Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Bioinformatics analysis of KIF20A, a potential therapeutic target for glioblastoma

Lu Han
Nanchang University Second Affiliated Hospital
Jiayang Wang (homesheep@qq.com)
Jiangxi Agricultural University

Research article

Keywords: glioblastoma, KIF20A, expression protein profiling, children brain tumor, microtubule cytoskeleton organization

Posted Date: January 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-138484/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Title page:

Bioinformatics analysis of KIF20A, a potential therapeutic target for glioblastoma

Running title: KIF20A, as a therapeutic target for GBM

Lu Han [a], Jiayang Wang [b, c*]

[a] Department of Cardiovascular Medicine, the Second Affiliated Hospital of Nanchang University, Nanchang, 330006, China;

[b] Software College, Jiangxi Agricultural University, Nanchang, 330045, China;

[c] School of information management, Jiangxi University of Financial and Economic, Nanchang, 330013, China;

* Corresponding author: Jiayang Wang, PhD, E-mail: jiayang-wang@qq.com, Tel: +86-18879166127; Mail Address: Jiangxi Agricultural University, Zhimin avenue No.1, Jiaoqiao town, Qingshanhu District, Nanchang of Jiangxi Province, 330045, China.

Bioinformatics analysis of KIF20A, a potential therapeutic target for glioblastoma

Lu Han [a], Jiayang Wang [b, c*]

Abstract

Background: Glioblastoma (GBM) is a malignant brain tumor with high mobility. The median survival time of GBM patients is 15 months. Currently, there is no effective treatment for improving the prognosis of the GBM due to a lack of prognostic markers.

Materials and methods: To predict core therapeutic targets for GBM, we analyzed four microarray datasets (GSE49810, GSE50161, GSE65624, and GSE90604) selected from the Gene Expression Omnibus (GEO) database and the other datasets obtained from The Cancer Genome Atlas (TCGA) database. Expression protein array of 227 GBM samples and 18 normal samples were clustered to summarize GBM tissue classification. Differentially expressed genes (DEGs) were analyzed by comparing GBM and normal brain tissues in each profile using the limma package of R software. GO function and KEGG pathway enrichment analysis was performed using the DAVID database. Overlapping DEGs were ranked based on protein expression ratios from the comparison between cancer and normal samples using robustRankaggreg package of R software and scored from high to low. Protein-protein interaction (PPI) network was visualized using CytoHubba and Cluego plugins in Cytoscape software. Core hub genes were analyzed by MCC, MNC, DMNC, and EPC methods. Besides, the GEPIA tool was used to create the survival curves and boxplots to evaluate the prognostic effect of hub genes for improving the diagnostic outcomes and treatment of GBM.

Results: A total of 2064 DEGs were analyzed (1400 downregulated DEGs and 1664 upregulated DEGs) in the GEO database. 3292 DEGs were found (1485 upregulated DEGs and 1807 downregulated DEGs) in TCGA. We selected 221 significant DEGs from four microarrays. Combining the GEO results with the results of TCGA, we found only 181 common DEGs by using Venn analysis. Further, expression levels of KIF20A selected from 10 hub genes closely associated with the survival rate.

Conclusion: Up-regulation of KIF20A has a pivotal role in controlling the prognosis of GBM in 2 years follow-up period; KIF20A should be considered as a potential therapeutic target for GBM.

Keywords: glioblastoma, KIF20A, expression protein profiling, children brain tumor, microtubule cytoskeleton organization

Introduction

Glioblastoma (GBM) is one of the brain cancer types in children, characterized by aggressive development. Even though it is uncommon, the morbidity of GBM in children is extremely high with a median survival time of only 15 months. At present, cerebral surgery, radiation, chemotherapy, and treatment with temozolomide are the common therapeutic options for GBM patients. However, there is a lack of effective therapies and good prognostic factors of GBM, especially in children [1]. GBM generally occurs in the supratentorial region (frontal, temporal, parietal, and occipital lobes) of the brain, and rarely originates in the cerebellum [2]. Currently, GBM accounts for 10-15% of the childhood brain tumors and originates from the glial cells in the central nervous system [3]. A recent study demonstrated that the underlying etiology of GBM is mainly associated with a complex microenvironment, which is regarded as a milieu that provides the condition of tumor cell transcriptomic adaptability and triggers disease development [4]. Besides, an imbalance of symmetric and asymmetric divisions in neural progenitor cells (NPCs) may play a critical role in brain tumor progression. Geng et al. reported that the increased levels of kinesin-like family protein 20A (KIF20A) and its relevant signaling pathways participate in the process of NPC division and decides the fate of brain cells [5]. Interestingly, it has already affirmed that the KIF20A gene is involved in the process of protecting cells' microenvironment [6]. However, the pathological mechanism of GBM remains unclear; therefore, efficient targeted-therapies for GBM are lacking. Researchers are collecting GBM cases and screening core genes to find a targeted-therapy for GBM.

The Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) database have large arrays of expression proteins for GBM samples. We revealed a series of differentially expressed genes (DEGs) *via* comparison of GBM and healthy brain tissues separately from two databases. Besides, overlapping DEGs were selected together from GEO and TCGA. We carried out Gene Ontology (GO) functional annotation analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, and constructed protein-protein interaction (PPI) network [7] to find the hub genes of GBM. We also evaluated the survival rate and created boxplots of gene expression levels by using the Gene Expression Profiling Interactive Analysis (GEPIA) [8].

Materials and methods

Data source

In this study, the gene expression profiles were selected from GEO (https://www.ncbi.nlm.nih.gov/geo/) and TCGA (https://portal.gdc.cancer.gov/)

databases. Through precise searching of the profiles that met the criteria of GBM, we retrieved 71 cases (58 tumors and 13 normal tissues) from 1024 series of studies; four gene expression profiles (GSE49810, GSE50161, GSE65624, and GSE90604) were included. The criteria were as follows: GSE49810 is based on GPL15648 platform [(HuGene-1_0-st) Affymetrix Human Gene 1.0 ST Array (HuGene10stv1_Hs_ENTREZG version 13.0.0)]; GSE50161 is based on the GPL4133 platform [(HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array]. GSE65624 is based on GPL81 platform [(MG_U74Av2) Affymetrix Murine Genome U74A Version 2 Array]; GSE90604 is based on GPL17692 platform [(HuGene-2_1-st) Affymetrix Human Gene 2.1 ST Array]. Besides, 174 samples were filtered *via* manual collection from the TCGA database according to the screening requirements, including project (TCGA-GBM), category of data (transcriptome profiling), type of data (gene expression quantification), and type of workflow (HTSeq-counts). 169 GBM and 5 normal samples were included.

DEGs data screening

R statistical software (https://www.r-project.org/; version 4.0.0) was used to calculate the DEGs *via* comparison of the tissues of GBM and normal groups along with Bioconductor/ limma package (code: install.packages.Bioconductor; install.packages.limma) [9]. The threshold of adjusted P-value (Padj) and log2-fold change (|logFC|) were decided. Only those genes that met the criteria of Padj < 0.05 and $|logFC| \geq 2.0$ were recorded as DEGs. Statistical analysis was carried out for each profile and the profiles were corrected and normalized for further calculation. In GEO, the ranking of the most significant protein expression was identified using RobustRankAggreg packages from R software (Padj = 0.05 and logFC = 1). Combining the results of DEGs from TCGA, we used the Venn website (http://bioinformatics.psb.ugent.be/webtools/Venn/) to find the intersecting DEGs between GEO and TCGA databases.

GO Function and KEGG pathway enrichment analysis of DEGs

GO analysis is a typical method for functional enrichment analysis. GO functions contain three categories including biological processes (BP), molecular functions (MF), and cellular components (CC) [10]. KEGG is a collection of manually drawn pathway maps that include knowledge of molecular interaction, reaction and relation networks for metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, human diseases, and drug development. We performed annotations, visualization, and integrated the discovery database (DAVID) tools (https://david.ncifcrf.gov/) to analyze GO and KEGG pathways [11]. In addition, $P < 0.01$ and gene count > 5 were considered as statistically significant.

PPI network construction and core gene identification

Cytoscape software (http://www.cytoscape.org/; version 3.8.0) was used to visualize PPI pairs, which were assessed using the Search Tool for the Retrieval of Interacting Genes (STRING) database (http://string-db.org/). The value of medium confidence in the PPI network was 0.4. Consequently, MCC, MNC, DMNC, and EPC methods were performed to evaluate the comprehensive score of node genes and assess the top 10 genes identified as hub genes. CotyHubba application was utilized to analyze the network of hub genes. Moreover, ClueGO, a plugin in Cytoscape [12-13], was used to depict the underlying pathway of overlapped DEGs from the two databases, and to draw the network of DEGs relative to upstream and downstream genes.

Survival curves

The GEIPA (http://gepia.cancer-pku.cn/) is a website to calculate cancer gene expression levels and to perform survival curve analysis [14]. The expression of DEGs in cancer and normal cells was compared for the discovery of specific genes. Then, GEIPA is used to perform a disease-free survival analysis. Besides, GEIPA provides survival plots in terms of pathological stages based on TCGA clinical annotation. Survival plots can also include Cox proportional hazard ratio (HR) and the 95% confidence intervals (CIs). All data should be corrected and adjusted. Probe IDs from each GSE database are presented in TableS1.

Receiver operating characteristic curves

ROC (Receiver operating characteristic) curve analysis can demonstrate whether a marker can distinguish between normal and cancer. Thus, a standard ROC curve approach can be used to determine the capability of the markers [15]. The area under the ROC Curve (AUC) provides an aggregate measure of performance across all possible classification thresholds. It aims to assess the probability that the model ranks a random tumor sample higher than a random normal sample.

Results

Identification of DEGs

Four gene expression profiles (GSE49810, GSE50161, GSE65624, and GSE90604) and TCGA databases were selected in this study. From TCGA, 3292 DEGs (1485 upregulated DEGs and 1807 downregulated DEGs) were identified from a comparison of 169 GBM and 5 normal samples. Among the outcomes of GEO, the GSE49810 dataset was found to contain 5 GBM samples and 1 normal tissue sample. 1340 DEGs were expressed in GSE50161 that contains 34 GBM and 2 normal specimens; GSE65626 contains 3 GBM and 3 adjacent normal specimens. GSE90604 contains 16 GBM and 7 normal samples (Table 1). Hierarchical clustering of the DEGs separated the case and the control samples with an accuracy rate of higher than 95% for

GSE49810 (Fig. 1A), GSE50161 (Fig. 1B), GSE65624 (Fig. 1C), GSE90604 (Fig. 1D), and TCGA profiles (Fig. 1E). Meanwhile, based on the criteria – $P < 0.05$ and $|logFC| > 2$– a total of 277 DEGs (71 upregulated genes and 206 downregulated genes) were identified from GSE49810. In GSE50161, 1,340 DEGs were identified, including 500 up-regulated genes and 840 downregulated genes. From GSE65626, 100 DEGs were screened (30 upregulated genes and 70 downregulated genes). In GSE90604, 347 DEGs met the criteria, including 63 upregulated genes and 284 downregulated genes (Fig. 2A). All DEGs were identified after the comparison of GBM samples and normal brain samples. RobustRankAggreg package of R software was used to rank the overlapping genes in four DEG profiles from GEO databases. The results showed that 211 DEGs were commonly expressed in four GSE chips and the results were statistically significant (Fig. 2B). Subsequently, 181 overlapping genes were identified from GEO and TCGA databases *via* Venn analysis (Fig. 2C). This group contained 103 up-regulated and 78 down-regulated DEGs. The heatmap of all the four GSE profiles and the top 35 genes of 3299 DEGs from TCGA is presented in Fig. 2D.

GO function and KEGG pathway analysis of DEGs

GO function and KEGG pathway enrichment analysis of DEGs were performed using DAVID (Table 2). The enriched GO terms were divided into CC, BP, and MF ontologies. The results of GO analysis indicated that 181 overlapping DEGs were mainly enriched in BPs including cell adhesion (GO: 0007155, P-value: 0.0072), and gamma-aminobutyric acid signaling pathway (GO: 0007214, P-value: 0.0015). MF analysis showed that the DEGs were significantly enriched in GABA-A receptor activity (GO: 0004890, P-value: 9.20E-04), and extracellular ligand-gated ion channel activity (GO: 0005230, P-value: 0.0032). For the cell potential core genes in the GBM component, DEGs were enriched in the synapse (GO: 0045202, P-value: 9.02E-04), and GABA-A receptor complex-dependent catabolic process (GO: 1902711, P-value: 7.26E-04). Besides, the results of the KEGG pathway analysis showed that DEGs were mainly involved in morphine (cfa05032, Padj value: 0.0059) and nicotine addictions (cfa05033, Padj value: 0.0039).

Selected Hub genes and relative network construction

Protein interactions among the 181 DEGs were predicted with STRING tools. A total of 181 nodes and 867 edges were included in the PPI network, as presented in Fig. 3A. The distribution of all genes was according to the yFiles Organic layout in Cytoscape software. Besides, the R square and correlation coefficient of the network was found to be 0.1903 and 0.33 respectively (Fig. 3B). Next, the top 10 hub genes were filtered and the network of related proteins was constructed *via* CytoHubba plugin (Fig. 3C; 3D). Besides, we annotated the participating pathway of 181 DEGs and the relative networks *via* the Cluego plugin (Fig. 3E). The network of 10 hub genes with their relative pathways is shown in Fig. 3F. All of these hub genes were found to be up-regulated in

GBM. Also, 10 hub gens evaluated by MCC, MNC, DMNC, and EPC methods in the PPI network were identified (Table 3). MCC analysis revealed that seven genes [Benzimidazole protein 1 (BUB1), Centromere Protein E (CENPE), Centromere Protein F (CENPF), Cyclin B1 (CCNB1), Kinesin Family Member 11 (KIF11), Kinesin Family Member 20A (KIF20A), Kinesin Family Member 15 (KIF15)] were having the highest connectivity score (5.35366489273773E+19), followed by rest of the genes [Xklp2 (TPX2), non-SMC condensin I complex subunit G (NCAPG), non-SMC condensin I complex subunit H (NCAPH); score = 5.35366489272975E+19; Table 3]. Although the score of CENPF, CCNB1, and KIF11 is slightly higher than the score of KIF20A in MNC, DMNC and EPC demonstrated a higher score of KIF20A as compared with that of CENPF, CCNB1, and KIF11. Therefore, KIF20A plays a pivotal role in the underlying mechanism of GBM.

Survival and ROC curve of selected core genes

To investigate the prognostic values of the 10 potential hub genes, we used the GEPIA bioinformatics analysis platform. A total of 370 GBM were available for analysis of overall survival. High expression of these hub genes, except KIF20A, was not found to be significantly associated with the overall survival of GBM patients (Fig. 4). Overexpression of KIF20A was regarded as a prognostic factor of relapse-free survival of GBM patients (logrank p=0.033, HR=2, p(HR)=0.039, n=58; Fig. 5). However, these findings were not relevant for overall survival analysis. We also analyzed ROC curves to evaluate sample recognition efficiency in the top 10 hub genes. CENPF was found to be highest in terms of AUROC value (AUC: 0.9271). Furthermore, boxplot analysis revealed that expression levels in the top 10 hub genes in GBM tissues were up-regulated as compared with normal tissues (Fig. 6).

Discussion

Glioblastomas (GBMs) are the most aggressive malignant primary tumors of the brain. Although common in adults [16], GBMs are also being detected in children [17]. Currently, lomustine and temozolomide following radiotherapy and concurrent temozolomide are the main therapeutic options available to the GBM patients. However, these treatments are still unable to effectively improve prognosis, especially in children [18]. Generally, children GBM patients only survive for 1-2.5 years [2, 19], with higher chances of relapse. Although the patterns of GBM in all patients are nearly identical, children have different clinical symptoms and post-therapy side-effects as compared with adults [2]. A study has already reported the genes involved and described the regulation of the pathological mechanisms of GBM [20].

In the present study, the genomic analysis was used to discover prognostic factors for the targeted-therapy of GBM. First, we selected four protein expression arrays from

GEO (GSE49810, GSE50161, GSE65624, and GSE90604), and TCGA databases. Besides, DEGs for each profile were detected separately. Subsequently, RobustRankAggreg package of R software was used to analyze the overlapping DEGs that met the criteria of $|logFC| > 2$ and Padj value < 0.05, commonly expressed in four arrays. The basis of ranking the core genes was as per the expression ratio of cancer/normal samples from highest to lowest, which were divided into up-regulation and down-regulation groups. Besides, 181 DEGs from GEO and TCGA databases were found to be intersecting and overlapping. This group included 78 upregulated and 103 down-regulated genes.

GO functions indicated that 181 DEGs participated in the process of cell adhesion, synapse, GABA-A receptor complex and its activity, extracellular ligand-gated ion channel activity, and gamma-aminobutyric acid signaling pathway. Besides, it was demonstrated that the DEGs were involved in the mechanisms of nicotine and morphine addictions *via* KEGG pathway analysis. These findings revealed that the pathological mechanisms associated with GBM involve multiple genes and relative biological processes. Furthermore, we created a PPI network of 181 overlapping DEGs *via* the STRING tool and the outcomes were 181 nodes, 867 edges, average node degree (9.63), and the P-value of enrichment (1.0e-16). CytoHubba and Cluego (Cytoscape plugins), not only identified up- and down-regulated DEGs and their interactive relationships but also revealed the pathways and biological processes in which the genes participate. Furthermore, four typical methods were applied to evaluate the comprehensive significance of 10 hub genes for GBMs: MCC, DMNC, EPC, and MNC. KIF20A, KIF11, KIF15, CENPE, CENPF, CCNB1, BUB1, NCAPG, NCAPH, and TPX2 were identified as hub genes. Among them, KIF20A, TPX2, KIF11, CENPE, and CCNB1 are involved in the process of microtubule cytoskeleton organization. Meanwhile, BUB1, NCAPG, NCAPH, CCNB1, and CENPF are involved in nuclear chromosome segregation. Spindle organization is mainly regulated by KIF11, TPX2, CENPE, and CCNB1. KIF15 was the only hub gene that takes part in the retrograde vesicle-mediated transport mechanism (Golgi to endoplasmic reticulum). CENPF was involved in the organelle localization. CENPE and CCNB1 also participate in organelle localization. Finally, we also found that the expression protein levels of KIF20A has the highest value as compared with other hub genes *via* four different methods.

Even though the role of KIF20A in the pathological mechanism of GBM is unknown, a previous study demonstrated that this protein is involved in the development and progression of human cancers [21]. Moreover, KIF20A is also regarded as an independent unfavorable prognostic factor, which can be a prospective therapeutic target for cervical cancer [21]. Besides that, multiple studies have also confirmed that KIF20A mediates the development of renal, liver, and pancreatic cancers [22-24]. Interestingly, Saito et al. discovered that KIF20A may be considered as a potential immunotherapeutic target of glioma [23]. Moreover, the process by which KIF20A

regulates the division modes of neural progenitor cells is also known [5]. It may be hypothesized that KIF20A mediates the pathological mechanism of GBM. As per a meta-analysis by Li et al., enhanced KIF20A expression levels cause poor prognosis and affect the cell growth and survival processes [23, 26]. KIF20A mainly participates in cancer cell proliferation, apoptosis, and migration through the regulation of various signaling pathways, especially the process of the spindle organization [20, 27, 28]. Further, KIF20A can also interact with KIF11 [6]. Combined with the results of KIF11 enriched in microtubule organization [29], it affirmed that the effect of KIF20A was associated microtubule cytoskeleton organization with modulate spindle organization, which was consistent with the outcomes that KIF20A exerted biological function in our study, called microtubule cytoskeleton organization.

Finally, we also analyzed the survival rate and created boxplots of gene expression levels of 10 hub genes *via* the GEPIA tool. Although the expression levels of all 10 genes in cancer tissues were significantly higher than the levels in the normal tissues, the survival curve analysis revealed that only KIF20A can be regarded as a predictor of GBM in terms of the survival rate in the 2 years follow-up period. As compared with GBM patients with lower KIF20A expression levels, GBM patients with higher KIF20A expression levels have a poor prognosis. Besides, AUC analysis was performed to evaluate the possibility of diagnosing GBM by these 10 hub genes. We further demonstrated that KIF20A is not only support tumor bioenergetics but also modulate the pathological progression of GBM. Therefore, KIF20A might be regarded as an unfavorable prognostic factor for GBM.

Conclusions

In conclusion, bioinformatics analyses were performed on the microarray datasets of GSE49810, GSE50161, GSE65624, GSE90604, and TCGA to explore the genetic mechanisms of GBM. The 181 overlapping DEGs from GEO and TCGA databases corresponding to GBM and normal samples were detected. KIF20A was found to play a critical role in the regulation of the microtubule cytoskeleton organization process, closely associated with GBM pathogenesis. KIF20A might play a role in targeted-therapy for GBM and the precise prognosis in clinical settings. However, further studies are needed to confirm these findings.

Acknowledgments

This study was supported by grants from the Science and Technology Project of the Jiangxi Provincial Department of Education (No: GJJ190240).

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Disclosure

The authors report no conflicts of interest in this work.

Competing of interest

The authors declare that they have no competing interests.

References

[1] Bryukhovetskiy I, Ponomarenko A, Lyakhova I, Zaitsev S, Zayats Y, Korneyko M, et al. Personalized regulation of glioblastoma cancer stem cells based on biomedical technologies: From theory to experiment (Review). Int J Mol Med. 2018;42(2):691-702.

[2] Fabian D, Guillermo Prieto Eibl MDP, Alnahhas I, Sebastian N, Giglio P, Puduvalli V, et al. Treatment of Glioblastoma (GBM) With the Addition of Tumor-Treating Fields (TTF): A Review. Cancers (Basel). 2019;11(2):174.

[3] Witthayanuwat S, Pesee M, Supaadirek C, Supakalin N, Thamronganantasakul K, Krusun S. Survival Analysis of Glioblastoma Multiforme. Asian Pac J Cancer Prev. 2018;19(9):2613-2617.

[4] Klopfenstein Q, Truntzer C, Vincent J, Ghiringhelli F. Cell lines and immune classification of glioblastoma define patient's prognosis. Br J Cancer. 2019;120(8):806-814.

[5] Geng A, Qiu R, Murai K, Liu J, Wu X, Zhang H, et al. KIF20A/MKLP2 regulates the division modes of neural progenitor cells during cortical development. Nat Commun. 2018;9(1):2707.

[6] Xu Y, Takahashi Y, Wang Y, Hama A, Nishio N, Muramatsu H, et al. Downregulation of GATA-2 and overexpression of adipogenic gene-PPARgamma in mesenchymal stem cells from patients with aplastic anemia. Exp Hematol. 2009;37(12):1393-1399.

[7] Ding J, Zhang Y. Analysis of Key GO Terms and KEGG Pathways Associated With Carcinogenic Chemicals. Comb Chem High Throughput Screen. 2017;10(20):1-20.

[8] Vella D, Marini S, Vitali F, Di Silvestre D, Mauri G, Bellazzi R. MTGO: PPI Network Analysis Via Topological and Functional Module Identification. Sci Rep. 2018;8(1):5499.

[9] Santiago ID, Carroll T. Analysis of ChIP-seq Data in R/Bioconductor. Methods Mol

Biol. 2018;1689:195-226.

[10] Luo Y, Yan Y, Zhang S, Li Z. Computational Approach to Investigating Key GO Terms and KEGG Pathways Associated With CNV. Biomed Res Int. 2018;2018:e8406857.

[11] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioin-formatics resources. Nat Protoc. 2009;4(1):44-57.

[12] Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25(8):1091-1093.

[13] Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics. 2013;29(5):661-663.

[14] 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(1):98-102.

[15] Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC Curve Analysis in Medical Research: Current Methods and Applications. BMC Med Res Methodol. 2017;17(1):53.

[16] Ostrom QT, Gittleman H, Fulop J, Liu M, Blanda R, Kromer C, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008–2012. Neuro Oncol. 2015;17(4):1–62.

[17] Venkatesh HS, Tam LT, Woo PJ, Lennon J, Nagaraja S, Gillespie SM, et al. Targeting Neuronal Activity-Regulated neuroligin-3 Dependency in High-Grade Glioma. Nature. 2017;549(7673):533-537.

[18] Jakacki RI, Cohen KJ, Buxton A, Krailo MD, Burger PC, Rosenblum MK, et al. Phase 2 Study of Concurrent Radiotherapy and Temozolomide Followed by Temozolomide and Lomustine in the Treatment of Children With High-Grade Glioma: A Report of the Children's Oncology Group ACNS0423 Study. Neuro Oncol. 2016;18(10):1442-1450.

[19] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010;17(1):98-110.

[20] Wu WD, Yu KW, Zhong N, Xiao Y, She ZY. Roles and Mechanisms of Kinesin-6 KIF20A in Spindle Organization during Cell Division. Eur J Cell Biol. 2019;98(2-4):74-80.

[21] Zhang W, He W, Shi Y, Gu H, Li M, Liu Z, Feng Y, et al. High Expression of KIF20A Is Associated with Poor Overall Survival and Tumor Progression in Early-Stage Cervical Squamous Cell Carcinoma. PLoS ONE 2016;11(12): e0167449.

[22] Saito K, Ohta S, Kawakami Y, Yoshida K, Toda M. Functional analysis of KIF20A, a potential immunotherapeutic target for glioma. J Neurooncol. 2017;132(1):63-74.

[23] Xiu G, Sui X, Wang Y, Zhang Z. FOXM1 Regulates Radiosensitivity of Lung Cancer Cell Partly by Upregulating KIF20A. Eur J Pharmacol. 2018;833:79-85.

[24] Wei W, Lv Y, Gan Z, Zhang Y, Han X, Xu Z. Identification of Key Genes Involved in the Metastasis of Clear Cell Renal Cell Carcinoma. Oncol Lett. 2019;17(5):4321-4328.

[25] Miyazawa M, Katsuda M, Maguchi H, Katanuma A, Ishii H, Ozaka M, et al. Phase II Clinical Trial Using Novel Peptide Cocktail Vaccine as a Postoperative Adjuvant Treatment for Surgically Resected Pancreatic Cancer Patients. Int J Cancer. 2017;140(4):973-982.

[26] Li X, Shu K, Wang Z, Ding D. Prognostic significance of KIF2A and KIF20A expression in human cancer a systematic review and meta-analysis Medicine (Baltimore). 2019;98(46):e18040.

[27] Gasnereau I, Boissan M, Margall-Ducos G, Couchy G, Wendum D, Bourgain-Guglielmetti F, et al. KIF20A mRNA and Its Product MKlp2 Are Increased During Hepatocyte Proliferation and Hepatocarcinogenesis. Am J Pathol. 2012;180(1):131–140.

[28] Wang K, Lin C, Wang C, Shao Q, Gao W, Song B, et al. Silencing Kif2a induces apoptosis in squamous cell carcinoma of the oral tongue through inhibition of the PI3K/Akt signaling pathway. Mol Med Rep 2014;9(1):273–278.

[29] Roostalu J, Rickman J, Thomas C, Nédélec F, Surrey T. Determinants of Polar versus Nematic Organization in Networks of Dynamic Microtubules and Mitotic Motors. Cell. 2018;175(3):796-808.

Figure 1

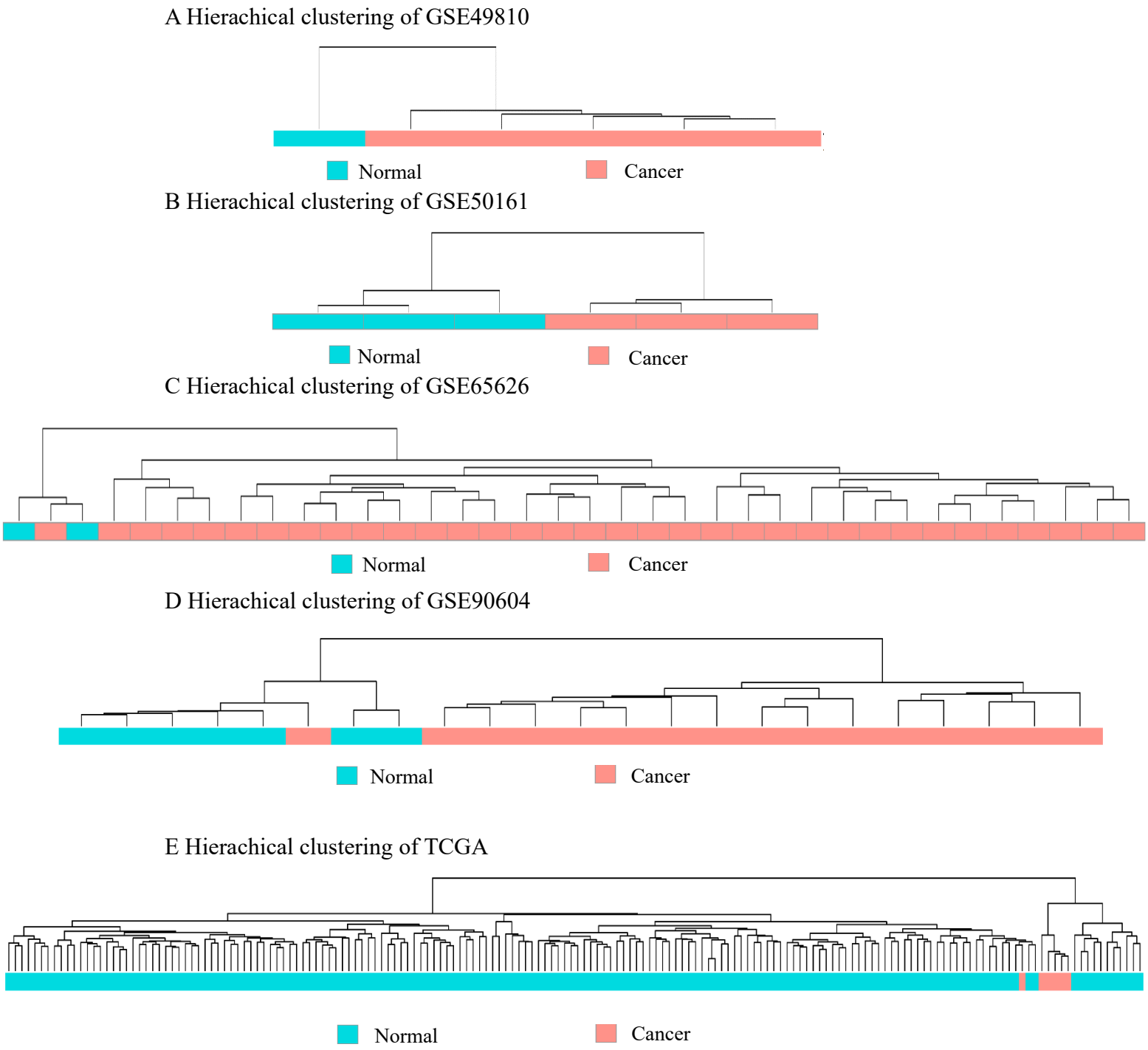
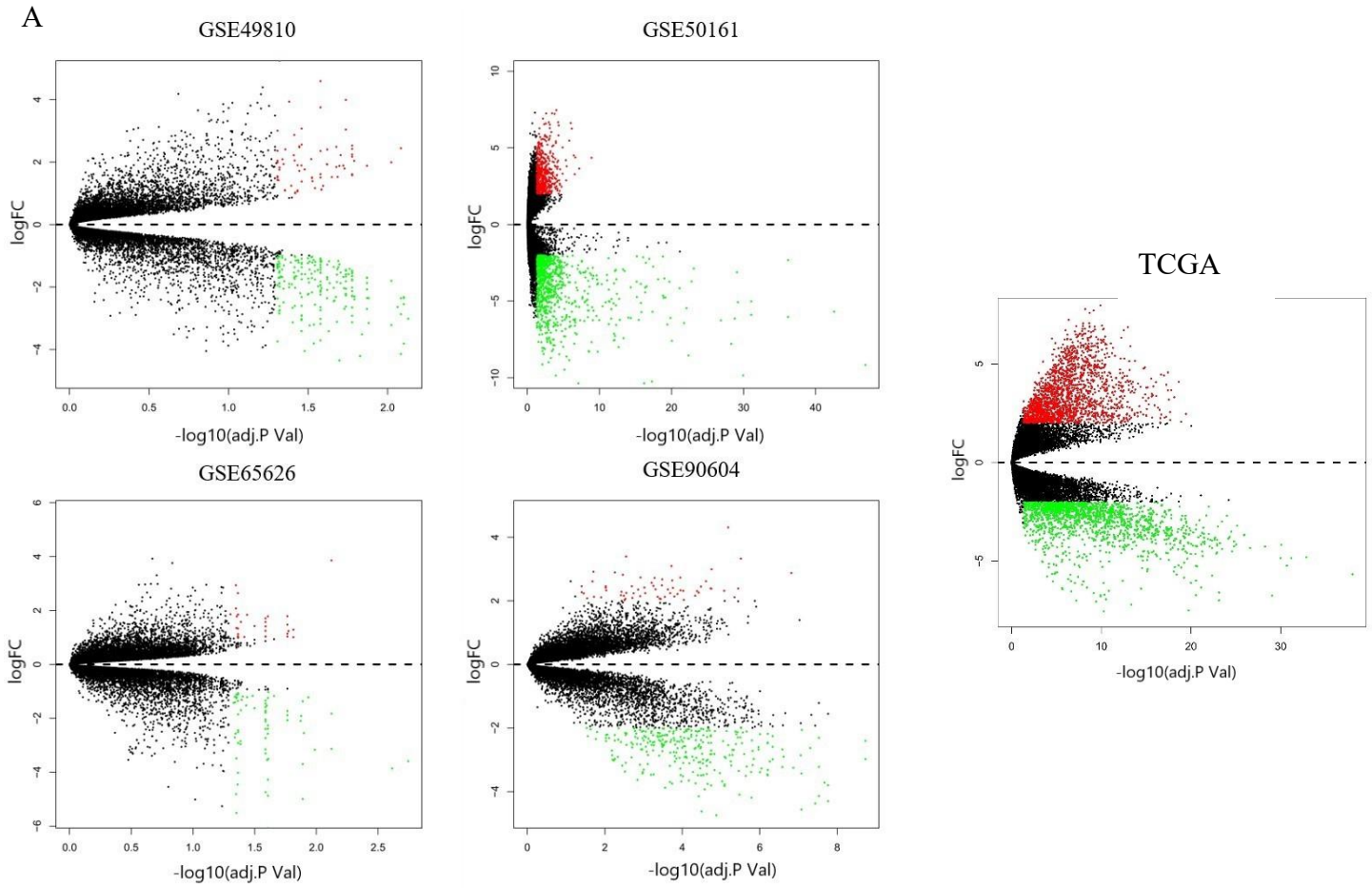
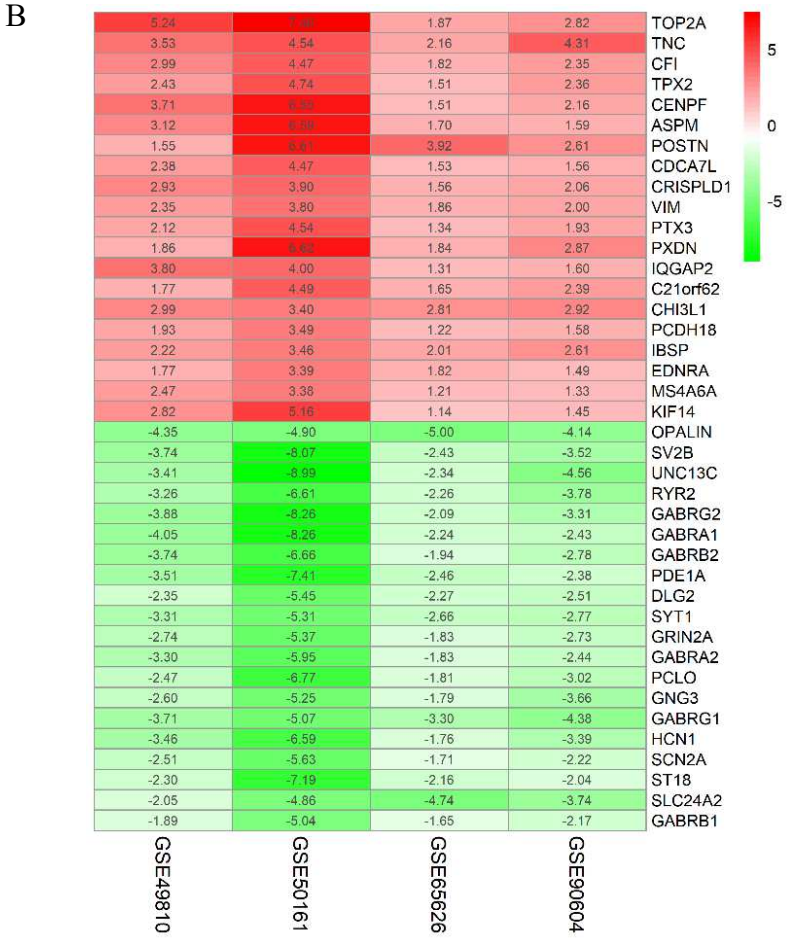


Fig. 1. Hierarchial clustering dendrograms for GSE49810 (A), GSE50161 (B), GSE65626 (C), GSE90604 (D) and TCGA (E). Blue and pink represent normal and GBM tumor samples, respectively.

Figure 2


A
GSE49810
GSE50161
TCGA
GSE65626
GSE90604
logFC
-log10(adj.P Val)


B
TOP2A 5.24 7.48 1.87 2.82
TNC 3.53 4.54 2.16 4.31
CFI 2.99 4.47 1.82 2.35
TPX2 2.43 4.74 1.51 2.36
CENPF 3.71 6.53 1.51 2.16
ASPM 3.12 6.59 1.70 1.59
POSTN 1.55 6.61 3.92 2.61
CDCA7L 2.38 4.47 1.53 1.56
CRISPLD1 2.93 3.90 1.56 2.06
VIM 2.35 3.80 1.86 2.00
PTX3 2.12 4.54 1.34 1.93
PXDN 1.86 6.62 1.84 2.87
IQGAP2 3.80 4.00 1.31 1.60
C21orf62 1.77 4.49 1.65 2.39
CHI3L1 2.99 3.40 2.81 2.92
PCDH18 1.93 3.49 1.22 1.58
IBSP 2.22 3.46 2.01 2.61
EDNRA 1.77 3.39 1.82 1.49
MS4A6A 2.47 3.38 1.21 1.33
KIF14 2.82 5.16 1.14 1.45
OPALIN -4.35 -4.90 -5.00 -4.14
SV2B -3.74 -8.07 -2.43 -3.52
UNC13C -3.41 -8.99 -2.34 -4.56
RYR2 -3.26 -6.61 -2.26 -3.78
GABRG2 -3.88 -8.26 -2.09 -3.31
GABRA1 -4.05 -8.26 -2.24 -2.43
GABRB2 -3.74 -6.66 -1.94 -2.78
PDE1A -3.51 -7.41 -2.46 -2.38
DLG2 -2.35 -5.45 -2.27 -2.51
SYT1 -3.31 -5.31 -2.66 -2.77
GRIN2A -2.74 -5.37 -1.83 -2.73
GABRA2 -3.30 -5.95 -1.83 -2.44
PCLO -2.47 -6.77 -1.81 -3.02
GNG3 -2.60 -5.25 -1.79 -3.66
GABRG1 -3.71 -5.07 -3.30 -4.38
HCN1 -3.46 -6.59 -1.76 -3.39
SCN2A -2.51 -5.63 -1.71 -2.22
ST18 -2.30 -7.19 -2.16 -2.04
SLC24A2 -2.05 -4.86 -4.74 -3.74
GABRB1 -1.89 -5.04 -1.65 -2.17
5
0
-5
GSE49810
GSE50161
GSE65626
GSE90604

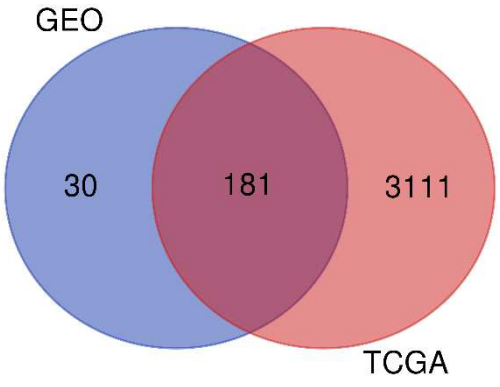

C
GEO
30
181
3111
TCGA

Figure 2

D

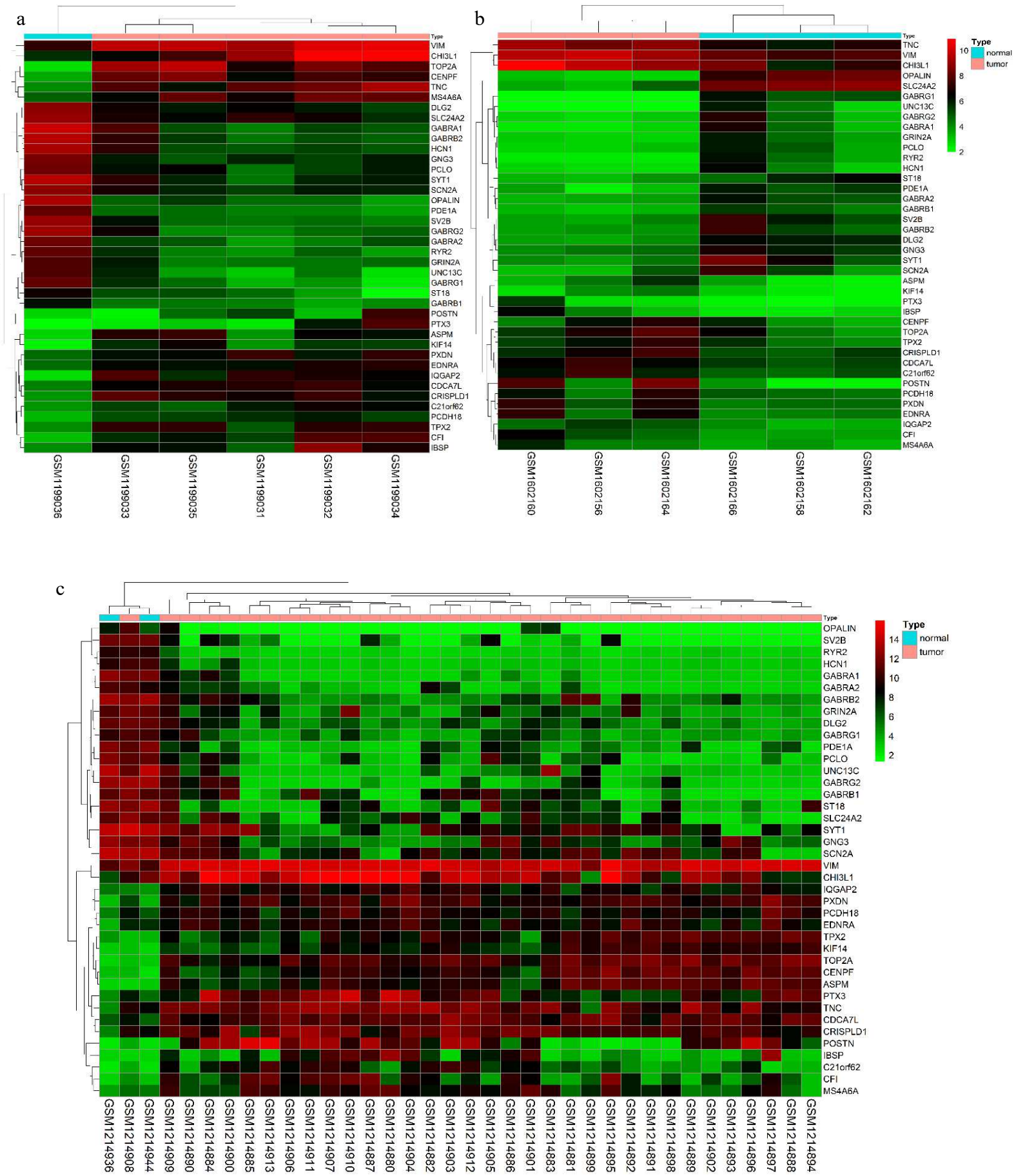
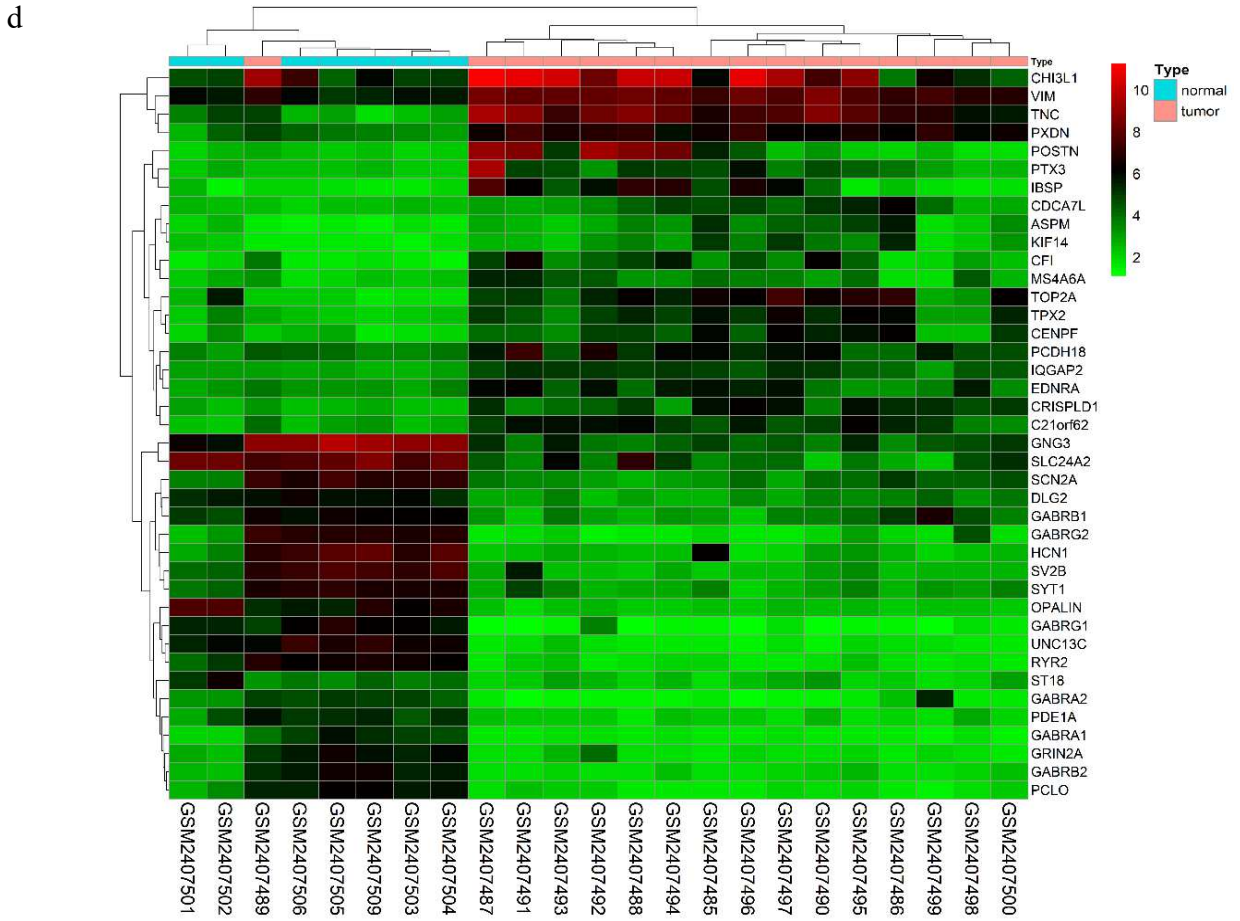

a
Type
VIM
CHI3L1
TOP2A
CENPF
TNC
MS4A6A
DLG2
SLC24A2
GABRA1
GABRB2
HCN1
GNG3
PCLO
SYT1
SCN2A
OPALIN
PDE1A
SV2B
GABRG2
GABRA2
RYR2
GRIN2A
UNC13C
GABRG1
ST18
GABRB1
POSTN
PTX3
ASPM
KIF14
PXDN
EDNRA
IQGAP2
CDCA7L
CRISPLD1
C21orf62
PCDH18
TPX2
CFI
IBSP
GSM1199036
GSM1199033
GSM1199035
GSM1199031
GSM1199032
GSM1199034
b
Type
TNC
VIM
CHI3L1
OPALIN
SLC24A2
GABRG1
UNC13C
GABRG2
GABRA1
GRIN2A
PCLO
RYR2
HCN1
ST18
PDE1A
GABRA2
GABRB1
SV2B
GABRB2
DLG2
GNG3
SYT1
SCN2A
ASPM
KIF14
PTX3
IBSP
CENPF
TOP2A
TPX2
CRISPLD1
CDCA7L
C21orf62
POSTN
PCDH18
PXDN
EDNRA
IQGAP2
CFI
MS4A6A
Type
normal
tumor
10
8
6
4
2
GSM1602160
GSM1602156
GSM1602164
GSM1602166
GSM1602158
GSM1602162
c
Type
OPALIN
SV2B
RYR2
HCN1
GABRA1
GABRA2
GABRB2
GRIN2A
DLG2
GABRG1
PDE1A
PCLO
UNC13C
GABRG2
GABRB1
ST18
SLC24A2
SYT1
GNG3
SCN2A
VIM
CHI3L1
IQGAP2
PXDN
PCDH18
EDNRA
TPX2
KIF14
TOP2A
CENPF
ASPM
PTX3
TNC
CDCA7L
CRISPLD1
POSTN
IBSP
C21orf62
CFI
MS4A6A
Type
normal
tumor
14
12
10
8
6
4
2
GSM1214936
GSM1214908
GSM1214944
GSM1214909
GSM1214890
GSM1214884
GSM1214900
GSM1214885
GSM1214913
GSM1214906
GSM1214911
GSM1214907
GSM1214910
GSM1214887
GSM1214880
GSM1214904
GSM1214882
GSM1214903
GSM1214912
GSM1214905
GSM1214886
GSM1214901
GSM1214883
GSM1214881
GSM1214899
GSM1214895
GSM1214892
GSM1214891
GSM1214898
GSM1214889
GSM1214902
GSM1214893
GSM1214896
GSM1214897
GSM1214888
GSM1214894

Figure 2



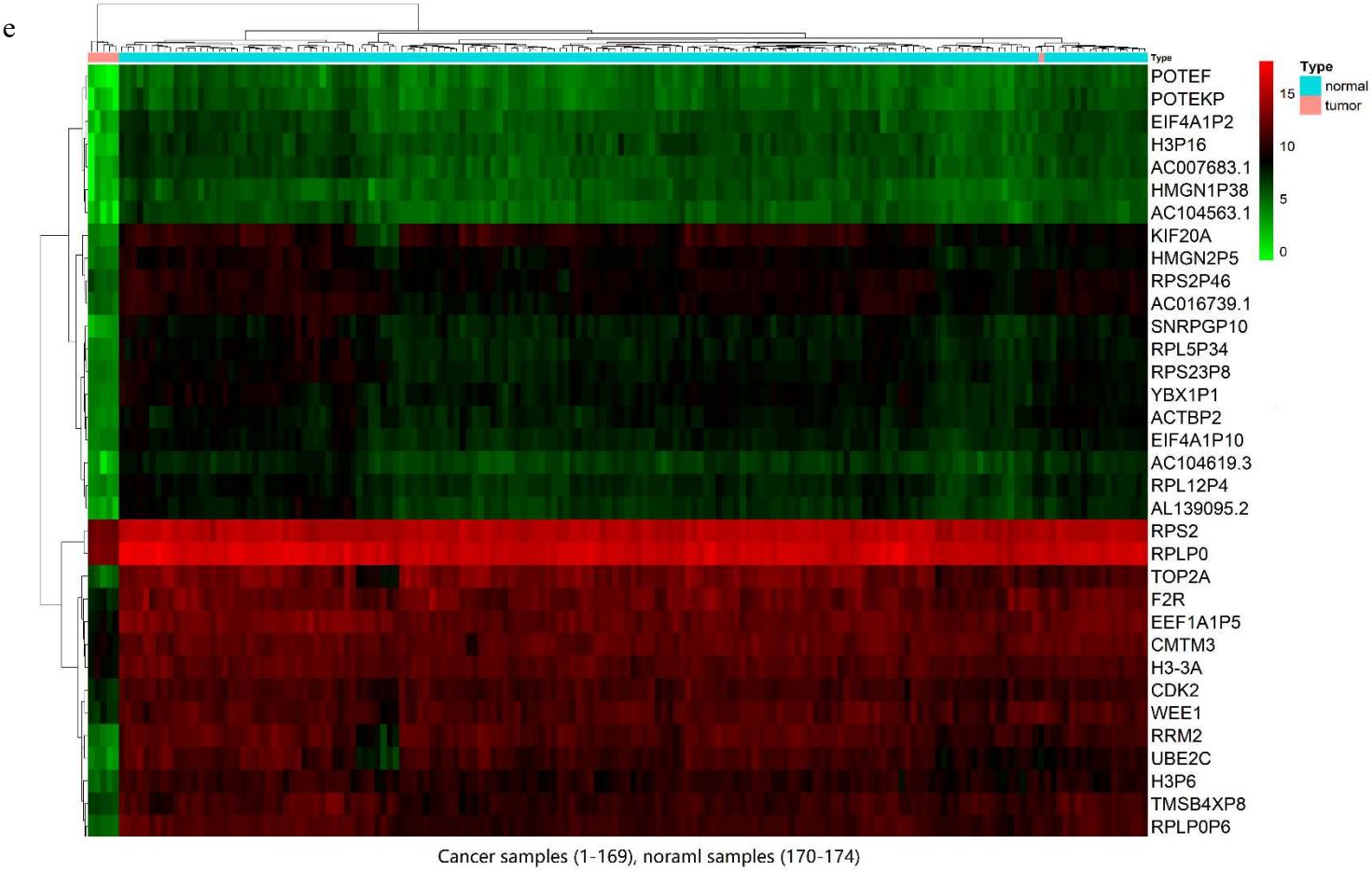
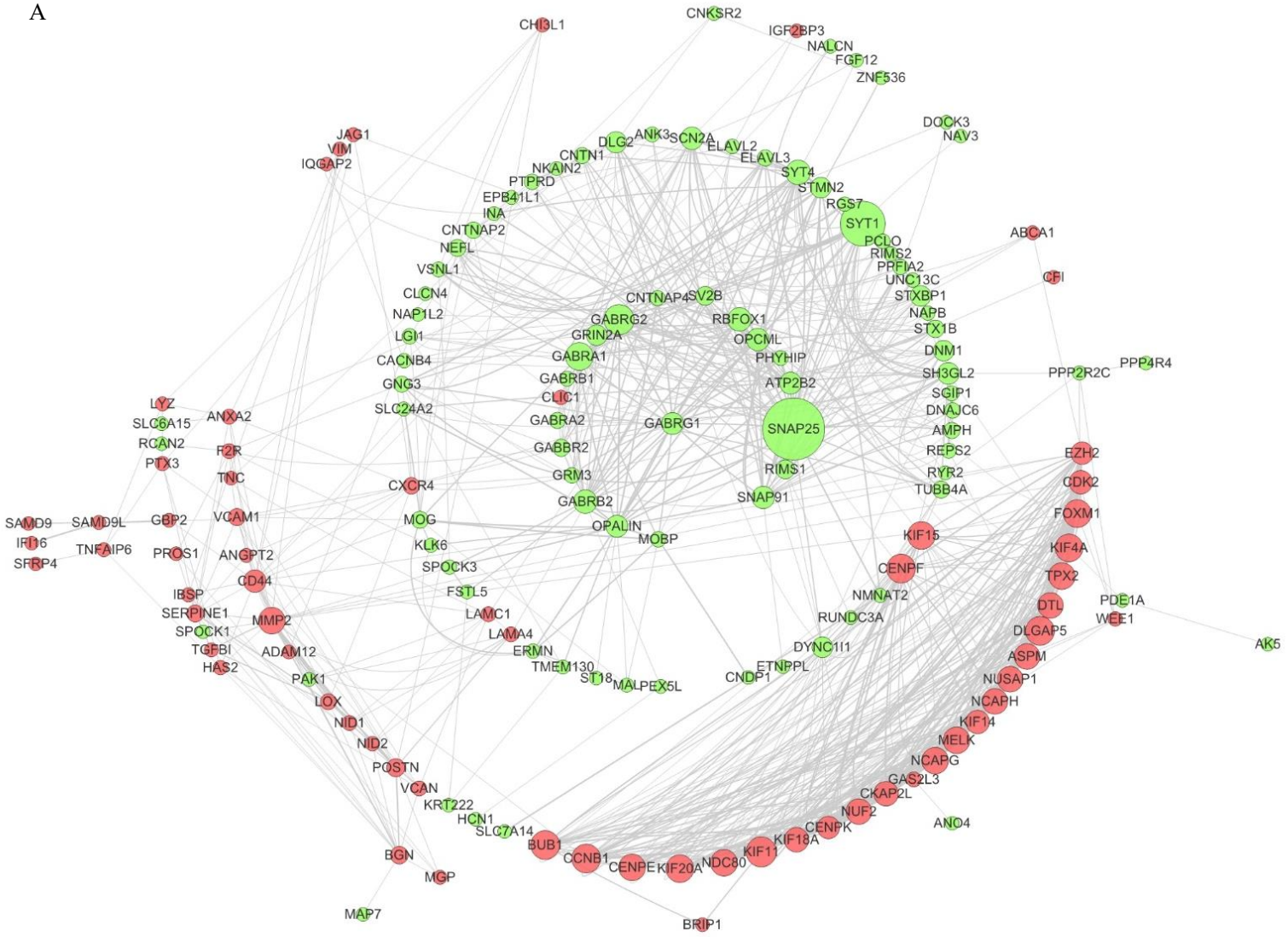


Fig. 2. A. Volcano plot of microarray. B. The expression level of core proteins of GBM in four expression arrays (most significance, top 20). C. Venn diagram of DEGs common to GEO and TCGA datasets. D. The top 20 protein expression levels of each samples in GSE49810 (a), GSE65625 (b), GSE50161 (c), GSE90604 (d); and the top 35 of TCGA (e) respectively.

Figure 3

A


CHI3L1
CNKSR2
IGF2BP3
NALCN
FGF12
ZNF536
DOCK3
NAV3
JAG1
VIM
IQGAP2
DLG2
ANK3
SCN2A
ELAVL2
ELAVL3
CNTN1
NKAIN2
SYT4
STMN2
PTPRD
EPB41L1
INA
RGS7
SYT1
CNTNAP2
NEFL
PCLO
RIMS2
PRRT1A2
UNC13C
ABCA1
CFI
VSNL1
CLCN4
CNTNAP4
SV2B
STXBP1
NAP1L2
GABRG2
RBFOX1
NAPB
STX1B
LGI1
GRIN2A
OPCML
DNM1
CACNB4
GABRA1
PHYHIP
PPP2R2C
PPP4R4
GNG3
GABRB1
ATP2B2
SH3GL2
SGIP1
SLC24A2
CLIC1
DNAJC6
LYZ
ANXA2
GABRA2
GABRG1
SNAP25
AMPH
SLC6A15
RCAN2
GABBR2
REPS2
EZH2
P2R
PTX3
RIMS1
RYR2
CDK2
TNC
CXCR4
GRM3
TUBB4A
SNAP91
GABRB2
FOXM1
SAMD9
SAMD9L
GBP2
VCAM1
MOG
OPALIN
KIF15
IFI6
TNFAIP6
PROS1
ANGPT2
KLK6
MOBP
KIF4A
SFRP4
SPOCK3
CENPF
CD44
TPX2
IBSP
FSTL5
NMNAT2
PDE1A
SERPINE1
LAMC1
RUNDC3A
DTL
WEE1
MMP2
SPOCK1
LAMA4
DLGAP5
AK5
TGFBI
ERMN
DYNC1I1
ASPM
ADAM12
TMEM130
ETNPPL
HAS2
ST18
MALPEX5L
CNDP1
NUSAP1
PAK1
NCAPH
LOX
KIF14
NID1
MELK
NID2
NCAPG
POSTN
GAS2L3
VCAN
CKAP2L
KRT222
NUF2
HCN1
ANO4
CENPK
SLC7A14
KIF18A
BUB1
KIF11
BGN
CCNB1
NDC80
CENPE
KIF20A
MGP
MAP7
BRIP1

B

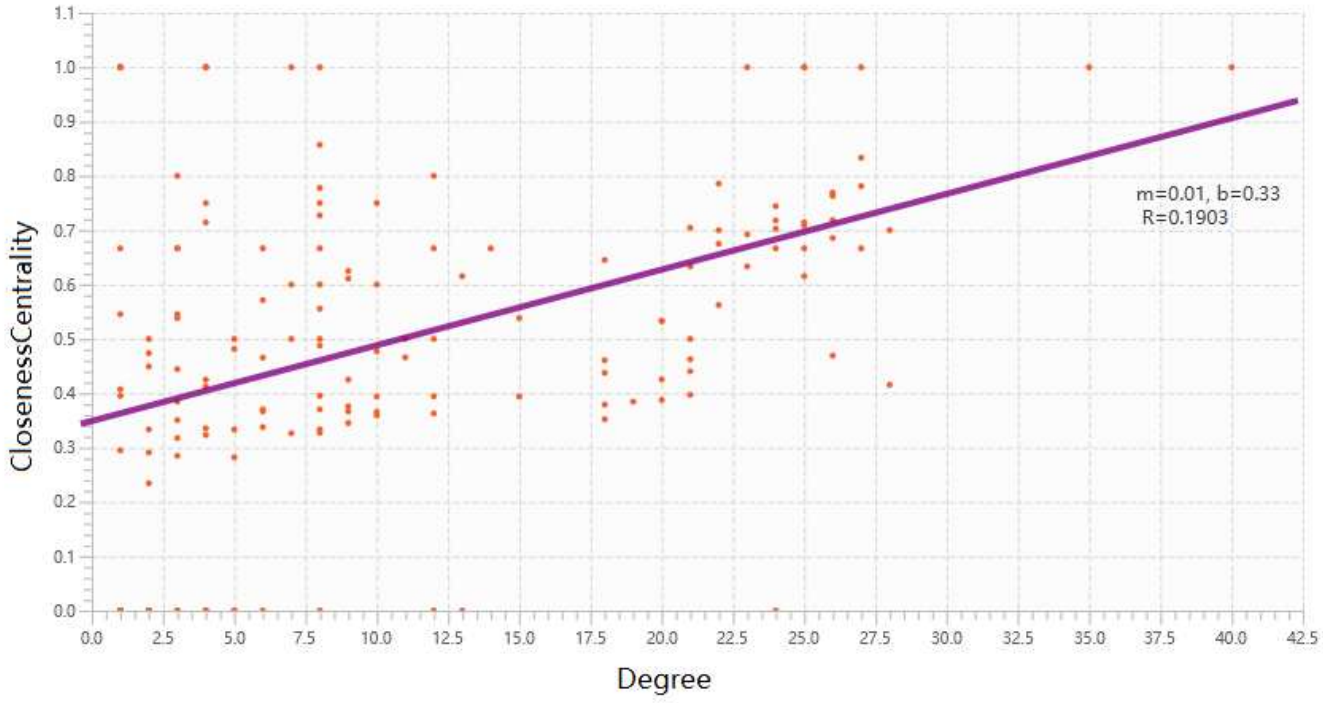
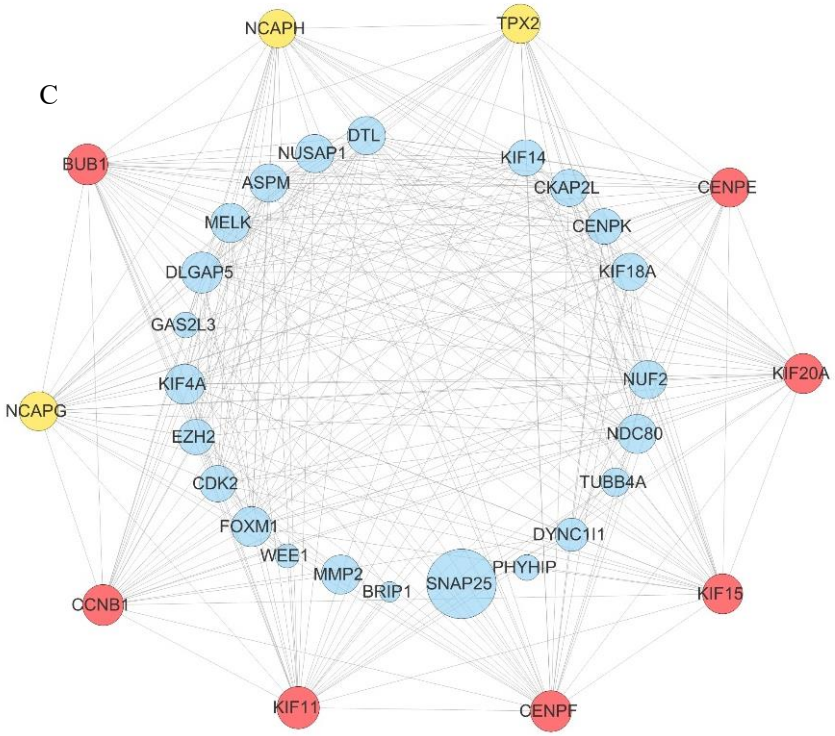
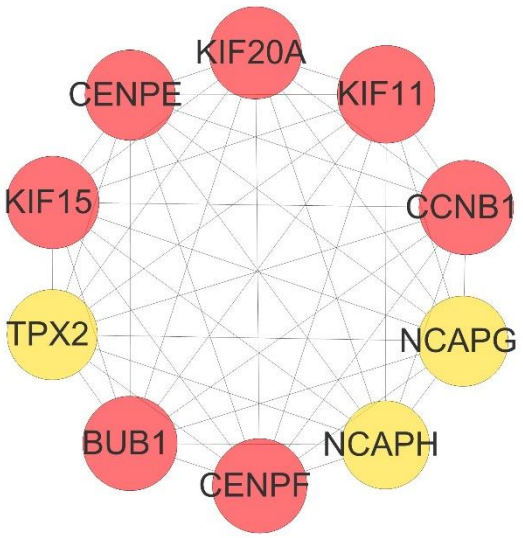

m=0.01, b=0.33
R=0.1903
ClosenessCentrality
Degree
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5
25.0
27.5
30.0
32.5
35.0
37.5
40.0
42.5

Figure 3


C
NCAPH
TPX2
BUB1
NUSAP1
DTL
ASPM
KIF14
CKAP2L
CENPE
MELK
CENPK
DLGAP5
KIF18A
GAS2L3
KIF4A
NUF2
KIF20A
NCAPG
EZH2
NDC80
CDK2
TUBB4A
FOXM1
DYNC1I1
WEE1
MMP2
PHYHIP
BRIP1
SNAP25
CCNB1
KIF15
KIF11
CENPF


D
KIF20A
CENPE
KIF11
KIF15
CCNB1
TPX2
NCAPG
BUB1
NCAPH
CENPF

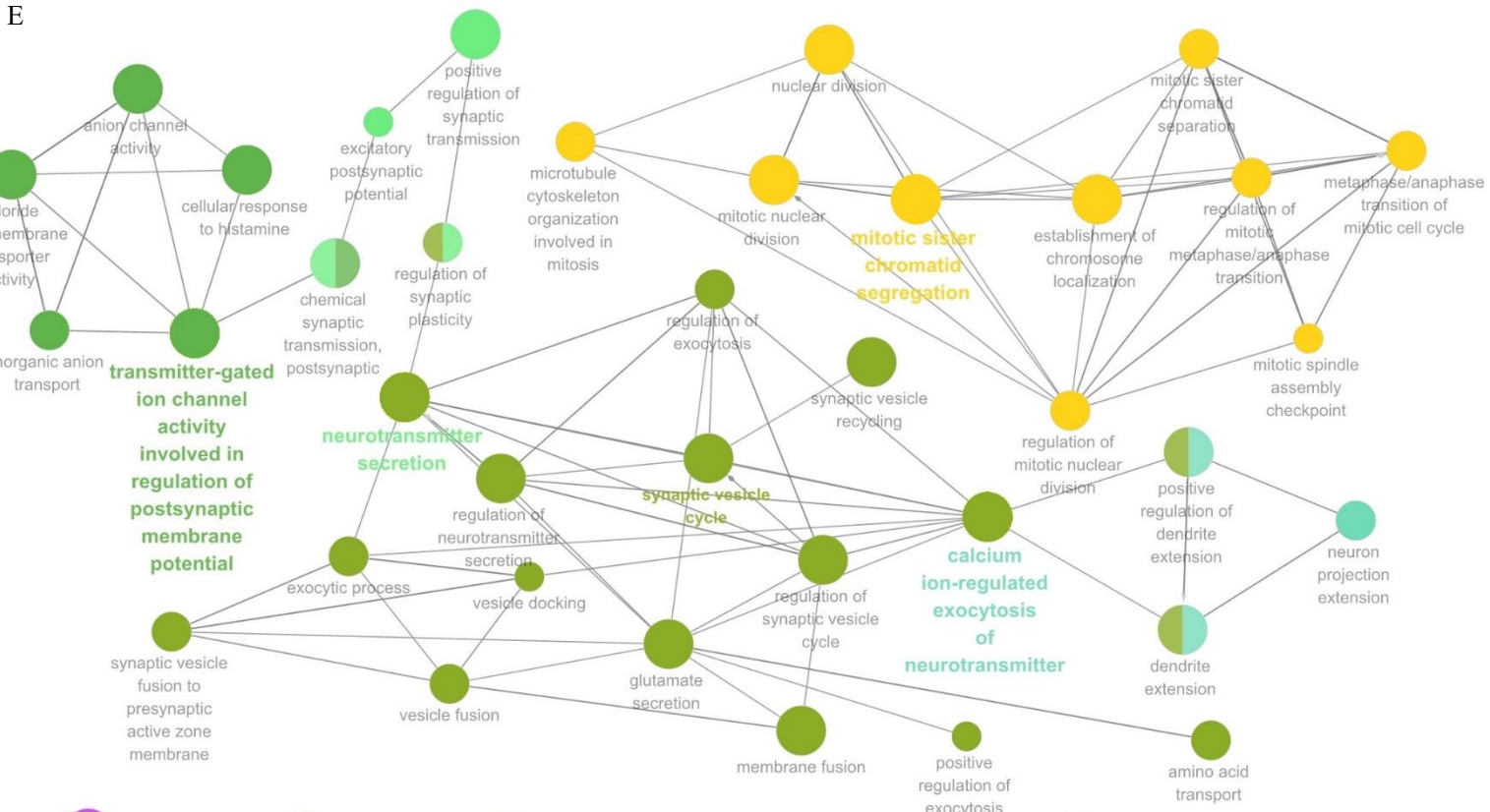

E
positive regulation of synaptic transmission
anion channel activity
excitatory postsynaptic potential
chloride transmembrane transporter activity
cellular response to histamine
chemical synaptic transmission, postsynaptic
regulation of synaptic plasticity
inorganic anion transport
transmitter-gated ion channel activity involved in regulation of postsynaptic membrane potential
nuclear division
microtubule cytoskeleton organization involved in mitosis
mitotic sister chromatid separation
mitotic nuclear division
mitotic sister chromatid segregation
establishment of chromosome localization
regulation of mitotic metaphase/anaphase transition
metaphase/anaphase transition of mitotic cell cycle
mitotic spindle assembly checkpoint
regulation of exocytosis
synaptic vesicle recycling
neurotransmitter secretion
regulation of mitotic nuclear division
synaptic vesicle cycle
regulation of neurotransmitter secretion
positive regulation of dendrite extension
neuron projection extension
exocytic process
vesicle docking
regulation of synaptic vesicle cycle
calcium ion-regulated exocytosis of neurotransmitter
synaptic vesicle fusion to presynaptic active zone membrane
glutamate secretion
dendrite extension
vesicle fusion
membrane fusion
positive regulation of exocytosis
amino acid transport
chemical synaptic transmission
trans-synaptic signaling
structural constituent of postsynapse
structural constituent of synapse
regulation of endocytosis
positive regulation of endocytosis
regulation of receptor-mediated endocytosis
protein localization to synapse
receptor localization to synapse
protein localization to postsynaptic membrane
regulation of blood coagulation
positive regulation of wound healing
regulation of microtubule polymerization or depolymerization
negative regulation of protein depolymerization
positive regulation of neuron projection development
extracellular matrix organization
basement membrane organization
smooth muscle cell migration
long-term synaptic potentiation
extracellular matrix assembly

Figure 3

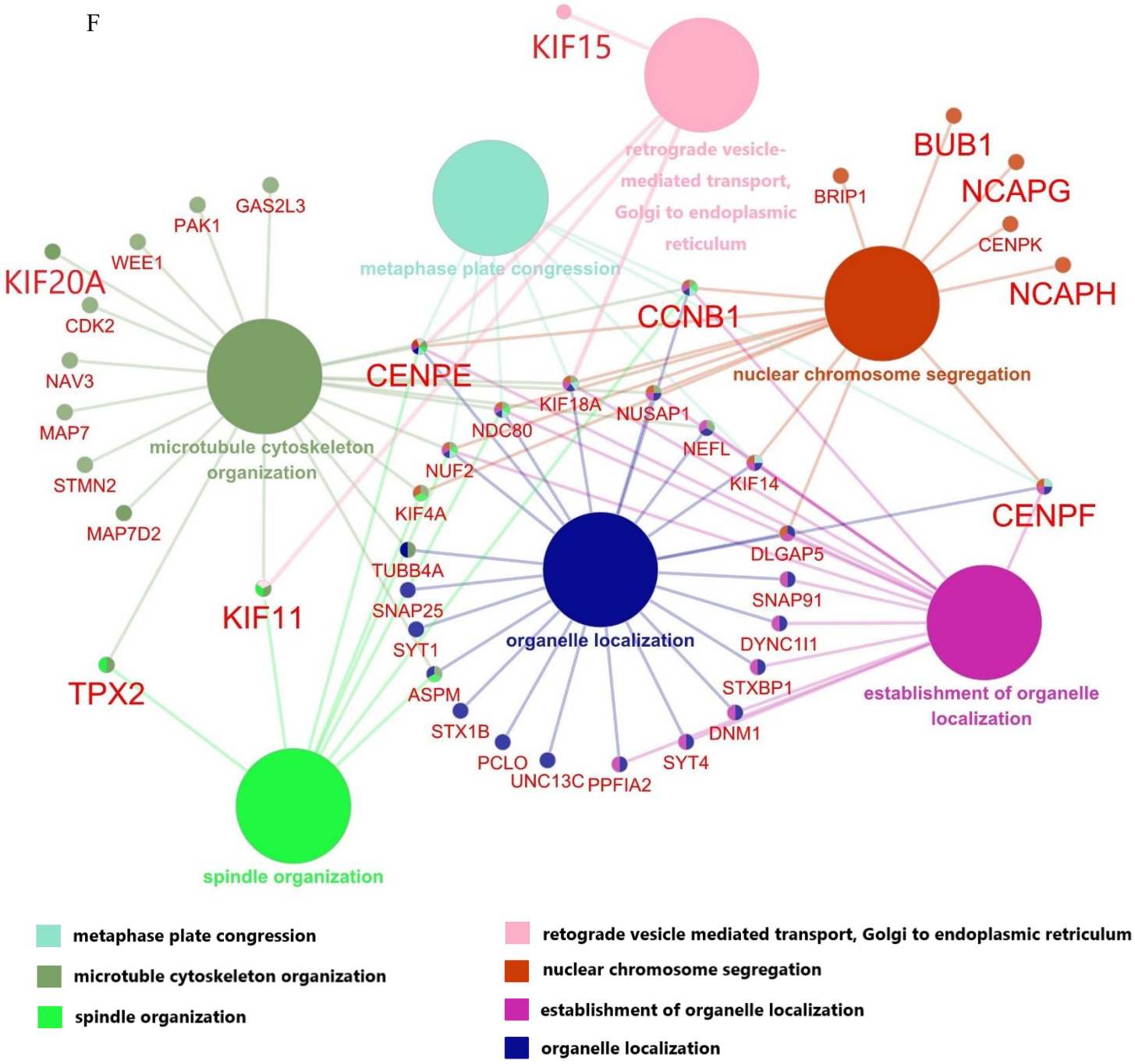


Fig. 3. A. The core proteins of GBM associated regulatory network. GBM: Glioblastoma. Red diamond represents up-regulated protein, Green diamond indicated down-regulated protein. Grey lines link target genes. B. The relationship between closeness centrality and degree of scatter plot of overlapped DEGs, R square=0.1903, and correlation = 0.33. C. Selected hub genes of GBM by MCC method in cytohubba package and its relative protein network. D. The hub genes of GBM. E. The participated pathway of 181 DEGs and relative networks performed by using Cluego plugin. F. The network of connecting ten hub genes with its relative pathway, and the colored ball refers to a pathway the gene node participated in.

Figure 4

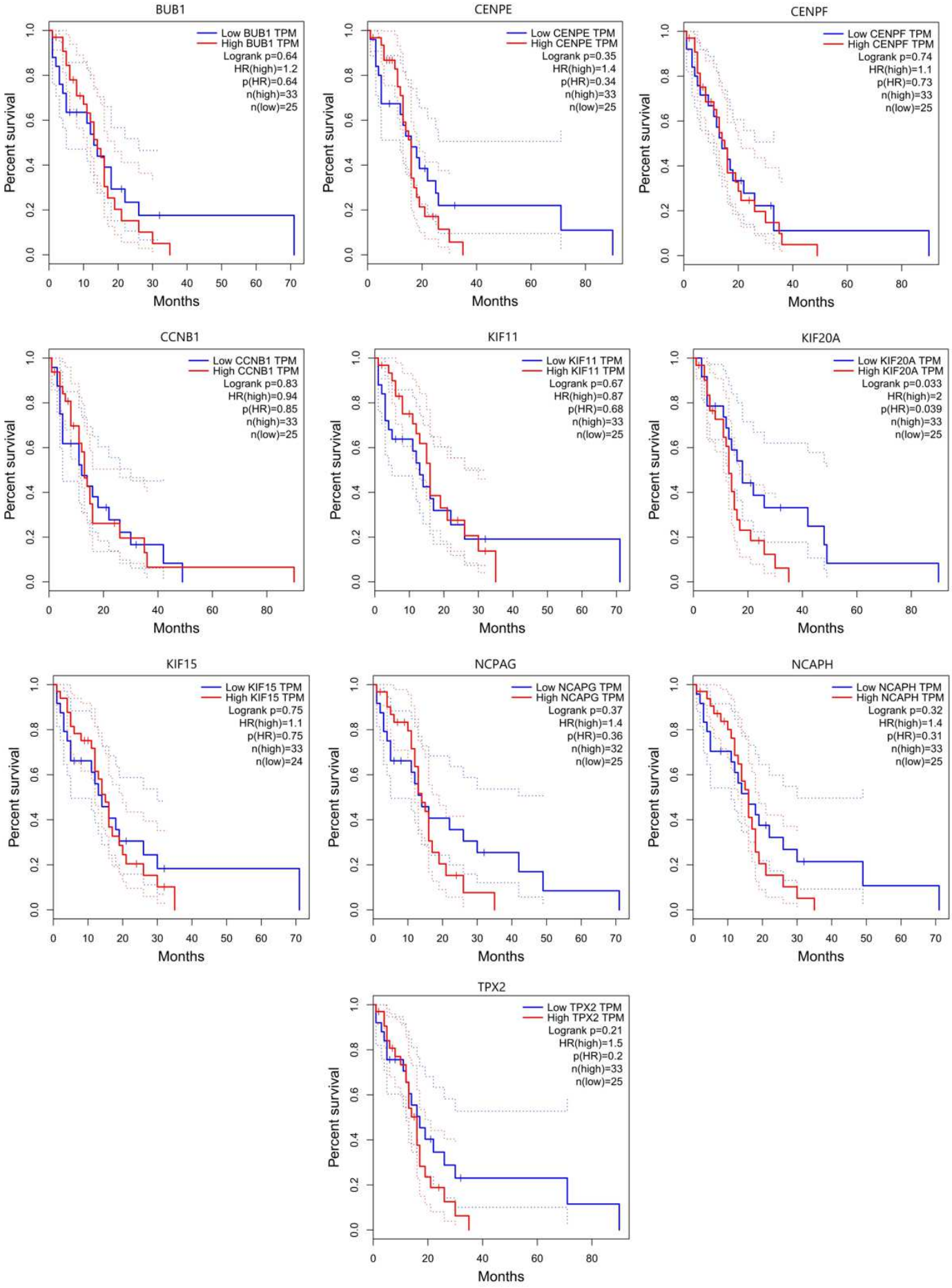


Fig. 4. Survival curves of ten hub genes expressed in GBM.

Figure 5

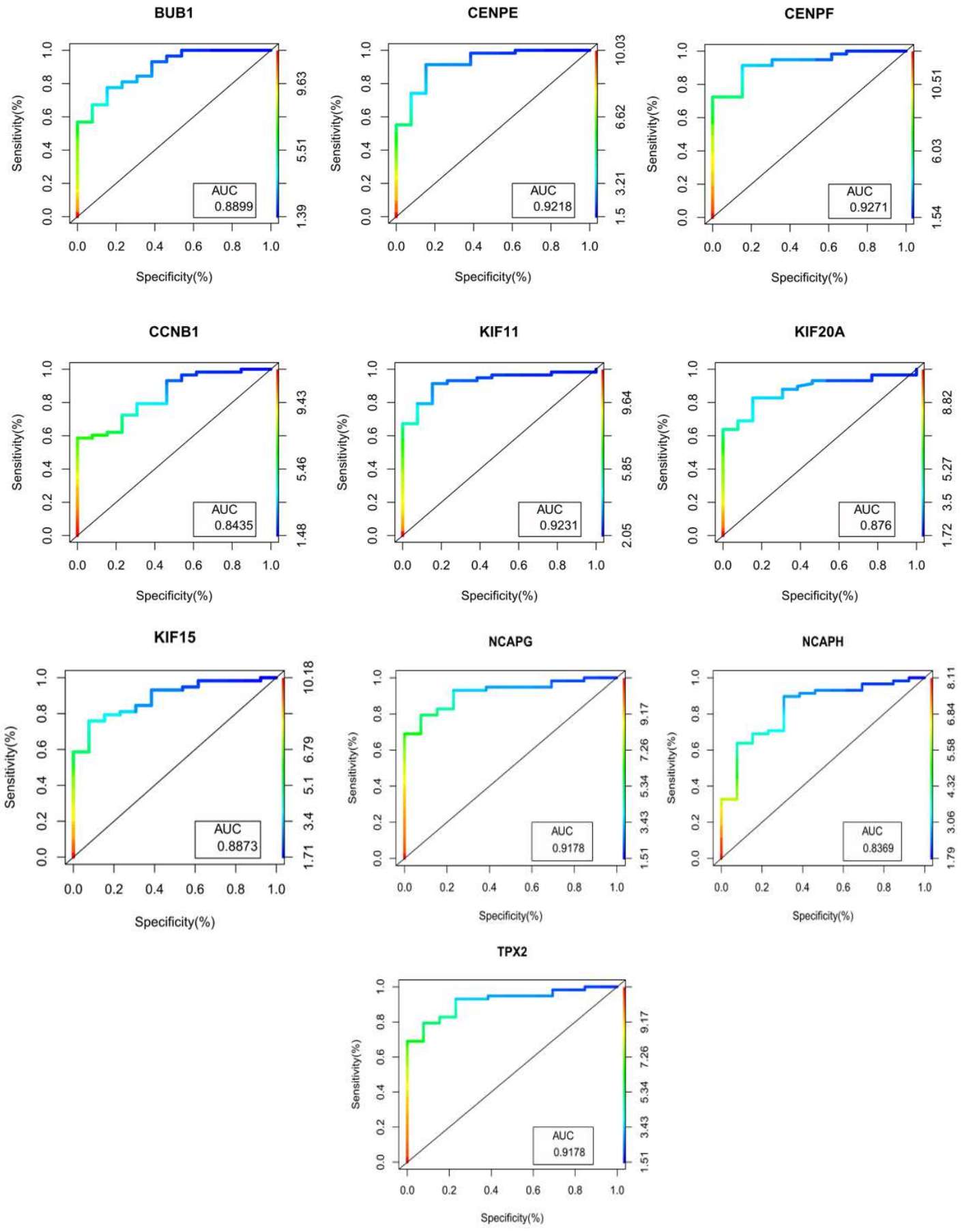


Fig. 5. The receiver operating characteristic curves of ten hub genes.

Figure 6

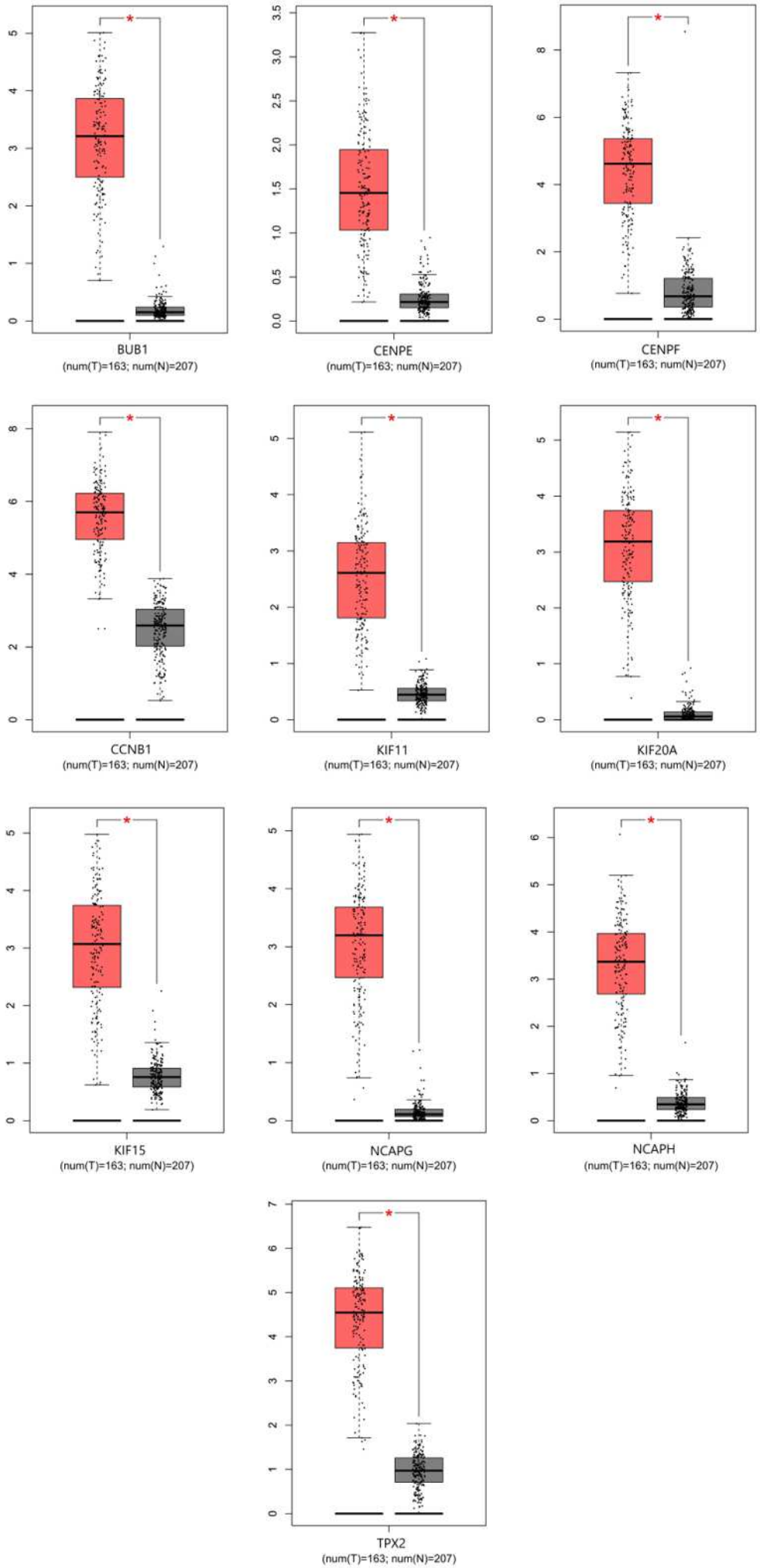


Fig. 6. Expression of ten hub genes using boxplot. Compared with the normal group, *p < 0.05.

Figures

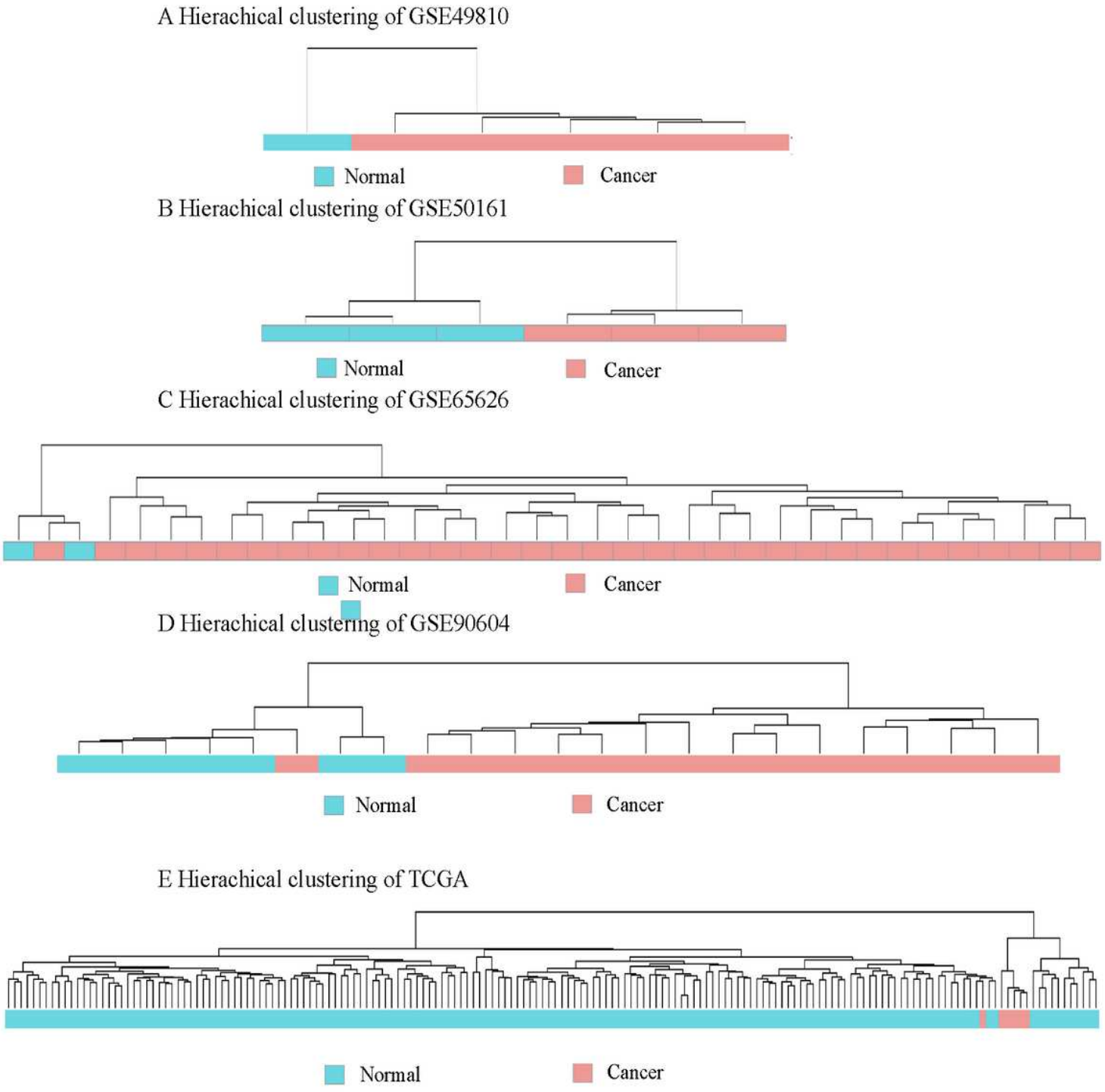


Figure 1

Hierarchial clustering dendrograms for GSE49810 (A), GSE50161 (B), GSE65626 (C), GSE90604 (D) and TCGA (E). Blue and pink represent normal and GBM tumor samples, respectively.

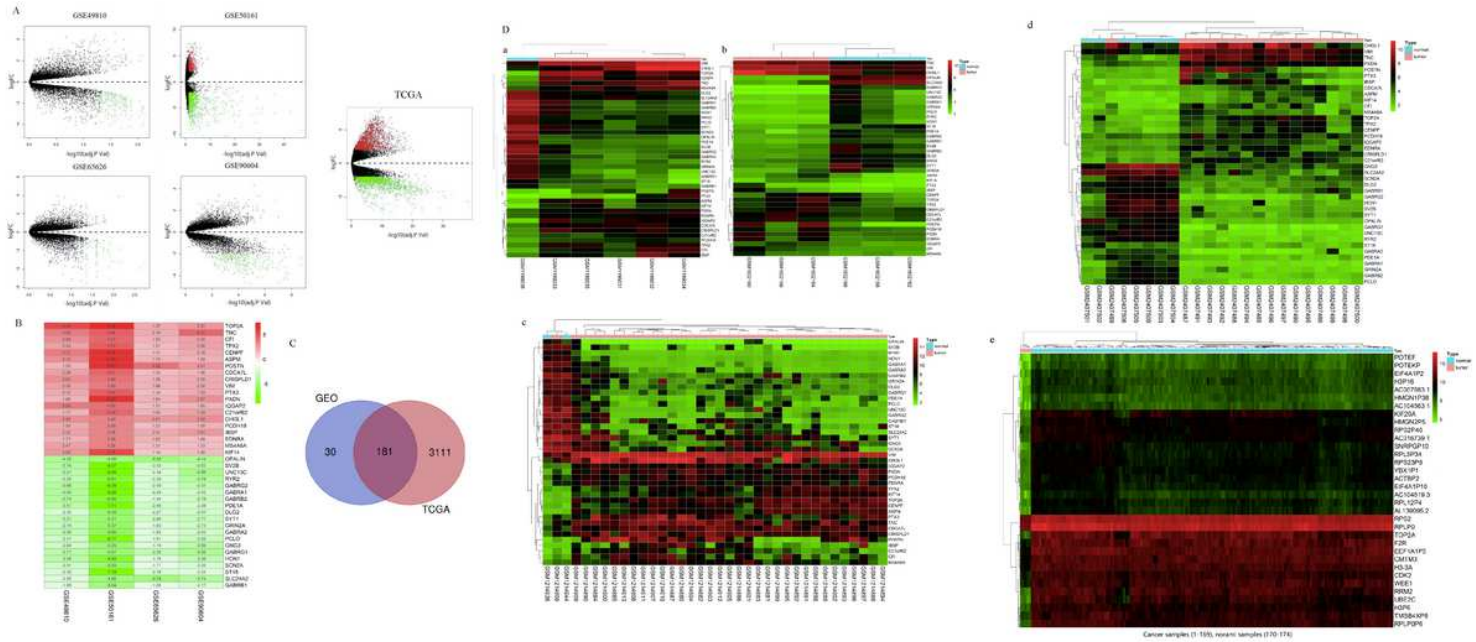


Figure 2

A. Volcano plot of microarray. B. The expression level of core proteins of GBM in four expression arrays (most significance, top 20). C. Venn diagram of DEGs common to GEO and TCGA datasets. D. The top 20 protein expression levels of each samples in GSE49810 (a), GSE65625 (b), GSE50161 (c), GSE90604 (d); and the top 35 of TCGA (e) respectively.

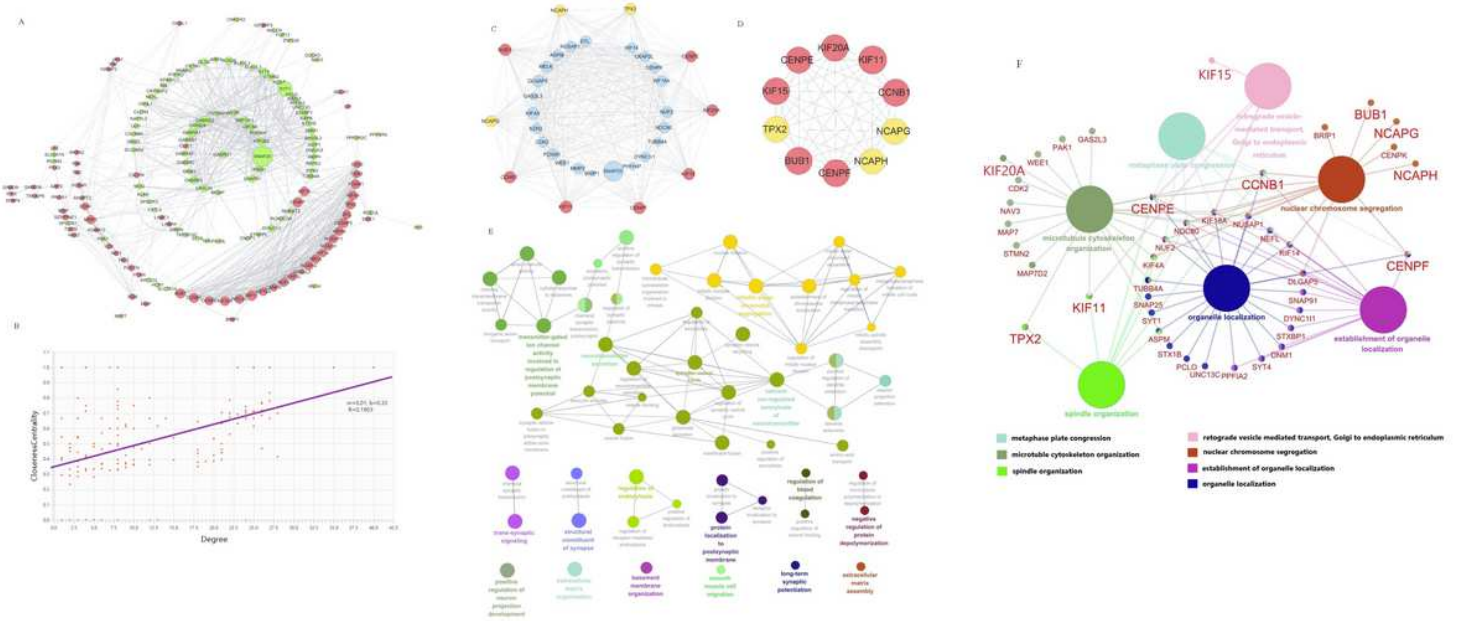


Figure 3

A. The core proteins of GBM associated regulatory network. GBM: Glioblastoma. Red diamond represents up-regulated protein, Green diamond indicated down-regulated protein. Grey lines link target genes. B. The relationship between closeness centrality and degree of scatter plot of overlapped DEGs, R

square=0.1903, and correlation = 0.33. C. Selected hub genes of GBM by MCC method in cytohubba package and its relative protein network. D. The hub genes of GBM. E. The participated pathway of 181 DEGs and relative networks performed by using Cluego plugin. F. The network of connecting ten hub genes with its relative pathway, and the colored ball refers to a pathway the gene node participated in.

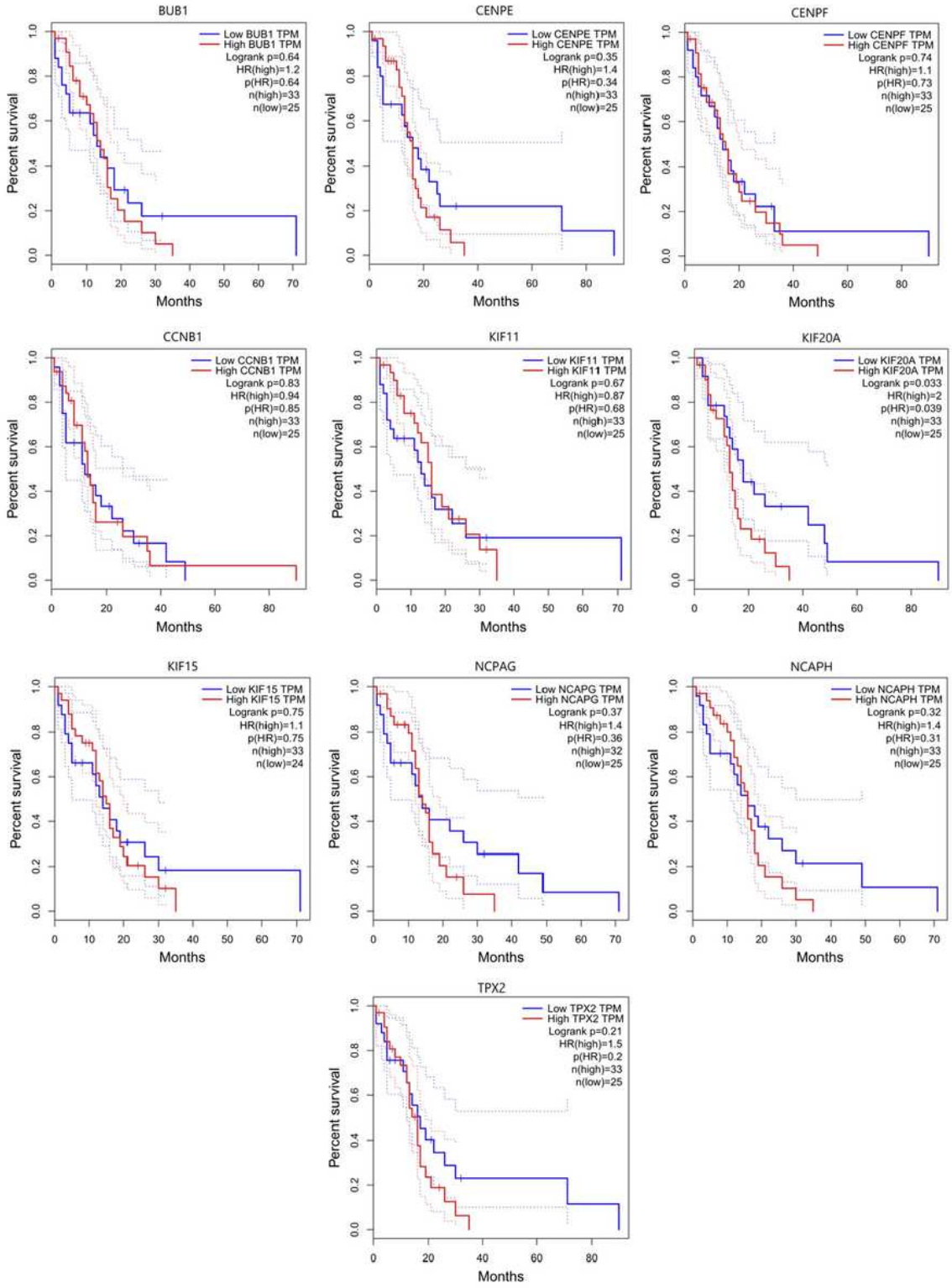


Figure 4

Survival curves of ten hub genes expressed in GBM.

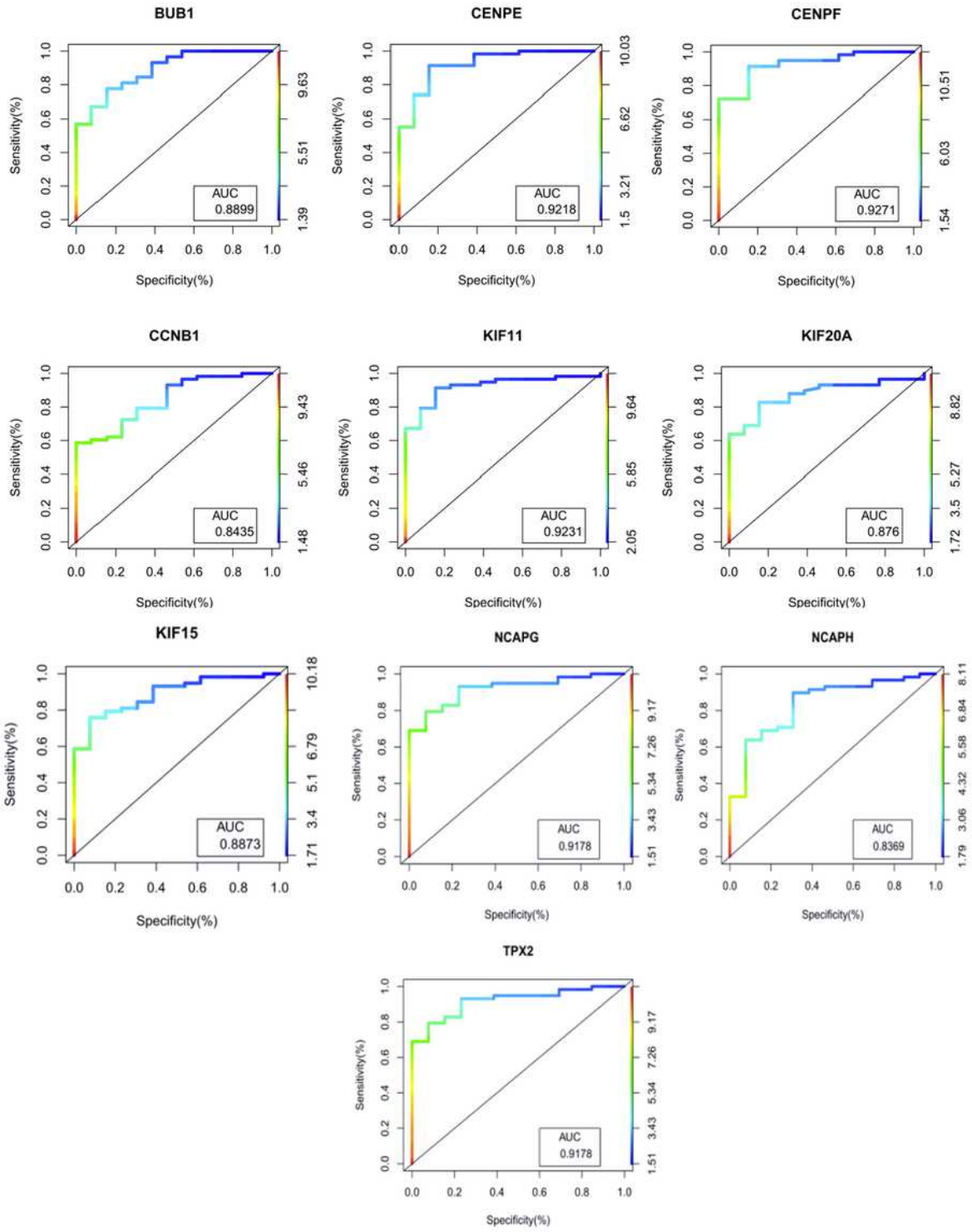


Figure 5

The receiver operating characteristic curves of ten hub genes.

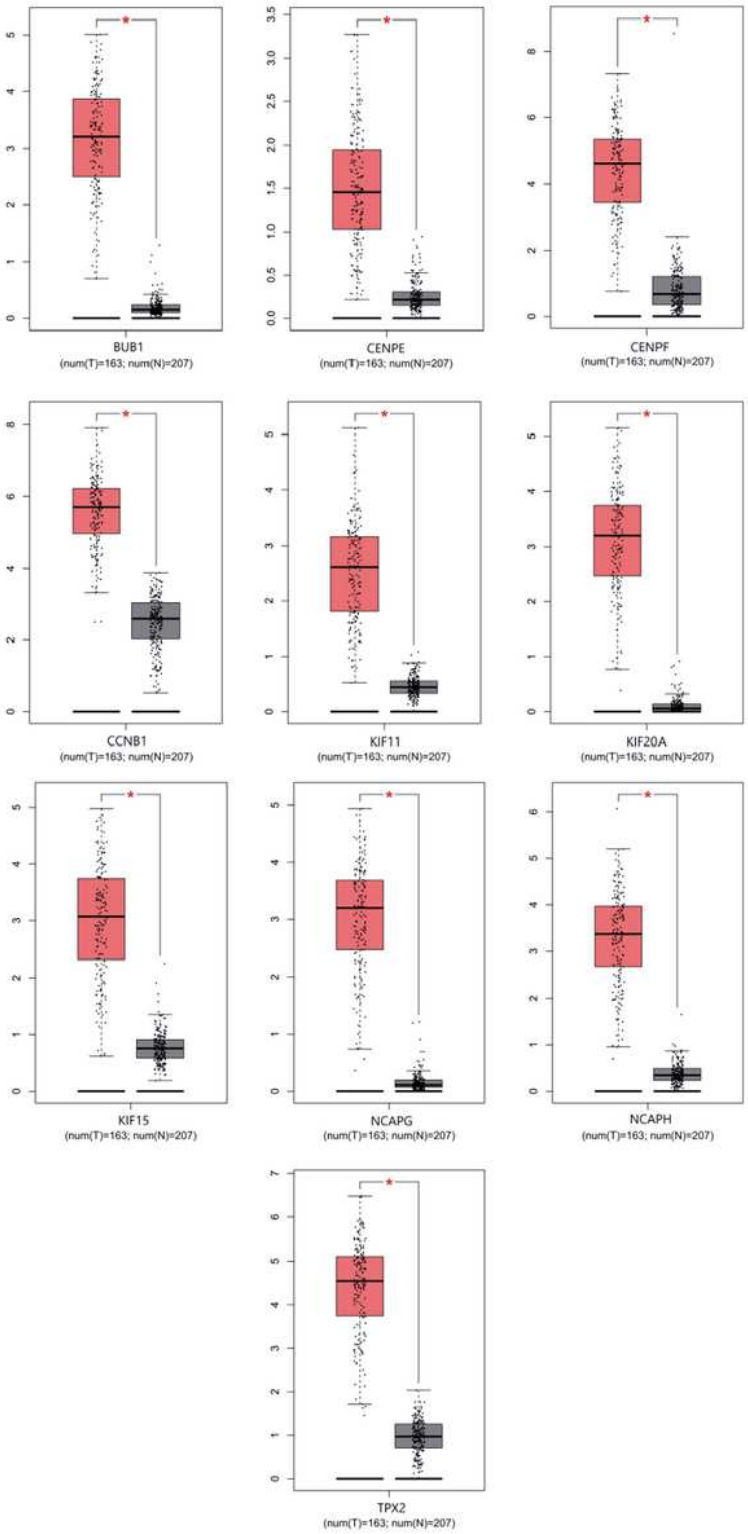


Figure 6

Expression of ten hub genes using boxplot. Compared with the normal group, *p < 0.05. The number of total samples is 380, including 163 tumors and 207 normal tissues.