

Gene Expression Analysis Reveals Novel Gene Prognostic Signatures Involved in Medulloblastoma Subgroups

Fangcheng Li (✉ sjwklfc@126.com)

Guangzhou Women and Children's Medical Center <https://orcid.org/0000-0003-0082-0095>

Hong-Yao Yuan

Guangzhou Women and Children's Medical Center

Cheng Chen

Guangzhou Women and Children's Medical Center

Jun-Ping Pan

Guangzhou Women and Children's Medical Center

Xin-Ke Xu

Guangzhou Women and Children's Medical Center

Cui-Yu Lin

Guangzhou Women and Children's Medical Center

Wei Chen

Guangzhou Women and Children's Medical Center

Yang Li

Guangzhou Women and Children's Medical Center

Jin-Rong Lin

Guangzhou Women and Children's Medical Center

Jun-Liang Li

Guangzhou Women and Children's Medical Center

Research Article

Keywords: Genomic analyses, Medulloblastoma, WGCNA, Prognostic Signatures

Posted Date: March 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-328561/v1>

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

[Read Full License](#)

Abstract

Purpose

Medulloblastoma is a malignant childhood tumor with four molecular subtypes: WNT, SHH, G3, and G4. The prognosis of these four molecular subtypes is different. WNT has the best prognosis, followed by SHH, and G3 and G4 subtypes have the worst prognosis. This study aimed to identify various molecular subtypes of medulloblastoma can independently predict the prognosis of patients and provide specific treatment way for them.

Methods

Based on the data in the GSE37418 dataset, the WGCNA method was used to find the genes most related to these molecular subtypes, and then the top ten hub genes in each subtype were found through the cytohubba plug-in of cytoscape. GO pathway enrichment of four interested modules was used, and then GSE85217 was used for clinical trials in single-factor Cox regression analysis .

Results

The information was subjected to single-factor Cox regression analysis, and twelve hub genes with the most significant prognostic effects on medulloblastoma were found, and then these genes were subjected to multi-factor Cox regression analysis on each molecular subtype, and finally GNG3 was determined .The combination of CALCB and GCGR can predict the development of SHH well ($p=0.0011$, $AUC=0.734$), SOCS3 and HOXC10 can better predict the development of G4 ($p=0.044$, $AUC=0.618$), and the combination of ADCY8 and LHX3 can predict G3 Development ($p=0.034$, $AUC=0.675$).

Conclusion

This report showed a possible evidence that OS-related features of various molecular subtypes of medulloblastoma can independently predict the prognosis of patients with each subtype of medulloblastoma, and provide new therapeutic targets for them.

Background

Medulloblastoma (WHO grade IV) is a highly malignant brain tumor arising in the posterior fossa in children, accounts for approximately 20% of all childhood brain tumors.(1) (2) It is an embryonic lesion and is thought to be caused by the progenitor cell population that exists during early brain development. Medulloblastoma is currently undergoing the most safe resection, chemotherapy, and skull radiation. Despite this radical multimodal therapy, about 30% of patients eventually die of the disease, and survivors can cope with long-term side effects of treatment, which will have a significant impact on their quality of life. Genomics has now greatly improved our understanding of medulloblastoma. Obviously, medulloblastoma contains at least four different molecular subtypes: WNT, SHH, group 3 and group 4. (3),(4)

These four subgroups have unique transcription profiles, copy number aberrations, somatic mutations, and clinical outcomes. Among them, WNT and SHH medulloblastoma can be clearly identified and separated in most transcriptome and methylation analysis studies, with little overlap with other molecular subtypes. (4) Different subtypes have different prognosis. The WNT subgroup has a good prognosis (5-year survival rate of pediatric patients > 95%), rarely metastasizes at the time of diagnosis (5%-10% of cases), and rarely recurs (5). WNT medulloblastoma is believed to originate from progenitor cells with the rhomboid lip in the developing lower brain stem.

(6,7) The SHH subgroup has a balanced sex ratio and bimodal age incidence, and is most common in infants and adults. About 30% of medulloblastomas are classified as SHH tumors. (8,9) The third group of medulloblastoma is mainly male and almost only occurs in infants. (9) Tumors in the Group3 subgroup account for approximately 25% of the overall medulloblastoma cases, and have a high metastasis rate (40%-45%) and the worst survival outcome (less than 60% at 5 years) at the time of diagnosis.

(9) Medulloblastoma in group 4 accounts for 35%-40% of all medulloblastoma diagnoses. It usually occurs in children and adolescents, more often in men (3:1 sex ratio). This subgroup also has frequent metastases (35%-40% of diagnoses), the survival outcome is moderate and its recurrence often occurs late. (9-11)

In order to explore dynamic changes and unbiasedly understand the system-level characteristics of the four types of medulloblastoma at the molecular level, we use a network method called Weighted Gene Correlation Network Analysis (WGCNA) for analysis. It can group functionally related genes into modules (12). These modules are constructed by calculating the correlation network analysis of a large number of high-dimensional datasets based on the correlation between genes (because they have similar expression profiles) and can be related to the phenotype (12). WGCNA's R software package has been successfully applied to various biological data analysis, such as cancer data (13)⁽¹⁴⁾, and brain imaging data analysis, and it can also be used to describe the correlation structure between gene expression profiles, image data, genetic marker data, proteomics data and other high-dimensional data (12).

The R package is available for free from Bioconductor. Although WGCNA analysis, we performed the WGCNA method by calculating the correlation between module features based on the GSE37418 public microarray dataset (including 73 samples). (15) This method identified co-expression modules that were highly related to SHH, G3, WNT, and G4, and these modules were significantly related to the molecular subtypes of normal medulloblastoma. In addition, using the clinical data of GSE85217 (16) (excluding some samples with obscure clinical data, 551 samples remain), we conducted a cox multivariate regression analysis and found that some gene combinations can predict the risk of each type well, and these have never been reported. This reveals that the various types of medulloblastoma may have deeper mechanisms.

Materials And Methods

Data Sources

First, the microarray dataset GSE37418 of RNA expression profile provided by Robinson G et al. (2012) was downloaded from NCBI Gene Expression Omnibus (GEO). (15) The dataset contains a total of 76 microarray spectra of pediatric medulloblastoma samples. The molecular subtypes of medulloblastoma are as follows: WNT, SHH, G4, G3. After removing the three samples of unknown origin, we also included 73 medulloblastoma samples and classified them according to the molecular subtype of the disease. In the GSE37418 dataset, the remaining samples include WNT (n = 8), SHH (n = 10), G4 (n = 39) and G3 (n = 16). In order to further verify our data, we use the GSE85217 data set for further survival analysis of the hub genes we selected. The GSE85217 dataset includes 763 medulloblastoma samples. The expression data of GSE85217 could be found in GEO, but its clinical data comes from relevant literature (16) (<https://ars.els-cdn.com/content/image/1-s2.0-S1535610817302015-mmc2.xlsx>). After excluding samples with unclear survival time, survival status, and gender in the GSE85217 dataset, there are 551 samples left. Including G4 (n=245), G3 (n=107), SHH (n=150), WNT (n=49). Other detailed experimental methods and instructions can be found in this original reference.

Establishment of weighted correlation network analysis

The WGCNA software package in R 4.02 was used to study the co-expression relationship based on gene expression. (12) Subsequent analysis of genes that usually represent noise with low expression or unchanged through variance filtering. To construct a weighted gene network, it is necessary to select the co-expression soft threshold ability β to increase the similarity to calculate the adjacency degree. The β value can be selected using the "pickSoftThreshold" function, which can then be used for network topology analysis and help selection based on the approximate scale-free topology. Appropriate soft threshold. Then convert the adjacency matrix to a topological overlap matrix (TOM). TOM can measure the network connectivity of a gene, which is defined as the sum of the adjacencies of the gene and all other genes, and then generates a network. In addition, in order to classify genes with similar expression patterns into gene modules, an average linkage hierarchical clustering is performed according to the TOM-based difference measure. Then, analyze the correlation between the gene expression of the module and the external clinical characteristics, and perform an in-model analysis to determine whether the gene expression is consistent with the clinical correlation.

Functional enrichment

In order to better understand the actual role of the genes in the WGCNA module, which are highly related to the molecular subtype of medulloblastoma, the "clusterProfiler" R package were performed the GO enrichment analysis and visualization. (17)

Protein-protein Interaction network creation and hub gene identification

For the genes of the key modules, pink, brown, green, and darkred modules, the protein-protein interaction analysis network generated by using the string database contains at least one protein subset, and each member forms a physical interaction. Then use cytoscape software was used for further drawing and analysis. (18) The mmc algorithm of the cytoHubba plug-in further identifies densely connected network

components.(19) Each module calculates the most critical 10 hub genes, and then performs further analysis.

Construction of the subtype of medulloblastoma related Prognostic Signature

We first obtained the prognosis of each molecular subtype of medulloblastoma through univariate Cox regression analysis, and then performed multiple Cox regression analysis and an optimized risk scoring model with a step function. The risk score is calculated as follows:

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

Among them, Coef is the coefficient, and x are the expression value of each selected gene. This formula is used to calculate the risk score of each medulloblastoma patient in each molecular subtype. Then, we performed Cox regression analysis to prove whether this feature is an independent prognostic factor for the molecular subtype of medulloblastoma.

Statistical Analysis

R software (version 4.0.2) was used for statistical analysis. The "survival" R package is used for cox regression analysis. Cox regression analysis (including univariate and multivariate) is used to screen the genes included in the risk scoring model. Medulloblastoma patients were divided into high-risk groups and low-risk groups according to the median risk score, and the OS of the patients was analyzed by Kaplan-Meier method and log-rank test. Use the "survival ROC" R package to perform ROC curve analysis. P<0.05 was considered statistically significant(20).

Results

Preprocessing of medulloblastoma dataset and construction of weighted gene co-expression network

The combined data set (GSE37418) contains a total of 73 samples [WNT (n = 8), SHH (n = 10), G4 (n = 39) and G3 (n = 16)] with clear molecular classification of medulloblastoma in this study(15). Then the same preprocessing is performed on the original data of each microarray data set for background correction and normalization. Firstly, the probes matching multiple genes are deleted from these datasets, and secondly, the average expression value of the genes measured by the multiple probes is calculated as the final expression value. Finally, we identified a total of 16,241 genes expressed. In addition, the construction of WGCNA also requires the best soft threshold capability(12), and the co-expression similarity is improved to calculate the adjacency. Therefore, we have conducted network topology analysis on various soft threshold power, so that WGCNA has relatively balanced scale independence and mean connectivity. As showed in Figure 1A and C, power 6 is the lowest power that selects the scale-free topology fitting index to reach 0.87 to generate a hierarchical clustering tree. Next, through dynamic tree cutting and merging dynamics, 28 different gene modules were generated in the hierarchical clustering tree from 73 samples, and each module was labeled differently (Figure 1C). Each tree branch constitutes

a module, and each leaf in the branch is a gene. The size of the modules varies from 56 (orange module) to 1783 (turquoise module) genes. In order to further quantify the co-expression similarity of the entire module, we calculated the characteristic gene adjacency of the entire module and the correlation of 19 modules (Figure 1D) based on the heat map. Each module also shows independent verification, and the gradually saturated blue and red indicate high co-expression interconnectivity (Figure 1D).

Identification of meta-modules related to four molecular subtypes of medulloblastoma

As we know, ME is the most important component of the module and can be considered as a representative of the gene expression profile of the module. The 28 MEs of 28 different modules are respectively related to the four molecular subtypes of medulloblastoma. Next, we assessed the relationship between each module and the molecular subtype of medulloblastoma by correlating the characteristic gene of each module. We found that, as expected, the pink module is highly correlated with the SHH molecular subtype, brown is highly correlated with G4, darkred is highly correlated with G3, and the green module is highly correlated with WNT (absolute $r > 0.81$, $p < 0.0001$). In addition, Figure 2C also shows four interesting modules based on ME expression profiles and 73 samples of medulloblastoma molecular subtype characteristics. These four modules are considered to be very meaningful modules and are highly related to their molecular subtypes

Differences in gene expression of molecular subtypes and GO enrichment

It is well known that there are differences in survival between different molecular subtypes of medulloblastoma. And this difference is probably due to the expression of different genes. In order to illustrate the difference in gene expression in the four modules, heat maps of the gene expression in medulloblastoma of pink, brown, green, and darkred modules were drawn. It can be seen that WNT, SHH, G4, and G3 molecular subtypes highly express the genes of green, pink, brown, and darkred modules, respectively. (Figure 3A). Besides to explore the enrichment of genes in these modules, R package clusterProfiler to analyze and visualize the GO enrichment of these genes. We found that the genes of the WNT module are mainly related to axonogenesis and regulation of neuron projection development, regulation of cell morphogenesis, and as expected, are also related to regulation of Wnt signaling pathway. (Figure 3B) The different gene of SHH is mainly related to the modulation of chemical synaptic transmission, regulation of trans-synaptic signaling, forebrain development. (Figure 3C) In addition, the differential genes of G4 are mainly concentrated in the synapse organization, regulation of membrane potential, modulation of chemical synaptic transmission, regulation of trans-synaptic signaling. (Figure 3D) What also surprised us is that G3-specific genes are mainly concentrated in Spinal cord development, and the top-ranked GO enrichment of the other three molecular subtypes are all related to the development of synapses and brain and nerves, which further reveals that the lethal mechanism of G3 may be different from the other three molecular subtypes. (Figure 3E)

Screening of hub gene

From the four modules of interest above, we further used cytoscape to analyze, drawing the protein interaction network (Figure 4) and the ten core genes in each module were taken out for further analysis.

Single factor COX regression analysis of hub gene and dataset verification

The corresponding clinical data in the literature cited by GSE85217 were used to perform survival analysis of these hub genes. In order to test the accuracy of the clinical data of GSE85217, we first performed the analysis of the cox regression of its molecular subtype and survival time (Figure5A), and found that it is consistent with most reports in the literature, that is, WNT has a better prognosis, while SHH , G3, G4 have worse prognosis. In addition, we also conducted a gender survival curve analysis, but the analysis showed that gender factors have no effect on the survival of medulloblastoma.

Multivariate COX regression analysis of molecular subtypes of medulloblastoma

The mortality rates of different molecular subtypes of medulloblastoma vary greatly (Figure 6B). Therefore, after screening out 12 hub genes with single-factor cox regression analysis (Figure 6A), we used these genes to perform multivariate cox regression survival curve analysis on 107 G3 samples, 245 G4 samples, and 150 SHH. (There are only 49 WNT samples, so these samples are not used to analyze WNT separately). Therefore, we are surprised to find that some genes can well predict the role of these three molecular subtypes. For example, among the G4 subtypes, the combination of SOCS3 and HOXC10 can better predict the progress of the G3 subtype. ($p=0.034$, $AUC=0.616$) (Figure6 A, D). In addition, in the G3 subtype, the combination of ADCY8 and LHX3 can better predict the progress of the G3 subtype($p=0.034$, $AUC=0.675$) (Figure5B,E). Besides, in the SHH subtype, the combination of GNG3, CALCB, and GCGR can better predict the progress of the SHH subtype. ($p=0.0011$, $AUC=0.734$) (Figure5C, F). The prediction of the SHH model is the most accurate, indicating that these three genes may play an important role in the occurrence and development of SHH molecular subtypes.

Discussion

Medulloblastoma (MB) is the most common pediatric malignant brain tumor, accounting for nearly 20% of all childhood central nervous system (CNS) malignancies(21). The 5-year survival rate of standard-risk children is more than 85% and the 5-year survival rate of high-risk children is about 70%.(22) As we all know, the four recognized molecular subtypes of medulloblastoma are WNT, SHH, G3, and G4. Current research shows that WNT accounts for about 10% of MB, SHH about 30% , Group3 about 20% , Group4 about 40%(21). This is similar to the proportion of the 551 four molecular subtypes of GSE85217 that exclude medulloblastomas with unclear clinical phenotypes. The 551 samples contained 49 WNT subtypes (8.9%), 107 G3 subtypes (19.4%), 245 G4 subtypes (44.4%), and 150 SHH subtypes (27.2%). This also shows that even after removing some samples with unclear clinical phenotypes, the dataset is still clearly representative. The distribution in GSE37418 is also relatively close. In addition, a large-scale study showed that gender factors have no significant effect on the survival rate of medulloblastoma(23). In addition, a large-scale study showed that although the incidence of medulloblastoma differs between men and women, about 2-3:1, gender factors have no significant effect on the survival rate of

medulloblastoma.(23) Our research shows (Figure 5C) that gender has no effect. In addition, using these samples, we found that even in these four subtypes, although males still account for a larger proportion than reported, there is no gender difference in survival rates (see attached picture)

The development of different molecular subtypes of medulloblastoma is different, so finding prognostic markers that determine each molecular subtype of medulloblastoma is indeed very important for predicting the occurrence and development of each molecular subtype of medulloblastoma. Previous studies have shown that the status of TP53 may be one of the important markers of medulloblastoma. (24,25) However, whether there are other markers that affect the occurrence and development of medulloblastoma and can better predict the molecular subtype of medulloblastoma alone is not yet known.

By using the WGCNA analysis method and reducing the complexity of the expression profile, 28 medulloblastoma-related modules were identified from 73 medulloblastoma samples of GSE37418. Among them, four modules were found to be highly correlated with four molecular subtypes. We have also conducted extensive research on the conservation of the four modules in different data sets. We confirm that based on WGCNA meta-modules, pink, brown, darkred, and green are indeed highly correlated with SHH, G4, G3, and WNT. The genes of these modules are the unique differential genes of these molecular subtypes. Therefore, the analysis of these differential genes can better reveal the essential differences between the development and deterioration of the four subtypes GO enrichment analysis shows that G3 basic genes are mainly concentrated in spinal cord development, while the first GO enrichment among SHH, WNT, and G4 subtypes are all related to the development of synapses, brain and nerves, which further reveals G3 The lethal mechanism may be different from the other three molecular subtypes (Figure 3). Perhaps this is related to the metastasis rate of medulloblastoma G3 subtype as high as 40%-45% at the time of diagnosis. (9,26)

And then we used univariate Cox regression analysis to calculate the hub genes calculated from the cytoscape algorithm and found that the overall effect on medulloblastoma was as significant twelve genes. Then use these twelve genes to perform COX multivariate regression analysis among the four subtypes of medulloblastoma. The higher the risk score, the worse the prognosis and the higher the degree of malignancy. Among them, we found that although GNG3, CALCB, and GCGR may not be significant as a single factor, the combination of the three genes can predict the occurrence and development of this subtype of SHH very well. In addition, these three genes belong to the hub genes of the pink module of SHH, which further shows that the mechanism of action between the three genes has a deeper impact on the occurrence and development of SHH. However, no one has reported the effect of the combination of these three genes on SHH. For these three genes, Guanine nucleotide-binding proteins (G proteins) are involved as a modulator or transducer in various transmembrane signaling systems. The beta and gamma chains are required for the GTPase activity, for replacement of GDP by GTP, and for G protein-effector interaction. Guanine nucleotide binding proteins are heterotrimeric signal-transducing molecules consisting of alpha, beta, and gamma subunits. The gamma subunit determines the specificity of which signaling pathways will be affected by this particular complex. The protein encoded by this gene

GNG3 represents the gamma subunit of both inhibitory and stimulatory complexes. (27,28)CGRP induces vasodilation. It dilates a variety of vessels including the coronary, cerebral and systemic vasculature. Its abundance in the CNS also points toward a neurotransmitter or neuromodulator role.(29) (28)And the protein encoded by GCGR is a glucagon receptor that is important in controlling blood glucose levels.(28) (30)In addition, we also found that the combination of SOCS3 and HOXC10 can predict G4, but the prediction efficiency may not be significant (AUC=0.616). The combination of ADCY8 and LHX3 can predict G3, but the prediction efficiency may still be low (AUC=0.676), but the two single factors may have a greater impact on G3.

Conclusions

In conclusion, we have constructed OS-related features of various molecular subtypes of medulloblastoma, which can independently predict the prognosis of patients with each subtype of medulloblastoma, especially those with SHH molecular subtypes, and provide new therapeutic targets for them. We have a deep understanding of the biological mechanism and clinical significance of the hub gene in each molecular subtype of medulloblastoma, but further experiments will be still needed to verify our findings in the future.

Abbreviations

G3 Group3,

G4 Group4

WGCNA Weighted Gene Correlation Network Analysis

GEO: Gene Expression Omnibus

TON: topological overlap matrix

MB: Medulloblastoma

CNS: central nervous system

G proteins: Guanine nucleotide-binding proteins

Declarations

Availability of data and materials

The datasets used during the current study are available from the corresponding author on a reasonable request. The data used to support the findings of this study are not shared. The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

The authors thank the participating investigators and patients and thank the NATIVE EE Company for English editing.

Funding

This research was funded by grants from the National Natural Science Foundation Committee of China (No. 81372713; 81572497; 81703011; 81873739), and the Guangdong Provincial Department of Science and Technology, China (No.2017A030313487;202002030079), This study was partially supported by the fund from Guangzhou Women and Children's Medical Center/Guangzhou Institute of Pediatrics (NO: YIP-2019-027,IP-2019-006; YIP-2019-006)

Author information

†Hong-Yao Yuan and Cheng Chen contributed equally to this work as the first authors.

Affiliations

Department of Neurosurgery, Guangzhou Women and Children's Medical Center, Guangzhou, China.

Hong-Yao Yuan✉Cheng Chen✉Jun-Ping Pan✉Xin-Ke Xu✉Cui-Yu Lin✉Wei Chen✉Yang Li✉Jin-Rong Lin✉Jun-Liang Li, Fang-Cheng Li

Guangdong Provincial Key Laboratory of Research in Structural Birth Defect Disease, Guangzhou Women and Children's Medical Center, Guangzhou, China.

Hong-Yao Yuan✉Cheng Chen✉Jun-Ping Pan✉Xin-Ke Xu✉Wei Chen✉Yang Li✉Jin-Rong Lin✉Jun-Liang Li, Fang-Cheng Li

Contributions

FCL conceived and designed the study. HYY, CC, CYL and JPP performed the experiments, analyzed the data and wrote the manuscript. JLL, JRL, YL, LFZ, PJP performed the analysis using bioinformatics, and CC, YL assisted in performing the research, and CYL provided language help and assisted in analyzing data.

Corresponding author

Correspondence to Fang-Cheng Li; Jun-Liang Li.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Conflict of interest statement

The authors declare no potential conflicts of interest

References

1. Morfouace M, Cheepala S, Jackson S, Fukuda Y, Patel YT, Fatima S, et al. ABCG2 Transporter Expression Impacts Group 3 Medulloblastoma Response to Chemotherapy. *Cancer research* 2015;75(18):3879-89 doi 10.1158/0008-5472.Can-15-0030.
2. Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. *Neuro-oncology* 2018;20(suppl_4):iv1-iv86 doi 10.1093/neuonc/noy131.
3. Hovestadt V, Remke M, Kool M, Pietsch T, Northcott PA, Fischer R, et al. Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays. *Acta neuropathologica* 2013;125(6):913-6 doi 10.1007/s00401-013-1126-5.
4. Lafay-Cousin L, Smith A, Chi SN, Wells E, Madden J, Margol A, et al. Clinical, Pathological, and Molecular Characterization of Infant Medulloblastomas Treated with Sequential High-Dose Chemotherapy. *Pediatric blood & cancer* 2016;63(9):1527-34 doi 10.1002/pbc.26042.
5. Cook Sangar ML, Genovesi LA, Nakamoto MW, Davis MJ, Knobluagh SE, Ji P, et al. Inhibition of CDK4/6 by Palbociclib Significantly Extends Survival in Medulloblastoma Patient-Derived Xenograft Mouse Models. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2017;23(19):5802-13 doi 10.1158/1078-0432.Ccr-16-2943.
6. Huang GH, Xu QF, Cui YH, Li N, Bian XW, Lv SQ. Medulloblastoma stem cells: Promising targets in medulloblastoma therapy. *Cancer science* 2016;107(5):583-9 doi 10.1111/cas.12925.
7. Vladioiu MC, El-Hamamy I, Donovan LK, Farooq H, Holgado BL, Sundaravadanam Y, et al. Childhood cerebellar tumours mirror conserved fetal transcriptional programs. *Nature* 2019;572(7767):67-73 doi 10.1038/s41586-019-1158-7.
8. Kool M, Korshunov A, Remke M, Jones DT, Schlanstein M, Northcott PA, et al. Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. *Acta neuropathologica* 2012;123(4):473-84 doi 10.1007/s00401-012-0958-8.
9. Taylor MD, Northcott PA, Korshunov A, Remke M, Cho YJ, Clifford SC, et al. Molecular subgroups of medulloblastoma: the current consensus. *Acta neuropathologica* 2012;123(4):465-72 doi 10.1007/s00401-011-0922-z.
10. Ramaswamy V, Remke M, Bouffet E, Faria CC, Perreault S, Cho YJ, et al. Recurrence patterns across medulloblastoma subgroups: an integrated clinical and molecular analysis. *The Lancet Oncology*

- 2013;14(12):1200-7 doi 10.1016/s1470-2045(13)70449-2.
11. Perreault S, Ramaswamy V, Achrol AS, Chao K, Liu TT, Shih D, et al. MRI surrogates for molecular subgroups of medulloblastoma. *AJNR American journal of neuroradiology* 2014;35(7):1263-9 doi 10.3174/ajnr.A3990.
 12. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* 2008;9:559 doi 10.1186/1471-2105-9-559.
 13. Heiland DH, Wörner J, Gerrit Haaker J, Delev D, Pompe N, Mercas B, et al. The integrative metabolomic-transcriptomic landscape of glioblastoma multiforme. *Oncotarget* 2017;8(30):49178-90 doi 10.18632/oncotarget.16544.
 14. Sun Q, Zhao H, Zhang C, Hu T, Wu J, Lin X, et al. Gene co-expression network reveals shared modules predictive of stage and grade in serous ovarian cancers. *Oncotarget* 2017;8(26):42983-96 doi 10.18632/oncotarget.17785.
 15. Robinson G, Parker M, Kranenburg TA, Lu C, Chen X, Ding L, et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature* 2012;488(7409):43-8 doi 10.1038/nature11213.
 16. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. *Cancer cell* 2017;31(6):737-54.e6 doi 10.1016/j.ccell.2017.05.005.
 17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* 2012;16(5):284-7 doi 10.1089/omi.2011.0118.
 18. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker TJB. Cytoscape 2.8: new features for data integration and network visualization. 2011;27(3):431-2.
 19. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-YJBSb. cytoHubba: identifying hub objects and sub-networks from complex interactome. 2014;8(S4):S11.
 20. Heagerty PJ, Saha-Chaudhuri P, Saha-Chaudhuri MP. Package 'survivalROC'. 2013.
 21. Khatua S, Song A, Citla Sridhar D, Mack SC. Childhood Medulloblastoma: Current Therapies, Emerging Molecular Landscape and Newer Therapeutic Insights. *Current neuropharmacology* 2018;16(7):1045-58 doi 10.2174/1570159x15666171129111324.
 22. Northcott PA, Buchhalter I, Morrissy AS, Hovestadt V, Weischenfeldt J, Ehrenberger T, et al. The whole-genome landscape of medulloblastoma subtypes. *Nature* 2017;547(7663):311-7 doi 10.1038/nature22973.
 23. Curran EK, Sainani KL, Le GM, Propp JM, Fisher PG. Gender affects survival for medulloblastoma only in older children and adults: a study from the Surveillance Epidemiology and End Results Registry. *Pediatric blood & cancer* 2009;52(1):60-4 doi 10.1002/pbc.21832.
 24. Zhukova N, Ramaswamy V, Remke M, Pfaff E, Shih DJ, Martin DC, et al. Subgroup-specific prognostic implications of TP53 mutation in medulloblastoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013;31(23):2927-35 doi 10.1200/jco.2012.48.5052.

25. Juraschka K, Taylor MD. Medulloblastoma in the age of molecular subgroups: a review. *Journal of neurosurgery Pediatrics* 2019;24(4):353-63 doi 10.3171/2019.5.Peds18381.
26. Schneider C, Ramaswamy V, Kulkarni AV, Rutka JT, Remke M, Tabori U, et al. Clinical implications of medulloblastoma subgroups: incidence of CSF diversion surgery. *Journal of neurosurgery Pediatrics* 2015;15(3):236-42 doi 10.3171/2014.9.Peds14280.
27. Schwindinger WF, Borrell BM, Waldman LC, Robishaw JDJAJoP-R, Integrative, Physiology C. Mice lacking the G protein $\gamma 3$ -subunit show resistance to opioids and diet induced obesity. 2009;297(5):R1494-R502.
28. Pruitt KD, Maglott DRJNar. RefSeq and LocusLink: NCBI gene-centered resources. 2001;29(1):137-40.
29. Van Rossum D, Hanisch U-K, QUIRION RJN, Reviews B. Neuroanatomical localization, pharmacological characterization and functions of CGRP, related peptides and their receptors. 1997;21(5):649-78.
30. Yang J, MacDougall ML, McDowell MT, Xi L, Wei R, Zavadoski WJ, et al. Polyomic profiling reveals significant hepatic metabolic alterations in glucagon-receptor (GCGR) knockout mice: implications on anti-glucagon therapies for diabetes. 2011;12(1):281.

Figures

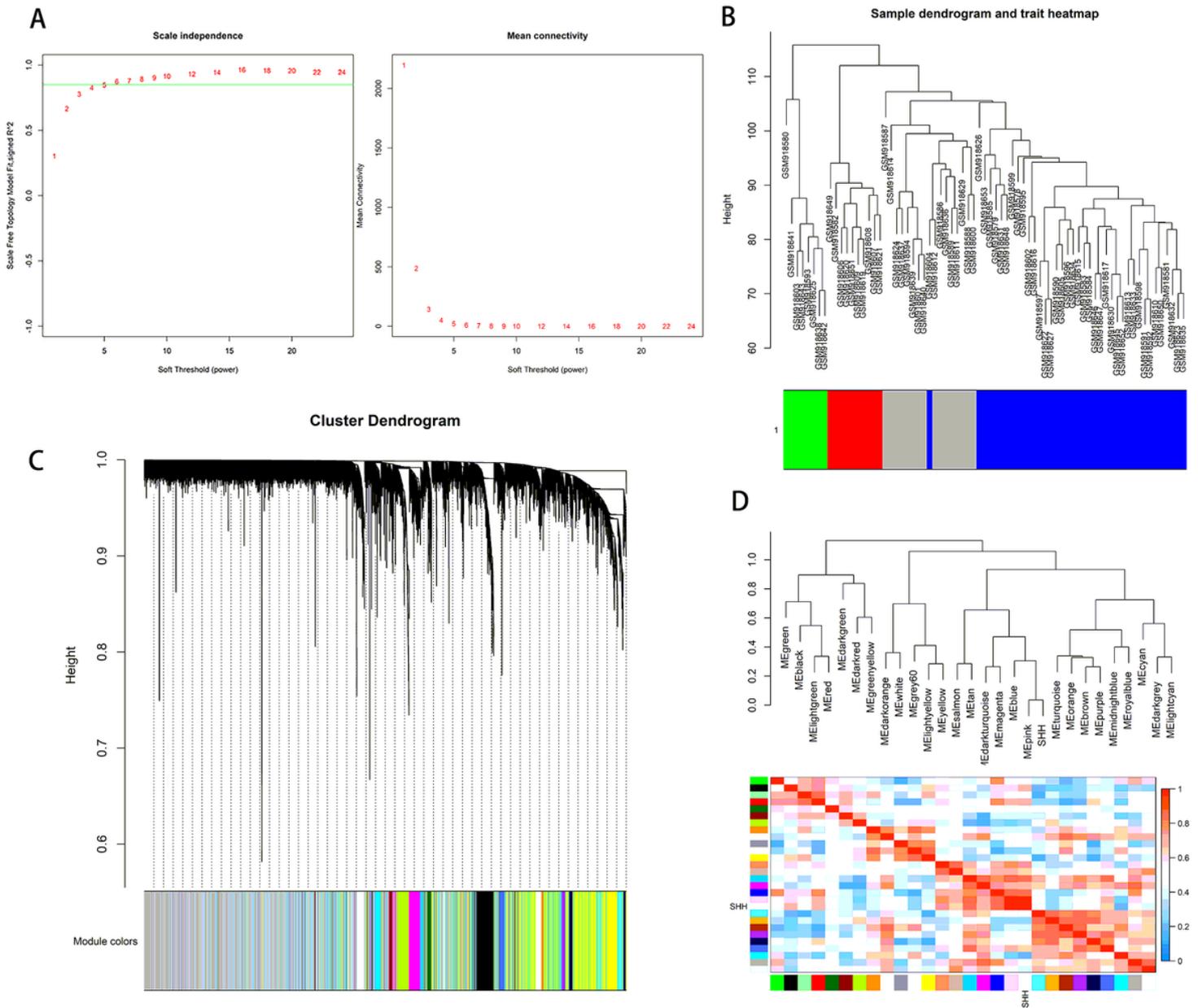


Figure 1

Construction of WGCNA network (A) The acquisition of soft threshold power. The left panel is a graph about the unscaled fit index and soft threshold function. And the right panel reveals average connectivity and soft threshold power. After choosing Power 6, the fitting exponential curve gradually becomes flat when it reaches a high value (>0.87). (B) Cluster tree diagrams and module distribution of modules produced by WGCNA. Genes are clustered by means of difference measurement (1-TOM). The branches correspond to modules of highly interconnected groups of genes. Colors in the horizontal bar represent the modules. 28 modules with 16241 transcripts were detected with WGCNA. (C,D) Identification of meta-module and module-module relationship. The modular network dendrogram is constructed by clustering the intrinsic gene distances of the modules. In the end, 28 different gene modules were identified. In the

heat map of the module-module relationship, from gradually saturated blue to red, the co-expression interconnectivity increases sequentially.

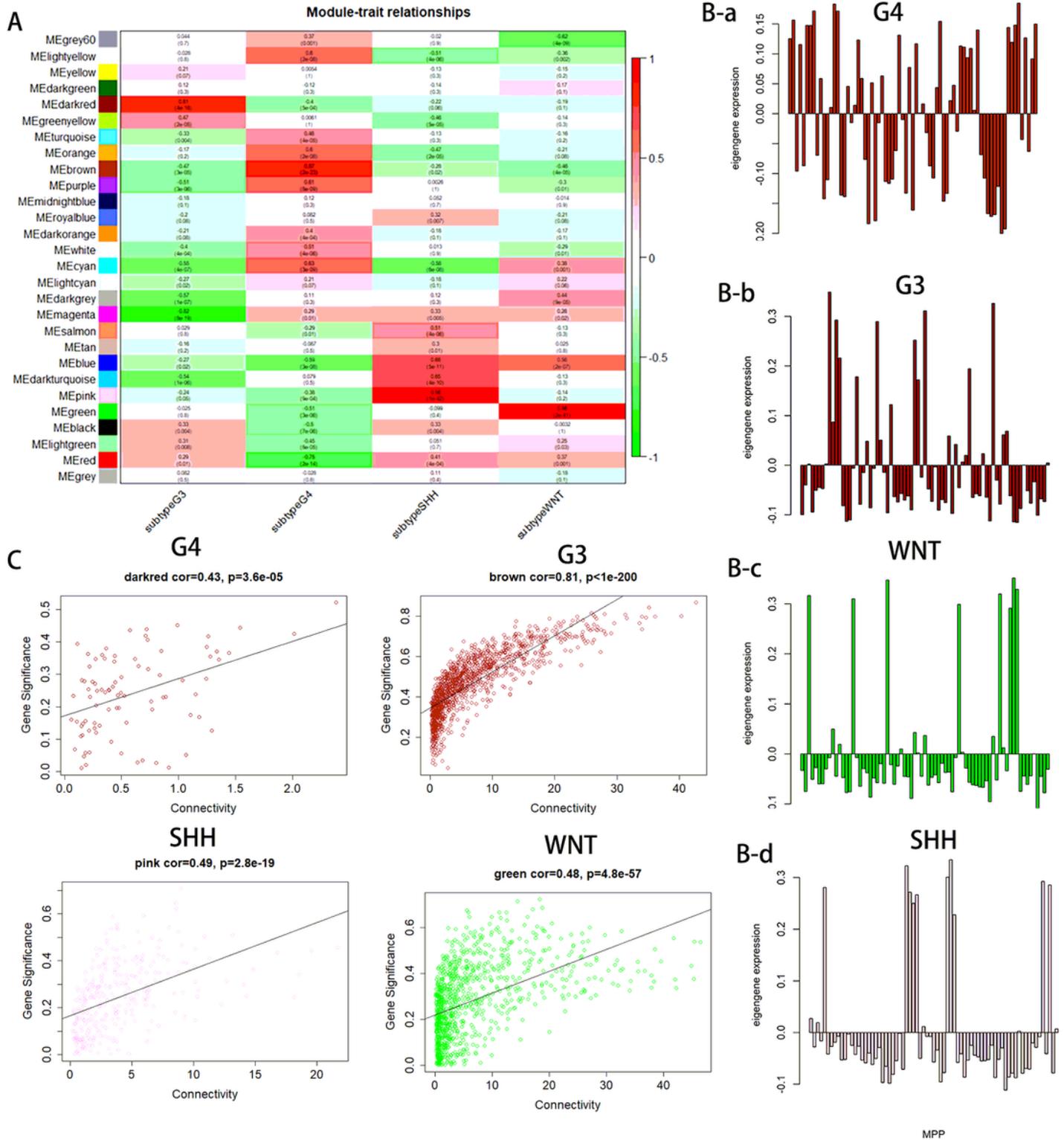


Figure 2

The characteristics of the module in medulloblastoma and the gene expression of the module of interest (A) Each cell contains a corresponding correlation. High correlation is marked as red, and low correlation is marked as green. (B) The histograms a-d sequentially described the eigengene expression of each

module from subtypes G3,G4,SHH,WNT (C) Relationship between gene significance and intramodular connectivity about 4 module.

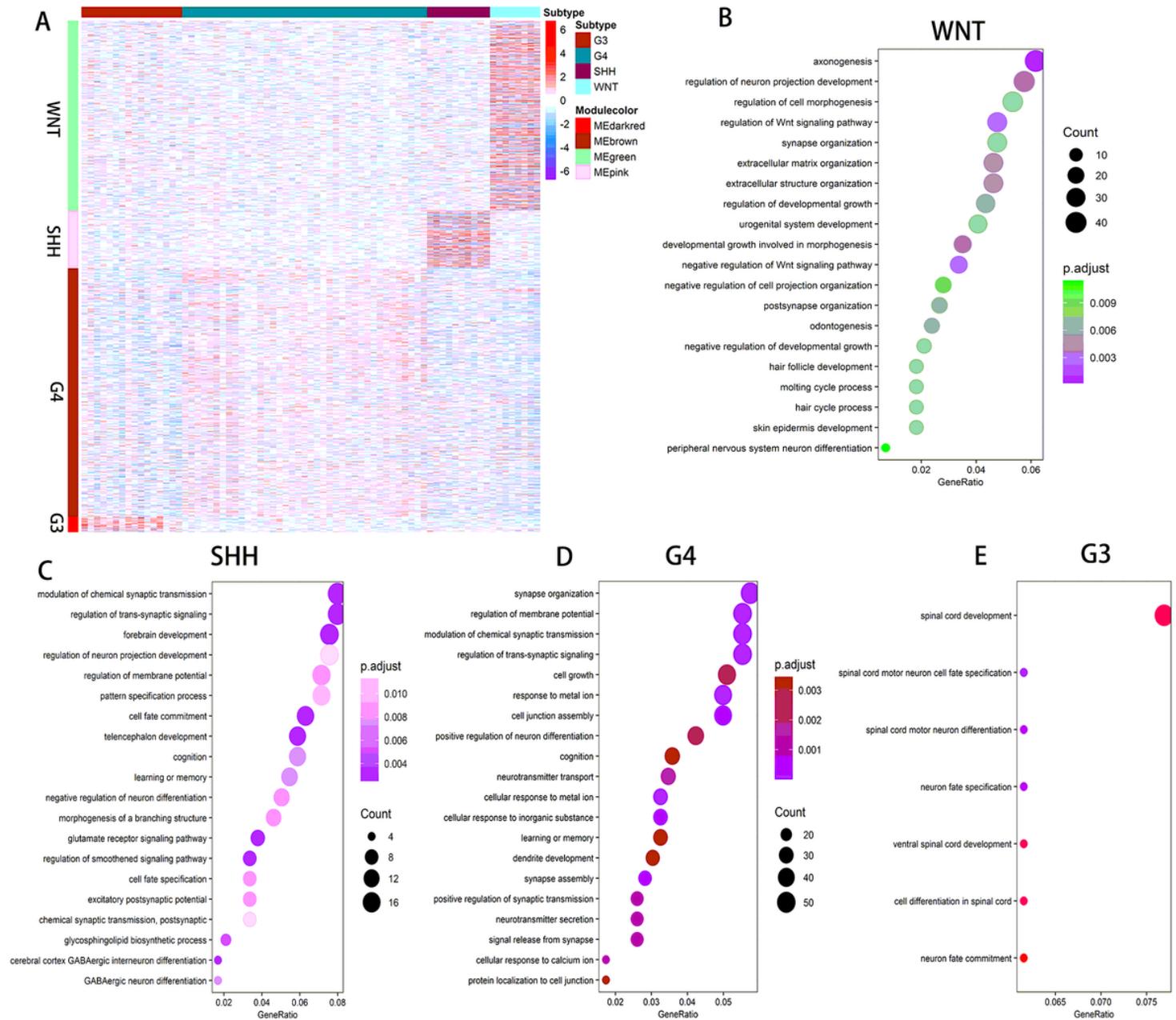


Figure 3

GO pathway enrichment of four interested modules (A) Hierarchical cluster analysis of six interested modules, based on the module-trait's correlation and p value (absolute $r > 0.8$, $P < 10^{-18}$), four modules (darkred,brown,green,pin) showed relatively high expression (red) and lower expression (purple) in the medulloblastoma population. (B)-(D) GO pathway enrichment of green, pink, brown and darkred modules

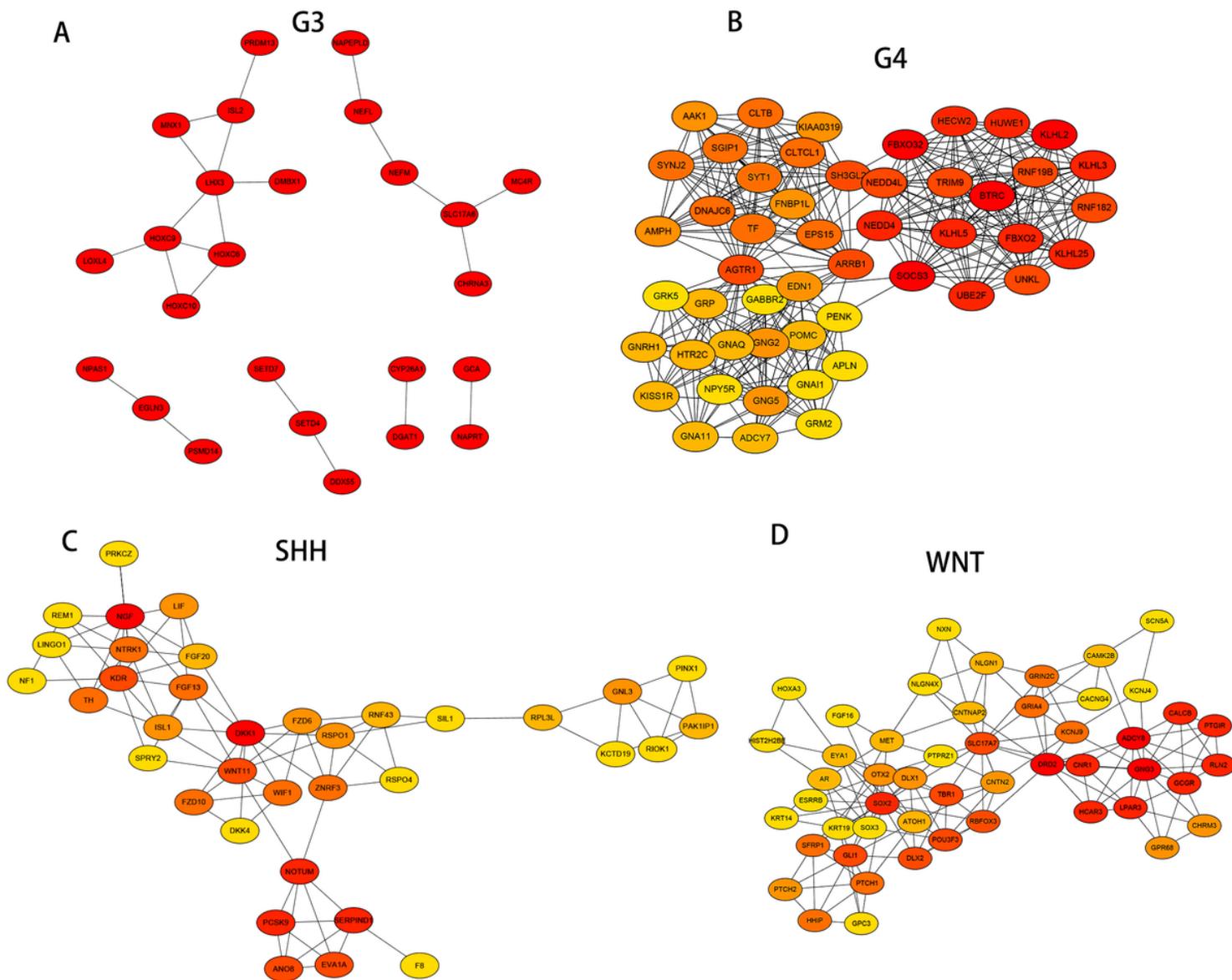


Figure 4

Obtain the hub gene (A) Interaction of gene co-expression patterns in the pink, green, brown, darkred module. The module was visualized using Cytoscape_3.6.0 software. The node colors coded from yellow to red (low to high).

A

<i>gene</i>	<i>HR</i>	<i>z</i>	<i>pvalue</i>
GNG3	0.812413676099547	-3.88669189316394	0.000101619601002968
HCAR3	0.681484288089742	-3.46252748717395	0.000535127251781804
ADCY8	0.755437285148181	-3.43907968577739	0.0005836953951016
DRD2	0.798033662714381	-3.2750525080878	0.0010564233599571
FBXO2	1.38302070182137	2.80219094615079	0.00507568220400271
SOCS3	1.38274024946006	2.75580469911299	0.00585479441103302
LHX3	1.63881013667134	2.72530076617241	0.00642429356921495
CALCB	0.785717390736341	-2.7216417845262	0.00649585129533493
SERPIND1	0.717590081259912	-2.69602033405753	0.00701733828345912
EGLN3	1.2795861510983	2.44793981765713	0.0143675651764183
HOXC10	1.16727713929543	2.43005657798032	0.0150964660208157
GCGR	0.625661707654417	-2.31533571996567	0.0205945663668737

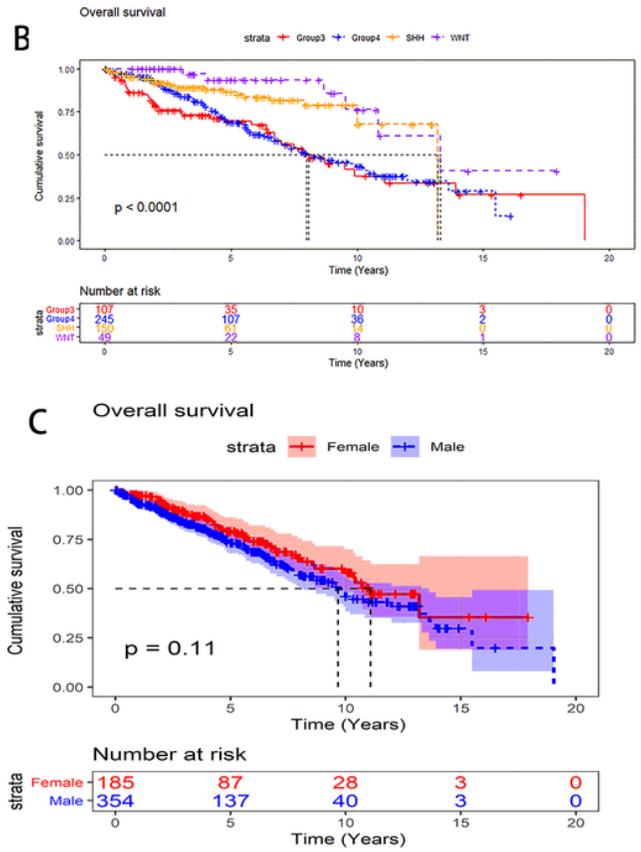


Figure 5

(A) 12 genes were obtained by univariate Cox regression analysis (B)-(C) Calculating the molecular subtype(B) or gender(C) and survival time of medulloblastoma by Cox regression analysis with the clinical data of the GSE85217 dataset

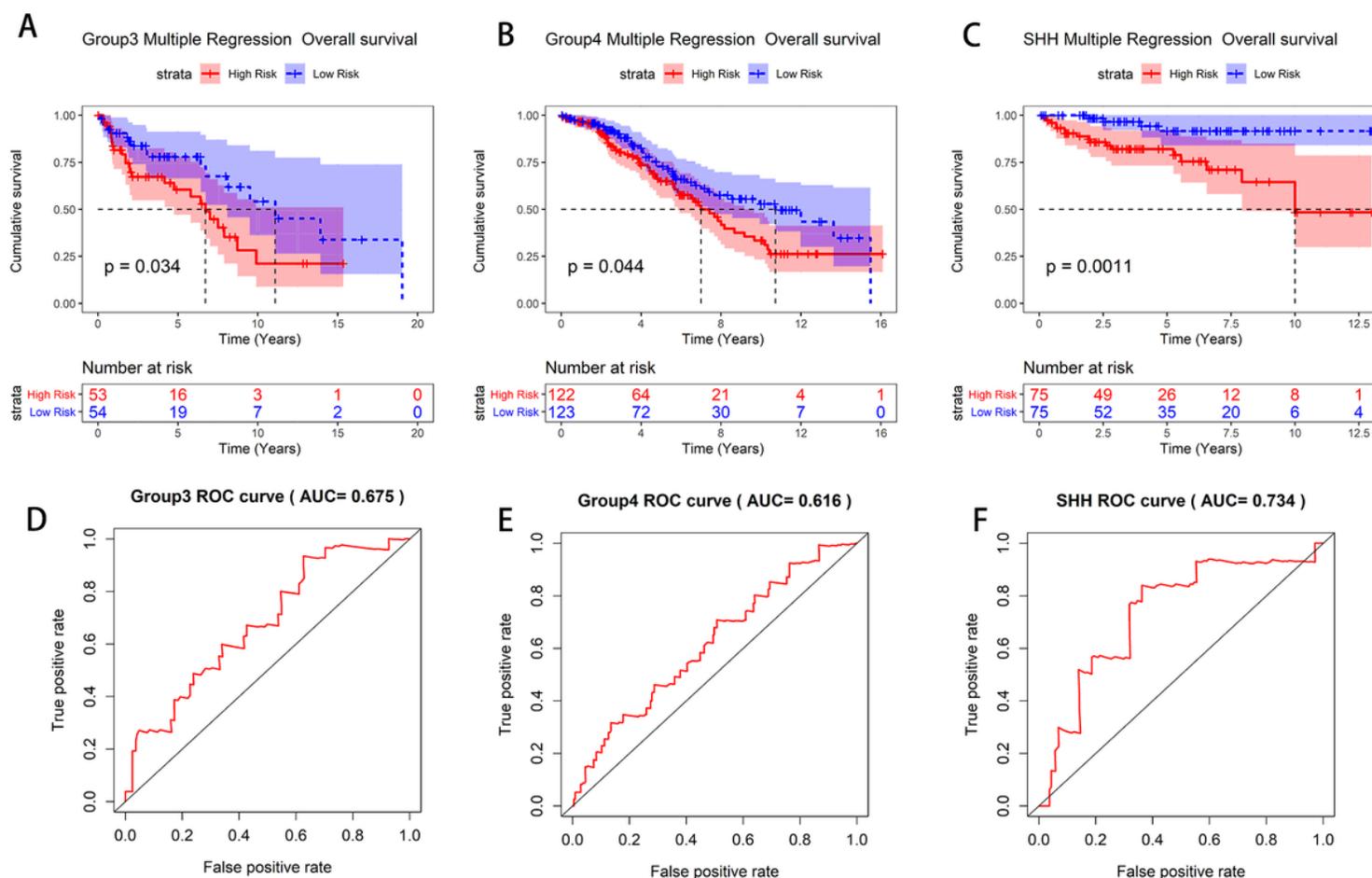


Figure 6

A-C The hub genes selected by 12 univariate COX regression analyses were used to perform multivariate COX regression analysis on the clinical data and expression data belonging to G3, G4 and SHH molecular subtypes in the GSE85217 data set. Multivariate COX regression analysis can show that the common expression of some genes can well predict the risk of the development of G3, G4, SHH, and further reveal that these genes have a high significance in the occurrence and development of these subtypes. Among them, blue represents low risk, and red represents high risk. B-D The ROC curve of the above multivariate cox regression model.