Identification of differentially expressed key genes between glioblastoma and low grade glioma by bioinformatics analysis

Yang Xu [1] , Rongxin Geng [2] , Fan'en Yuan [2] , Qian Sun [2] , Baohui Liu [Corresp., 2] , Qianxue Chen [Corresp. 2]

[1] Renmin Hospital of Wuhan University, Wuhan, China

[2] Renmin Hospital of Wuhan University, Wuhan, Hubei, China

Corresponding Authors: Baohui Liu, Qianxue Chen
Email address: bliu666@whu.edu.cn, chenqx666@whu.edu.cn

Gliomas are a very diverse group of brain tumors that are most commonly primary tumor and difficult to cure in central nervous system. It's necessary to distinguish low-grade tumors from high-grade tumors by understanding the molecular basis of different grades of glioma, which is an important step in defining new biomarkers and therapeutic strategies. We have chosen gene expression profile of GSE52009 from GEO database to detect important differential genes. GSE52009 contains 120 samples, including 60 WHO II samples and 24 WHO IV samples, were selected in our analysis. We used GEO2R tool to pick out differently expressed genes (DEGs) between low-grade glioma and high-grade glioma, and then we used The Database for Annotation, Visualization and Integrated Discovery (DAVID) to perform gene ontology (GO) analysis and Kyoto Encyclopedia of Gene and Genome (KEGG) pathway analysis. Furthermore, Cytoscape with Search Tool for the Retrieval of Interacting Genes (STRING) and Molecular Complex Detection (MCODE) plug-in were applied to achieve the visualization of protein-protein interaction (PPI). We selected 15 hub genes with higher degree of connectivity, including TIMP1 and SAA1 and we used GSE53733 containing 70 glioblastoma samples to conduct Gene Set Enrichment Analysis on. In conclusion, DEGs and hub genes were demonstrated in our bioinformatics analysis, that they might be defined as new biomarkers for diagnosis and to guide the therapeutic strategies of glioblastoma.

Identification of differentially expressed key genes between glioblastoma and low-grade glioma by bioinformatics analysis

Yang Xu [1,2], Rongxin Geng [1,2], Fan'en Yuan [1,2], Qian Sun [1,2], Baohui Liu [1,2*], Qianxue Chen [1,2*]

[1]*Department of Neurosurgery, Renmin Hospital of Wuhan University, Hubei 430060, China*

[2]*Brain Tumor Clinical Center of Wuhan, Hubei 430060, China*

Correspondence should be addressed to Baohui Liu; bliu666@whu.edu.cn and Qianxue Chen; chenqx666@whu.edu.cn

Abstract

Gliomas are a very diverse group of brain tumors that are most commonly primary tumor and difficult to cure in central nervous system. It's necessary to distinguish low-grade tumors from high-grade tumors by understanding the molecular basis of different grades of glioma, which is an important step in defining new biomarkers and therapeutic strategies. We have chosen gene expression profile of GSE52009 from GEO database to detect important differential genes. GSE52009 contains 120 samples, including 60 WHO II samples and 24 WHO IV samples, were selected in our analysis. We used GEO2R tool to pick out differently expressed genes (DEGs) between low-grade glioma and high-grade glioma, and then we used The Database for Annotation, Visualization and Integrated Discovery (DAVID) to perform gene ontology (GO) analysis and Kyoto Encyclopedia of Gene and Genome (KEGG) pathway analysis. Furthermore, Cytoscape with Search Tool for the Retrieval of Interacting Genes (STRING) and Molecular Complex Detection (MCODE) plug-in were applied to achieve the visualization of protein-protein

27   interaction (PPI). We selected 15 hub genes with higher degree of connectivity, including TIMP1

28   and SAA1 and we used GSE53733 containing 70 glioblastoma samples to conduct Gene Set

29   Enrichment Analysis on. In conclusion, DEGs and hub genes were demonstrated in our

30   bioinformatics analysis, that they might be defined as new biomarkers for diagnosis and to guide

31   the therapeutic strategies of glioblastoma.

33   **1. Introduction**

34

35   Gliomas are a very diverse group of brain tumors that are most commonly primary tumor and

36   difficult to cure in central nervous system (Louis et al. 2016; Ostrom et al. 2017). Despite the

37   multiple therapeutic strategies, including surgery, radiation, and chemotherapy, the average

38   survival time of glioblastoma patients is less than 15 months (Liu et al. 2016). Additionally,

39   approximately 70% of low-grade glioma patients develop to grade IV glioblastoma multiforme

40   (GBM) within 5–10 years as ending (Furnari et al. 2007). With the development of molecular

41   pathology in gliomas, several biomarkers routinely applied to evaluate gliomas include MGMT

42   promoter methylation, EGFR alterations, IDH1 or IDH2 mutations, and 1p19q co-deletion and

43   many of these markers have become standard of care for molecular testing and prerequisites for

44   clinical trial enrollment(Rodriguez et al. 2016). Therefore, it's necessary to differentiate low-grade

45   tumors to glioblastomas by understanding the molecular basis of different grades of glioma, which

46   is an important step in defining new biomarkers and therapeutic strategies.

47   Gene expression profiling analysis is a useful method with broad clinical application for

48   identifying tumor-related genes in various types of cancer, from molecular diagnosis to

49   pathological classification, from therapeutic evaluation to prognosis prediction, and from drug

50   sensitivity to neoplasm recurrence(De Preter et al. 2010; Freije et al. 2004; Kim et al. 2011;

51   Kulasingam & Diamandis 2008). In recent years, large scales of gene profiling have been made to

52   identify the overwhelming number of genes by the use of microarrays in clinical practice, and

53   complicated and systemic statistical analyses should be made to provide both repeatability and

54 independent validation (Cheng et al. 2016).

55   In this analysis, GEO2R online tool was applied to look for the differentially expressed genes

56 (DEGs) according to GSE52009 from Gene Expression Omnibus (GEO). Followed by, we

57 produced a heatmap and picked out 15 genes with higher degree of connectivity from the DEGs

58 selected. Furthermore, we analyzed cellular component (CC), biological process (BP), molecular

59 function (MF) and KEGG pathways of the DEGs. In addition, the overall survival (OS) analysis

60 and expression of these hub genes were mad e online. Then, we established PPI network of the

61 DEGs and managed a GSEA using GBM patient gene profiling data (GSE53733).

62

63 **2. Materials and Methods**

64

65 *2.1. Data of Microarray.* Gene expression profile of GSE52009 and GSE53733 were downloaded

66 from GEO database, which is a public and freely accessible database. Based on Agilent GPL6480

67 platform (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F), GSE52009

68 dataset included 120 samples, containing 60 WHO II samples and 24 WHO IV samples. And

69 GSE53733 was based on GPL570 platform ([HG-U133_Plus_2] Affymetrix Human Genome

70 U133 Plus 2.0 Array), which contained70 glioblastoma samples.

71

72 *2.2. Screen Genes of Differential Expression.* Differentially expressed genes between low grade

73 glioma and high-grade glioma was detected by GEO2R, which was an online analysis tool based

74 on R language (Davis & Meltzer 2007). We set the adjust P value < 0.05 and |logFC| ≥ 2 as the

75 selection criteria to decrease the false positive rate and false discovery rate. Furthermore, the top

76 15 genes with higher degree of connectivity were selected as hub genes among the 133 discovered

77 DEGs which includes 56 down-regulated genes and 77 up-regulated genes. In addition, we used

78 visual hierarchical cluster analysis to show the two groups by Morpheus online analysis software

79 (https://software.broadinstitute.org/morpheus/) and volcano plot of two groups by ImageGP

80 (http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html) after the relative raw data

81 of TXT files were downloaded.

82

83 *2.3. Gene Ontology and KEGG Pathway Analysis of DEGs.* With functions including molecular

84 function, biological pathways, and cellular component, gene ontology (GO) analysis we annotated

85 genes and gene products (2006). KEGG. comprises a set of genome and enzymatic approaches

86 and abiological chemical energy online database (Kanehisa & Goto 2000). It is a resource for

87 systematic analysis of gene function and related high-level genome functional information.

88 DAVID (https://david.ncifcrf.gov/) can provide systematic and comprehensive biological function

89 annotation information for high-throughput gene expression (Dennis et al. 2003). Therefore, we

90 applied GO and KEGG pathway analyses to the DEGs by using DAVID online tools at functional

91 level. We considered P < 0.05 had significant differences. In addition, we used visual analysis to

92 show   GO   Enrichment   plot   of   two   groups   by   ImageGP

93 (http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html) after the relative raw data

94 of TXT files were downloaded(Geng et al. 2018).

95

96 *2.4. PPI network and module analysis.* The online tool, Search Tool for the Retrieval of Interacting

97 Genes   (STRING),   is   designed   to   demonstrate   the   interaction   between   different

98 proteins(Szklarczyk et al. 2015). STRING in Cytoscape was applied and mapped the DEGs into

99 STRING to detect the possible relationship among the selected DEGs. We set the confidence score

100 ≥ 0.4, maximum number of interactors = 0 as the selection criteria. In addition, the Molecular

101 Complex Detection (MCODE) was used to screen modules of PPI network in Cytoscape with

102 degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max. depth = 100. DAVID was used to

103 perform the signal pathway analysis of genes in the module. 15 hub genes were also mapped into

104 STRING with confidence score $\geq$ 0.4, maximum number of interactors $\leq$ 5. The potential

105 information was explored through GO and KEGG pathway analysis.

106

107 *2.5. Comparing the expression level of the hub genes.* The GlioVis (http://gliovis.bioinfo.cnio.es/)

108 is a user-friendly web application for data visualization and analysis to explore brain tumors

109 expression datasets, which was used to analyze the gene expression data of brain tumors and

110 normal samples based on the TCGA datasets. (Bowman et al. 2017). The customizable functions

111 are provided such as analyzing the differences of expression levels between glioblastoma and low-

112 grade glioma, so the expression of these genes was demonstrated. And the relationship could be

113 visualized through the boxplot. All values are presented as the mean±SD. All statistical analyses

114 were performed by SPSS 19.0 software. A difference of P < 0.05 was considered statistically

115 significant.

116

117 *2.6. Gene expression profile and gene set enrichment analysis.* The expression profiles of

118 GSE53733 was downloaded from the Gene Expression Omnibus (GEO). And we used gene set

119 enrichment analysis (GSEA) (http://www.broadinstitute.org/gsea) to detect the potential genes

120 influenced by SAA1 and TIMP1 through the Java programming. According to their hub genes

121 expression level (top 50%: high vs bottom 50%: low), we divided the patients into two groups and

122 GSEA was conducted to analyze the effects of selected genes expression level on different

123 biological process. And we set P value of <0.05 and false discovery rates of <0.25 as selection

124 criteria to confirm significant gene sets.

125

126 **3. Results**

127

128 *3.1. Identification of DEGs and hub genes.* 60 WHO II samples and 24 WHO IV samples from

129 GSE52009 were selected in this study. DEGs were detected by applying the GEO2R online

130 analysis tool, setting adjust P value < 0.05 and |logFC| ≥ 2 as selection criteria. 133 differential

131 expressed genes, containing 77 up-regulated genes and 56 down-regulated genes, were detected

132 after the analysis of GSE52009. In addition, we selected 15 hub genes with higher degree of

133 connectivity (Table 1). The results were validated with a DEG expression heatmap and volcano

134 plot of the all downregulated genes and upregulated genes (Figure 1).

135

136   *3.2. GO function and KEGG pathway enrichment analysis.* To explore more particular knowledge

137   of the selected DEGs, we used DAVID to gain the results of GO function and KEGG pathway

138   enrichment analysis. All DEGs were imported to DAVID software and GO analysis results

139   demonstrated that up-regulated and down-regulated DEGs were particularly enriched in the

140   following biological processes (BP): cell migration, locomotion and leukocyte migration, cell

141   motility for up-regulated DEGs, and for downregulated DEGs nervous system development, brain

142   development and regulation of cell projection organization (Table 2, Figure2a, Figure2b). The

143   upregulated DEGs were enriched in phospholipase A2 inhibitor activity, growth factor binding,

144   extracellular matrix structural construction, receptor binding, and the down-regulated DEGs were

145   enriched in calcium ion binding, structural construction of myelin sheath, and protein complex

146   binding for molecular function (MF) (Table 2, Figure2). Moreover, GO cell component (CC)

147   analysis showed that the up-regulated DEGs were enriched in the proteinaceous extracellular

148   matrix, extracellular matrix and cytoplasmic membrane-bounded vesicle lumen, and down-

149   regulated DEGs enriched in neuron part, myelin sheath, and internode region of axon (Table 2,

150   Figure2a, Figure2b).

151   Interestingly, the most significantly enriched of KEGG pathway only showed in up-regulated

152   pathway, including adhesion, ECM-receptor interaction, amoebiasis and PI3K-Akt signaling

153   pathway (Table3, Figure2c).

154

155   *3.3. Hub Genes and Module Screening from PPI Network.* PPI network of the top 15 hub genes

156   with higher degree of connectivity was made based on the information in the STRING protein

157   query from public databases (Figure 3A). The top module was selected by using MCODE plug-in

158   in the PPI network. (Figure 3B).

159

160   *3.4. The Kaplan-Meier Plotter of hub genes.* The website, http:// gepia.cancer-pku.cn/, freely

161   provides the prognostic data of the hub genes. It was found that expression of VEGFA (HR 4.2, p

162   = 0) was associated with worse overall survival (OS) for glioblastoma patients, as well as NDC80

163   (HR 5.8, p = 0), CENPA (HR 5.3, p = 0), CENPF (HR 3.9, p = 0), NCAPG (HR 5.6 p = 0), ASPM

164   (HR 5, p = 0), ITGA2 (HR 3, p = 0), TIMP1 (HR 7, p = 0)and SAA1 (HR 4.8, p = 0). (Figure4).

165

166   *3.5. Expression Level and relationship with molecular pathologic diagnosis of Hub genes.* We

167   used data from GlioVis to detect the hub gene expression level between glioblastoma and low-

168   grade glioma including astrocytoma, oligodendroglioma and oligoastrocytoma reflect that

169   compared to low-grade glioma, the expression level of SAA1 and TIMP1 significantly increased

170   in glioblastomas (Figure 5A, 5C). And the expression level of SAA1 have no significant difference

171   in three kinds LGG (Figure 5B). However, the expression level of TIMP1 is significantly higher

172   in astrocytoma than oligodendroglioma and oligoastrocytoma (Figure 5D). We also detect the

173   relationship between expression level and molecular pathologic diagnosis of hub genes. We found

174   both SAA1 and TIMP1 increase in both IDH mutant IDH wild type. The same results were found

175   in MGMT promoter and non-deletion of chromosome 1p.19q. Because of the limited samples in

176   the datasets, we didn't the result of co-deletion of chromosome 1p.19q. (Figure 5E,5F) What's

177   more, we also found that both SAA1 and TIMP1 played important roles in MES-like in the IDH

178   wild type. (Figure 5G,5H)

179

180   *3.6. Gene expression profile and gene set enrichment analysis.* We managed a GSEA by using

181   GBM patient gene profiling data (GSE53733), and showed in figure 5, gene set differences in

182   SAA1 in low versus high glioma patients indicated that SAA1 regulates biology process mainly

183   associated with inflammatory response processes (P<0.001 FDR=0.012) and cytokine mediated

184   signaling pathway (P<0.001 FDR=0.012) (Figure5I, 5J). We considered that SAA1 may

185   negatively regulate inflammatory response and might promote the survival of cancer cells.

186   Seemingly we concluded that TIMP1 might negatively regulates adaptive immune response based

187   on somatic recombination of immune receptors built from leucine-rich (P<0.001 FDR=0.021) and

188   response to interferon (P<0.001 FDR=0.027) may promote the survival of cancer cells. (Figure5K,

189   5L)

190

4. Discussion

192

193   In this study, we identified 15 DEGs between glioblastoma and low-grade and used a series

194   of bioinformatics analyses to screen the key genes and pathways associated with glioma.

195   GSE52009 dataset contains 60 WHO II samples and 24 WHO IV samples. In order to improve the

196   statistical power of DEG, we defined that the absolute value of the logarithm (base 2) fold change

197   (logFC) greater than 2 and 133 DEGs were obtained. Bioinformatics analysis on DEGs, including

198   GO enrichment, KEGG pathway, PPI network, and survival analysis, expression level, gene set

199   enrichment analysis, found glioblastoma related genes and pathways that play important roles in

200   glioma development.

201   The GO analysis showed that the up-regulated DEGs were mainly associated with cell

202   migration, extracellular matrix structural construction, cell motility and down-regulated DEGs

203   were involved central nervous system development, calcium ion binding and internode region of

204   axon. Additionally, the KEGG pathways of up-regulated DEGs regulate focal adhesion, ECM-

205   receptor interaction, PI3K-Akt signaling pathway. Among these DEGs, we selected 15 hub genes

206   with higher degree of connectivity. In addition, we found several hub genes with worse overall

207   survival (OS) and higher expression level in glioma patients, including VEGFA, CXCL8, TIMP1,

208   SAA1, PLAT and NCAPG and we firstly found relationship of SAA1 ,TIMP1 and molecular

209   pathology in glioblastoma. We could hypothesize that these genes might contribute to the

210   malignance of glioma.

211   Serum amyloid A1 (SAA1) which was secreted by liver is an acute-phase high density lipo-

212   protein in immune response. Injury, inflammation, and brain trauma can elevate the plasma levels

213   of SAA1 (Lu et al. 2014; Villapol et al. 2015). What's more, it has long been suspected that SAA1

214   might be a prognostic marker and predictor of cancer risk. Elevated levels of SAA1 in the serum

215   of cancer patients directly correlate with poor prognosis and tumor aggressiveness in various types

216 of cancer, including lung cancer(Cho et al. 2010), cell renal carcinoma(Kosari et al. 2005),

217 melanomas(Findeisen et al. 2009) and so on. Normal brain does not express SAA1(Liang et al.

218 1997), though an in vitro study demonstrated that microglia and astrocytes are responsive to SAA

219 (Yu et al. 2014). Recently, it has been reported that SAA1 the expression levels in GBM patients

220 are upregulated on both miRNA and protein in human glioblastoma, and SAA1 involves in

221 angiogenesis via HIF-1α and tumor associated macrophages(Knebel et al. 2017). Consistently,

222 SAA1 has been reported as a molecular/metabolic signature that can help identify patients are at

223 high risk of malignant GBM and promotes glioma cells migration and invasion through integrin

224 aVb3(Lin et al. 2018). What's more, although it's unknown why SAA1 upregulated in GBM and

225 other malignant cancers, it has been speculated that SAA proteins play a primary role in the

226 regulation of immunity and invasion processes (Moshkovskii 2012), which is consistent with the

227 result of our study. Thus, we hypothesis SAA1 could be a biomarker of GBM and predict the

228 prognosis of GBM patients. And the mechanism of SAA1 regulate in GBM need further research.

229   The tissue inhibitors of metalloproteinases (TIMPs, including TIMP-1, TIMP-2, TIMP-3,

230 TIMP-4) are well known play critical roles in both metastasis and invasion through extracellular

231 matrix remodeling which are controlled by the activity of matrix metalloproteinases (MMPs)

232 (Jackson et al. 2017; Ries 2014). Among the four well-known TIMPs characterized so far, most

233 focus has been on TIMP-1 defined as a naturally occurring inhibitor of most MMPs, a family of

234 zinc dependent endopeptidases essential for degrading components of the ECM(Aaberg-Jessen et

235 al. 2009). In addition, TIMP1 shows protease-independent function including anti-apoptotic,

236 antiangiogenic, and differentiation activities in cells (Bridoux et al. 2013; Mandel et al. 2017).

237 Over the past year, more and more studies focus on the influence of TIMP1 in cancers. Serum or

238 urine levels of TIMP1 are also considered as a diagnostic predictor in pancreatic ductal carcinomas

239 containing extensive desmoplasia(Jenkinson et al. 2015; Roy et al. 2014). Increased levels of

240 cytosolic TIMP1 in pretreatment tumor tissue is associated with a significantly shorter overall

241 survival in patients with breast cancer receiving standard adjuvant chemotherapy(Dechaphunkul

242 et al. 2012). And it has been reported that low expression of TIMP-1 in glioblastoma patient

243 predicts longer survival. The shorter survival of glioblastoma patients with a high tumor TIMP-1

244 level may be explained by the antiapoptotic effect of TIMP-1 preventing apoptosis induced by

245 radiation and chemotherapy(Aaberg-Jessen et al. 2009). More recently, Aaberg-Jessen C et al

246 demonstrated that Co-expression of TIMP-1 and CD63 might have effects in glioblastoma

247 stemness and may predict the poor prognosis of patients through influencing tumor aggressiveness

248 and resistance of therapy(Aaberg-Jessen et al. 2018). We consider TIMP-1 can be identified as a

249 future biomarker for prognosis or monitoring patients' treatment response. However, all these

250 studies didn't demonstrate the specific mechanism, which is the direction for further researches.

251

252 5.conclusion,

253 We presumed these key genes identified by a series of bioinformatics analyses on DEGs between

254 glioblastoma samples and low-grade glioma samples, probably related to the development of

255 glioma. These hub genes could also affect the survival time of glioma patients. These identified

256 genes and pathways provide a more detailed molecular mechanism for underlying glioma initiation

257 and development. According to the study, SAA1 and TIMP1 can be considered as biomarkers or

258 therapeutic targets or monitoring patients' treatment response for glioblastoma. However, further

259 molecular and biological experiments are required to confirm the functions of the key genes in

260 glioblastoma.

261

262 Data Availability

263

264 The data used to support the findings of this study are available from the corresponding author

265 upon request.

266

267 Acknowledgements

268

269   This work was supported by the National Natural Science Foundation of China (No. 81572489,

270   81372683, and 81502075).

271

Conflict of interest

273

274   The authors declare that they have no conflict of interest

275

References

277

278   2006. The Gene Ontology (GO) project in 2006. *Nucleic Acids Res* 34:D322-326. 10.1093/nar/gkj021

279   Aaberg-Jessen C, Christensen K, Offenberg H, Bartels A, Dreehsen T, Hansen S, Schroder HD, Brunner N, and
280   Kristensen BW. 2009. Low expression of tissue inhibitor of metalloproteinases-1 (TIMP-1) in glioblastoma
281   predicts longer patient survival. *J Neurooncol* 95:117-128. 10.1007/s11060-009-9910-8

282   Aaberg-Jessen C, Sorensen MD, Matos A, Moreira JM, Brunner N, Knudsen A, and Kristensen BW. 2018. Co-
283   expression of TIMP-1 and its cell surface binding partner CD63 in glioblastomas. *BMC Cancer* 18:270.
284   10.1186/s12885-018-4179-y

285   Bowman RL, Wang Q, Carro A, Verhaak RGW, and Squatrito M. 2017. GlioVis data portal for visualization and
286   analysis of brain tumor expression datasets. *Neuro Oncol* 19:139-141.

287   Bridoux L, Etique N, Lambert E, Thevenard J, Sowa ML, Belloy N, Dauchez M, Martiny L, and Charpentier E. 2013.
288   A crucial role for Lyn in TIMP-1 erythroid cell survival signalling pathway. *FEBS Lett* 587:1524-1528.
289   10.1016/j.febslet.2013.03.032

290   Cheng W, Ren X, Zhang C, Cai J, Liu Y, Han S, and Wu A. 2016. Bioinformatic profiling identifies an immune-
291   related risk signature for glioblastoma. *Neurology* 86:2226-2234. 10.1212/wnl.0000000000002770

292   Cho WC, Yip TT, Cheng WW, and Au JS. 2010. Serum amyloid A is elevated in the serum of lung cancer patients
293   with poor prognosis. *Br J Cancer* 102:1731-1735. 10.1038/sj.bjc.6605700

294   Davis S, and Meltzer PS. 2007. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor.
295   *Bioinformatics* 23:1846-1847. 10.1093/bioinformatics/btm254

296   De Preter K, Vermeulen J, Brors B, Delattre O, Eggert A, Fischer M, Janoueix-Lerosey I, Lavarino C, Maris JM,
297   Mora J, Nakagawara A, Oberthuer A, Ohira M, Schleiermacher G, Schramm A, Schulte JH, Wang Q,
298   Westermann F, Speleman F, and Vandesompele J. 2010. Accurate outcome prediction in neuroblastoma
299   across independent data sets using a multigene signature. *Clin Cancer Res* 16:1532-1541. 10.1158/1078-
300   0432.ccr-09-2607

301   Dechaphunkul A, Phukaoloun M, Kanjanapradit K, Graham K, Ghosh S, Santos C, and Mackey JR. 2012. Prognostic
302   significance of tissue inhibitor of metalloproteinase-1 in breast cancer. *Int J Breast Cancer* 2012:290854.
303   10.1155/2012/290854

304   Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, and Lempicki RA. 2003. DAVID: Database for

305   Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4:P3.

306 Findeisen P, Zapatka M, Peccerella T, Matzk H, Neumaier M, Schadendorf D, and Ugurel S. 2009. Serum amyloid A
307   as a prognostic marker in melanoma identified by proteomic profiling. *J Clin Oncol* 27:2199-2208.
308   10.1200/jco.2008.18.0554

309 Freije WA, Castro-Vargas FE, Fang Z, Horvath S, Cloughesy T, Liau LM, Mischel PS, and Nelson SF. 2004. Gene
310   expression profiling of gliomas strongly predicts survival. *Cancer Res* 64:6503-6510. 10.1158/0008-
311   5472.can-04-0452

312 Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, Stegh A, Hahn WC, Ligon KL, Louis DN, Brennan C,
313   Chin L, DePinho RA, and Cavenee WK. 2007. Malignant astrocytic glioma: genetics, biology, and paths to
314   treatment. *Genes Dev* 21:2683-2710. 10.1101/gad.1596707

315 Geng R-X, Li N, Xu Y, Liu J-h, Yuan F-e, Sun Q, Liu B-H, and Chen Q-X. 2018. Identification of Core Biomarkers
316   Associated with Outcome in Glioma: Evidence from Bioinformatics Analysis. *Disease Markers* 2018:16.
317   10.1155/2018/3215958

318 Jackson HW, Defamie V, Waterhouse P, and Khokha R. 2017. TIMPs: versatile extracellular regulators in cancer. *Nat
319   Rev Cancer* 17:38-53. 10.1038/nrc.2016.115

320 Jenkinson C, Elliott V, Menon U, Apostolidou S, Fourkala OE, Gentry-Maharaj A, Pereira SP, Jacobs I, Cox TF,
321   Greenhalf W, Timms JF, Sutton R, Neoptolemos JP, and Costello E. 2015. Evaluation in pre-diagnosis
322   samples discounts ICAM-1 and TIMP-1 as biomarkers for earlier diagnosis of pancreatic cancer. *J
323   Proteomics* 113:400-402. 10.1016/j.jprot.2014.10.001

324 Kanehisa M, and Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27-30.

325 Kim YW, Liu TJ, Koul D, Tiao N, Feroze AH, Wang J, Powis G, and Yung WK. 2011. Identification of novel
326   synergistic targets for rational drug combinations with PI3 kinase inhibitors using siRNA synthetic lethality
327   screening against GBM. *Neuro Oncol* 13:367-375. 10.1093/neuonc/nor012

328 Knebel FH, Uno M, Galatro TF, Belle LP, Oba-Shinjo SM, Marie SKN, and Campa A. 2017. Serum amyloid A1 is
329   upregulated in human glioblastoma. *J Neurooncol* 132:383-391. 10.1007/s11060-017-2386-z

330 Kosari F, Parker AS, Kube DM, Lohse CM, Leibovich BC, Blute ML, Cheville JC, and Vasmatzis G. 2005. Clear cell
331   renal cell carcinoma: gene expression analyses identify a potential signature for tumor aggressiveness. *Clin
332   Cancer Res* 11:5128-5139. 10.1158/1078-0432.ccr-05-0073

333 Kulasingam V, and Diamandis EP. 2008. Strategies for discovering novel cancer biomarkers through utilization of
334   emerging technologies. *Nat Clin Pract Oncol* 5:588-599. 10.1038/ncponc1187

335 Liang JS, Sloane JA, Wells JM, Abraham CR, Fine RE, and Sipe JD. 1997. Evidence for local production of acute
336   phase response apolipoprotein serum amyloid A in Alzheimer's disease brain. *Neurosci Lett* 225:73-76.

337 Lin CY, Yang ST, Shen SC, Hsieh YC, Hsu FT, Chen CY, Chiang YH, Chuang JY, Chen KY, Hsu TI, Leong WC,
338   Su YK, Lo WL, Yeh YS, Patria YN, Shih HM, Chang CC, and Chou SY. 2018. Serum amyloid A1 in
339   combination with integrin alphaVbeta3 increases glioblastoma cells mobility and progression. *Mol Oncol*
340   12:756-771. 10.1002/1878-0261.12196

341 Liu B, Dong H, Lin X, Yang X, Yue X, Yang J, Li Y, Wu L, Zhu X, Zhang S, Tian D, Wang J, Cai Q, Mao S, Chen
342   Q, and Chang J. 2016. RND3 promotes Snail 1 protein degradation and inhibits glioblastoma cell migration
343   and invasion. *Oncotarget* 7:82411-82423. 10.18632/oncotarget.12396

344 Louis DN, Perry A, Reifenberger G, Deimling AV, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD,
345   Kleihues P, and Ellison DW. 2016. The 2016 World Health Organization Classification of Tumors of the

346   Central Nervous System: a summary. *Acta Neuropathologica* 131:803-820.

347 Lu J, Yu Y, Zhu I, Cheng Y, and Sun PD. 2014. Structural mechanism of serum amyloid A-mediated inflammatory
348   amyloidosis. *Proc Natl Acad Sci U S A* 111:5189-5194. 10.1073/pnas.1322357111

349 Mandel ER, Uchida C, Nwadozi E, Makki A, and Haas TL. 2017. Tissue Inhibitor of Metalloproteinase 1 Influences
350   Vascular Adaptations to Chronic Alterations in Blood Flow. *J Cell Physiol* 232:831-841. 10.1002/jcp.25491

351 Moshkovskii SA. 2012. Why do cancer cells produce serum amyloid A acute-phase protein? *Biochemistry (Mosc)*
352   77:339-341. 10.1134/s0006297912040037

353 Ostrom QT, Gittleman H, Liao P, Vecchione-Koval T, Wolinsky Y, Kruchko C, and Barnholtz-Sloan JS. 2017.
354   CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United
355   States in 2010-2014. *Neuro Oncol* 19:v1-v88. 10.1093/neuonc/nox158

356 Ries C. 2014. Cytokine functions of TIMP-1. *Cell Mol Life Sci* 71:659-672. 10.1007/s00018-013-1457-3

357 Rodriguez FJ, Vizcaino MA, and Lin MT. 2016. Recent Advances on the Molecular Pathology of Glial Neoplasms in
358   Children and Adults. *J Mol Diagn* 18:620-634. 10.1016/j.jmoldx.2016.05.005

359 Roy R, Zurakowski D, Wischhusen J, Frauenhoffer C, Hooshmand S, Kulke M, and Moses MA. 2014. Urinary TIMP-
360   1 and MMP-2 levels detect the presence of pancreatic malignancies. *Br J Cancer* 111:1772-1779.
361   10.1038/bjc.2014.462

362 Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A,
363   Tsafou KP, Kuhn M, Bork P, Jensen LJ, and von Mering C. 2015. STRING v10: protein-protein interaction
364   networks, integrated over the tree of life. *Nucleic Acids Res* 43:D447-452. 10.1093/nar/gku1003

365 Villapol S, Kryndushkin D, Balarezo MG, Campbell AM, Saavedra JM, Shewmaker FP, and Symes AJ. 2015. Hepatic
366   expression of serum amyloid A1 is induced by traumatic brain injury and modulated by telmisartan. *Am J*
367   *Pathol* 185:2641-2652. 10.1016/j.ajpath.2015.06.016

368 Yu Y, Liu J, Li SQ, Peng L, and Ye RD. 2014. Serum amyloid a differentially activates microglia and astrocytes via
369   the PI3K pathway. *J Alzheimers Dis* 38:133-144. 10.3233/jad-130818

370

Differentiallyexpressed gene expression heat map and volcano plot of glioma.

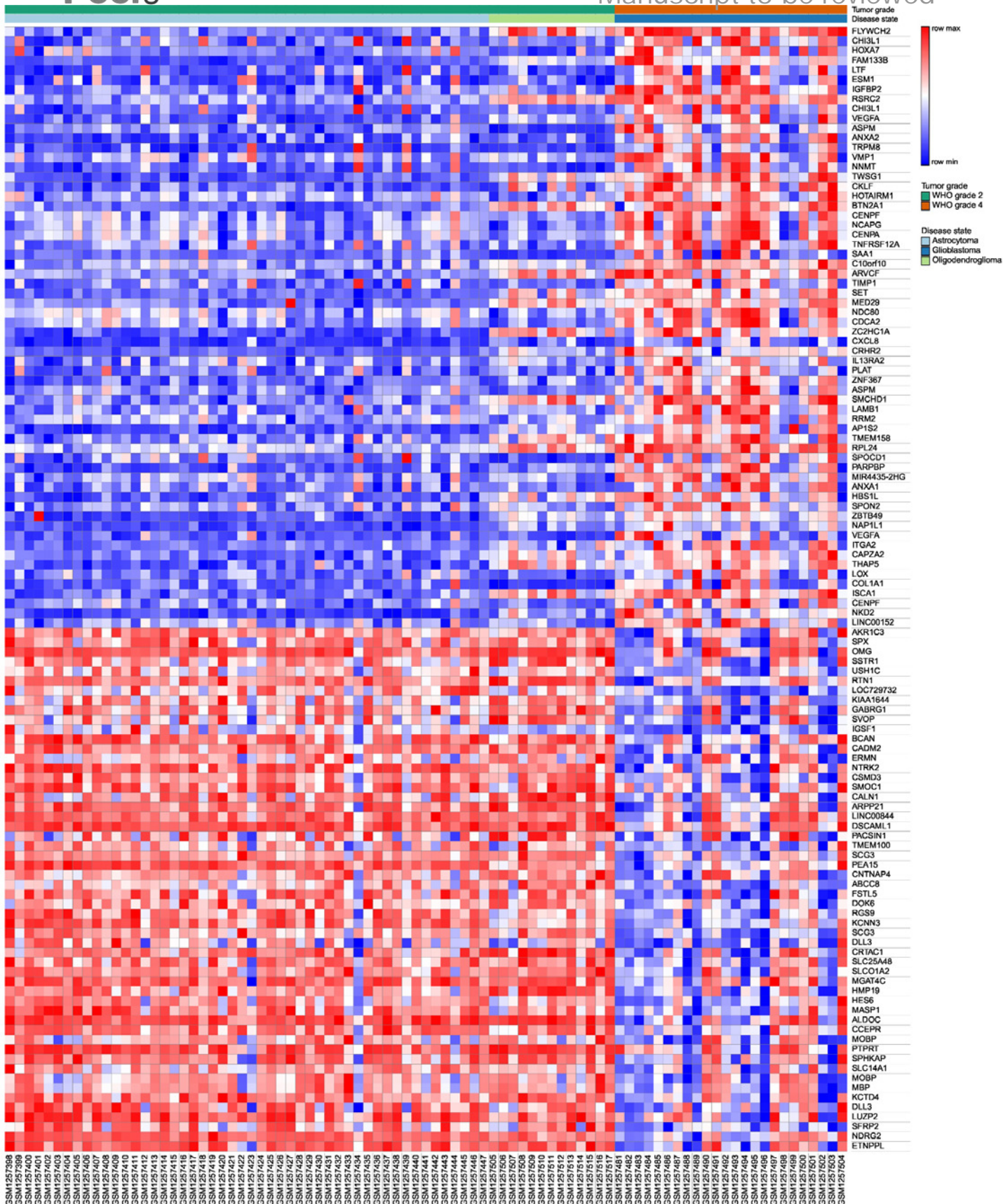
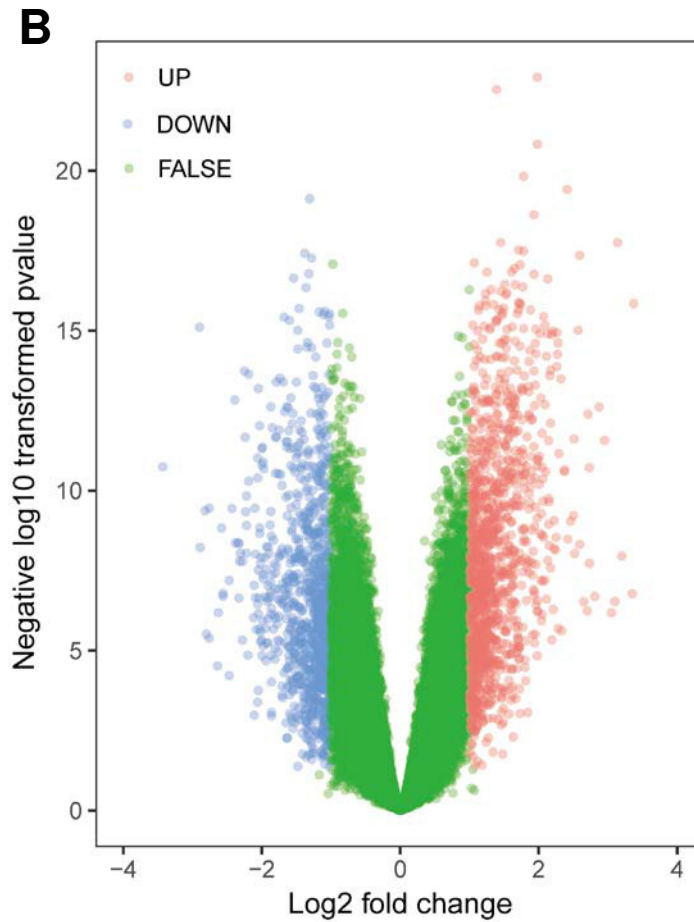
Figure 1 : (A) Differentially expressed gene expression heat map of glioma (all upregulated and downregulated genes). (B) Differentially expressed genes were selected by volcano plot filtering (fold change ≥ 1 and p value ≤ 0.05). The blue point in the plot represents the differentially expressed genes with statistical significance.

GOanalysis results showed that upregulated DEGs

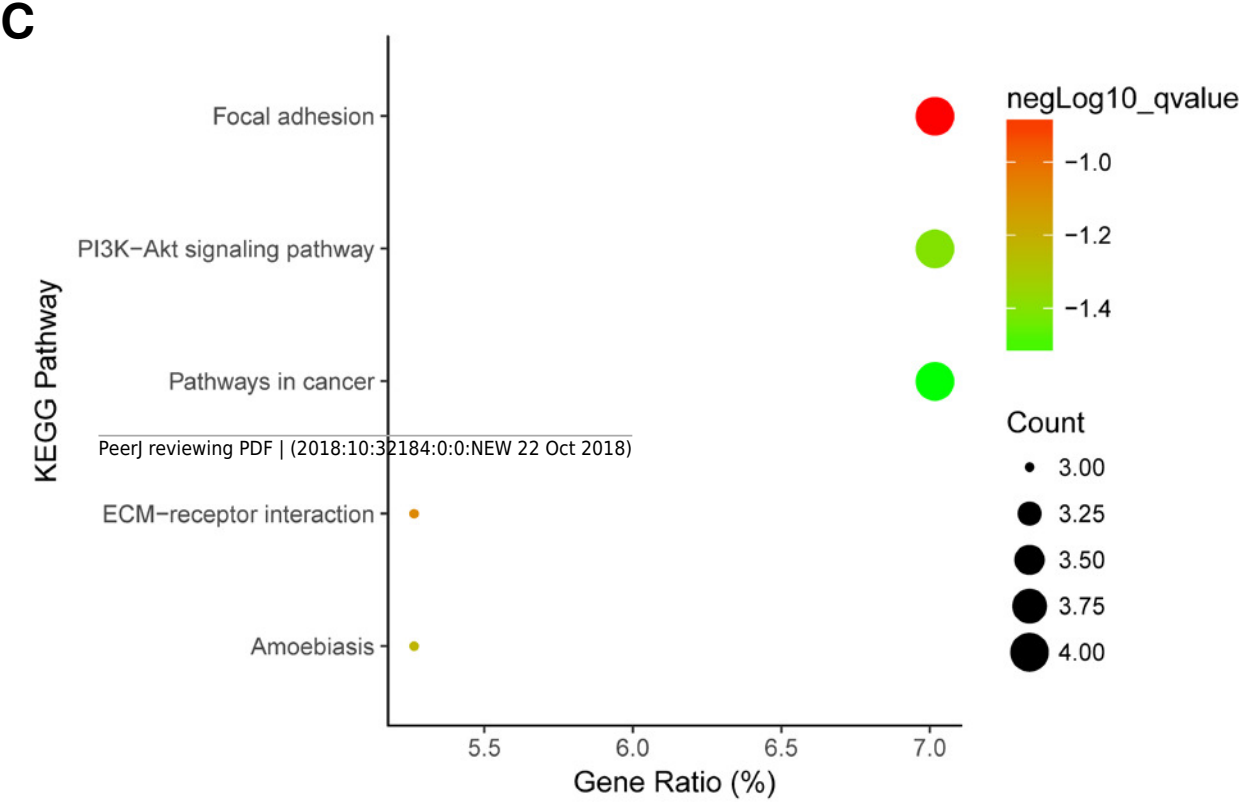
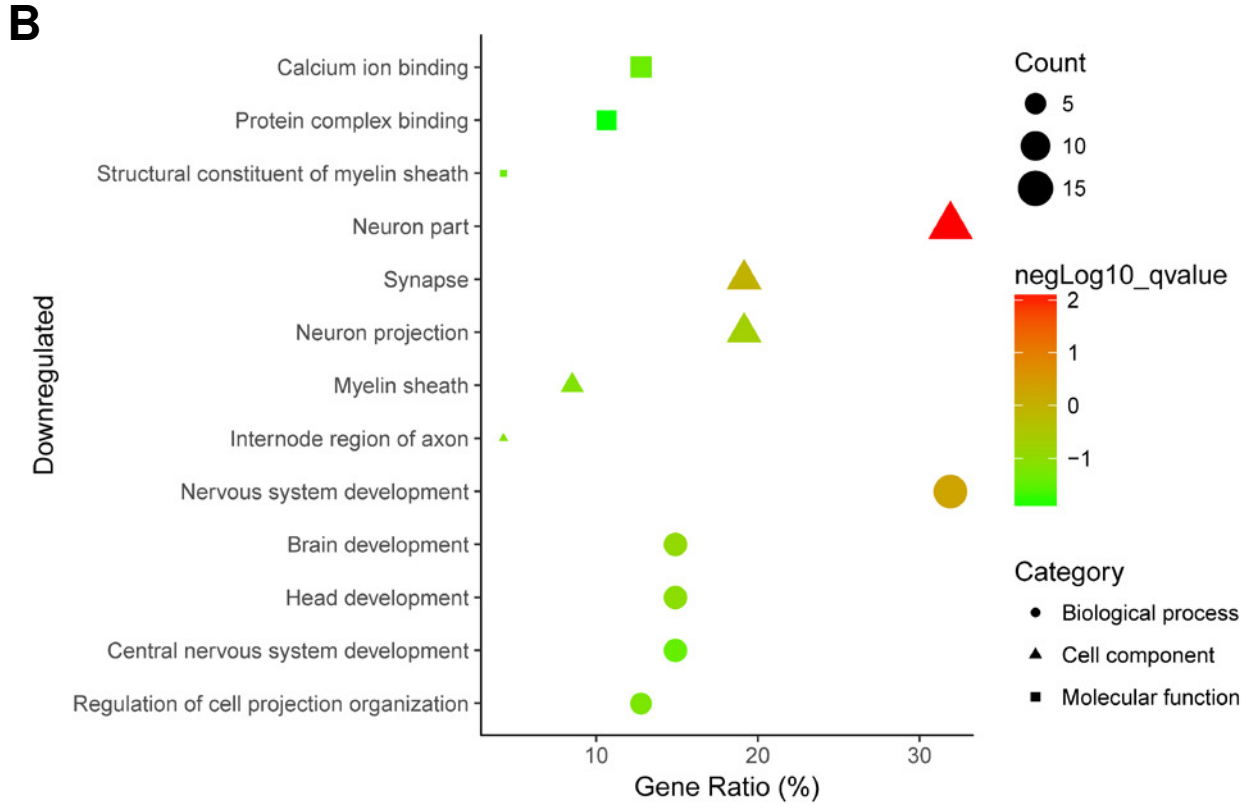
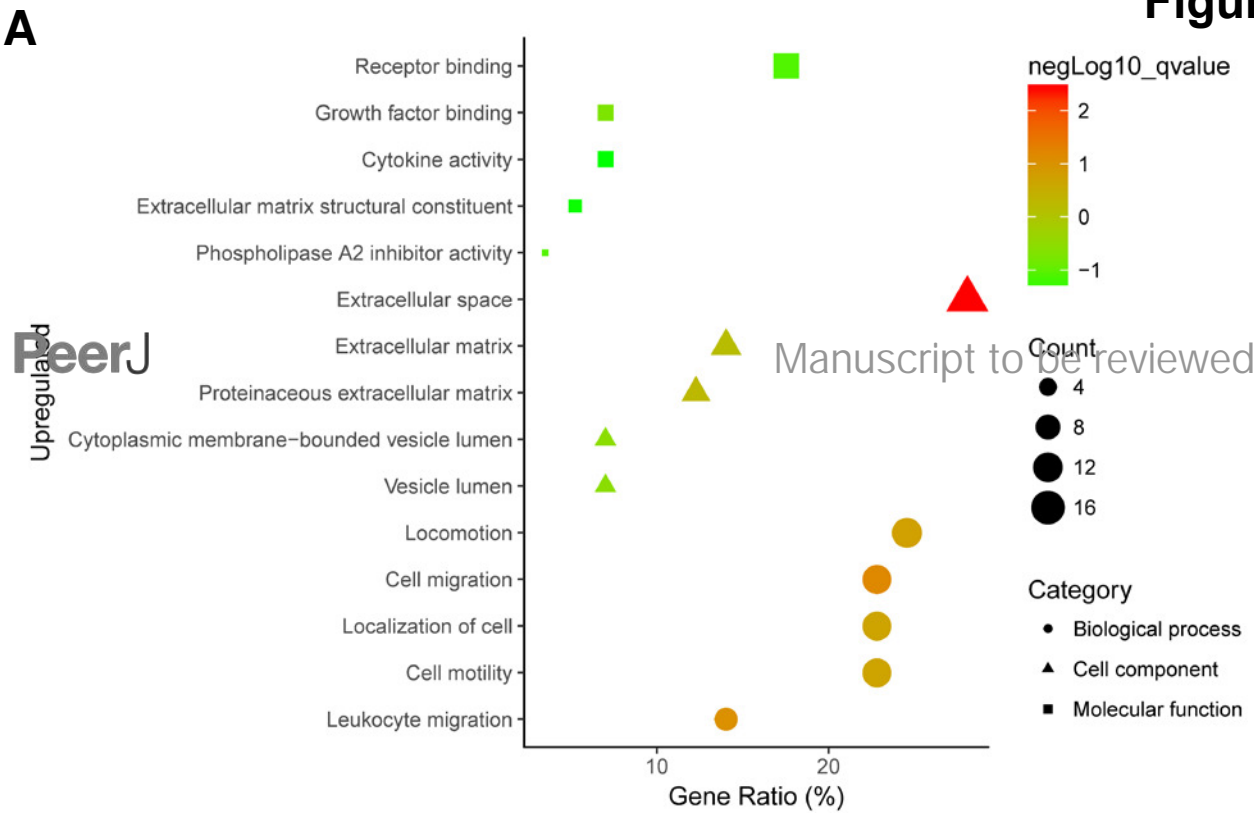
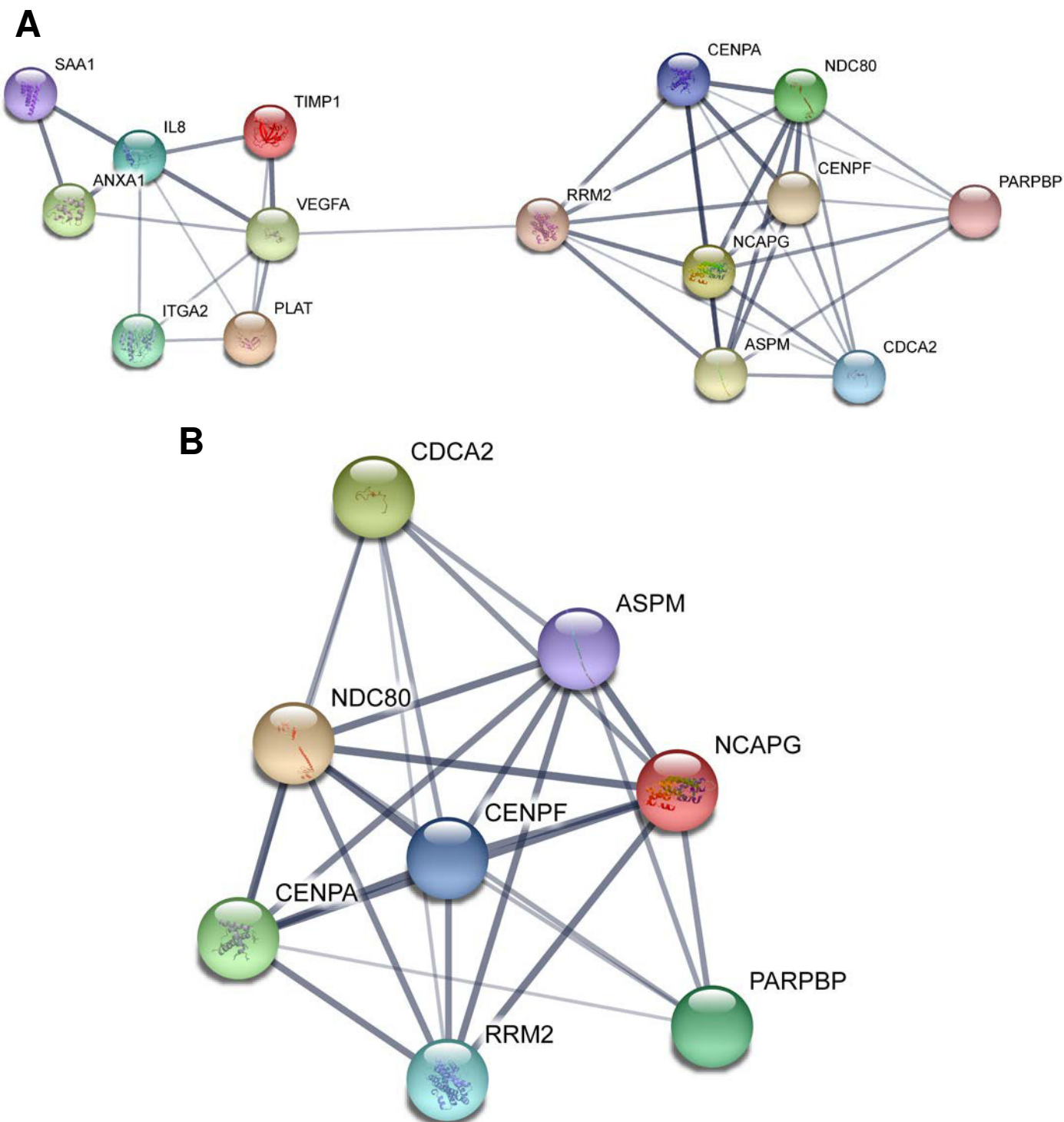
Figure 2 : GO analysis results showed that upregulated DEGs (A) and downregulated DEGs (B) were particularly enriched in BP, MF, and CC. (C) The most significantly enriched KEGG pathway of the upregulated DEGs. GO: gene ontology; BP: biological process; MF: molecular function; CC: cell component; KEGG: Kyoto Encyclopedia of Genes and Genomes.

protein-proteininteraction network and top module of hub genes

Figure 3 : (A) The protein-protein interaction network of top 15 hub genes. (B) Top module from the protein-protein interaction network.

A

B

Prognosticvalue of hub genes (VEGFA, NDC80, CENPA, CENPF, NCAPG, ASPM, ITGA2, TIMP1, andSAA1) in glioma patients.

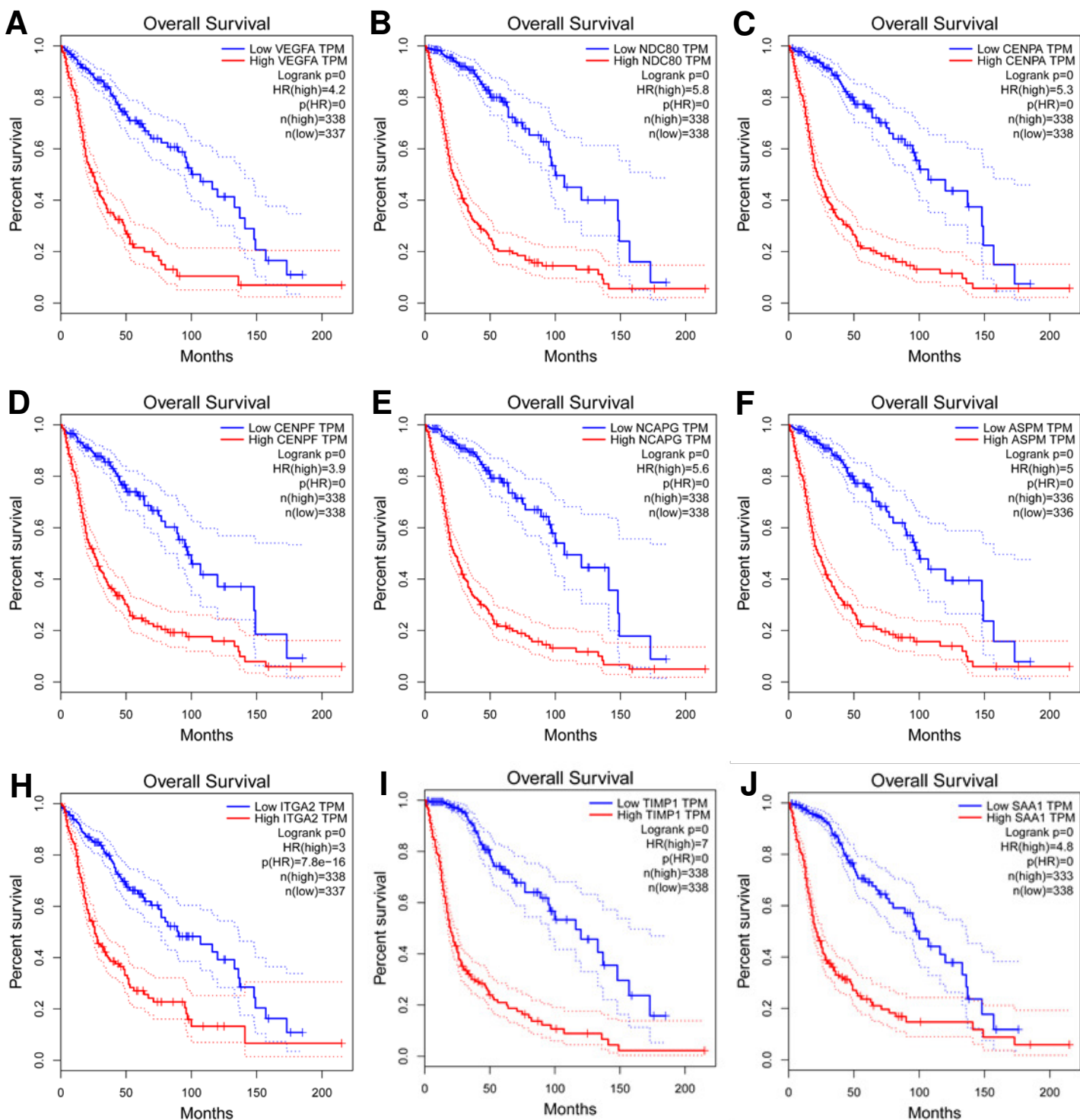
Figure 4 : Prognostic value of hub genes (VEGFA, NDC80, CENPA, CENPF, NCAPG, ASPM, ITGA2, TIMP1, and SAA1) in glioma patients. HR: hazard ratio.

the expression level and potential function of SAA1 and TIMP1

Figure 5 : (A) SAA1 significantly increased in glioblastomas; (B) The expression level of SAA1 have no significant difference in LGG; (C) TIMP1 significantly increased in glioblastomas; (D) TIMP1 is significantly higher in astrocytoma than oligodendroglioma and oligoastrocytoma; (E, F) SAA1 and TIMP1 increase in both IDH mutant IDH wild type. The same results were found in MGMT promoter and non-deletion of chromosome 1p.19q; G, H. SAA1 and TIMP1 played important roles in MES-like in the IDH wild type; (I, J) SAA1 regulates biology process associated with inflammatory response processes and cytokine mediated signaling pathway; (K, L). TIMP1 negatively regulates adaptive immune response based on somatic recombination of immune receptors built from leucine-rich and response to interferons.

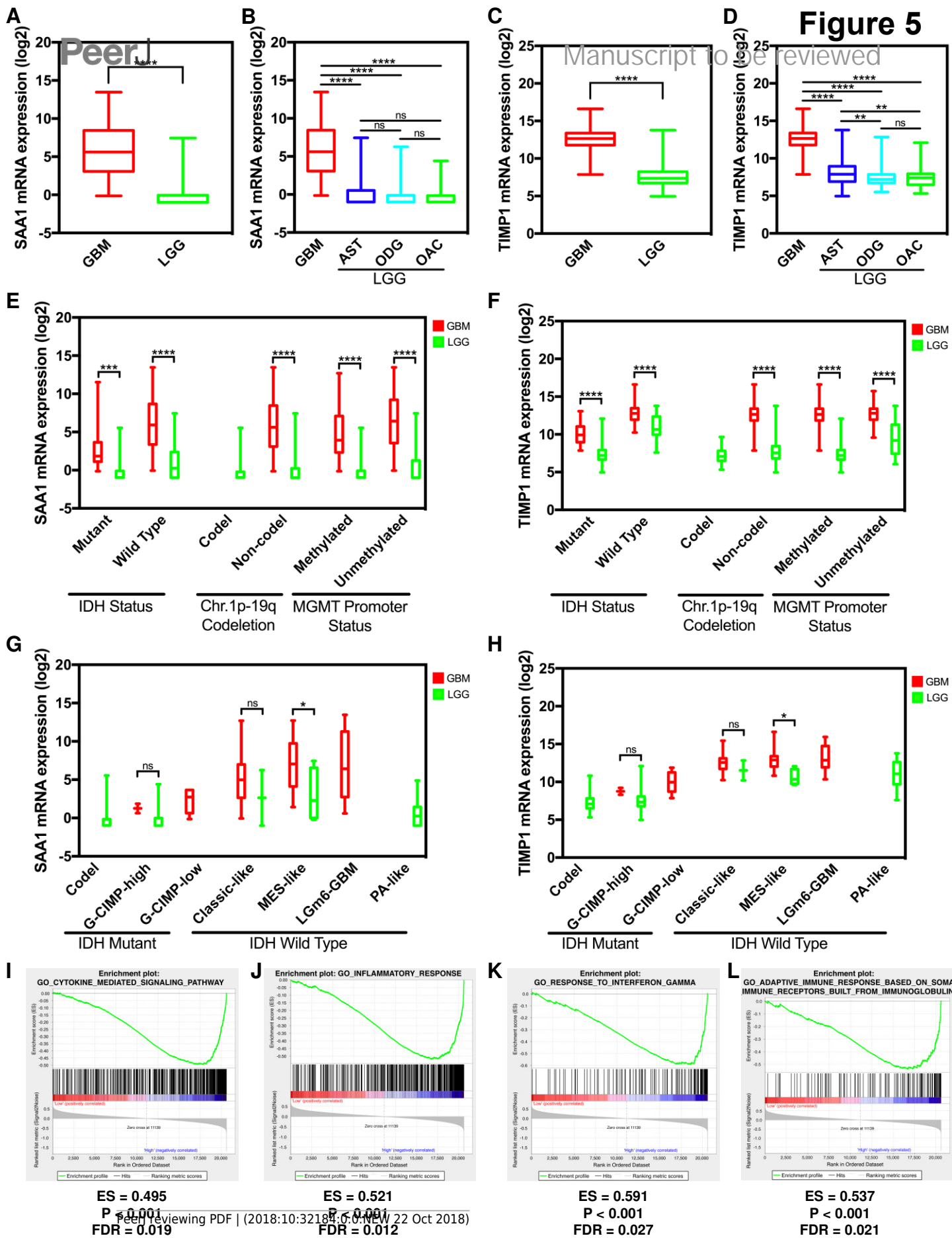



Figure 5

ES = 0.495
P < 0.001
FDR = 0.019

ES = 0.521
P < 0.001
FDR = 0.012

ES = 0.591
P < 0.001
FDR = 0.027

ES = 0.537
P < 0.001
FDR = 0.021

Table 1. Top 15 hub genes with higher degree of connectivity

Table 1. Top 15 hub genes with higher degree of connectivity

Table 1. Top 15 hub genes with higher degree of connectivity

Gene	Degree	P-value
VEGFA	13	7.55E-07
NDC80	8	7.16E-09
IL8	8	5.15E-07
CENPA	7	3.31E-09
CENPF	7	6.06E-10
NCAPG	7	9.09E-10
ASPM	7	1.86E-08
RRM2	7	1.13E-09
ITGA2	6	2.45E-09
ANXA1	6	6.32E-08
CDCA2	6	6.42E-09
PLAT	5	2.64E-08
PARPBP	5	6.34E-15
TIMP1	4	2.06E-06
SAA1	4	2.40E-06

Table 2. Gene ontology analysis of differentially expressed genes associated with GBM and LGG

Table 2. Gene ontology analysis of differentially expressed genes associated with GBM and LGG

Expression	Category	Term	Count	Ratio (%)	P Value	FDR
Up-regulated	Biological Process	GO:0016477~cell migration	13	22.80702	3.95E-05	0.067618
		GO:0050900~leukocyte migration	8	14.03509	5.28E-05	0.090452
		GO:0040011~locomotion	14	24.5614	1.09E-04	0.185762
		GO:0051674~localization of cell	13	22.80702	1.22E-04	0.208673
		GO:0048870~cell motility	13	22.80702	1.22E-04	0.208673
	Molecular Function	GO:0019838~growth factor binding	4	7.017544	0.004366	5.428929
		GO:0019834~phospholipase A2 inhibitor activity	2	3.508772	0.010298	12.37092
		GO:0005102~receptor binding	10	17.54386	0.010815	12.95228
		GO:0005201~extracellular matrix structural constituent	3	5.263158	0.018986	21.69308
		GO:0005125~cytokine activity	4	7.017544	0.019477	22.19193
	Cell Component	GO:0005615~extracellular space	16	28.07018	3.09E-06	0.003877
		GO:0005578~proteinaceous extracellular matrix	7	12.2807	4.24E-04	0.531176
		GO:0031012~extracellular matrix	8	14.03509	5.41E-04	0.676623
		GO:0060205~cytoplasmic membrane-bounded vesicle lumen	4	7.017544	0.002985	3.680925
		GO:0031983~vesicle lumen	4	7.017544	0.003066	3.779192

Down-regulated	Biological Process	GO:0007399~nervous system development	15	31.91489	3.29E-04	0.533895
		GO:0007420~brain development	7	14.89362	0.00548	8.553316
		GO:0060322~head development	7	14.89362	0.007003	10.80691
		GO:0031344~regulation of cell projection organization	6	12.76596	0.011466	17.11031
		GO:0007417~central nervous system development	7	14.89362	0.019918	27.91963
	Molecular Function	GO:0005509~calcium ion binding	6	12.76596	0.021579	24.70502
		GO:0019911~structural constituent of myelin sheath	2	4.255319	0.02239	25.5132
		GO:0032403~protein complex binding	5	10.6383	0.092825	71.83703
	Cell Component	GO:0097458~neuron part	15	31.91489	6.61E-06	0.007993
		GO:0045202~synapse	9	19.14894	8.96E-04	1.078808
		GO:0043005~neuron projection	9	19.14894	0.00424	5.011629
		GO:0033269~internode region of axon	2	4.255319	0.01097	12.49482
		GO:0043209~myelin sheath	4	8.510638	0.01206	13.65505

Table 3. KEGG pathway analysis of differentially expressed genes associated with GBM and LGG

Table 3. KEGG pathway analysis of differentially expressed genes associated with GBM and LGG

Table 3. KEGG pathway analysis of differentially expressed genes associated with GBM and LGG

Expression	Term	Count	Ratio (%)	P Value	Genes	FDR
Up-regulated	hsa04510: Focal adhesion	4	7.017544	0.007454	LAMB1, VEGFA, ITGA2, COL1A1,	7.094772
	hsa04512:ECM-receptor interaction	3	5.263158	0.012926	LAMB1, ITGA2, COL1A1	12.01164
	hsa05146:Amoebiasis	3	5.263158	0.018811	LAMB1, CXCL8, COL1A1	17.03775
	hsa04151:PI3K-Akt signaling pathway	4	7.017544	0.02979	LAMB1, VEGFA, ITGA2, COL1A1	25.72978
	hsa05200:Pathways in cancer	4	7.017544	0.041599	LAMB1, VEGFA, ITGA2, CXCL8	34.15806