenase (IDH1) mutation and 1p/19q codeletion are essential biomarkers that characterize the classification and prognosis of gliomas and other molecular biomarkers, including TERT promoter mutation, isocitrate dehydrogenase mutation, MGMT gene promoter methylation, and 1p19q codeletion. By associating clinical data with molecular mechanisms at the genome-wide level, the new 2016 WHO classification of CNS tumours provides more precise diagnosis, treatment guidance and prognostic evaluation for gliomas.

miRNAs are small noncoding RNAs that are approximately 19 to 25 nucleotides in length. miRNAs regulate target genes by binding to target mRNAs to repress their translation. An increasing number of studies have shown that miRNAs play an important role in the occurrence and development of cancer, and miRNAs are recognized as potential biomarkers for the diagnosis and prognosis of human cancer. For example, miRNA-21 is upregulated in gliomas and promotes invasion by targeting specific inhibitors of matrix metalloproteases.[7] miRNAs–mRNAs constitute networks to participate in many cellular pathways, such as proliferation, apoptosis and differentiation.[8,9] A significant number of studies have searched for deregulated miRNA–mRNA networks in cancers. However, research on the miRNA–mRNA regulatory network mechanisms in the pathogenesis of low-grade gliomas

has been limited, hindering the comprehension and treatment of this disease.

To investigate the regulatory mechanisms of mRNA–miRNA networks in LGG, a comprehensive analysis of the genomic landscape between LGG and normal brain tissues was first conducted to identify differentially expressed genes (DEGs) in LGG in the present study. Following a series of analyses of those DEGs with clinical phenotypes, the potential key genes were screened and identified. Furthermore, an miRNA–mRNA network was constructed, and important miRNAs were identified. Finally, the potential hub genes and miRNAs were validated in solid specimens.

Materials and Methods
Data Collection and Differentially Expressed Gene and miRNA Identification

The datasets used in the present study were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih. gov/geo/) and The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). Two microarray datasets (GSE68848 and GSE4290) from the GEO datasets were used to screen DEGs in LGG tissues and normal brain tissues. The platform for the GSE68848 and GSE4290 datasets was GPL570 (Affymetrix Human Genome U113 Plus 2.0 Array). GSE68848 contained 215 primary LGG samples and 28 normal samples, and GSE4290 contained 76 primary LGG samples and 23 normal samples. Meanwhile, the LGG-related transcriptome profiles containing 512 primary LGG samples and 5 normal samples in the TCGA database were downloaded for differential expressed gene and miRNA analysis. The clinical data were also downloaded from TCGA to analyse the relationship between DEGs and clinical characteristics. Genes with |logFC| (fold change) >2 and p < 0.05 were identified as DEGs between LGG samples and normal samples, and miRNAs with |logFC| (fold change) >1 and p< 0.05 were identified as differentially expressed miRNAs. The DEGs were analysed by the limma package of R software for GEO datasets and the "edgeR" package for TCGA profiles. The R package "ggplot2," was used to generate volcano plots for visualization of DEGs and differentially expressed miRNAs. The intersection of DEGs from the GEO and TCGA datasets was identified as the common DEGs.

Gene Ontology and Pathway Enrichment Analysis Analyses of DEGs

The Gene Ontology (GO) database (http://www.geneontology.org/) is widely used in functional annotation and enrichment analysis and describes the three major components of gene function: biological process (BP), cellular component (CC) and molecular function (MF).[10] The Kyoto Encyclopedia of Genes and Genomes pathway (KEGG) (http://www.kegg.jp), a comprehensive knowledge database, plays an important role in both functional interpretation and practical application of genomic information.[11] Through the Database for Annotation, Visualization and Integrated Discovery (DAVID) tools[12] (https://david.ncifcrf.gov/), we carried out GO terms and significant pathway enrichment analysis of DEGs. P<0.01 and gene counts ≥10 were considered statistically significant terms.

Protein-protein Interaction Network Construction and Key Module Analysis

To further understand the protein-protein interactions (PPIs) between DEGs, a PPI network analysis was performed on the STRING (http://string-db.org) website.[13] An interaction with a combined score >0.4 was deemed significant, and Cytoscape was used to visualize the PPI network for these significant interactions. Furthermore, the MCODE plug-in for Cytoscape was used to identify the key clustering modules in PPI networks. The criteria for the most significant module selection were as follows: MCODE scores >10, degree cutoff=2, node score cutoff=0.2, max depth=100 and k-score=2. Moreover, functional and pathway enrichment analyses were performed for DEGs in the key module.

Weighted Gene Correlation Network Analysis of DEGs

WGCNA is a weighted gene coexpression network analysis that aims to find coexpressed gene modules and explore the association between gene networks and phenotypes of concern, as well as the core genes in the network. The R package WGCNA based on the DEGs was applied for this analysis.[14] In brief, a coexpression network can be constructed based on the expression profile and the estimated best soft thresholding power. Average linkage hierarchical clustering was then performed to identify modules of densely interconnected genes. Network modules were identified using a dynamic tree cut algorithm with

a minimum cluster size of 30. Finally, the correlation between the modules and the clinical parameters was calculated, and the gene set module most related to certain clinical parameters of interest was identified.

Hub Genes Related to LGG and Clinical Validation

To narrow the scope of key genes, Venn analysis was used to integrate the genes obtained by WGCNA with the key module genes and differentially expressed miRNA target genes. A heat map was used to visualize the expression differences of key genes between LGG and normal brain tissues. Subsequently, the mRNA expression levels of key genes were validated in the GEPIA[15] (http://gepia.cancer-pku.cn/index.html) web server, and survival analyses were performed using clinical data downloaded from the TCGA database.

Construction of the miRNA–mRNA Network and Validation

The interactions between the differentially expressed miRNAs and the hub genes related to LGG were predicted using miRWalk 3.0 (http://mirwalk.umm. uni-heidelberg. de/), which integrated the prediction results of both TargetScan and miRDB,[16] and a score ≥ 0.95 was considered the cutoff criterion for the prediction analysis in miRWalk. Eventually, Cytoscape software was employed to establish the miRNA–mRNA network. Only the interactions of miRNAs and mRNAs with opposite expression were included for subsequent analysis. Pearson's correlation analysis was performed in CGGA[17] (Chinese Glioma Genome Atlas) datasets (http://www.cgga.org.cn/) to validate miRNA–mRNA regulation.

Clinical Samples and RT-qPCR Analysis

In total, ten LGG tumour tissues from glioma patients and five normal brain tissues from traumatic haematoma patients were obtained during neurosurgical procedures at the Second Xiangya Hospital of Central South University. The samples taken during the surgery were immediately frozen at −80 °C and then used for RNA isolation. The collection of human tissues was approved by the Ethical Committee of the Second Xiangya Hospital of Central South University (2019080), and written informed consent was obtained from all patients and their families. This study was conducted in accordance with the Declaration of Helsinki. Total RNA was extracted from LGG tumour

and normal brain tissues using an RNA extraction kit (Bioteke Corporation, Beijing, China). For miRNA expression detection, miRNA was reverse transcribed with a miScript II RT Kit (Qiagen). The relative expression levels of miRNA were determined with the miScript SYBR Green PCR Kit (Qiagen) using a QuantStudio TM6 Flex real-time PCR system (Life Technologies). For mRNA expression detection, the GoScript Reverse Transcription System (Promega Corporation, Madison, USA) was used to reverse transcribe RNA templates, and the relative expression levels of mRNAs were determined with GoTaq qPCR Master mix (Promega Corporation) using a QuantStudio TM6 Flex real-time PCR system. The relative expression levels (fold change) were calculated using the 2-$\Delta\Delta$Cq method, with GAPDH and U6 used for normalization. The primers used were GAPDH F: 5′-GGTGGTCTCCTCTGACTTCAACA-3′ and R: 5′-GTTGCTGTAGCCAAATTCGTTGT-3′, TOP2A F: 5′-CATTGAAGACGCTTCGTTATGG-3′ and R: 5′-CAGAAGAGAGGGCCAGTT GTG-3′, MELK F: 5′-GCTGCAAGGTATAATTGATGGA-3′ and R: 5′-CAGTAACATAATGACAGATGGGC-3′.

Statistical Analyses

SPSS 20.0 software (SPSS IBM, Armonk, NY, USA) and GraphPad 7.0 (La Jolla, CA, USA) were used for data analyses. The data are presented as the mean ± standard deviation (SD) for the relative expression levels of mRNA and miRNA. The survival analysis between two groups used the long rank test. A two-tailed Student's *t*-test was used to assess differences between the tumour group and the normal group. Pearson correlation tests were conducted to assess the correlation between the expression of miRNAs and targeted mRNAs. A p-value < 0.05 was considered statistically significant.

Results

Identification of DEGs and miRNAs Based on TCGA and GEO

After a series of analyses according to the calculation criteria mentioned above, a total of 487 DEGs were identified in the GSE4290 dataset (122 upregulated and 365 downregulated, Figure 1A), and 811 DEGs were identified in the GSE68848 dataset (152 upregulated and 649 downregulated, Figure 1B). Furthermore, a total of 2097 DEGs (1590 upregulated and 507 downregulated, Figure 1C) were also identified from TCGA datasets. Only DEGs

that commonly appeared in those three sets were chosen as candidate DEGs. As shown in Figures1D, 338 DEGs (200 upregulated and 138 downregulated) were finally identified. For differentially expressed miRNAs, 246 DEmiRNAs (121 upregulated and 125 downregulated) were identified from TCGA.

Functional Annotation and Pathway Enrichment Analysis

GO and KEGG enrichment analyses were performed to explore the potential functions of the DEGs, which included three categories: cellular component (CC), biological process (BP) and molecular function (MF). In the BP group, the GO terms were mainly associated with chemical synaptic transmission, negative regulation of neuron apoptotic process and positive regulation of cell proliferation (e 1E). In the CC group, the GO terms included mainly cell junctions, axons and dendrites (Figure 1F). In the MF group, the GO terms involved mainly calmodulin binding and calcium ion binding (Figure 1G). In addition, KEGG pathway analyses indicated that the most significant pathways were the calcium signalling pathway, pathways in cancer and neuroactive ligand-receptor interactions (Figure 1H). The top 10 GO terms and enriched pathways of targets of DEGs are shown in the bubble chart.

Protein–Protein Interaction Network Construction and Key Model Analysis

To explore the relationship between these CDEGs, a PPI network of DEGs was constructed using the STRING online database and visualized by Cytoscape. As shown in Figure 2A, there were 234 nodes and 1683 edges in the PPI network. Three key clusters were identified using the MCODE plug-in according to the cut-off value described above. As shown in Figure 2A and B, model 1 was the most significant module (MCODE score=13.8), which was constructed with 14 nodes and 90 edges (Figure 2A). Further pathway enrichment analysis showed that the 14 genes in the most significant module were associated mainly with inflammatory regulation of the TRP channel, Ras signalling pathway, MAPK signalling pathway, and pathways in cancer (Figure 2C).

Weighted Gene Correlation Network Analysis

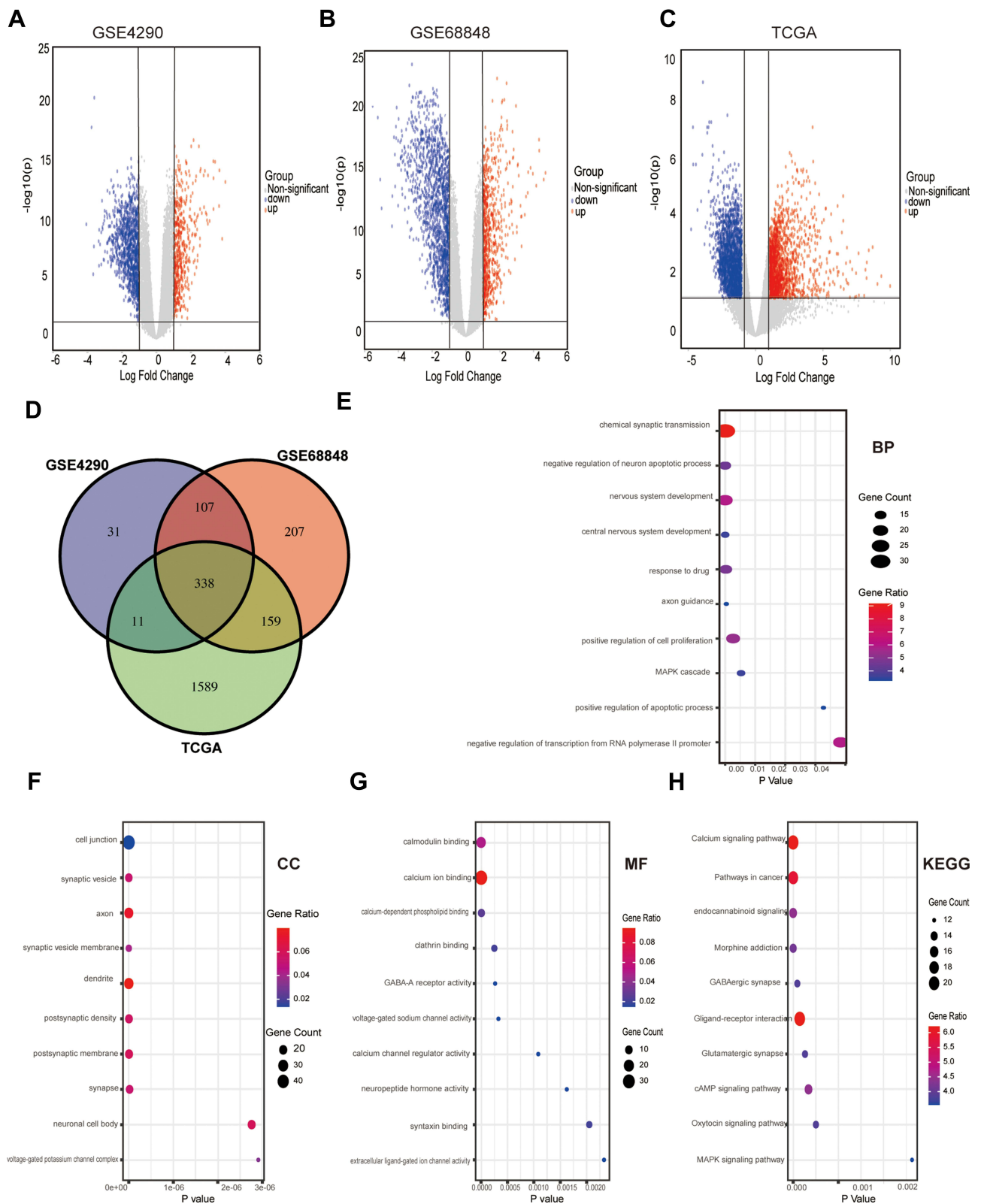We analysed the coexpression of DEGs from the TCGA dataset by the WGCNA package. A coexpression network

Figure 1 Differential expression of data between three sets of samples, Venn diagram, GO and KEGG analysis of DEGs. (**A-C**) Volcano plot of the genes between normal brain and low-grade glioma samples from the GSE4290, GSE68848 and TCGA datasets. Red dots indicate upregulated genes, and blue dots indicate downregulated genes. Grey dots show the genes with expression of |log2FC|<2. The Y axis represents an FDR, and the X axis represents the value of log2FC. (**D**) Venn diagram of overlapping DEGs obtained by the GSE4290, GSE68848 and TCGA datasets. (**E-H**) GO and KEGG analysis of the overlapping DEGs in LGG: (**E**) Biological processes (**F**) Cellular components (**G**) Molecular function (**H**) KEGG pathway.
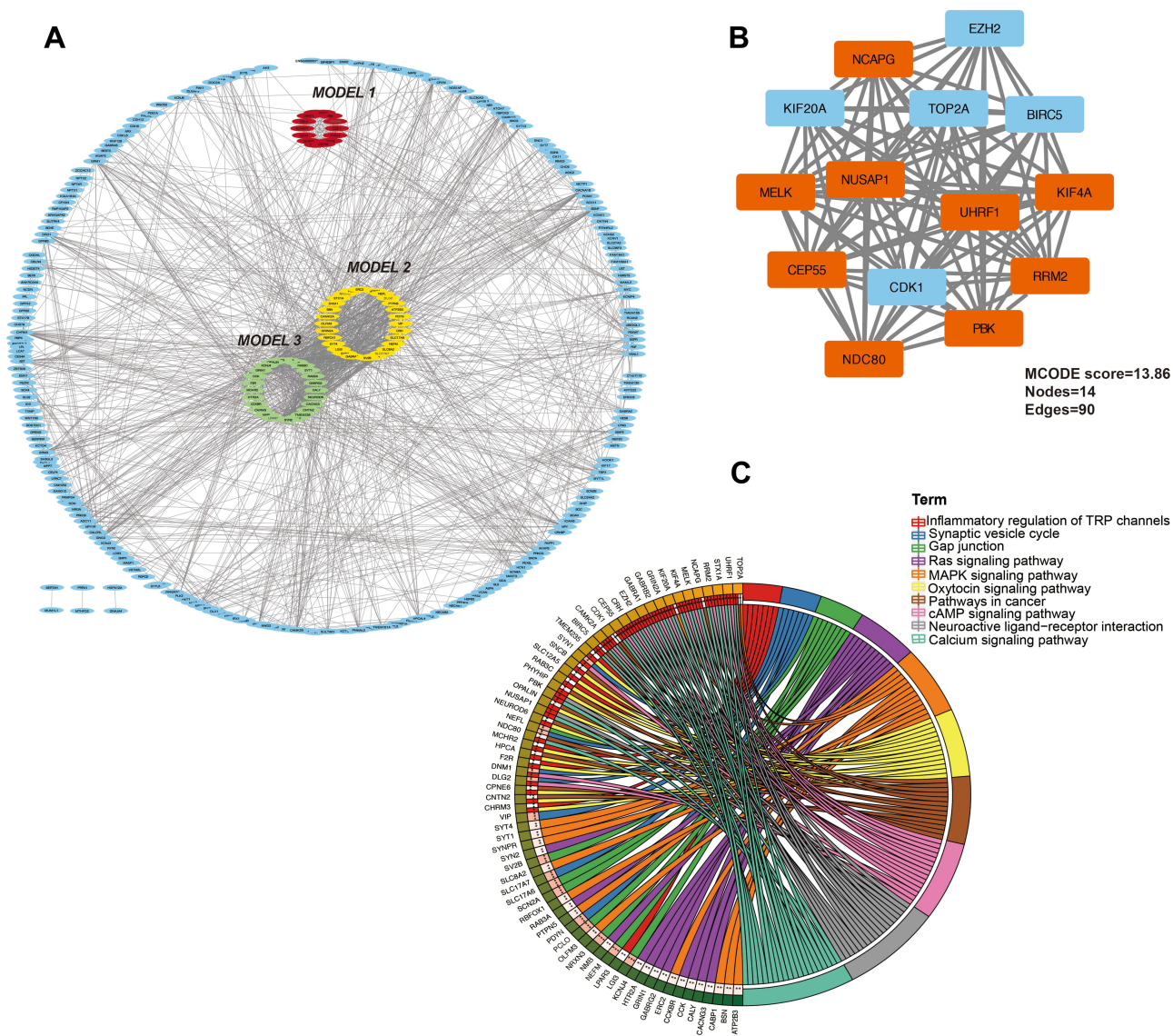
Figure 2 PPI network and the most significant model of DEGs. (**A**) Protein-protein interaction network construction and module analysis. The red, yellow, and green parts represent the top 3 modules of the PPI network. (**B**) The most significant mode from the PPI network constructed by the DEGs in LGG. Red dots indicate downregulated genes, and blue dots indicate upregulated genes. (**C**) KEGG pathway analysis of the most significant module in PPI analysis.

was constructed when the soft thresholding was set to five. Through the merger of similar modules, all DEGs were finally clustered into 15 modules (Figure 3A, B, D). Different gene coexpression modules and the patient's age, sex, survival status, survival time, and recurrence were further selected to draw a clustering heat map (Figure 3C). As shown in Figure 3C, a heatmap with correlation coefficient (R) and significant difference (P-value) showed the correlations between the modules and phenotypes. Finally, we screened out the yellow coexpression module (R=0.43, p<0.01) in which the DEGs in this module were significantly correlated with LGG patient survival status. As shown in Figure 3E, the correlation between genes in the yellow

module and the survival status was greater than 0.6, and the significance p value was less than 0.05, which indicated that this module can be determined to be an important module. There were 109 DEGs in the yellow coexpression module, and KEGG analysis (Figure 4A) and GO (Figure 4B) analysis results show that the yellow module genes were associated mainly with the cell cycle, which is important for the proliferation of tumour cells.

Hub Gene Screen and Clinical Significance Analysis

To further narrow the range of key genes, the common genes in the WGCNA, key module genes, and miRNA
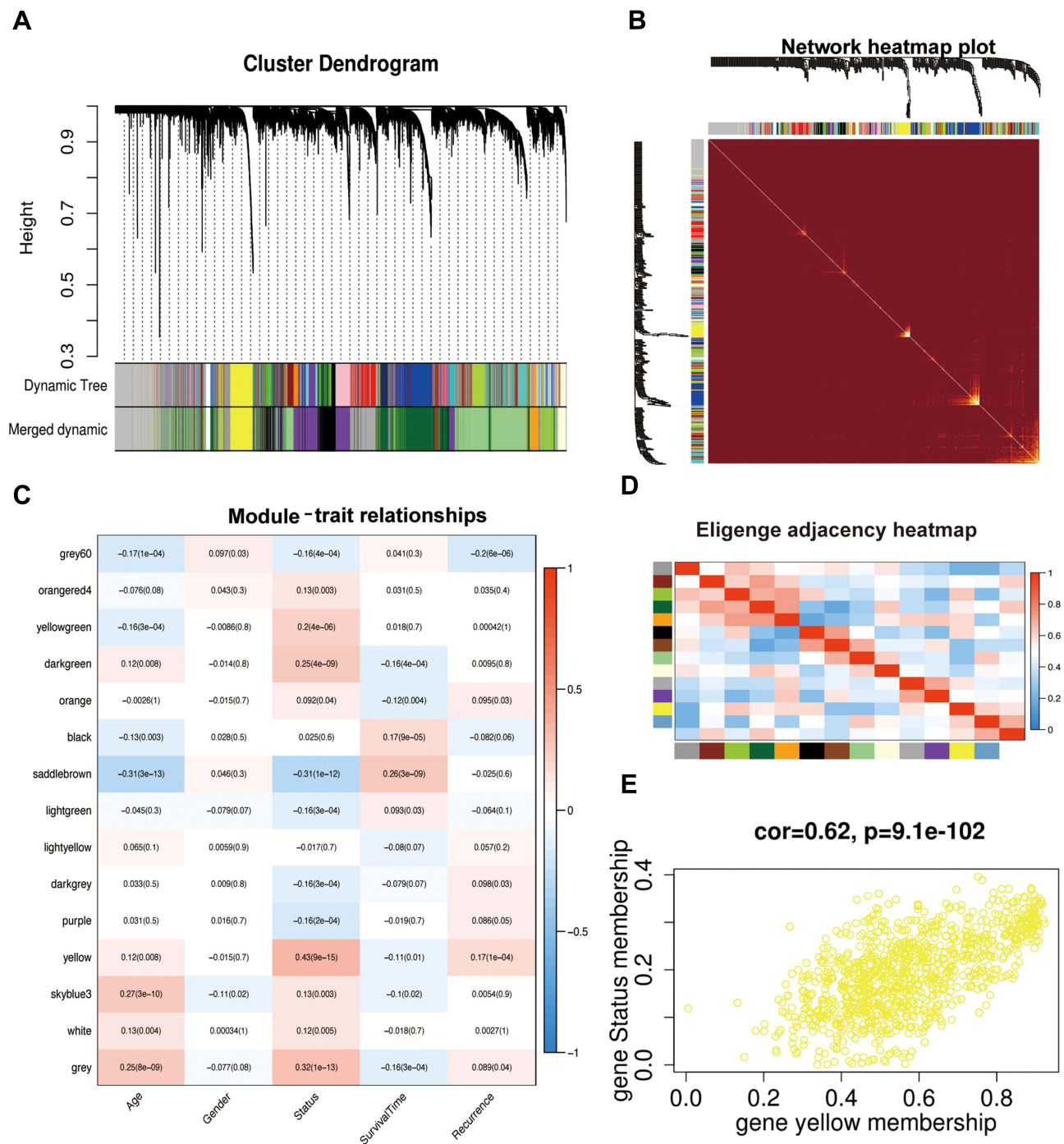
Figure 3 WGCNA of DEGs. (**A**) WGCNA dendrogram indicating the expression of different gene modules in all LGG samples. (**B**) Network heatmap plot to visualize the genetic correlation within the modules. (**C**) Module-sample feature correlation analysis. Sample features included age, sex, status, survival time and recurrence. (**D**) Correlation analysis between 15 coexpression modules. (**E**) Gene correlation scatter plots of the yellow model.

target genes were obtained from Venn analysis. Finally, six hub genes (MELK, NCAPG, KIF4A, NUSAP1, CEP55, TOP2A) were identified (Figure 5A), and the heat maps of these hub genes suggest that they can obviously be distinguished between the LGG group and the normal group (Figure 5B). Subsequently, we verified the mRNA

expression levels of these six hub genes in GEPIA, and the results revealed that only the expression levels of MELK, NUSAP1 and TOP2A were significantly increased in the LGG group compared to the normal group, with a p value <0.05 (Figure 5C). To test whether the hub genes had clinical significance, we verified gene expression and
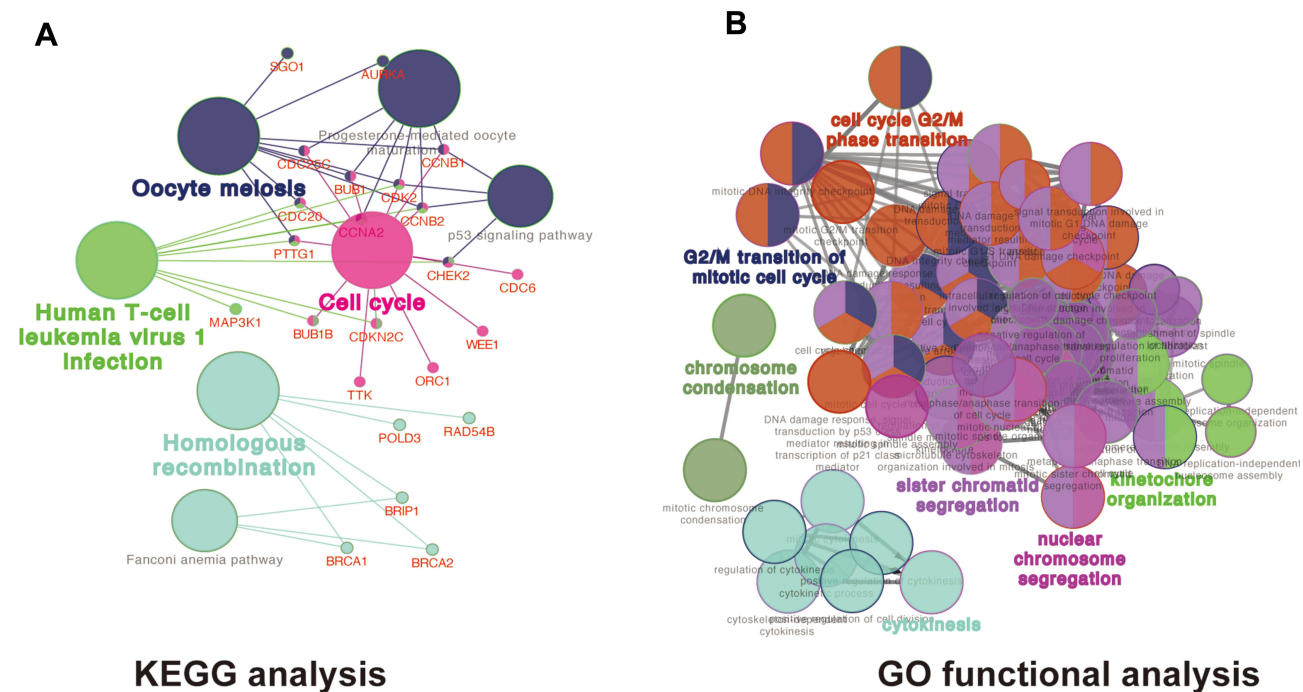
Figure 4 GO and KEGG analysis of the most significant coexpression module. (**A**)The KEGG functional annotation analysis of the yellow model genes was performed by ClueGO and CluePedia in Cytoscape software. Different colours of nodes refer to the functional annotation of ontologies, each line reflects the correlation between the terms, the network shows the relationship between the yellow model genes and KEGG terms. (**B**) The biological process functional annotation analysis of hub genes was performed by ClueGO and CluePediain Cytoscape software. Different colours of nodes refer to the functional annotation of ontologies, each line reflects the correlation between the terms, the network shows the relationship between the yellow model genes and Go terms. A corrected p value <0.01 was considered statistically significant.

its relationship with patient overall survival (OS) by using clinical data downloaded from TCGA. As shown in Figure 5D, high expression of these six genes was associated with worse overall survival (OS) in LGG patients.

Identification and Validation of Potential miRNA–mRNA Regulatory Pathways in LGG

To further explore the molecular mechanisms in LGG, target miRNAs of hub genes were predicted using miRWalk3.0. As shown in Figure 6, we identified six potential miRNA–mRNA networks associated with the pathogenesis of LGG. However, as verified above, there was no significant difference in the expression levels of CEP55, KIF4A and NCAPG between LGG and normal tissues, so we finally selected CEP55, MELK, NUSAP1 and their negatively correlated miRNAs as the potential miRNA–mRNA networks (NUSAP1-hsa-let-7b-5p, MELK-hsa-miR770-5p-hsa-miR-1224-3p and Top2A-hsa-miR-218-5p/hsa-miR-495-3p/hsa-miR-7-1-3p/hsa-miR-203a-3p) for further verification. Pearson's correlation analysis was performed on the mRNA and miRNA expression profiles downloaded from CGGA datasets to validate

the potential miRNA–mRNA regulation. As shown in Figure 7A-F, only the expression of hsa-miR-495-3p and hsa-miR-1224-3p was markedly associated with TOP2A (r=0.339, P=0.004) and MELK (r=0.301, P=0.047). To validate these potential regulatory networks, the expression levels of these two miRNAs and their target genes were further detected in LGG clinical tissue samples and normal brain tissues by PCR. As shown in Figure 7G-J, in accordance with our previous analytic results, TOP2A and MELK were significantly upregulated, whereas hsa-miR-1224-3p and hsa-miR-495-3p were markedly downregulated in LGG compared to matched normal brain samples.

Discussion

Given that glioma is one of the most common tumours in the central nervous system, it is urgent to thoroughly explore the molecular pathogenesis of LGG. The dysregulation of the miRNA–mRNA regulatory network is believed to lead to a variety of human diseases, including cancers.[18–20] During the past few years, many studies on miRNA and mRNA expression profiles in LGG have been published. However, to date, comprehensive research on the miRNA–mRNA regulatory network in the pathogenesis of LGG remains largely insufficient. In the present
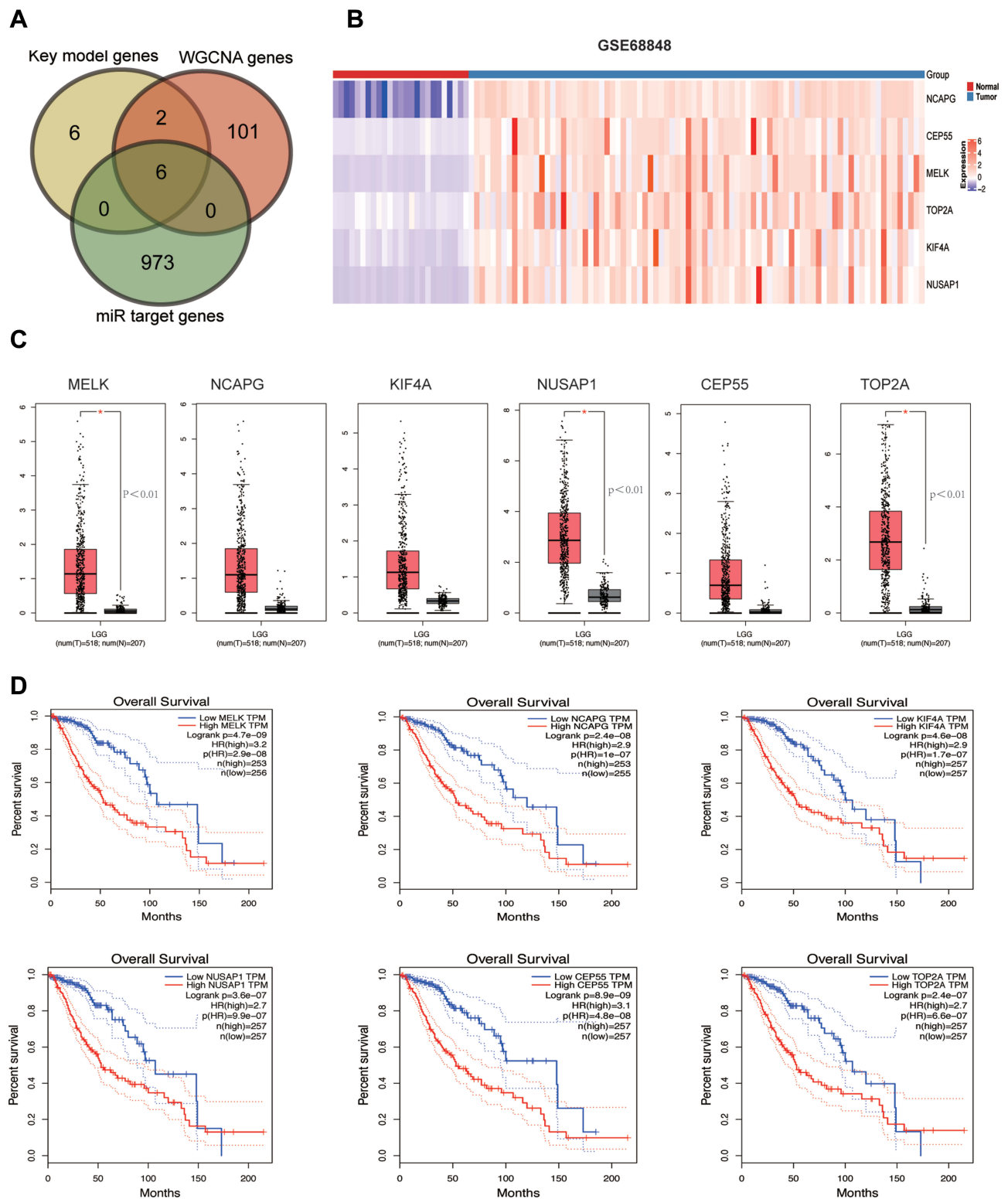
Figure 5 Hub genes validation and clinical analysis. (**A**) Venn analysis of differentially expressed miRNA target genes and key model genes in the PPI network and yellow model genes in WGCNA. Sex candidate genes were calculated, including MELK, NCAPG, KIF4A, NUSAP1, CEP55 and TOP2A. (**B**) Heatmap of the six hub genes in LGG tissues and normal brain tissues from the GES68848 dataset. Colouring illustrates the high expression (red) and low expression (blue) of genes. (**C**) The expression levels of 6 hub genes between LGG tissues and normal brain tissues validated according to the GEPIA database. (**D**) The expression level of hub genes and their prognostic significance were validated according to clinical datasets downloaded from TCGA. Survival curves of hub genes were generated in GraphPad Prism software, and the red and black curves represent the high and low expression groups, respectively. *<0.05.
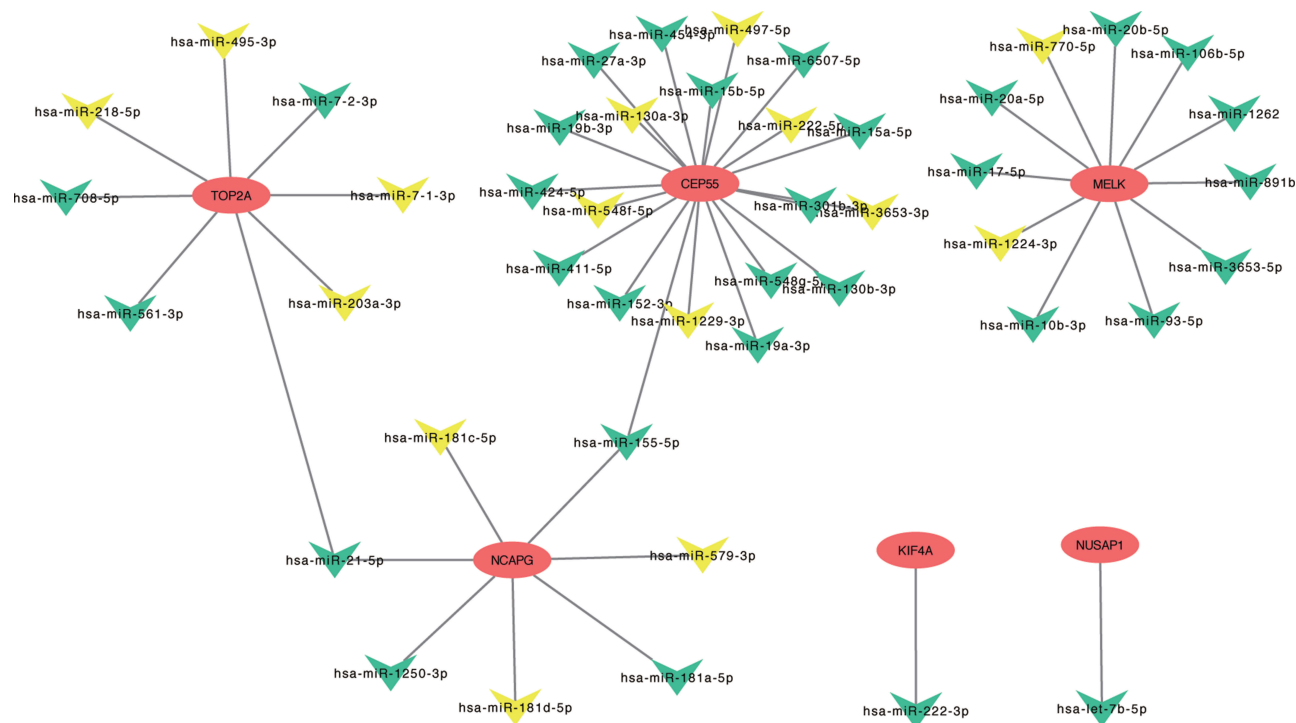
Figure 6 miRNA–mRNA regulate network construction. Identified potential miRNA–mRNA regulatory network contributing to the pathogenesis of LGG: Oval represents mRNA, and triangle represents miRNA. All shapes in red and blue represent upregulation, and yellow represents downregulation.

study, we used bioinformatics analysis of the LGG-associated miRNA–mRNA network to study the molecular pathogenesis of LGG. Ultimately, two regulatory pathways, miR-495-3p-TOP2A and miR-1224-3p-MELK, related to the pathogenesis of LGG, were identified.

Topoisomerase (DNA) II alpha (TOP2A) is involved in processes such as chromosome condensation, chromatid separation, and the relief of torsional stress that occurs during DNA transcription and replication. TOP2A is aberrantly expressed in many human cancers, including lung cancer, breast cancer, and colorectal cancer.[21–23] Tianmin et al[24] reported that TOP2A was highly expressed in glioma tissues compared with normal brain tissues. They also found that TOP2A overexpression was highly correlated with glioma grade stage, KI67 positive percentage, IDH1 mutation and poor prognosis. Similar results were also found in the present study. TOP2A was significantly upregulated in LGG tissues compared with normal brain tissues, which was validated by GEO, TCGA, GEPIA and experimental PCR. Moreover, survival analysis found that higher expression of TOP2A in LGG was associated with worse OS. All these results concluded that TOP2A can be a potential biomarker for the prognosis of LGG. An increasing number of studies have shown that miRNAs play an important role in the occurrence and development

of cancer. MiRNAs regulate target genes by binding to target mRNAs to repress their translation. To identify the pathway involved in the regulation of TOP2A expression in LGG, we constructed a miRNA–mRNA network and found that miR-495-3p may regulate the expression of TOP2A by targeting these oncogenes in LGG. miR-495-3p has been reported to possibly have oncogenic or anticancer functions in several cancers. For instance, miR-495-3p inhibits proliferation and invasion in clear cell renal cell carcinoma cells.[25] However, miR-495-3p can promote colon cancer cell proliferation via the Wnt/β-catenin signalling pathway.[26] In addition, miR-495-3p inhibits the proliferation, invasion and migration of osteosarcoma cells by targeting C1q/TNF-related protein 3[27]. Nevertheless, the roles and underlying mechanisms of miR-495-3p in low-grade glioma are unclear. The miRWALK website is a comprehensive miRNA target gene prediction database that integrates the prediction results of both TargetScan and miRDB.[16] We found that many miRNAs may target TOP2A from miRWALK3.0 datasets, but only miR-495-3p was statistically downregulated in LGG and was significantly correlated with the expression of TOP2A in LGG. The higher expression of TOP2A and lower expression of miR-495-3p in LGG tissues than in normal brain tissues were also validated
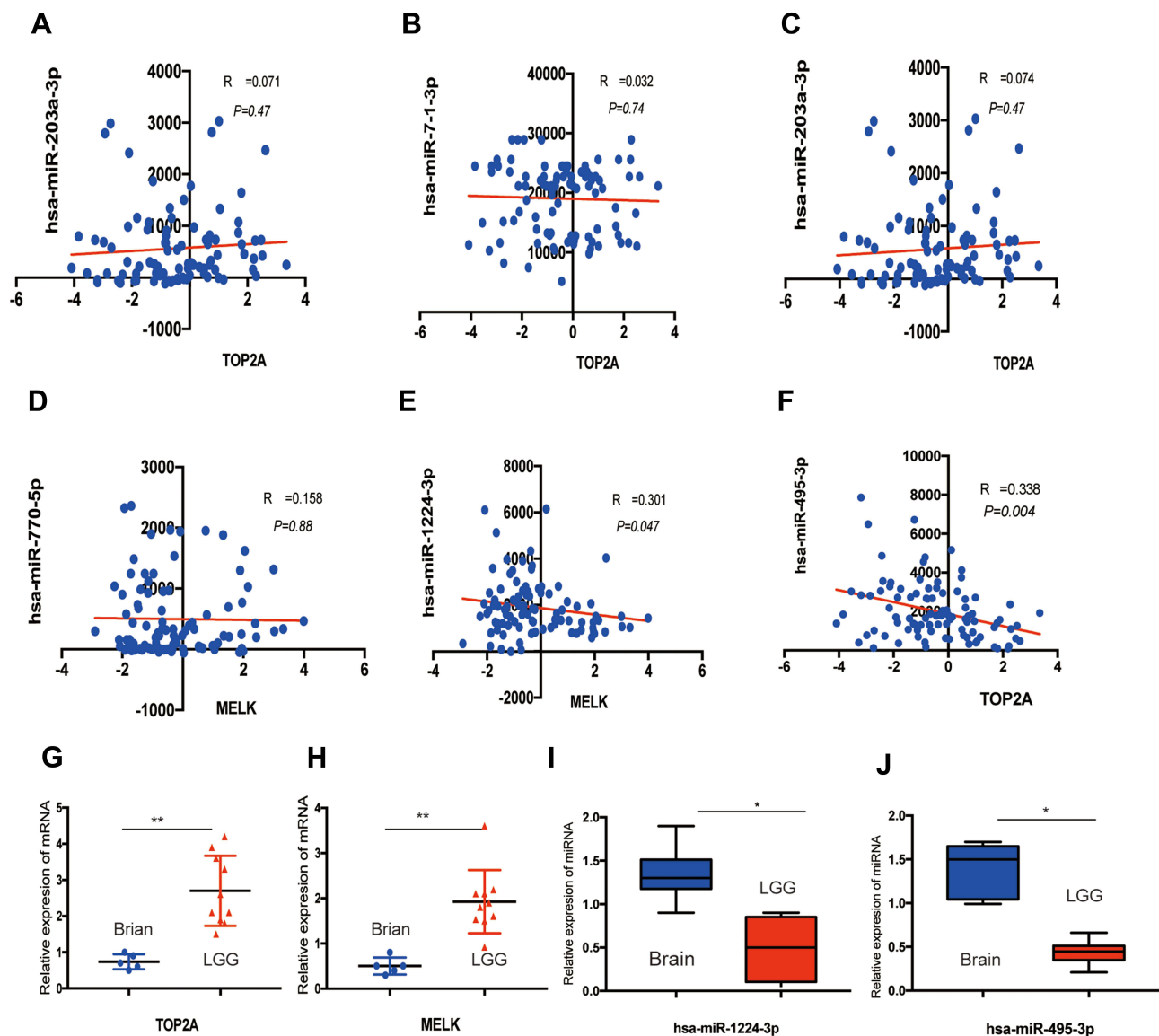
Figure 7 Validation of the significant miRNA–mRNA pathway. The correlation of potential DE-miRNAs and target genes validated in a patient cohort from CGGA datasets: (**A**) For TOP2A and hsa-miR-203a-3p (**B**) for TOP2A and hsa-miR-7-1-3p (**C**) for TOP2A and hsa-miR-203a-3p (**D**) for MELK and hsa-miR-770-5p (**E**) for MELK and hsa-miR-1224-3p (**F**) for TOP2A and hsa-miR-495-3p. (**G, H**) The relative mRNA expression of TOP2A and MELK between LGG tissues and normal brain tissues by using qRT-PCR, *<0.05, **<0.01. (**I, J**) The relative miRNA expression of hsa-miR-1224-3p and hsa-miR-495-3p between LGG tissues and normal brain tissues by using qRT-PCR, *<0.05.

by PCR experiments. All these findings support our analytical results that the miR-495-3p-TOP2A pathway may contribute to the pathogenesis of LGG. Therefore, future experimental validation of our current analytical result—the promotion of the miR-495-3p-TOP2A pathway in carcinogenesis of LGG is of great significance.

Our analysis suggested that miR-1224-3p-MELK is a tumour-suppressive pathway in LGG, and patients with higher expression of miR-1224-3p and lower expression of MELK always had a favourable OS. Maternal embryonic leucine zipper kinase (MELK) has been characterized as an

important kinase in development and has been implicated in a range of biological functions, including mitotic progression, proliferation, apoptosis, differentiation and stem cell phenotypes, and tumourigenesis.[28,29] Many investigations have verified that MELK plays an oncogenic role in the development of many malignant tumours. For example, MELK promotes endometrial carcinoma progression by activating the mTOR signalling pathway,[30] and MELK accelerates the progression of colorectal cancer by activating the FAK/Src pathway.[31] Regarding the role of MELK in glioma, several studies have proven its important

function in glioma stem-like cells.[32,33] Kaushal et al proved that MELK leads to the phosphorylation and activation of FOXM1, which results in a subsequent increase in mitotic regulatory genes in GSCs33 and the progression of glioma. MELK is considered to be an attractive molecular target for the development of new types of drugs due to its critical roles in cancer stem cell maintenance.[34] Our results indicated that MELK is obviously upregulated and significantly associated with a poor prognosis. Therefore, combined with these findings and our results, MELK may be selected as a potential target for the treatment of LGG. The role of miR-1224-3p in human tumours is not clear. The literature about miR-1224-3p is fairly sparse, and only two papers about miR-1224-3p in human cancer have been published to date. Zuo, et al illustrated that miR-1224-3p plays an antitumour role in lung adenocarcinoma via the miR-1224-3p/ETV1 signalling pathway.[35] Miah et al pointed out that miR-1224-3p may correlate with the prognosis of bladder cancer, but the molecular mechanism is unclear.[36] Our results indicated that miR-1224-3p correlated with a good prognosis of LGG patients but was always downregulated in LGG. The CGGA datasets, which contain both mRNA and miRNA profiles from 94 patients, revealed that the expression of MELK was negatively correlated with the expression of miR-1224-3p, which implies that miR-1224-3p may inhibit the growth of LGG cells by targeting MELK. However, this hypothesis needs to be verified by our further in vitro and in vivo experiments.

Although we performed a comprehensive analysis and some experimental validation of the miRNA–mRNA regulatory network involved in LGG and successfully identified several potential miRNA–mRNA pathways that may be linked to the development of LGG, there were also some limitations in this study: (1) The clinical sample size used for validation here was relatively small. (2) The specific molecular mechanism of miR-495-3p-TOP2A and miR-1224-3p-MELK in LGG has not been studied. (3) The dual-luciferase reporter assay was absent for the validation of miRNA and targeted mRNA in the present study. Therefore, future investigations with larger clinical samples, cell functional experiments and animal validation experiments to elucidate the signalling pathways of LGG are urgently required.

Conclusions

In conclusion, this study identified the hub genes and related miRNA–mRNA regulatory pathways contributing to the carcinogenesis of LGG, which may help us reveal the mechanisms underlying the development of LGG and provide more guidance for its prognosis. We hope that our findings are helpful for developing treatments for LGG that target the established miRNA–mRNA regulatory network, thus improving the prognosis of LGG patients in the future.

Funding

Disclosure

The authors declare that they have no conflicts of interest.

References

1. Jansen M, Yip S, Louis DN. Molecular pathology in adult gliomas: diagnostic, prognostic, and predictive markers. *Lancet Neurol*. 2010;9 (7):717–726. doi:10.1016/S1474-4422(10)70105-8
2. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol*. 2007;114(2):97–109.
3. Wirsching HG, Galanis E, Weller M. Glioblastoma. *Handb Clin Neurol*. 2016;134:381–397.
4. Morshed RA, Young JS, Hervey-Jumper SL, Berger MS. The management of low-grade gliomas in adults. *J Neurosurg Sci*. 2019;63 (4):450–457.
5. Reddy RRS, Ramanujam MV. High throughput sequencing-based approaches for gene expression analysis. *Methods Mol Biol*. 2018;1783:299–323.
6. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol*. 2016;131(6):803–820.
7. Luo G, Luo W, Sun X, et al. MicroRNA-21 promotes migration and invasion of glioma cells via activation of Sox2 and β-catenin signaling. *Mol Med Rep*. 2017;15(1):187–193.
8. Lou W, Liu J, Ding B, et al. Identification of potential miRNA-mRNA regulatory network contributing to pathogenesis of HBV-related HCC. *J Transl Med*. 2019;17(1):7.
9. Xie B, Zhao R, Bai B, et al. Identification of key tumorigenesis-related genes and their microRNAs in colon cancer. *Oncol Rep*. 2018;40(6):3551–3560.
10. Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res*. 2015;43(Databaseissue):D1049–D1056.
11. Du J, Yuan Z, Ma Z, Song J, Xie X, Kegg-path: CY. Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model. *Mol Biosyst*. 2014;10(9):2441–2447.
12. Jr DG, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol*. 2003;4(5): P3.
13. Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45(D1):D362–D368.
14. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559.
15. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–W102.
16. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. *PLoS One*. 2018;13(10):e0206239.

17. Zhao Z, Zhang KN, Wang Q, et al. Chinese Glioma Genome Atlas (CGGA): a Comprehensive Resource with Functional Genomic Data from Chinese Gliomas. *Genomics Proteomics Bioinformatics*. 2021; S1672-0229(21).

18. Cheng J, Zhuo H, Xu M, et al. Regulatory network of circRNA-miRNA-mRNA contributes to the histological classification and disease progression in gastric cancer. *J Transl Med*. 2018;16 (1):216.

19. Yang J, Song H, Cao K, Song J, Zhou J. Comprehensive analysis of Helicobacter pylori infection-associated diseases based on miRNA-mRNA interaction network. *Brief Bioinform*. 2019;20 (4):1492–1501.

20. Lou W, Liu J, Ding B et al. Identification of potential miRNA-mRNA regulatory network contributing to pathogenesis of HBV-related HCC. *J Transl Med*. 2019;17(1):7.

21. Wang TL, Ren YW, Wang HT, Yu H, Zhao YX. Association of Topoisomerase II (TOP2A) and Dual-Specificity Phosphatase 6 (DUSP6) single nucleotide polymorphisms with radiation treatment response and prognosis of lung cancer in Han Chinese. *Med Sci Monit*. 2017;23:984–993.

22. Dimas-González J, Maldonado-Lagunas V, Díaz-Chávez J, et al. Overexpression of p53 protein is a marker of poor prognosis in Mexican women with breast cancer. *Oncol Rep*. 2017;37 (5):3026–3036.

23. Hendricks A, Gieseler F, Nazzal S, et al. Prognostic relevance of topoisomerase II α and minichromosome maintenance protein 6 expression in colorectal cancer. *BMC Cancer*. 2019;19(1):429.

24. Zhou T, Wang Y, Qian D, Liang Q, Wang B. Over-expression of TOP2A as a prognostic biomarker in patients with glioma. *Int J Clin Exp Pathol*. 2018;11(3):1228–1237.

25. Wang LN, Zhu XQ, Song XS, Xu Y. Long noncoding RNA lung cancer associated transcript 1 promotes proliferation and invasion of clear cell renal cell carcinoma cells by negatively regulating miR-495-3p. *J Cell Biochem*. 2018;119(9):7599–7609.

26. Liu P, Shen JK, Hornicek FJ, Liu F, Duan Z. Wnt inhibitory factor 1 (WIF1) methylation and its association with clinical prognosis in patients with chondrosarcoma. *Sci Rep*. 2017;7(1):1580.

27. Zhao G, Zhang L, Qian D, Sun Y, Liu W. miR-495-3p inhibits the cell proliferation, invasion and migration of osteosarcoma by targeting C1q/TNF-related protein 3. *Onco Targets Ther*. 2019;12: 6133–6143.

28. Jiang P, Zhang D. Maternal embryonic leucine zipper kinase (MELK): a novel regulator in cell cycle control, embryonic development, and cancer. *Int J Mol Sci*. 2013;14(11):21551–21560.

29. Ganguly R, Mohyeldin A, Thiel J, Kornblum HI, Beullens M, Nakano I. MELK-a conserved kinase: functions, signaling, cancer, and controversy. *Clin Transl Med*. 2015;4:11.

30. Xu Q, Ge Q, Zhou Y, et al. MELK promotes Endometrial carcinoma progression via activating mTOR signaling pathway. *EBioMedicine*. 2020;51:102609.

31. Liu G, Zhan W, Guo W, et al. MELK accelerates the progression of colorectal cancer via activating the FAK/Src pathway. *Biochem Genet*. 2020;58(5):771–782.

32. Joshi K, Banasavadi-Siddegowda Y, Mo X et al. MELK-dependent FOXM1 phosphorylation is essential for proliferation of glioma stem cells. *Stem Cells*. 2013;31(6):1051–1063.

33. Ganguly R, Hong C s, Smith L g, Kornblum H i, Nakano I. Maternal embryonic leucine zipper kinase: key kinase for stem cell phenotype in glioma and other cancers. *Mol Cancer Ther*. 2014;13(6): 1393–1398.

34. Chung S, Nakamura Y. MELK inhibitor, novel molecular targeted therapeutics for human cancer stem cells. *Cell Cycle*. 2013;12 (11):1655–1656.

35. Zuo Y, Shen W, Wang C, Niu N, Circular P J. RNA Circ-ZNF609 promotes lung adenocarcinoma proliferation by modulating miR-1224-3p/ETV1 signaling. *Cancer Manag Res*. 2020;12: 2471–2479.

36. Miah S, Dudziec E, Drayton RM, et al. An evaluation of urinary microRNA reveals a high sensitivity for bladder cancer. *Br J Cancer*. 2012;107(1):123–128.