Identification of hub genes and small-molecule compounds in medulloblastoma by integrated bioinformatic analyses

Zhendong Liu [Equal first author, 1, 2], **Ruotian Zhang** [Equal first author, 1, 2], **Zhenying Sun** [1, 2], **Jiawei Yao** [1, 2], **Penglei Yao** [1, 2], **Xin Chen** [1, 2], **Xinzhuang Wang** [1, 2], **Ming Gao** [1, 2], **Jinzhao Wan** [1, 2], **Yiming Du** [1, 2], **Shiguang Zhao** [Corresp. 1, 2]

[1] Department of Neurosurgery, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang Province, People's Republic of China

[2] Institute of Brain Science, Harbin Medical University, Harbin, Heilongjiang Province, People's Republic of China

Corresponding Author: Shiguang Zhao
Email address: guangsz@hotmail.com

Background. Medulloblastoma (MB) is the most common intracranial malignant tumor in children. The genes and pathways involved in the pathogenesis of MB are relatively unknown. We aimed to identify potential biomarkers and small-molecule drugs for MB.

Methods. Gene expression profile data sets were obtained from the Gene Expression Omnibus (GEO) database and the differentially expressed genes (DEGs) were identified using the Limma package in R. Functional annotation, and cell signaling pathway analysis of DEGs was carried out using DAVID and Kobas. A protein-protein interaction network was generated using STRING. Potential small-molecule drugs were identified using CMap.

Result. We identified 104 DEGs (29 upregulated; 75 downregulated). Gene ontology analysis showed enrichment in the mitotic cell cycle, cell cycle, spindle, and DNA binding. Cell signaling pathway analysis identified cell cycle, HIF-1 signaling pathway, and phospholipase D signaling pathway as key pathways. SYN1, CNTN2, FAIM2, MT3, and SH3GL2 were the prominent hub genes and their expression level were verified by RT-qPCR. Vorinostat, resveratrol, trichostatin A, pyrvinium, and prochlorperazine were identified as potential drugs for MB. The five hub genes may be targets for diagnosis and treatment of MB, and the small-molecule compounds are promising drugs for effective treatment of MB.

Conclusion. In this study we obtained five hub genes of MB, SYN1, CNTN2, FAIM2, MT3, and SH3GL2 were confirmed as hub genes. Meanwhile, Vorinostat, resveratrol, trichostatin A, pyrvinium, and prochlorperazine were identified as potential drugs for MB.

1   **Identification of hub genes and small-molecule compounds in medulloblastoma by**

2   **integrated bioinformatic analyses**

3   Zhendong Liu[1,2]*, Ruotian Zhang[1,2]*, Zhenying Sun[1,2], Jiawei Yao[1,2], Penglei Yao[1,2], Xin

4   Chen[1,2], Xinzhuang Wang[1,2], Ming Gao[1,2], Jinzhao Wan[1,2], Yiming Du[1,2], Shiguang

5   Zhao[1,2]

7   [1]Department of Neurosurgery, The First Affiliated Hospital of Harbin Medical University, No. 23 Youzhe

8   ng Street, Nangang District, 150001 Harbin, Heilongjiang Province, People's Republic of China

9   [2]Institute of Brain Science, Harbin Medical University, No. 23 Youzheng Street, Nangang District, 15000

10   1 Harbin, Heilongjiang Province, People's Republic of China

12   Corresponding Author:
13   Shiguang Zhao
14   Department of Neurosurgery, The First Affiliated Hospital

15   of Harbin Medical University, No. 23 Youzheng Street, Nangang District,150001

16   Harbin, Heilongjiang Province, People's Republic of China.

17   Email address: guangsz@hotmail.com
18
19   *These authors contributed equally: Zhendong Liu and Ruotian Zhang

21   **Abstract**

22   ***Background.*** Medulloblastoma (MB) is the most common intracranial malignant tumor

23   in children. The genes and pathways involved in the pathogenesis of MB are relatively

24   unknown. We aimed to identify potential biomarkers and small-molecule drugs for MB.

25   ***Methods.*** Gene expression profile data sets were obtained from the Gene Expression

26   Omnibus (GEO) database and the differentially expressed genes (DEGs) were identified

27   using the Limma package in R. Functional annotation, and cell signaling pathway

28   analysis of DEGs was carried out using DAVID and Kobas. A protein-protein interaction

29   network was generated using STRING. Potential small-molecule drugs were identified

30   using CMap.

31   ***Result.*** We identified 104 DEGs (29 upregulated; 75 downregulated). Gene ontology

32   analysis showed enrichment in the mitotic cell cycle, cell cycle, spindle, and DNA

33   binding. Cell signaling pathway analysis identified cell cycle, HIF-1 signaling pathway,

34   and phospholipase D signaling pathway as key pathways. SYN1, CNTN2, FAIM2, MT3,

35   and   SH3GL2   were   the   prominent   hub   genes   and

36   their expression levels were verified by RT-qPCR. Vorinostat, resveratrol, trichostatin A,

37   pyrvinium, and prochlorperazine were identified as potential drugs for MB. The five hub

38   genes   may   be   targets   for   diagnosis   and   treatment   of   MB, and   the   small-molecule

39   compounds are promising drugs for effective treatment of MB.

40   ***Conclusion.*** In this study we obtained five hub genes of MB, SYN1, CNTN2, FAIM2,

41   MT3, and SH3GL2 were confirmed as hub genes.   Meanwhile, Vorinostat, resveratrol,

42   trichostatin A, pyrvinium, and prochlorperazine were identified as potential drugs for

43   MB.

Introduction

45   Medulloblastoma (MB) is one of the most common intracranial malignant tumors in

46   children. Standardized treatment options including maximal surgical resection, radiation

47   therapy, and chemotherapy can improve overall five-year survival rates; however, the

48   prognosis of patients with advanced disease is still unsatisfactory (Kim et al. 2010) (1).

49   Despite   extensive   studies   on   the   pathogenesis   and   progression   of   MB, the

50   pathophysiology of disease development is still unclear. The pathogenesis of MB is

51   closely related to many factors such as mutations in genes, abnormalities in cellular

52   immunity, and changes in environmental factors (Wang et al. 2018) (2). A deep

53   understanding of the changes in protein expression involved in the pathogenesis of MB is

54   critical for the development of better treatment strategies.

55   Gene chip is a genetic testing technology that has been in use for more than a

56   decade. The gene chip technology can rapidly detect changes in gene expression in the

57   sample and is suitable for screening differentially expressed genes (DEGs) (Vogelstein et

58   al. 2013) (3). In recent years, gene chip technology has been widely used, and data from

59   numerous microarray studies have been stored in free public databases. These databases

60   provide valuable data for further research. So far, several gene chip data have been

61   analyzed, and hundreds of different genes involved in the development of several central

62   nervous system cancers such as gliomas have been identified (Xi et al. 2017; Zeng et al.

63   2018) (4,5). However, most gene chip data is a mixture of both MB and glioma and does

64   not accurately represent changes in MB, and therefore does not generate satisfactory

65   results.

66   There have been reports of microarray data to find DEGs in medulloblastoma.(Shaabanpo

67   ur   Aghamaleki   et   al.

68   2019) However, the reliability of the analysis results was controversial due to the heterog

69   eneity of the sample from a single cohort study. Therefore, most single-chip analyses are

70   not sufficient to identify effective biomarkers in MB. To overcome this, we used data

71   derived from four gene chips to identify effective biomarkers for MB. This method is

72   more accurate and does not have the disadvantages associated with the single-chip

73   analysis.

74   In this study, we downloaded four MB microarray datasets – GSE42656,

75   GSE74195, GSE109401, and GSE50161 from free public Gene Expression

76   Omnibus database (GEO). A total of 77 MB samples and 39 normal brain samples were

77   included in this study. We identified the DEGs using the Limma package in R software

78   (version 3.5.0) from the four gene expression profiles and subsequently, used the Venny

79   online tool for further integrated analyses. We then employed the DAVID databases and

80   Kobas online tool to identify the functions of the identified DEGs and the key cell

81   signaling pathways involved. The network of protein-protein interaction (PPI) was

82   generated using the STRING database. Finally, we used the CMap database to explore

83   potential small-molecule compounds that can be used for treating MB.

84   **Materials & Methods**
85   *Microarray data information*

86   We downloaded the gene expression profile data for GSE42656, GSE74195,

87   GSE109401, and GSE50161 from the GEO database (http://www.ncbi.nlm.nih.gov/geo).

88   The GSE42656 data set was based on the GPL6947 Platforms (Illumina HumanHT-12

89   V3.0 expression beadchip) and contained nine MB samples and 16 normal brain samples

90   (Henriquez et al. 2013)(6). The GSE74195 data set was based on the GPL570 Platform

91   (Affymetrix Human Genome U133 Plus 2.0 Array) and contained 27 MB and five normal

92   brain tissue samples (de Groot et al. 2011) (7). The GSE50161 data set was based on the

93   GPL570 Platform (Affymetrix Human Genome U133 Plus 2.0 Array) and included 22

94   MB samples and 13 normal brain tissues (Griesinger et al. 2013) (8). The GSE109401

95   data set was based on GPL16686 Platforms Affymetrix Human Gene 2.0 ST Array

96   [transcript (gene) version]) and included 19 medulloblastoma samples and five normal

97   brain samples (Rivero-Hinojosa et al. 2018) (9). We selected these four gene expression

98   profiles for further integrated analyses to avoid racial differences and errors in individual

99   experiments.

Identification of DEGs in Medulloblastoma

101   The DEGs were identified based on the series matrix file using Limma package in R

102   software (version 3.5.0) according to the cut-off standard ($p < 0.05$ and logFC > 1) (Fang

103   et al. 2017) (10). The four sets of differential expression data were respectively divided

104   into upregulated DEmRNAs and downregulated DEmRNAs and were uploaded to Venny

105   (http://bioinfogp.cnb.csic.es/tools/venny/) for integrated analyses. A heat map was

106   generated using the pheatmap package in R. The GSE74195 dataset was used as the

107   reference to generate the heat map. The heatmap showed 29 upregulated DEGs, and 75

108   downregulated DEGs.

GO and KEGG enrichment analysis of DEGs

110   The DAVID online database and Kobas online tool were used for functional and

111   pathway enrichment analyses (Huang da et al. 2009) (11). We uploaded the DEGs to

112   DAVID for GO functional annotation analysis and Kobas for KEGG pathway enrichment

113   analysis. P < 0.05 was used as the cut-off value.

Building a PPI network

115   The STRING public database, which provides PPI network analysis, can evaluate

116   direct and indirect links between DEGs (http://string-db.org) (Franceschini et al. 2013)

117   (12). The DEGs were uploaded to STRING to obtain the PPI network using an

118   Interaction score of 0.2. Cytoscape was used for visualizing the PPI network and for

119   identifying the hub genes based on the degree of connectivity between DEGs.

Identification of small-molecule compounds for the treatment of MB

121 CMap, an online tool, can be based on the gene expression profile of a disease to

122 mine potential therapeutic drugs (Lamb et al. 2006) (13). We divided the DEGs of MB

123 into upregulated DEGs and downregulated DEGs, and uploaded them to CMap to explore

124 the drugs. P<0.001 and Enrichment <-0.8 was used as the cut-off criteria. The 3D

125 structures of the small-molecule compounds are available from PubChem

126 (https://pubchem.ncbi.nlm.nih.gov/).

127 **RNA isolation and reverse transcription**

128 **quantitative polymerase chain reaction (RT-qPCR) analysis**

129 Total

130 RNA is extracted from in normal human brain tissue and medulloblastoma cell lines (Dao

131 y and D283) using Tri-

132 Reagent (Sigma, USA) according to the manufacturer 's instructions. NanoDrop One spe

133 ctrophotometer (Thermo Fisher Scientific, USA) was used to evaluate the quality and qua

134 ntity of the RNA. The Transcriptor First Stand cDNA Synthesis Kit (Roche, USA) was us

135 ed to reverse transcribe total RNA into cDNA.The FastStart Universal SYBR Green Mast

136 er (ROX) (Roche, Germany) and the QuantStudio software (Thermo Fisher Scientific, U

137 SA) were reserved for RT-

138 qPCR based on the manufacturer 's instructions.The GAPDH gene was used as an endog

139 enous reference.

140 The primer sequences were as follows: 5'-GGACACGTGCTCAGAGATT-

141 3′ (sense) and 5'-TCTACGATGAGCTGTTTGTCTTC-3′(antisense) for SYN1,5'-

142 GGGGTGATGTTGCCCTGTAA-3′ (sense) and 5'-AGGTCTGAGGCATTGGTTCG-

143 3′(antisense) for CNTN2, 5'-CTGATTCTCCTGACCGTCTTTAC-3′ (sense) and 5'-

144   GAACTTGGTCTGGAAGCTGAA-3′(antisense) for FAIM2,5'-

145   CAAGTGCGAGGGATGCAAAT-3′ (sense) and 5'-TGGCACACTTCTCACACTCC-

146   3′(antisense) for MT3, 5'-CTCAGCCTAGAAGGGAATATCAAC-3′ (sense) and 5'-

147   CAGCAGGGCTGATCCATTT-3′(antisense) for SH3GL2, and 5'-

148   CACCCACTCCTCCACCTTTGA-3′ (sense) and 5'-ACCACCCTGTTGCTGTAGCCA-

149   3′ (anti- sense) for GAPDH. The results were analyzed using the -

150   ΔCT method with an unpaired t-test, and a P-

151   value <0.05 was considered a meaningful result.

152   **Results**

153   ***Screening of DEGs in MB***

154   We downloaded the gene expression datasets – GSE42656, GSE109401,

155   GSE50161, and GSE74195 from GEO. From the GSE42656 dataset, we

156   identified 869 DEGs, of which 222 were upregulated, and 647 were downregulated.

157   A total of 764 DEGs were identified from the GSE109401 data set –274 upregulated

158   and 490 downregulated; 5494 DEGs from GSE50161 – 2798 upregulated and 2696

159   downregulated, and 1000 DEGs were identified from GSE74195 – 422

160   upregulated and 578 downregulated. Integrated analyses revealed that 104 DEGs were

161   consistently expressed in the four data sets (Figure. 1A and 1B); these included 29

162   upregulated DEGs, and 75 downregulated DEGs in MB tissue compared to the normal

163   brain tissue (Table 1). A heat map of the DEG distribution was generated using the

164   GSE74195 dataset as a reference(Figure S1).

165   ***Gene Ontology Analysis of DEGs***

166   The identified DEGs were analyzed using DAVID using p <0.05 as the cut-off

167   standard to identify the functions associated with the DEGs. The GO function annotation

168   is divided into three functional groups – cell component (CC), molecular function

169   (MF), and biological process (BP). As shown in Figure. 2 and Table 2, within the BPs the

170   upregulated DEGs were closely related to mitotic cell cycle, cell cycle phase-M phase,

171   cell cycle process, cell cycle, spindle organization, and microtubule-based process and

172   the downregulated DEGs were closely related to vesicle-mediated transport. Within CC

173   the upregulated DEGs were closely related to the spindle-microtubule, and cytoskeleton,

174   whereas the downregulated DEGs were closely related to the synapse, synapse part,

175   clathrin-coated vesicle, synaptic vesicle, cytoplasmic vesicle, coated vesicle, cytoplasmic

176   membrane-bounded vesicle, vesicle, and membrane-bounded vesicle. Within MF, the

177   upregulated DEGs showed a close relationship with DNA binding and the downregulated

178   DEGs did not show any association.

179   ***KEGG pathway analysis***

180   The 104 DEGs were grouped as upregulated and downregulated DEGs and the

181   cellular signaling pathways represented were analyzed by KEGG (Figure. 3 and Table

182   3). The upregulated genes are mainly involved in the cell cycle, ubiquitin-mediated

183   proteolysis, viral carcinogenesis, one-carbon pool by folate, and microRNAs in cancer.

184   The downregulated genes are mainly involved in the synaptic vesicle cycle, bacterial

185   invasion of epithelial cells, insulin secretion, HIF-1 signaling pathway, phospholipase D

186   signaling pathway, and cell adhesion molecules (CAMs).

187

188   ***PPI network generation and identification of the hub gene***

189   The interrelationships between the 104 DEGs were analyzed using STRING. We found

190   that 100 of the 104 DEGs were related to each other and were visualized using Cytoscape

191   – 100 nodes and 776 edges are included in the network of PPI and the hub gene

192   based on the degree of connectivity between genes were identified (Figure.4A).

193   Among the hub genes, the most significant hub genes were *SYN1, CNTN2, FAIM2, MT3,*

194   and *SH3GL2* (Figure.4B).

195   ***Potential therapeutic drugs for MB***

196   Of the 104 DEGs, only 85 were eventually converted to IDs from 22,214 probes in

197   the Affymetrix platform. After that, we used CMap and identified six potential

198   therapeutic drugs based on set criteria (Table 4). A PubMed literature search revealed that

199   two drugs – vorinostat and resveratrol – have been reported to have a therapeutic

200   effect in MB. The other three drugs, trichostatin A, pyrvinium, and prochlorperazine,

201   have been reported to have therapeutic effects in other cancers. The 3D structures of

202   these potential therapeutic drugs are available from PubChem (Figure.5).

203   **Experiments verify the five hub DEmRNAs in MB**

204   To verify the expression levels of the five hub genes in the PPI network, RT-

205   qPCR was used to detect their expression levels in normal human brain tissue and medull

206   oblastoma cell lines (Daoy and D283). The five hub genes in the PPI network ,including

207   SYN1,CNTN2, FAIM2, MT3, and SH3GL2, were all down-

208   regulated DEGs in MB, and this result was confirmed by RT-qPCR

209   (Figure 6).This shows that our analysis results were completely credible.

210   **Discussion**

211   Although several studies have been carried out to understand the pathogenesis of MB,

212   there is no effective strategy to reduce the incidence or mortality of MB. This may be

213   because most studies till date have focused on a single genetic event contributing to MB

214   pathogenesis. Therefore, in order to identify effective molecular markers, we used data

215   derived from four cohorts profile datasets obtained from the GEO database (GSE42656,

216   GSE74195, GSE109401, GSE50161). Integrated analyses of the four data sets revelated

217   104 DEGs including 29 upregulated, and 75 downregulated DEGs. Enrichment of these

218   genes identified certain cellular signaling pathways which may provide novel insights to

219   the understanding of the pathogenesis of MB. We also used CMap, a drug development

220   tool, and identified six potential drugs that can be used for the treatment of MB. To our

221   knowledge, this is the first study to use integrated bioinformatical analysis for studying

222   the pathogenesis of MB.

223   GO enrichment analysis showed that the upregulated DEGs were mainly involved in

224   the spindle, cell cycle phase, M phase, cell cycle, spindle organization, microtubule

225   cytoskeleton and the downregulated DEGs were mainly closely related to clathrin-coated

226   vesicle, synaptic vesicle, and cytoplasmic vesicle. These results are consistent with

227   previous literature reports that spindle defects play a vital role in the pathological process

228   of MB (Abdelfattah et al. 2018) (14). LCL161 an inhibitor of apoptosis proteins (IAP)

229   combined with chemotherapy can slow MB cell proliferation by inducing G2/M phase

230   arrest (Chen et al. 2018) (15). Patupilone, a microtubule stabilizer, can reduce clonogenic

231   survival and enhance the therapeutic efficacy of radiotherapy effect in MB (Oehler et al.

232   2011) (16). α-synuclein binds to cytoplasmic vesicles to change the surface

233   morphology of in U251 glioblastoma cells (Duan et al. 2017) (17). Synaptic vesicle

234   protein 2A is a predictor of the efficacy of levetiracetam in glioma patients (de Groot et al.

235   2011)(18). The internalization of clathrin-coated vesicles can reduce the sensitivity of

236   metabotropic receptors in C6 cells (Luis Albasanz et al. 2002) (19).

237   KEGG pathway analysis demonstrated that the upregulated genes have significant

238   enrichment in pathways including one-carbon pool by folate, microRNAs in cancer,

239   DNA replication, and basal transcription factors. Previous studies have shown that these

240   signaling pathways are mainly involved in the development of cancer. For example, the

241   basal transcription factors such as oncogene Orthodenticle Homeobox 2, play an

242   important role in cell migration and proliferation in MB (Wortham et al. 2014) (20).

243   Folate receptor (Folr1) participates in the pathway network of the one-carbon pool by

244   folate, which is related to clinical, pathological, neuroimaging features, and prognosis of

245   MB patients (Liu et al. 2017) (21). MicroRNA-31 inhibits DNA replication by targeting

246   minichromosome maintenance complex component 2 (MCM2), which has a strong

247   inhibitory effect on MB growth (Jin et al. 2014) (22). The downregulated genes were

248   mainly involved in HIF-1 signaling pathway, phospholipase D signaling pathway, and

249   phosphatidylinositol signaling system. These signaling pathways are involved in the

250   pathological process of cancer. For example, phosphatidylinositol signaling system, such

251   as the phosphatidylinositol 3-kinase (PI3K) pathway is associated with reduced survival

252   in patients and is one of the most common signaling pathway-related abnormality in

253   gliomas (Tuncel & Kalkan 2018) (23). Phospholipase D signaling pathway specifically

254   inhibits autophagic flux and decreases GBM cell viability (Bruntz et al. 2014) (24). The

255   hypoxia-inducible factor-1 signaling pathway is vital for the invasion and metastasis of

256   glioma cells (Yaghi et al. 2016) (25). Together, our results show that the cellular

257   pathways of DEGs identified in this study are closely related to the development of

258   cancer.

259   The PPI network of DEGs represents an overview of their functional relations. Most

260   of the hub genes selected in this network have been reported to be closely related to the

261   pathological process of cancers. For example, *SYN1*, one of the neuronal genes, is related

262   to the formation and maintenance of synaptic contacts and the expression of REST.

263   RCOR1 may cause deregulated expression of *SYN1*, which contributes to the

264   maintenance of glioblastoma stem-like cells (GSC) (Yucebas et al. 2016) (26). *CNTN2*, a

265   cell adhesion protein, is a downstream protein of RACK1, which affects the growth and

266   differentiation of glioma cells through the RTK/Ras/MAPK signaling pathway (Yan &

267   Jiang 2016) (27). *FAIM2* is an anti-apoptotic molecule that promotes tumor cell growth

268   through Fas-mediated mechanisms. Knock-down of FAIM2 can significantly affect

269   tumor cell proliferation in small-cell lung cancer (Kang et al. 2016) (28). *MT3* is a tumor

270   suppressor gene, and its expression is significantly reduced in AML samples. The

271   overexpression of *MT3* can inhibit cell proliferation and promote tumor cell apoptosis in

272   pediatric acute myeloid leukemia (Tao et al. 2014) (29). *SH3GL2*, as a suppressor for

273   tumors, and has reduced expression in glioma tissues promoting migration and infiltration

274   of glioma cells by enhancing STAT3/MMP2 signaling (Zhu et al. 2017) (30). These five

275   hub genes have important regulatory effects in the pathophysiology of cancers, but their

276   role has not been reported in MB. These hub genes are potential targets for the treatment

277   of MB.

278   CMap is an online tool that can be used to identify potential therapeutic drugs based on

279   DEGs in various disease (Lamb et al. 2006) (13). In the current study, we identified six

280   drugs using CMap analysis. Two of these drugs have been previously reported to have a

281   therapeutic effect on MB. The combination of vorinostat and MLN8237 can significantly

282 inhibit the proliferation of MB cells (Muscal et al. 2013) (31). Resveratrol inhibits the

283 growth of cancer cells by regulating the Notch signaling pathway to promote apoptosis

284 and differentiation of MB cells (Wang et al. 2008) (32). Three drugs – trichostatin A,

285 pyrvinium, and prochlorperazine – have not been reported for the treatment of MB;

286 however, several studies have reported their treatment efficacies in other cancers.

287 Prochlorperazine and trichostatin A were used for treating glioblastoma and pyrvinium

288 for the treatment of ovarian cancer (Horing et al. 2013; Otreba & Buszman 2018; Zhang

289 et al. 2017) (33-35). Although the six small-molecule compounds

290 obtained have not been adequately studied in MB. CMap is an online tool that provides r

291 esearchers with an index based on disease DEGs to search for potential therapeutic drugs.

292 Recently, some traditional drugs have been found to have new therapeutic effects. The d

293 evelopment of new uses of old drugs can more quickly understand the pharmacological p

294 roperties of drugs, which will help their early application in clinical treatment. For examp

295 le, atorvastatin is a well-known traditional lipid-

296 lowering drug, but in recent years, it has been found to have obvious therapeutic effects o

297 n chronic subdural hematoma (Jiang et al. 2018). Metformin is a traditional an anti-

298 diabetic drug. In recent years, new therapeutic effects have been discovered in human dis

299 eases such as anti-tumor effects, neuroprotective effects, etc (Podhorecka et al. 2017;

300 Saewanee et al. 2019). So we have reason to speculate that the small-molecule drugs

301 identified by CMap analysis may be potential drugs for the treatment of MB.

302 **Conclusions**

303 In summary, most single data set analyses are limited by the small sample size, high

304 experimental error, and lack of ethnic differences, and therefore cannot reliably identify

305 the important genes and pathways involved in the pathogenesis of diseases. In this study,

306 the above problems were well avoided by using multiple data sets and integrated analysis

307 to improve the reliability and accuracy of the results. We identified 104 DEGs from four

308 groups of gene expression profiling data of MB and analyzed them comprehensively. The

309 hub genes identified by PPI network analysis include *SYN1, CNTN2, FAIM2, MT3, and*

310 *SH3GL2*, which are involved in the pathogenesis of different cancers. We also identified

311 several small-molecule compounds that may have potential therapeutic effects on MB.

312 These findings provide new insights into the pathogenesis of MB and provide a basis for

313 treatment. however, further experimental verification is needed.

314 **Acknowledgements**

315 Not applicable.

316 **References**

317 Abdelfattah N, Rajamanickam S, Panneerdoss S, Timilsina S, Yadav P, Onyeagucha BC,
318   Garcia M, Vadlamudi R, Chen Y, Brenner A, Houghton P, and Rao MK. 2018. MiR-584-
319   5p potentiates vincristine and radiation response by inducing spindle defects and DNA
320   damage in medulloblastoma. *Nat Commun* 9:4541. 10.1038/s41467-018-06808-8
321 Bruntz RC, Taylor HE, Lindsley CW, and Brown HA. 2014. Phospholipase D2 mediates survival
322   signaling through direct regulation of Akt in glioblastoma cells. *J Biol Chem* 289:600-616.
323   10.1074/jbc.M113.532978
324 Chen SM, Lin TK, Tseng YY, Tu CH, Lui TN, Huang SF, Hsieh LL, and Li YY. 2018. Targeting
325   inhibitors of apoptosis proteins suppresses medulloblastoma cell proliferation via G2/M
326   phase arrest and attenuated neddylation of p21. *Cancer Med* 7:3988-4003.
327   10.1002/cam4.1658
328 de Groot M, Aronica E, Heimans JJ, and Reijneveld JC. 2011. Synaptic vesicle protein 2A
329   predicts response to levetiracetam in patients with glioma. *Neurology* 77:532-539.
330   10.1212/WNL.0b013e318228c110
331 Duan J, Ying Z, Su Y, Lin F, and Deng Y. 2017. alpha-Synuclein binds to cytoplasmic vesicles
332   in U251 glioblastoma cells. *Neurosci Lett* 642:148-152. 10.1016/j.neulet.2017.01.067
333 Fang C, Huang Y, Pei Y, Zhang HH, Chen X, Guo H, Li S, Ji X, and Hu J. 2017. Genome-wide
334   gene expression profiling reveals that CD274 is up-regulated new-onset type 1 diabetes
335   mellitus. *Acta Diabetol* 54:757-767. 10.1007/s00592-017-1005-y
336 Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork
337   P, von Mering C, and Jensen LJ. 2013. STRING v9.1: protein-protein interaction
338   networks, with increased coverage and integration. *Nucleic Acids Res* 41:D808-815.
339   10.1093/nar/gks1094
340 Griesinger AM, Birks DK, Donson AM, Amani V, Hoffman LM, Waziri A, Wang M, Handler MH,
341   and Foreman NK. 2013. Characterization of distinct immunophenotypes across pediatric
342   brain tumor types. *J Immunol* 191:4880-4888. 10.4049/jimmunol.1301966

343 Henriquez NV, Forshew T, Tatevossian R, Ellis M, Richard-Loendt A, Rogers H, Jacques TS,
344   Reitboeck PG, Pearce K, Sheer D, Grundy RG, and Brandner S. 2013. Comparative
345   expression analysis reveals lineage relationships between human and murine gliomas
346   and a dominance of glial signatures during tumor propagation in vitro. *Cancer Res*
347   73:5834-5844. 10.1158/0008-5472.CAN-13-1299
348 Horing E, Podlech O, Silkenstedt B, Rota IA, Adamopoulou E, and Naumann U. 2013. The
349   histone deacetylase inhibitor trichostatin a promotes apoptosis and antitumor immunity
350   in glioblastoma cells. *Anticancer Res* 33:1351-1360.
351 Huang da W, Sherman BT, and Lempicki RA. 2009. Systematic and integrative analysis of large
352   gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44-57.
353   10.1038/nprot.2008.211
354 Jiang R, Zhao S, Wang R, Feng H, Zhang J, Li X, Mao Y, Yuan X, Fei Z, Zhao Y, Yu X, Poon
355   WS, Zhu X, Liu N, Kang D, Sun T, Jiao B, Liu X, Yu R, Zhang J, Gao G, Hao J, Su N,
356   Yin G, Zhu X, Lu Y, Wei J, Hu J, Hu R, Li J, Wang D, Wei H, Tian Y, Lei P, Dong JF, and
357   Zhang J. 2018. Safety and Efficacy of Atorvastatin for Chronic Subdural Hematoma in
358   Chinese Patients: A Randomized ClinicalTrial. *JAMA Neurol* 75:1338-1346.
359   10.1001/jamaneurol.2018.2030
360 Jin Y, Xiong A, Zhang Z, Li S, Huang H, Yu TT, Cao X, and Cheng SY. 2014. MicroRNA-31
361   suppresses medulloblastoma cell growth by inhibiting DNA replication through
362   minichromosome maintenance 2. *Oncotarget* 5:4821-4833. 10.18632/oncotarget.2043
363 Kang HC, Kim JI, Chang HK, Woodard G, Choi YS, Ku JL, Jablons DM, and Kim IJ. 2016.
364   FAIM2, as a novel diagnostic maker and a potential therapeutic target for small-cell lung
365   cancer and atypical carcinoid. *Sci Rep* 6:34022. 10.1038/srep34022
366 Kim SY, Sung KW, Hah JO, Yoo KH, Koo HH, Kang HJ, Park KD, Shin HY, Ahn HS, Im HJ, Seo
367   JJ, Lim YJ, Lee YH, Shin HJ, Lim DH, Cho BK, Ra YS, and Choi JU. 2010. Reduced-
368   dose craniospinal radiotherapy followed by high-dose chemotherapy and autologous
369   stem cell rescue for children with newly diagnosed high-risk medulloblastoma or
370   supratentorial primitive neuroectodermal tumor. *Korean J Hematol* 45:120-126.
371   10.5045/kjh.2010.45.2.120
372 Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP,
373   Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ,
374   Clemons PA, Wei R, Carr SA, Lander ES, and Golub TR. 2006. The Connectivity Map:
375   using gene-expression signatures to connect small molecules, genes, and disease.
376   *Science* 313:1929-1935. 10.1126/science.1132939
377 Liu H, Sun Q, Zhang M, Zhang Z, Fan X, Yuan H, Li C, Guo Y, Ning W, Sun Y, Song Y, and Yu
378   C. 2017. Differential expression of folate receptor 1 in medulloblastoma and the
379   correlation with clinicopathological characters and target therapeutic potential.
380   *Oncotarget* 8:23048-23060. 10.18632/oncotarget.15480
381 Luis Albasanz J, Fernandez M, and Martin M. 2002. Internalization of metabotropic glutamate
382   receptor in C6 cells through clathrin-coated vesicles. *Brain Res Mol Brain Res* 99:54-66.
383   10.1016/s0169-328x(02)00103-1
384 Muscal JA, Scorsone KA, Zhang L, Ecsedy JA, and Berg SL. 2013. Additive effects of vorinostat
385   and MLN8237 in pediatric leukemia, medulloblastoma, and neuroblastoma cell lines.
386   *Invest New Drugs* 31:39-45. 10.1007/s10637-012-9831-9
387 Oehler C, von Bueren AO, Furmanova P, Broggini-Tenzer A, Orlowski K, Rutkowski S, Frei K,
388   Grotzer MA, and Pruschy M. 2011. The microtubule stabilizer patupilone (epothilone B)
389   is a potent radiosensitizer in medulloblastoma cells. *Neuro Oncol* 13:1000-1010.
390   10.1093/neuonc/nor069
391 Otreba M, and Buszman E. 2018. Perphenazine and prochlorperazine induce concentration-
392   dependent loss in human glioblastoma cells viability. *Pharmazie* 73:19-21.
393   10.1691/ph.2018.7806

394 Podhorecka M, Ibanez B, and Dmoszynska A. 2017. Metformin - its potential anti-cancer and
395   anti-aging effects. *Postepy Hig Med Dosw (Online)* 71:170-175.
396   10.5604/01.3001.0010.3801
397 Rivero-Hinojosa S, Lau LS, Stampar M, Staal J, Zhang H, Gordish-Dressman H, Northcott PA,
398   Pfister SM, Taylor MD, Brown KJ, and Rood BR. 2018. Proteomic analysis of
399   Medulloblastoma reveals functional biology with translational potential. *Acta Neuropathol*
400   *Commun* 6:48. 10.1186/s40478-018-0548-7
401 Saewanee N, Praputpittaya T, Malaiwong N, Chalorak P, and Meemon K. 2019.
402   Neuroprotective effect of metformin on dopaminergic neurodegeneration and alpha-
403   synuclein aggregation in C. elegans model of Parkinson's disease. *Neurosci Res*.
404   10.1016/j.neures.2019.12.017
405 Shaabanpour Aghamaleki F, Mollashahi B, Aghamohammadi N, Rostami N, Mazloumi Z,
406   Mirzaei H, Moradi A, Sheikhpour M, and Movafagh A. 2019. Bioinformatics Analysis of
407   Key Genes and Pathways for Medulloblastoma as a Therapeutic Target. *Asian Pac J*
408   *Cancer Prev* 20:221-227. 10.31557/APJCP.2019.20.1.221
409 Tao YF, Xu LX, Lu J, Cao L, Li ZH, Hu SY, Wang NN, Du XJ, Sun LC, Zhao WL, Xiao PF, Fang
410   F, Li YH, Li G, Zhao H, Li YP, Xu YY, Ni J, Wang J, Feng X, and Pan J. 2014.
411   Metallothionein III (MT3) is a putative tumor suppressor gene that is frequently
412   inactivated in pediatric acute myeloid leukemia by promoter hypermethylation. *J Transl*
413   *Med* 12:182. 10.1186/1479-5876-12-182
414 Tuncel G, and Kalkan R. 2018. Receptor tyrosine kinase-Ras-PI 3 kinase-Akt signaling network
415   in glioblastoma multiforme. *Med Oncol* 35:122. 10.1007/s12032-018-1185-5
416 Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., and Kinzler KW. 2013.
417   Cancer genome landscapes. *Science* 339:1546-1558. 10.1126/science.1235122
418 Wang J, Garancher A, Ramaswamy V, and Wechsler-Reya RJ. 2018. Medulloblastoma: From
419   Molecular Subgroups to Molecular Targeted Therapies. *Annu Rev Neurosci* 41:207-232.
420   10.1146/annurev-neuro-070815-013838
421 Wang Q, Li H, Liu N, Chen XY, Wu ML, Zhang KL, Kong QY, and Liu J. 2008. Correlative
422   analyses of notch signaling with resveratrol-induced differentiation and apoptosis of
423   human medulloblastoma cells. *Neurosci Lett* 438:168-173. 10.1016/j.neulet.2008.04.012
424 Wortham M, Guo C, Zhang M, Song L, Lee BK, Iyer VR, Furey TS, Crawford GE, Yan H, and
425   He Y. 2014. Chromatin accessibility mapping identifies mediators of basal transcription
426   and retinoid-induced repression of OTX2 in medulloblastoma. *PLoS One* 9:e107156.
427   10.1371/journal.pone.0107156
428 Xi Y, Tang W, Yang S, Li M, He Y, and Fu X. 2017. Mining the glioma susceptibility genes in
429   children from gene expression profiles and a methylation database. *Oncol Lett* 14:3473-
430   3479. 10.3892/ol.2017.6579
431 Yaghi L, Poras I, Simoes RT, Donadi EA, Tost J, Daunay A, de Almeida BS, Carosella ED, and
432   Moreau P. 2016. Hypoxia inducible factor-1 mediates the expression of the immune
433   checkpoint HLA-G in glioma cells through hypoxia response element located in exon 2.
434   *Oncotarget* 7:63690-63707. 10.18632/oncotarget.11628
435 Yan Y, and Jiang Y. 2016. RACK1 affects glioma cell growth and differentiation through the
436   CNTN2-mediated RTK/Ras/MAPK pathway. *Int J Mol Med* 37:251-257.
437   10.3892/ijmm.2015.2421
438 Yucebas M, Yilmaz Susluer S, Onur Caglar H, Balci T, Dogan Sigva ZO, Akalin T, Oktar N,
439   Dalbasti T, Biray Avci C, and Gunduz C. 2016. Expression profiling of RE1-silencing
440   transcription factor (REST), REST corepressor 1 (RCOR1), and Synapsin 1 (SYN1)
441   genes in human gliomas. *J BUON* 21:964-972.
442 Zeng WJ, Yang YL, Liu ZZ, Wen ZP, Chen YH, Hu XL, Cheng Q, Xiao J, Zhao J, and Chen XP.
443   2018. Integrative Analysis of DNA Methylation and Gene Expression Identify a Three-

444   Gene Signature for Predicting Prognosis in Lower-Grade Gliomas. *Cell Physiol Biochem*
445   47:428-439. 10.1159/000489954
446 Zhang C, Zhang Z, Zhang S, Wang W, and Hu P. 2017. Targeting of Wnt/beta-Catenin by
447   Anthelmintic Drug Pyrvinium Enhances Sensitivity of Ovarian Cancer Cells to
448   Chemotherapy. *Med Sci Monit* 23:266-275. 10.12659/msm.901667
449 Zhu Y, Zhang X, Wang L, Ji Z, Xie M, Zhou X, Liu Z, Shi H, and Yu R. 2017. Loss of SH3GL2
450   promotes the migration and invasion behaviours of glioblastoma cells through activating
451   the STAT3/MMP2 signalling. *J Cell Mol Med* 21:2685-2694. 10.1111/jcmm.13184
452

Figure 1

Using the Venny map to obtain common DEGs in MB, and the cross areas represented the commonly DEGs of the four datasets (GSE42656, GSE109401, GSE50161, and GSE74195).

A) Commonly up-regulated genes(29 DEGs); B) Commonly down-regulated genes(75 DEGs).

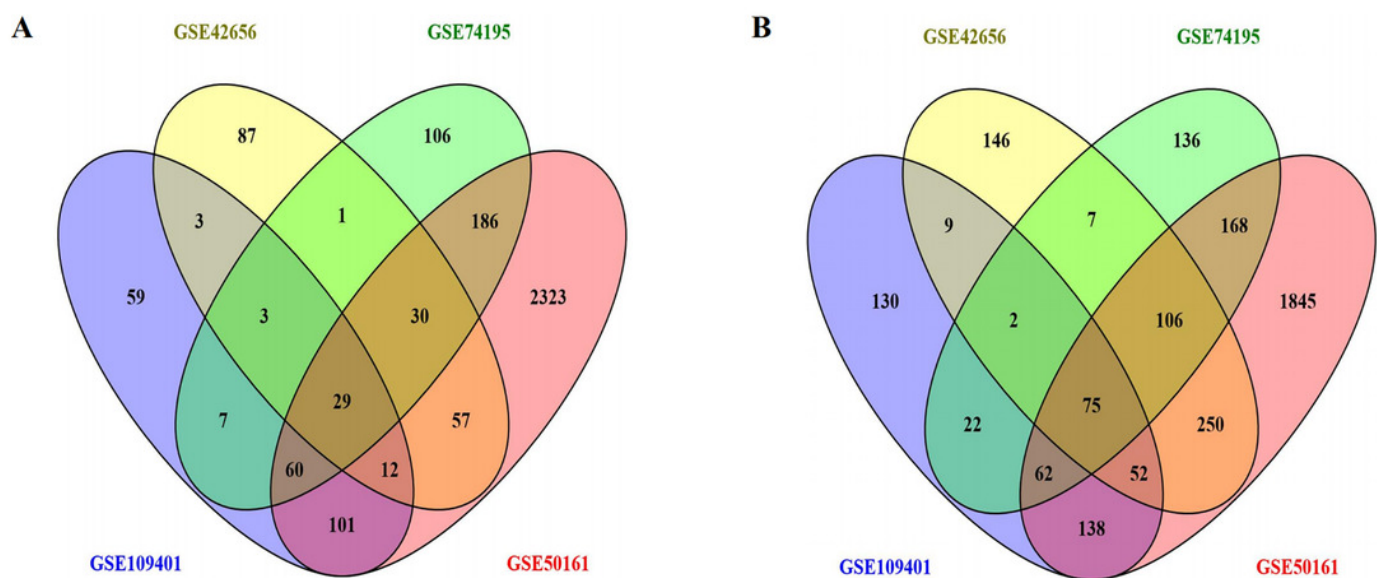



Figure 2

Gene Ontology (GO) functional annotation analysis and significant terms of DEGs in MB
and GO analysis divided into three parts.

cyan for cell component , pink for molecular function , and blue for biological process

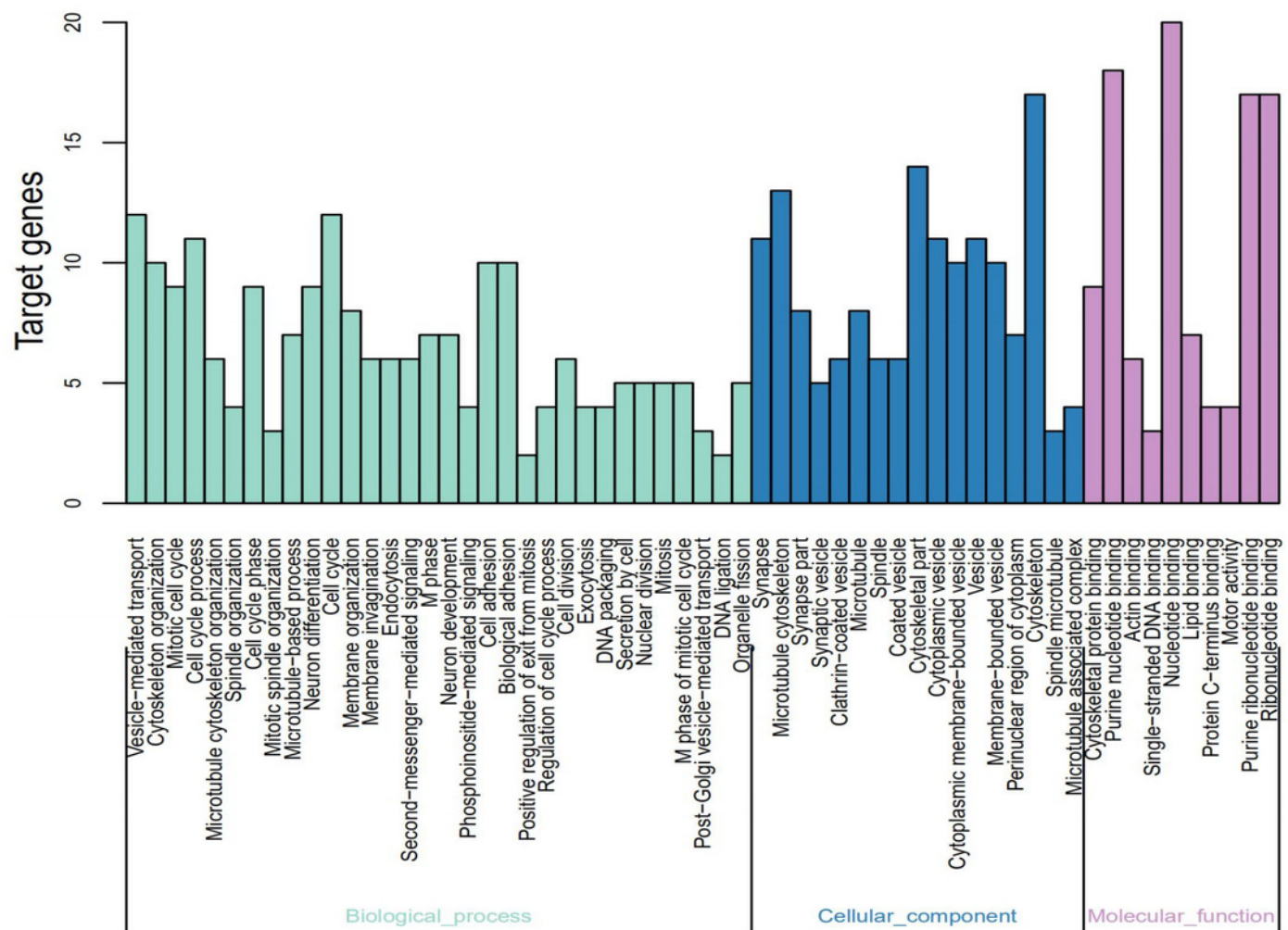



Figure 3

Kyoto Encyclopedia of Genes and Genomes (KEGG) signal pathway analysis of DEGs in MB.

Red represent the common up-regulated genes and green represent the common down-regulated genes

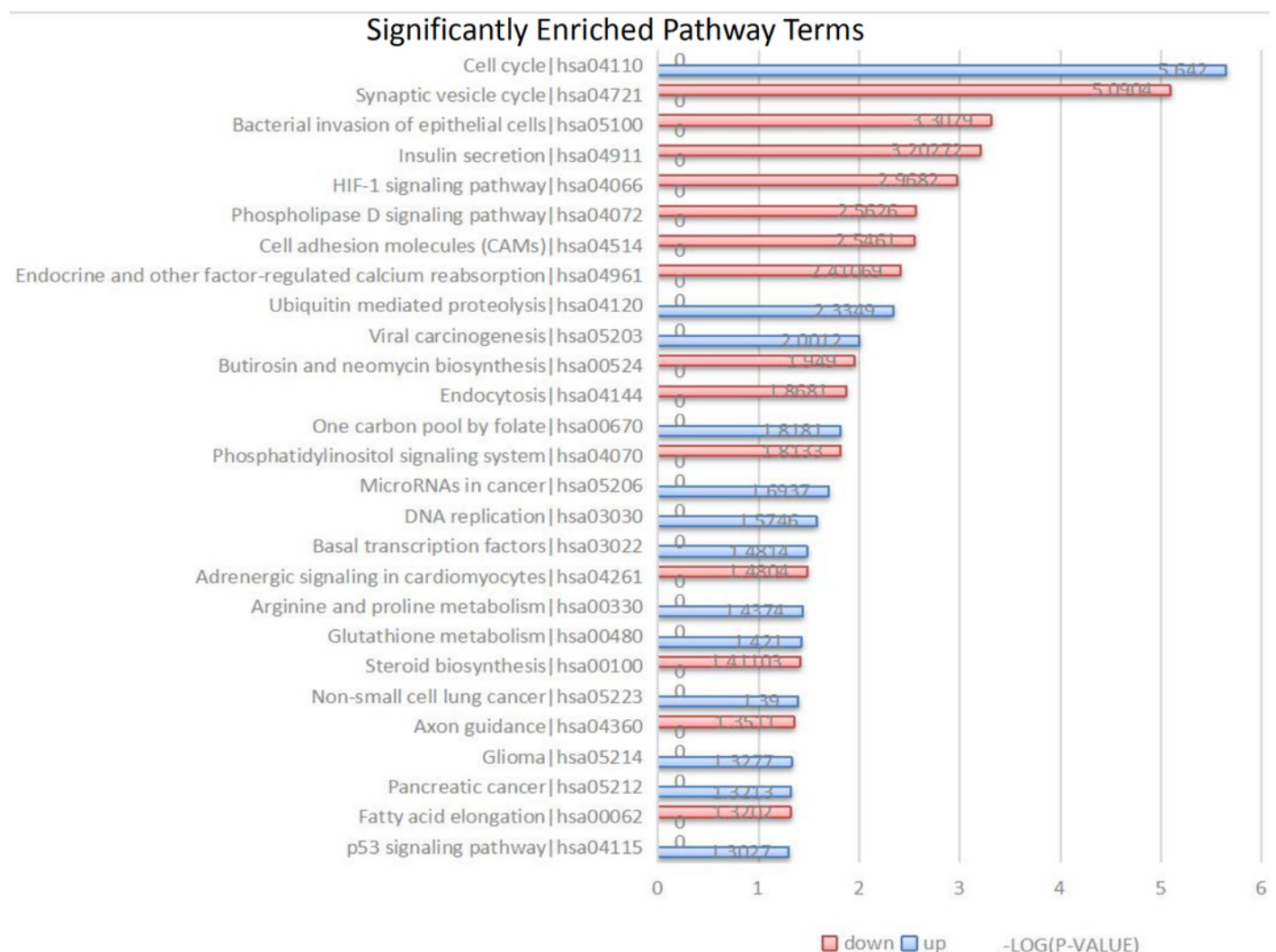



Figure 4

The PPI network of DEGs.

(A) 100 DEGs were incorporated into the network consisted of 29 upregulated (red) and 71 downregulated (green) genes; (B) Hub genes determined by the degree of connectivity between DEGs.

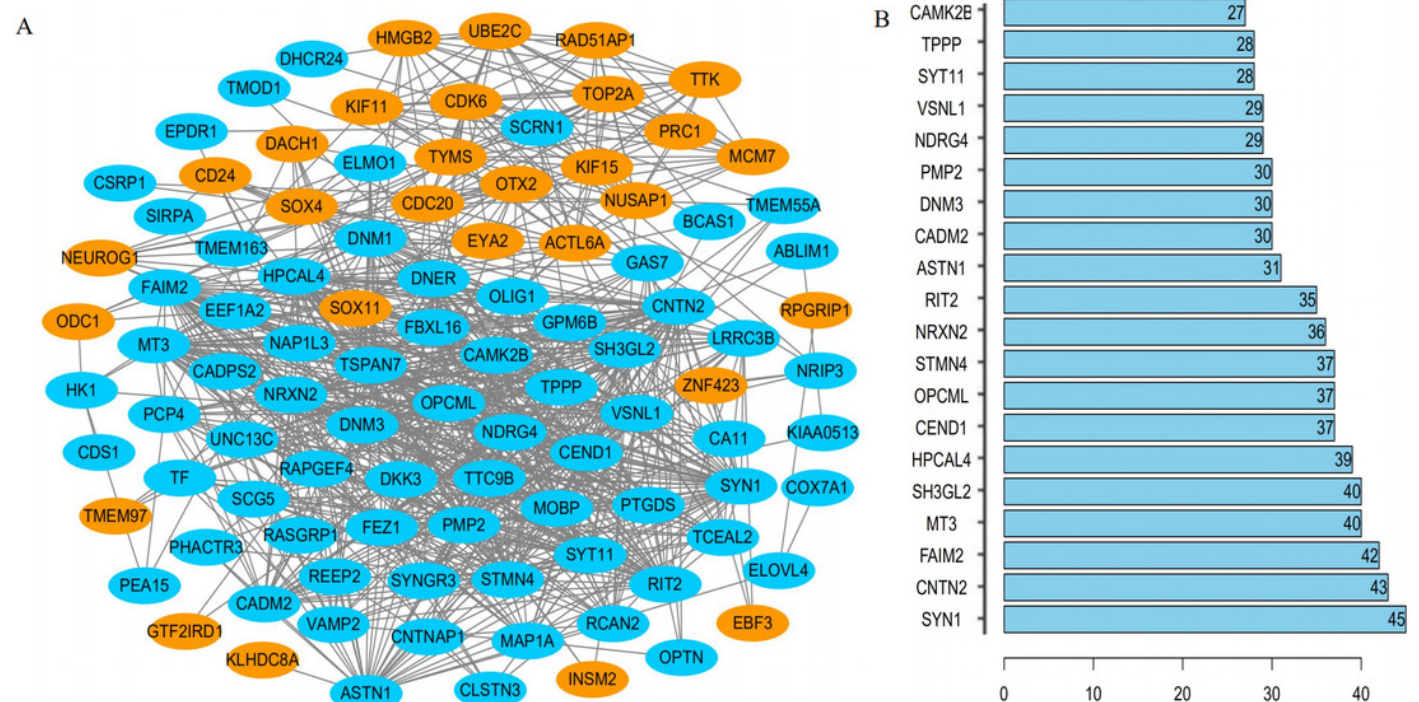



Figure 5

The 3D structures of five small-molecule compounds identified as potential drugs for MB treatment.

(A) Vorinostat; (B) Resveratrol; (C) Trichostatin A; (D) Pyrvinium, and (E) Prochlorperazine.

A   Vorinostat
MolecularFormula:$C_{14}H_{20}N_2O_3$
PubChem CID: 5311

B   Resveratrol
Molecular Formula:$C_{14}H_{12}O_3$
PubChem CID: 445154

C   Trichostatin A
Molecular Formula:$C_{17}H_{22}N_2O_3$
PubChem CID: 444732

D   Pyrvinum
Molecular Formula:$C_{26}H_{28}N_3^+$
PubChem CID:5281035

E   Prochlorperazine
Molecular Formula:$C_{20}H_{24}C_1N_3S$
PubChem CID:4917

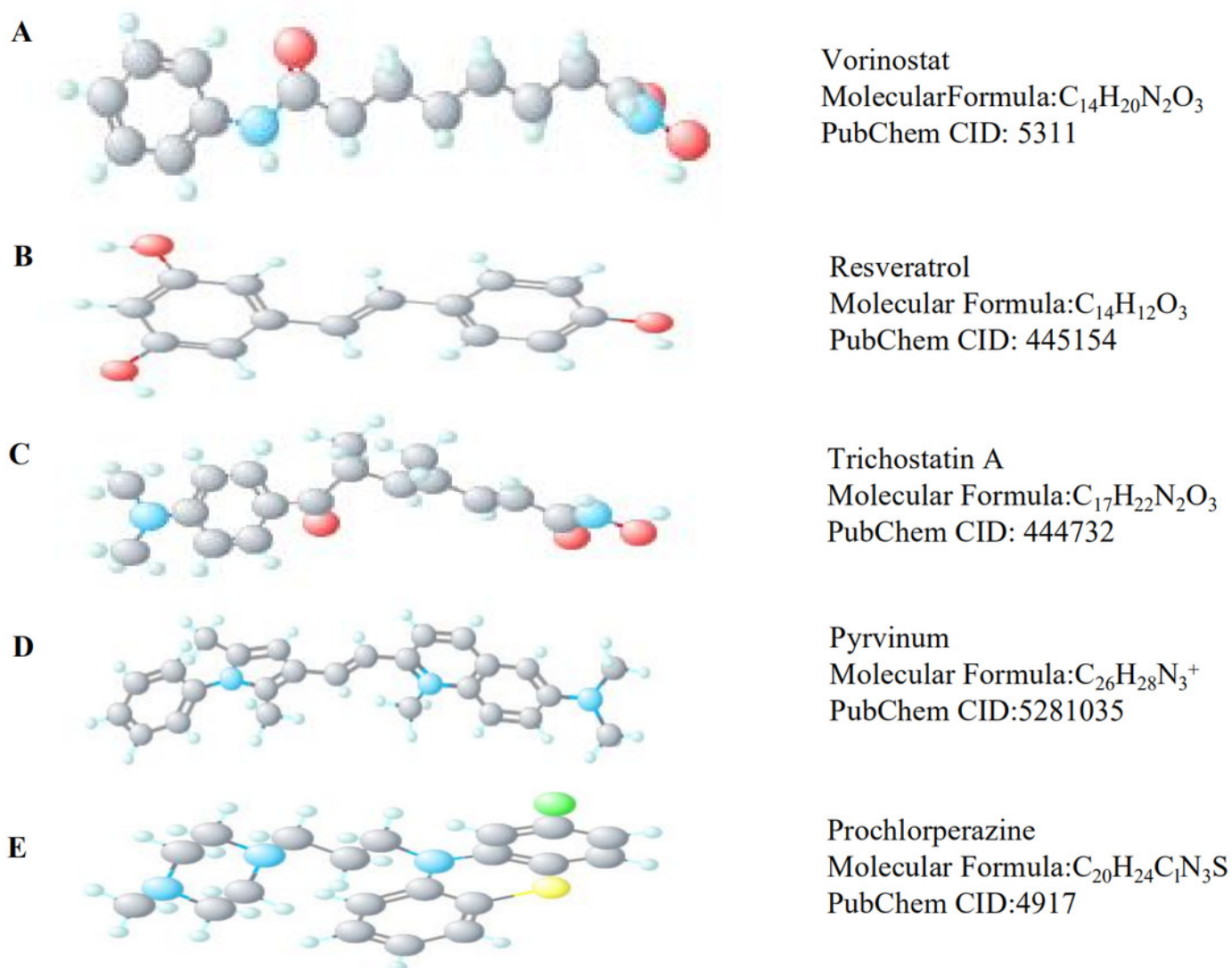
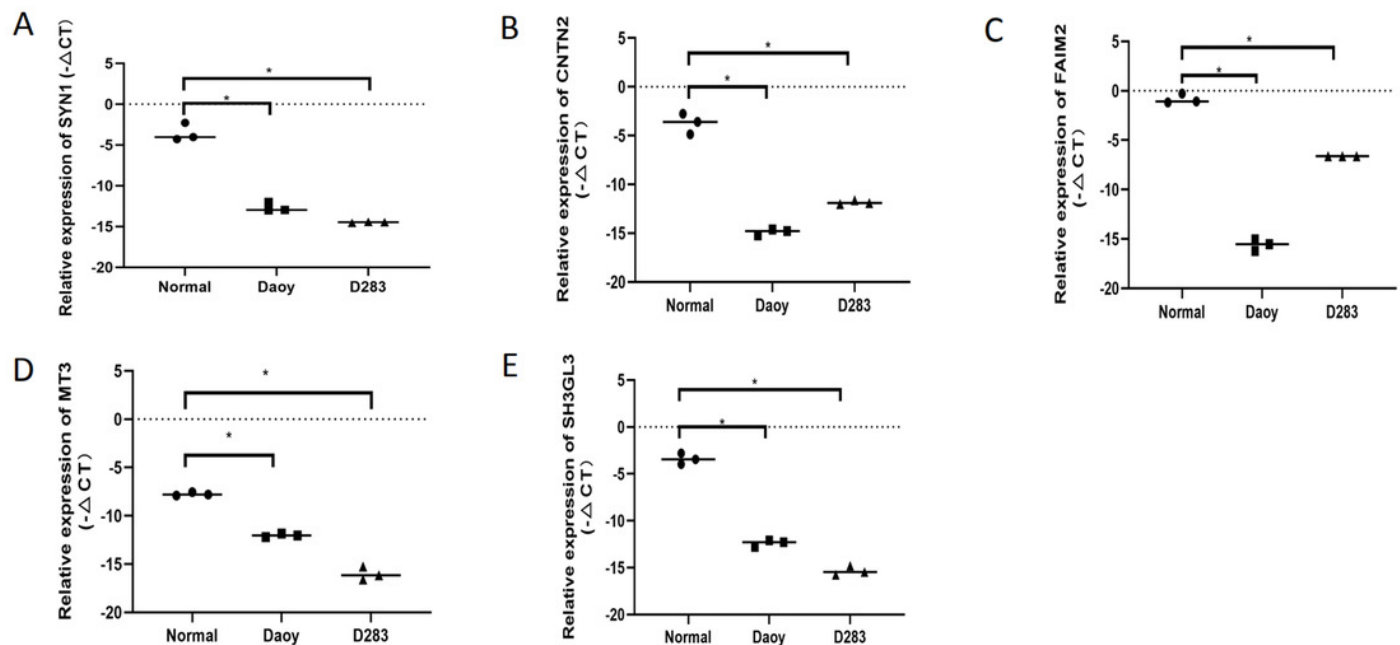



Figure 6

TheexpressionlevelsofthefivehubgenesinthePPInetworkweredetectedby RT-qPCR.

TheseincludedSYN1,CNTN2,FAIM2,MT3,andSH3GL2(P<0.05).

The 104 differentially expressed genes (DEGs), including 29 upregulated genes and 75 downregulated genes were identified in the medulloblastoma tissues from four profile datasets using normal brain tissues as a reference.

Abbreviation:DEGs,differentially expressed genes.

1 **Table** I .The 104 differentially expressed genes (DEGs), including 29 upregulated genes and 75
2 downregulated genes were identified in the medulloblastoma tissues from four profile datasets
3 using normal brain tissues as a reference.

DEGs	Gene Name
Up-regulated 29	ACTL6A,UBE2C,TMEM97,GTF2IRD1,INSM2,CDC20,PRC1,ZNF423,HMGB2,ODC1,MCM7,KLHDC8A,SOX11,RAD51AP1,KIF15,CDK6,EBF3,EYA2,TYMS,TTK,CD24,DACH1,SOX4,NUSAP1,RPGRIP1,KIF11,NEUROG1,OTX2,TOP2A
Down-regulated 75	CADPS2,TMEM163,UNC13C,DNER,CNTNAP1,DNM3,TMOD1,FEZ1,RAPGEF4,RIT2,NDRG4,TPPP,TMEM55A,MT3,FAIM2,ABLIM1,RCAN2,MAP1A,DHCR24,NRXN2,PTGDS,CDS1,RASGRP1,PCP4,NRIP3,HPCAL4,PTCHD1,GAS7,KIAA0513,PMP2,PHACTR3,TF,CADM2,CNTN2,VSNL1,BCAS1,DKK3,SCRN1,FBXL16,ELOVL4,OPCML,DNM1,EEF1A2,TTC9B,CA11,CEND1,MOBP,SH3GL2,SCG5,EPDR1,LRRC3B,COX7A1,NAP1L3,TSPAN7,CAMK2B,OPTN,STMN4,CLSTN3,PEA15,SYN1,RNF175,REEP2,CSRP1,SYT11,TCEAL2,GPM6B,TPRG1L,ASTN1,MAP7D2,SIRPA,OLIG1,SYNGR3,ELMO1,HK1,VAMP2

4 Abbreviation:DEGs,differentially expressed genes.

DEGs divided into upregulated genes and downregulated genes and their first ten meaningful GO enrichment analyses.

Abbreviation: GO, Gene Ontology; DEGs, differentially expressed genes.

1 **Table Ⅱ.** DEGs divided into upregulated genes and downregulated genes and their first ten

2 meaningful GO enrichment analyses.

Term	Description	Count	P-Value
Up-regulated			
GO:0005819	Spindle	6	1.92E-06
GO:0000278	Mitotic cell cycle	8	5.99E-06
GO:0003677	DNA binding	15	1.12E-05
GO:0022403	Cell cycle phase	8	1.25E-05
GO:0000279	M phase	7	3.80E-05
GO:0022402	Cell cycle process	8	9.14E-05
GO:0007049	Cell cycle	9	9.45E-05
GO:0007051	Spindle organization	4	9.51E-05
GO:0015630	Microtubule cytoskeleton	7	1.03E-04
GO:0007017	Microtubule-based process	6	1.27E-04
Down-regulated			
GO:0045202	Synapse	11	5.27E-06
GO:0016192	Vesicle-mediated transport	11	5.69E-05
GO:0044456	Synapse part	8	1.45E-04
GO:0030136	Clathrin-coated vesicle	6	3.77E-04
GO:0008021	Synaptic vesicle	5	4.37E-04
GO:0031410	Cytoplasmic vesicle	11	7.37E-04
GO:0030135	Coated vesicle	6	8.80E-04
GO:0016023	Cytoplasmic membrane-bounded vesicle	10	9.86E-04
GO:0031982	Vesicle	11	0.001022664
GO:0031988	Membrane-bounded vesicle	10	0.001235649

3 Abbreviation: GO, Gene Ontology; DEGs, differentially expressed genes.

KEGG pathway analysis of upregulated and downregulated genes in MB

Abbreviation: KEGG, Kyoto Encyclopedia of Genes and Genomes; ID,identification.

1 **Table** III. KEGG pathway analysis of upregulated and downregulated genes in MB

Database	ID	Term	Gene Count	P-Value	Gene Names
Up-regulated					
KEGG PATHWAY	hsa04110	Cell cycle	4	2.28E-06	TTK、CDC20、MCM7、CDK6
KEGG PATHWAY	hsa04120	Ubiquitin mediated proteolysis	2	0.004624353	CDC20、UBE2C
KEGG PATHWAY	hsa05203	Viral carcinogenesis	2	0.009970546	CDC20、CDK6
KEGG PATHWAY	hsa00670	One carbon pool by folate	1	0.015198488	TYMS
KEGG PATHWAY	hsa05206	MicroRNAs in cancer	2	0.020243092	CDK6、SOX4
KEGG PATHWAY	hsa03030	DNA replication	1	0.026628344	MCM7
KEGG PATHWAY	hsa03022	Basal transcription factors	1	0.033001239	GTF2IRD1
KEGG PATHWAY	hsa00330	Arginine and proline metabolism	1	0.03652431	ODC1
KEGG PATHWAY	hsa00480	Glutathione metabolism	1	0.037930065	ODC1
KEGG PATHWAY	hsa05223	Non-small cell lung cancer	1	0.040735637	CDK6
Down-regulated					
KEGG PATHWAY	hsa04721	Synaptic vesicle cycle	4	8.12E-06	DNM3、UNC13C、DNM1、VAMP2
KEGG PATHWAY	hsa05100	Bacterial invasion of epithelial cells	3	0.000492126	DNM3、DNM1、ELMO1
KEGG PATHWAY	hsa04911	Insulin secretion	3	0.000627009	RAPGEF4、CAMK2B、VAMP2
KEGG PATHWAY	hsa04066	HIF-1 signaling pathway	3	0.001075892	HK1、TF、CAMK2B
KEGG PATHWAY	hsa04072	Phospholipase D signaling pathway	3	0.002737401	DNM3、RAPGEF4、DNM1
KEGG PATHWAY	hsa04514	Cell adhesion molecules (CAMs)	3	0.002843803	NRXN2、CNTN2、CNTNAP1
KEGG PATHWAY	hsa04961	Endocrine and other Factor-regulated calcium reabsorption	2	0.003884233	DNM3、DNM1
KEGG PATHWAY	hsa00524	Butirosin and neomycin biosynthesis	1	0.011244262	HK1
KEGG PATHWAY	hsa04144	Endocytosis	3	0.01354801	DNM3、DNM1、SH3GL2
KEGG PATHWAY	hsa04070	Phosphatidylinositol signaling system	2	0.0153706	CDS1、TMEM55A

3 Abbreviation: KEGG, Kyoto Encyclopedia of Genes and Genomes; ID,identification.

4 .

Six small-molecule compounds identified as potential drugs for MB treatment by CMap analysis

Abbreviation: CMap, connectivity map.

1 **Table Ⅳ.** Six small-molecule compounds identified as potential drugs for MB treatment by

2 CMap analysis

Term	Enrichment	P-Value
Vorinostat	-0.842	0
Resveratrol	-0.827	0
Trichostatin A	-0.568	0
Pyrvinium	-0.727	0.00077
Prochlorperazine	-0.475	0.00082
0175029-0000	-0.718	0.00099

3 Abbreviation: CMap, connectivity map.