Identification of novel genes associated with diagnosis and prognosis in glioblastoma

Gaowei Li
Department of neurosurgery, West China Hospital, West China school of Medicine, Sichuan University
Linjun Cai
Department of Neurology, West China Hospital, West China School Of Medicine,Sichuan University
Xin Tang
Department of Neurosurgery, West China Hospital, West China School Of Medicine, Sichuan University
Jianhan Huang
Department of Neurosurgery, West China Hospital, West China School Of Medicine, Sichuan University
Liangxue Zhou (zhlxlll@163.com)
Sichuan University West China Hospital

Research article

Keywords: GBM; GSEA; prognostic biomarker; diagnostic biomarker;

Posted Date: June 29th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-35310/v1

Abstract

Background: Glioblastoma (GBM) is the most lethal primary brain cancer and its survival rate is very low. Comprehensive genomic characteristics and high degree of heterogeneity of GBM are main causes that contribute to the absence of effective therapeutic targets and prognostic markers and eventual treatment failures. Here, Gene set enrichment analysis (GSEA) was used to explore comprehensive genomic characteristics and high degree of heterogeneity of GBM. Our study will help explain potential tumorigenesis mechanism and contribute to development of targeted therapeutics and prediction of patient prognosis.

Methods: Gene expression profile of GSE50161 was downloaded from Gene Expression Omnibus (GEO) database and GSEA was performed to evaluate the microarray data at level of gene sets of GO, KEGG and hallmark of Molecular Signatures Database (MSigDB). Core enrichments of selected hallmark gene sets were applied for clinical prognosis verification in Gene Expression Profiling Interactive Analysis (GEPIA). Genes associated with significant overall survival (OS) and unreported before were recognized as novel; the expression levels of novel genes in GBM samples and every novel gene between normal brain and GBM samples were compared. Receiver operating characteristic (ROC) and Pearson correlation analysis were also performed to evaluate the diagnostic biomarkers of GBM and correlations among novel genes respectively.

Results: We obtained 511 and 494 GO gene sets, 12 and 10 KEGG gene sets, and 2 and 8 hallmark gene sets in normal and GBM samples respectively. Five novel genes SERPINA5, TENM2, ARAP3, THBD, TCF19 associated with GBM prognosis were selected and four novel genes SERPINA5, TENM2, ARAP3, TCF19 performed well for GBM diagnosis prediction were obtained. In addition, we also found that four novel genes SERPINA5, ARAP3, THBD, TCF19 that high expressed in GBM samples were all positively correlated with each other and correlation between ARAP3 and TCF19 was the strongest.

Conclusions: Our study will help deeper understand the molecular heterogeneity and develop of targeted therapeutics of GBM; five novel related to prognosis genes (of which four are also related to diagnosis) may be applied to more accurate clinical decision-making process.

Introduction

GBM is the most lethal primary brain cancer making up 57% of all gliomas and malignant central nervous system tumors[1]. Despite multidisciplinary treatments of resection, radiotherapy and adjuvant chemotherapy, the median survival for patients with GBM is only 15 months and 5-year survival rates remain < 10%[2, 3]. Comprehensive genomic characteristics and high degree of heterogeneity of GBM are main causes that contribute to the absence of effective therapeutic targets and prognostic markers and eventual treatment failures. Therefore, understanding the molecular alterations of GBM will help explain potential tumorigenesis mechanism and contribute to development of targeted therapeutics and prediction of patient prognosis[4]. Now, high-throughput genomic technology (such as microarrays) combined with bioinformatics analysis make it possible to reveal the biologic heterogeneity and molecular alterations of GBM.

GSEA, a Java based software, is a robust computational method that assesses whether an a-priori defined set of genes shows statistically significant, concordant differences between two groups. Initial application of GSEA is to microarray experiments, and now the applications distribute many fields, including RNA-seq gene expression experiments, genome-wide associations studies, proteomics and metabolomics studies[5–8]. Here, gene expression profile data GSE50161 was downloaded from GEO database and GSEA was performed to evaluate the microarray data at level of gene sets of Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and hallmark of Molecular Signatures Database (MSigDB)[9]. Based on the set cut-off value, hub gene sets of GO, KEGG and hallmark were screened in GSEA software and genes of core enrichment of hallmark gene sets were selected for clinical prognosis verification in a web server of Gene Expression Profiling Interactive Analysis (GEPIA)[10]. In short, our study will help deeper understand the molecular heterogeneity and develop of targeted therapeutics of GBM; five novel related to prognosis genes (of which four are also related to diagnosis) may be applied to more accurate clinical decision-making process.

Methods

Microarray data preparation

Series Matrix file of GSE50161 including 13 normal brain samples, 34 GBM samples and 83 other brain tumors was downloaded from GEO database. Perl language was used to map the probe data to gene symbols based on GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array) annotation information[11]. Subsequently, K-nearest neighbor (KNN) method of impute R package was used to fill the probe missing value; background correction, quartile normalization and probe summarization were done using the limma R package[12, 13]. Finally, normal brain and GBM samples were selected for study.

Identification of gene sets of GO, KEGG and hallmark of MSigDB

MSigDB is one of the most widely used and comprehensive databases of gene sets and can provide the benefits of better representation and coverage of biological processes[14]. GO, KEGG and hallmark gene sets were downloaded from MSigDB database (version:7.1) and the obtained annotation files of c5.all.v7.1.symbols.gmt, c2.cp.kegg.v7.1.symbols.gmt, h.all.v7.1.symbols.gmt were input into GSEA software to detect the gene sets in normal brain and GBM samples. “1,000” times of permutations, “False” of collapse dataset to gene symbols and “phenotype” of permutation type were set. Besides, normalized enrichment scores (|NES|) > 1.0, false discovery rate (FDR) $q < 0.25$ and nominal $p < 0.05$ were set as cut-off criteria to select hub gene sets.

novel genes selection, expression level and survival analysis

Based on the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) data, GEPIA is a web-based tool to deliver fast and customizable functionalities [10]. Therefore, in GEPIA server, GBM patients were split into two groups by the median expression of genes (high vs. low expression) and the overall survival (OS) was assessed by Kaplan–Meier survival analysis(log-rank p value < 0.05); subsequently, all the core enrichments of hub hallmark gene sets were input into GEPIA to determine whether a gene have different prognosis. Genes associated with significant OS and unreported before were recognized as novel and the expression levels of novel genes were compared in GBM samples, in addition, expression level comparations of every novel gene were also performed between normal brain and GBM samples.

Evaluation of diagnostic biomarkers and correlations

Novel genes were applied for further analysis. Based on the microarray expression values, receiver operating characteristic (ROC) analyses were performed in GraphPad Prism 8 software to evaluate the specificity and sensitivity of novel genes for GBM prediction, a gene associated with an area under the curve (AUC) value greater than 0.8 was considered good. Besides, in GEPIA server, four novel genes that were high expressed in GBM samples were also correlated by Pearson correlation analysis.

Result

Differential expression levels of gene symbols

A total of 19171 gene symbols that had complete expression values in normal brain and GBM samples were obtained. Some genes showed significantly differential expression in normal brain and GBM samples (Fig. 1).

Hub gene sets of GO, KEGG and hallmark

There were 511 and 494 selected GO gene sets, 12 and 10 selected KEGG gene sets, and 2 and 8 selected hallmark gene sets in normal and GBM samples, respectively. According to the |NES| of each gene sets, *GO_SODIUM_ION_TRANSMEMBRANE_TRANSPORT*, *KEGG_CARDIAC_MUSCLE_CONTRACTION*, and *HALLMARK_PANCREAS_BETA_CELLS* were the highest of GO, KEGG, and hallmark gene sets in normal brain samples; *GO_ANAPHASE_PROMOTING_COMPLEX_DEPENDENT_CATABOLIC_PROCESS, KEGG_CYTOSOLIC_DNA_SENSING_PATHWAY*, and *HALLMARK_MITOTIC_SPINDLE* were the highest of GO, KEGG, and hallmark gene sets in GBM samples, respectively, here, we showed top five gene sets of GO and KEGG and all the gene sets of hallmark (Table 1,2; Fig. 2,3). All gene sets of GO and KEGG of normal brain and GBM were presented as additional files [see Additional file 1,2].

Table.1 Top five gene sets of GO and KEGG were showed. According to the |NES|, GO gene sets were all

enriched in GBM samples; KEGG gene sets were enriched in GBM samples and normal brain samples.

PHENOTYPE	NAME	SIZE	NES	NOM p-val	FDR q-val
GBM	GO_ANAPHASE_PROMOTING_COMPLEX_DEPENDENT_CATABOLIC_PROCESS	76	1.949	0.000	0.222
	GO_SPINDLE_MICROTUBULE	56	1.919	0.000	0.186
	GO_DNA_PACKAGING_COMPLEX	24	1.889	0.002	0.192
	GO_MITOTIC_SPINDLE_ORGANIZATION	96	1.818	0.000	0.246
	GO_MICROTUBULE_CYTOSKELETON_ORGANIZATION_INVOLVED_IN_MITOSIS	122	1.813	0.000	0.233
GBM	KEGG_CYTOSOLIC_DNA_SENSING_PATHWAY	50	1.734	0.008	0.163
	KEGG_CELL_CYCLE	122	1.692	0.006	0.193
NORMAL	KEGG_CARDIAC_MUSCLE_CONTRACTION	70	-1.847	0.000	0.070
	KEGG_TASTE_TRANSDUCTION	44	-1.772	0.004	0.086
	KEGG_LONG_TERM_POTENTIATION	67	-1.730	0.002	0.100

Table.2 A total of ten hallmark gene sets were selected, eight gene sets were enriched in GBM samples and two gene sets were enriched in normal brain samples.

PHENOTYPE	NAME	SIZE	NES	NOM p-val	FDR q-val
GBM	HALLMARK_MITOTIC_SPINDLE	197	1.820	0.000	0.010
	HALLMARK_G2M_CHECKPOINT	184	1.671	0.000	0.069
	HALLMARK_E2F_TARGETS	188	1.568	0.000	0.110
	HALLMARK_NOTCH_SIGNALING	31	1.535	0.010	0.131
	HALLMARK_MTORC1_SIGNALING	185	1.628	0.012	0.088
	HALLMARK_INTERFERON_ALPHA_RESPONSE	90	1.482	0.039	0.134
	HALLMARK_ANGIOGENESIS	34	1.507	0.040	0.130
	HALLMARK_SPERMATOGENESIS	132	1.380	0.046	0.177
NORMAL	HALLMARK_PANCREAS_BETA_CELLS	40	-1.559	0.002	0.193
	HALLMARK_KRAS_SIGNALING_DN	188	-1.515	0.000	0.154

survival analyses and expression levels of novel genes

A total of 347 core enrichments of hub hallmark gene sets were obtained [see Additional file 3] and 5 unreported genes SERPINA5, TENM2, ARAP3, THBD, TCF19 that had survival differences were selected (Fig. 4a-e). Comparation of expression levels of novel genes in GBM samples showed that ARAP3 had the highest expression value, TENM2 had the lowest expression value (Fig. 5a). In addition, we also showed every expression level of novel genes in normal and GBM samples and found that all genes except TENM2 were highly expressed in tumor samples, TENM2 was highly expressed in normal samples (Fig. 5b-f).

diagnostic biomarkers and correlations analyses

Four novel genes SERPINA5, TENM2, ARAP3, TCF19 performed well for diagnosis prediction ($p < 0.001$) and ROC analyses displayed an AUC of 0.8303, 0.8937, 0.9638, and 0.9819 respectively; however, THBD was not suitable for diagnosis prediction for an AUC of 0.6719 ($p = 0.0707$) (Fig. 4f-g). Since TENM2 was low expressed in GBM samples, we correlated the other four novel genes that were high expressed in GBM samples by Pearson correlation analysis. We found that the four novel genes were positively correlated with each other and correlation between ARAP3 and TCF19 was the strongest (Fig. 6).

Discussion

when no significant expression level differences, simply looking for differentially expressed genes that exist between normal and tumor tissues to associate biological phenotypes with their underlying molecular mechanisms may lose some vital information. Unlike this, GESA, based on a given list of genes, focuses on evaluating whether an a-priori defined set of genes shows statistically significant, concordant differences between two groups[15]. In our study, we performed GSEA to evaluate the microarray data at level of gene sets of GO, KEGG and hallmark of MSigDB, the aims were to select some important and meaningful gene sets distributed in normal brain and GBM tissues and hub genes that would help deeper understand the molecular heterogeneity of GBM and were used to GBM diagnosis and treatment. In total, 511 GO gene sets, 12 KEGG gene sets and 2 hallmark gene sets were selected in normal brain tissues. 494 GO gene sets, 10 KEGG gene sets and 8 hallmark gene sets were selected in GBM tissues. In normal brain samples, *GO_SODIUM_ION_TRANSMEMBRANE_TRANSPORT*, *KEGG_CARDIAC_MUSCLE_CONTRACTION*, and *HALLMARK_PANCREAS_BETA_CELLS* had the highest |NES| in GO, KEGG, and hallmark gene sets, respectively; In GBM samples, *GO_ANAPHASE_PROMOTING_COMPLEX_DEPENDENT_CATABOLIC_PROCESS, KEGG_CYTOSOLIC_DNA_SENSING_PATHWAY*, and *HALLMARK_MITOTIC_SPINDLE* had the highest |NES| in GO, KEGG, and hallmark gene sets, respectively.

In addition, five novel and related to GBM patients' prognosis genes (SERPINA5, TENM2, ARAP3, THBD, TCF19) were also selected from the 347 core enrichments in hallmark gene sets. SERPINA5 (serpin peptidase inhibitor clade A member 5), a member of the alpha-1-antitrypsin clade (clade A) of serpins, was originally identified as an inhibitor of the anticoagulant protease activated protein C (aPC) and was therefore named protein C inhibitor (PCI). However, SERPINA5 was not only an inhibitor of activated protein C in plasma, but a serpin with broad protease reactivity and wide tissue distribution, as a result, the biological role of SERPINA5 had not yet been define[16]. As for the role in tumors, although Ying Jing and Bijsmans respectively found that SERPINA5 inhibited tumor cell migration in hepatocellular carcinoma and loss of SERPINA5 protein expression was associated with advanced-stage serous ovarian tumors[17, 18], Kunihiro Asanuma also found that Protein C inhibitor could inhibit breast cancer cell growth, metastasis and angiogenesis[19], to the best of our knowledge, studies on SERPINA5 in intracranial tumors had not been reported. Our analyses showed that SERPINA5 was high expressed in GBM tissues and may be a risk factor for poor prognosis. Teneurins with four paralogs (teneurin-1, teneurin-2, teneurin-3,teneurin-4) were type II transmembrane glycoproteins and mainly involved in morphogenesis and development of the central nervous system[17, 18]. As one of the four paralogs, Annemarie Ziegler found that TENM2(teneurin-2)was primarily expressed in breast and cervix cancer and in neuroblastoma cells, but almost no expression in gastric cancer cells. Besides, the authors also found that the expression of TENM2 in ovarian cancer had a negative relationship with tumor differentiation and survival time[19]. In our analyses, we found that the expression level of TENM2 in GBM tissues was lower than that in normal brain tissues, which was also a special case we encountered in five novel genes. However, it was worth noting that different expression level of TENM2 still had prognostic significance. ARAP3 (ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 3) encoded a phosphoinositide binding protein containing ARF-GAP, RHO-GAP, RAS-associating, and pleckstrin homology domains, in which ARF-GAP and RHO-GAP domains cooperated in mediating rearrangements in the cell cytoskeleton and cell shape. Deficiency of ARAP3 in cells may cause defects of cell–autonomous and lamellipodia producing[23–26]. Yagi reported that ARAP3 inhibited peritoneal dissemination of scirrhous gastric carcinoma cells by regulating cell adhesion and invasion and Qing-Xuan Wang found that downregulation of ARAP3 led to ability inhibition of cell proliferation, migration and invasion of papillary thyroid carcinoma[27, 28]. Our study found that the expression level of ARAP3 in GBM was the highest, and its expression level was clearly related to the prognosis of patients. THBD (thrombomodulin) was a membrane receptor that bound thrombin resulting in activation of protein C, degradation clotting factors Va and VIIIa and reduction of thrombin[29]. Studies had found that THBD was expressed in bladder cancer, lung cancer

and other tumors, and its expression level was related to tumorigenesis, invasiveness and metastasis[30–32], specifically, the expression of THBD was inversely correlated with tumor progression. There were no reports of the expression level of THBD in brain tumors, our study showed that THBD was highly expressed in GBM patients, and patients with different expression levels had inequable prognosis. Finally, several studies had found that in non-small cell lung cancer, hepatocellular carcinoma, and colorectal cancer TCF19 (transcription factor 19) could promote cell proliferation and increase tumor invasiveness through some pathways, but there were no reports about its role in GBM[33–36]. Our study found that TCF19 was highly expressed in GBM patients, and patients with different expression levels also had inequable prognosis.

Besides, novel genes were performed ROC and Pearson correlation analyses in GraphPad Prism 8 software and GEPIA server respectively. We found that four novel genes SERPINA5, TENM2, ARAP3, TCF19 performed well for diagnosis prediction for an AUC of 0.8303, 0.8937, 0.9638, and 0.9819 respectively ($p < 0.001$); however, THBD was not suitable for diagnosis prediction for an AUC of 0.6719 ($p = 0.0707$). Four novel genes SERPINA5, ARAP3, THBD, TCF19 that high expressed in GBM samples were correlated by Pearson correlation analysis and the results showed that they were positively correlated with each other.

In conclusion, GSEA analyses were conducted to evaluate gene sets between normal brain tissues and GBM tissues. Hub gene sets of GO, KEGG, and hallmark of normal brain and GBM tissues as well as five novel genes (SERPINA5, TENM2, ARAP3, THBD, TCF19) related to GBM patients' prognosis and diagnosis were selected. Our study would help deeper understand the molecular heterogeneity and develop of targeted therapeutics of GBM; five novel related to prognosis genes (of which four are also related to diagnosis) may be applied to more accurate clinical decision-making process.

Conclusions

Our study will help deeper understand the molecular heterogeneity and develop of targeted therapeutics of GBM; five novel related to prognosis genes (of which four are also related to diagnosis) may be applied to more accurate clinical decision-making process.

List Of Abbreviations

AUC, area under the curve;

FDR, false discovery rate;

GBM, Glioblastoma;

GEO, Gene Expression Omnibus;

GEPIA, Gene Expression Profiling Interactive Analysis;

GO, Gene Ontology;

GSEA, Gene set enrichment analysis;

GTEx, Genotype-Tissue Expression;

KEGG, Kyoto Encyclopedia of Genes and Genomes;

MSigDB, Molecular Signatures Database;

NES, normalized enrichment scores;

OS, overall survival;

ROC, Receiver operating characteristic;

TCGA, the Cancer Genome Atlas;

Declarations

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

These data we used were derived from the following resources available in the public domain: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE50161.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by the National Natural Science Foundation of China under Grant 81772693.

Authors' contributions

Conception and study design: G.W.L., L.J.C., X.T. and J.H.H. Data collection: G.W.L. and X.T. Data analysis: G.W.L. and J.H.H. Drafting the article: G.W.L., L.J.C., X.T. and J.H.H. All authors read and approved the final manuscript. LX.Z. conceived and designed the study.

Acknowledgements

Not applicable

References

1. 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 Oncol. 2018;20: 1–86. doi:10.1093/neuonc/noy131
2. Koshy M, Villano JL, Dolecek TA, Howard A, Mahmood U, Chmura SJ, et al. Improved survival time trends for glioblastoma using the SEER 17 population-based registries. J Neurooncol. 2012;107: 207–212. doi:10.1007/s11060-011-0738-7
3. Tan AC, Ashley DM, López GY. Management of Glioblastoma : State of the Art and Future Directions. CA CANCER J CLIN. 2020;0: 1–14. doi:10.3322/caac.21613
4. Aldape K, Zadeh G, Mansouri S, Reifenberger G, Deimling A Von. Glioblastoma : pathology , molecular mechanisms and markers. Acta Neuropathol. 2015;129: 829–848. doi:10.1007/s00401-015-1432-1
5. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL. Gene set enrichment analysis : A knowledge-based approach for interpreting genome-wide. Proc Natl Acad Sci. 2005;102: 15545–15550.
6. Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: A web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. Nucleic Acids Res. 2010;38: 90–95. doi:10.1093/nar/gkq324
7. Lavallée-Adam M, Rauniyar N, McClatchy DB, Yates JR. PSEA-quant: A protein set enrichment analysis on label-free and label-based protein quantification data. J Proteome Res. 2014;13: 5496–5509. doi:10.1021/pr500473n
8. Xia J, Wishart DS. MSEA: A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. Nucleic Acids Res. 2010;38: 71–77. doi:10.1093/nar/gkq329
9. Liberzon A, Birger C, Helga T, Ghandi M, Jill P, Mesirov, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1: 417–425. doi:10.1016/j.cels.2015.12.004
10. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;1. doi:10.1093/nar/gkx247
11. Morris JA, Gayther SA, Jacobs IJ, Jones C. A suite of Perl modules for handling microarray data. Bioinformatics. 2008;24: 1102–1103. doi:10.1093/bioinformatics/btn085
12. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics. 2010;26: 2363–2367. doi:10.1093/bioinformatics/btq431
13. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic Acids Res. 2008;36: 1–9. doi:10.1093/nar/gkm1075
14. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. BIOINFORMATICS. 2011;27: 1739–1740. doi:10.1093/bioinformatics/btr260
15. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P : a desktop application for Gene Set Enrichment Analysis. BIOINFORMATICS. 2007;23: 3251–3253. doi:10.1093/bioinformatics/btm369
16. Yang H, Geiger M. Cell penetrating SERPINA5 (Protein C inhibitor , PCI): More questions than answers. Semin Cell Dev Biol. 2017;62: 187–193. doi:10.1016/j.semcdb.2016.10.007

17. Jing Y, Jia D, Wong C, Ng IO. SERPINA5 inhibits tumor cell migration by modulating the fibronectin e integrin b 1 signaling pathway in hepatocellular carcinoma. Mol Oncol. 2014;8: 3 6 6 e37 7.
18. Bijsmans ITGW, Smits KM, Graeff P De, Wisman GBA, Zee AGJ Van Der, Slangen BF, et al. Loss of SerpinA5 protein expression is associated with advanced-stage serous ovarian tumors. Mod Pathol. 2011;24: 463–470. doi:10.1038/modpathol.2010.214
19. Asanuma K, Yoshikawa T, Hayashi T, Akita N, Nakagawa N, Hamada Y, et al. Protein C inhibitor inhibits breast cancer cell growth, metastasis and angiogenesis independently of its protease inhibitory activity. Int J Cancer. 2007;121: 955–965. doi:10.1002/ijc.22773
20. Li J, Shalev-benami M, Sando R, Su TC, Li J, Shalev-benami M, et al. Structural Basis for Teneurin Function in Circuit- Wiring : A Toxin Motif at the Synapse. Cell. 2018;173: 735–748. doi:10.1016/j.cell.2018.03.036
21. Ziegler A, Corvalán A, Roa I, Brañes JA, Wollscheid B. Teneurin protein family: An emerging role in human tumorigenesis and drug resistance. Cancer Lett. 2012;326: 1–7. doi:10.1016/j.canlet.2012.07.021
22. Graumann R, Capua GAD, Oyarzún JE, Vásquez MA, Liao C, Brañes JA, et al. Expression of teneurins is associated with tumor differentiation and patient survival in ovarian cancer. PLoS One. 2017;12: 1–24. doi:10.1371/journal.pone.0177244
23. Krugmann S, Anderson KE, Ridley SH, Risso N, McGregor A, Coadwell J, et al. Identification of ARAP3, a novel PI3K effector regulating both Arf and Rho GTPases, by selective capture on phosphoinositide affinity matrices. Mol Cell. 2002;9: 95–108. doi:10.1016/S1097-2765(02)00434-3
24. Blighe K, Kenny L, Patel N, Guttery DS, Page K, Gronau JH, et al. Whole Genome Sequence Analysis Suggests Intratumoral Heterogeneity in Dissemination of Breast Cancer to Lymph Nodes. PLoS One. 2014; 1–11. doi:10.1371/journal.pone.0115346
25. Gambardella L, Hemberger M, Hughes B, Zudaire E, Andrews S, Vermeren S. PI3K Signaling Through the Dual GTPase – Activating Protein ARAP3 Is Essential for Developmental Angiogenesis. Sci Signal. 2010;3.
26. Krugmann S, Andrews S, Stephens L, Hawkins PT. ARAP3 is essential for formation of lamellipodia after growth factor stimulation. J Cell Sci. 2006;119: 425–432. doi:10.1242/jcs.02755
27. Yagi R, Tanaka M, Sasaki K, Kamata R, Nakanishi Y, Kanai Y, et al. ARAP3 inhibits peritoneal dissemination of scirrhous gastric carcinoma cells by regulating cell adhesion and invasion. Oncogene. 2011;30: 1413–1421. doi:10.1038/onc.2010.522
28. Wang Q, Chen E, Zheng Z, Wang Y, Zhang X, Wang O. Next-generation sequence detects ARAP3 as a novel oncogene in papillary thyroid carcinoma. Onco Targets Ther. 2016;9: 7161–7167.
29. Esmon CT. The roles of protein C and thrombomodulin in the regulation of blood coagulation. J Biol Chem. 1989;264: 4743–4746. doi:10.2491/jjsth1970.15.607
30. Wu C Te, Chang YH, Lin PY, Chen WC, Chen MF. Thrombomodulin expression regulates tumorigenesis in bladder cancer. BMC Cancer. 2014;14: 1–11. doi:10.1186/1471-2407-14-375
31. Liu PL, Tsai JR, Chiu CC, Hwang JJ, Chou SH, Wang CK, et al. Decreased expression of thrombomodulin is correlated with tumor cell invasiveness and poor prognosis in nonsmall cell lung cancer. Mol Carcinog. 2010;49: 874–881. doi:10.1002/mc.20663
32. Palmowski J. Thrombomodulin. J Thromb Haemost. 2003;1: 1515–1524. doi:10.1191/026635501701526957
33. Zhou ZH, Chen G, Deng C, Tang JM, Xie L, Zhou HY, et al. TCF19 contributes to cell proliferation of non-small cell lung cancer by inhibiting FOXO1. Cell Biol Int. 2019;43: 1416–1424. doi:10.1002/cbin.11189
34. C. X. ZENG1, S. B. FU2, W. S. FENG, J. Y. ZHAO, F. X. LI PG. TCF19 enhances cell proliferation in hepatocellular carcinoma by activating the ATK/FOXO1 signaling pathway. Neoplasma. 2019;66: 46–53. doi:10.4149/neo_2018_171227N845
35. Mondal P, Sen S, Klein BJ, Tiwary N, Gadad SS, Kutateladze TG, et al. TCF19 Promotes Cell Proliferation through Binding to the Histone H3K4me3 Mark. Biochemistry. 2020;59: 389–399. doi:10.1021/acs.biochem.9b00771
36. Du WB, Huang Z, Luo L, Tong SP, Li HQ, Li X, et al. TCF19 aggravates the malignant progression of colorectal cancer by negatively regulating WWC1. Eur Rev Med Pharmacol Sci. 2020;24: 655–663.

Figures

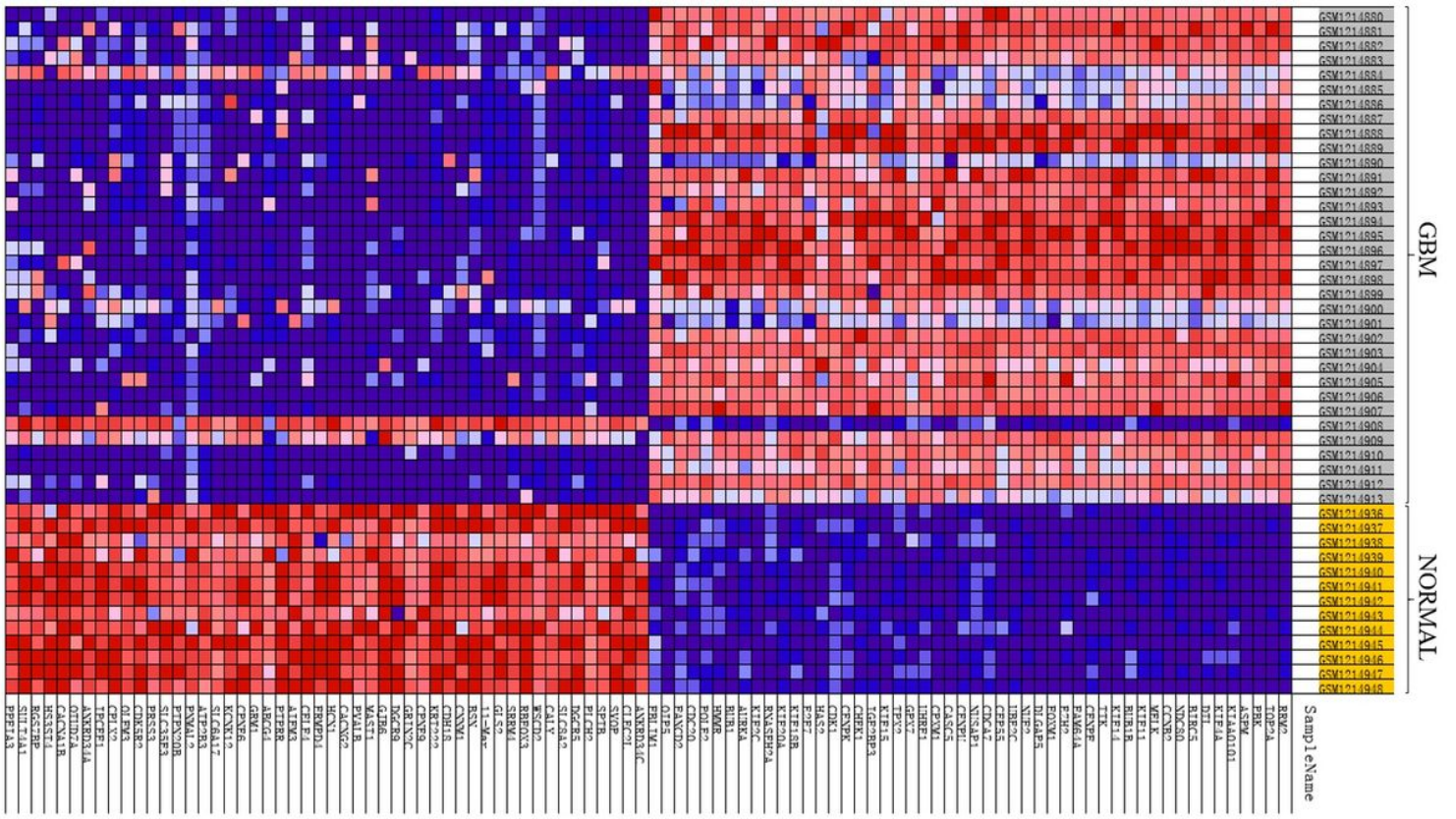

Figure 1

Some genes had a clear role in the difference of normal brain and GBM samples

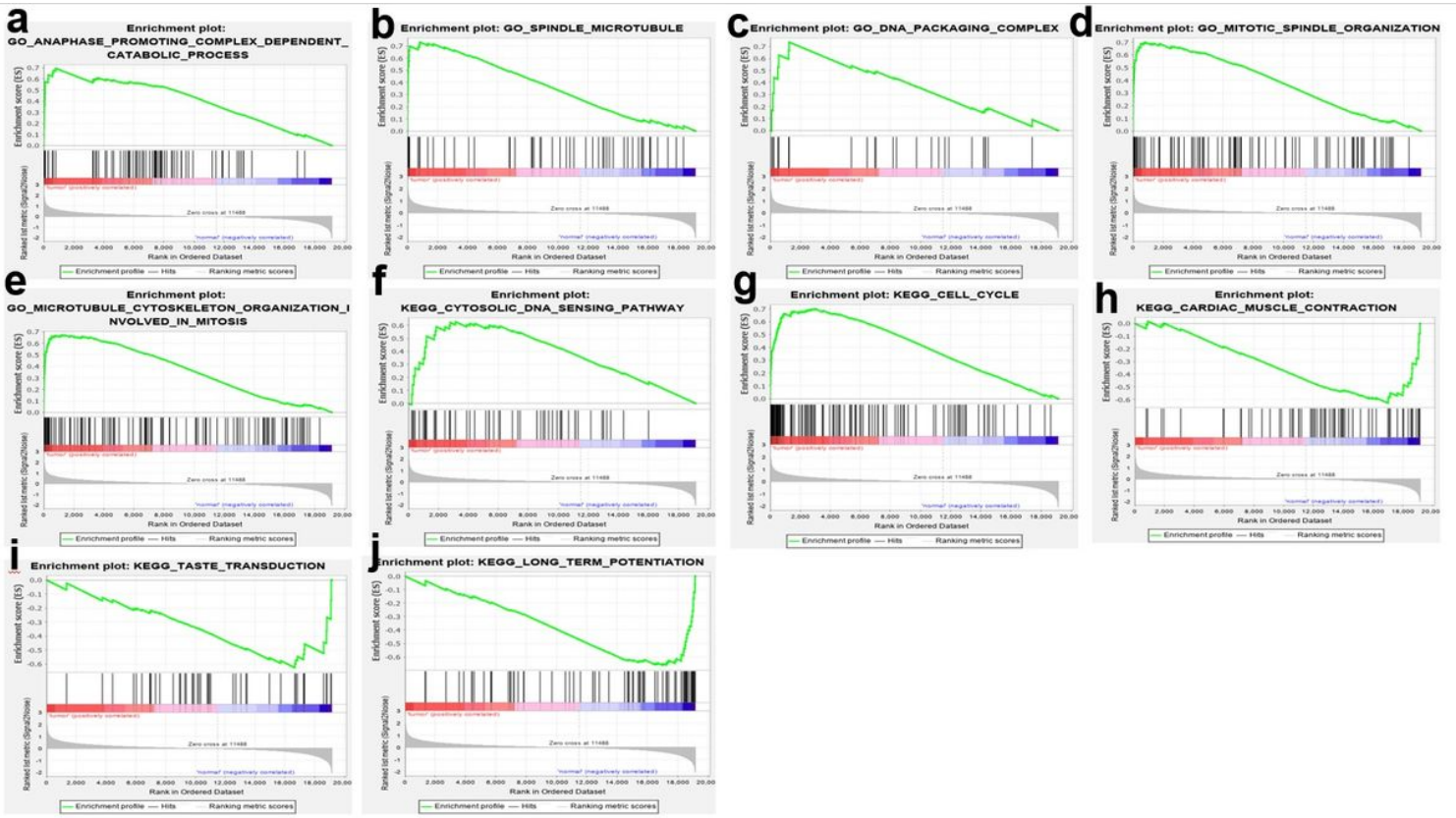

Figure 2

Top five gene sets of GO and KEGG were showed. According to the |NES|, all five GO gene sets were enriched in GBM samples (a-e); two and three KEGG gene sets were enriched in GBM samples (f,g) and normal brain samples (h-j) respectively.

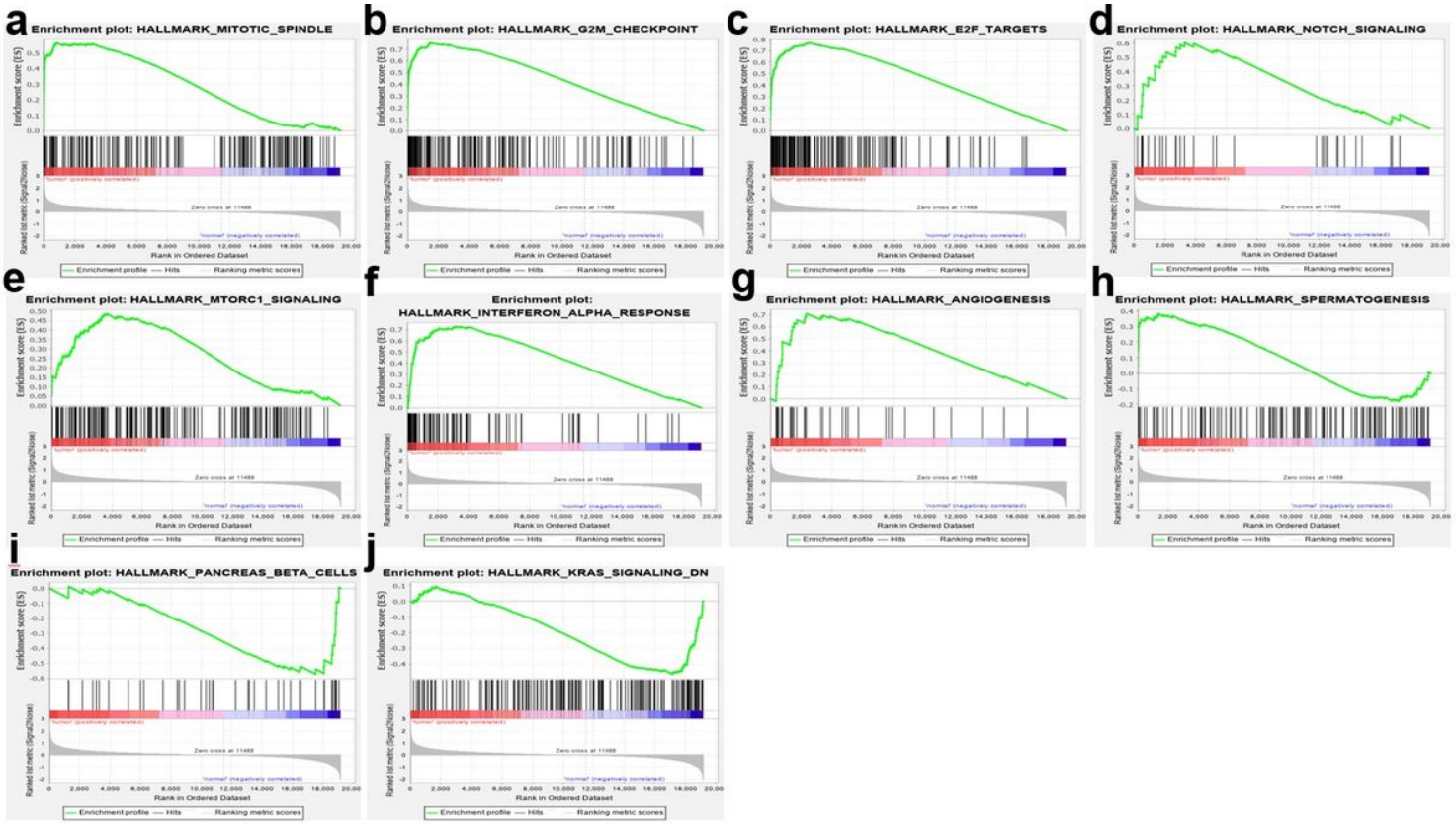

Figure 3

A total of ten hallmark gene sets were selected, eight gene sets were enriched in GBM samples(a-h), two gene sets were enriched in normal brain samples (i-j).

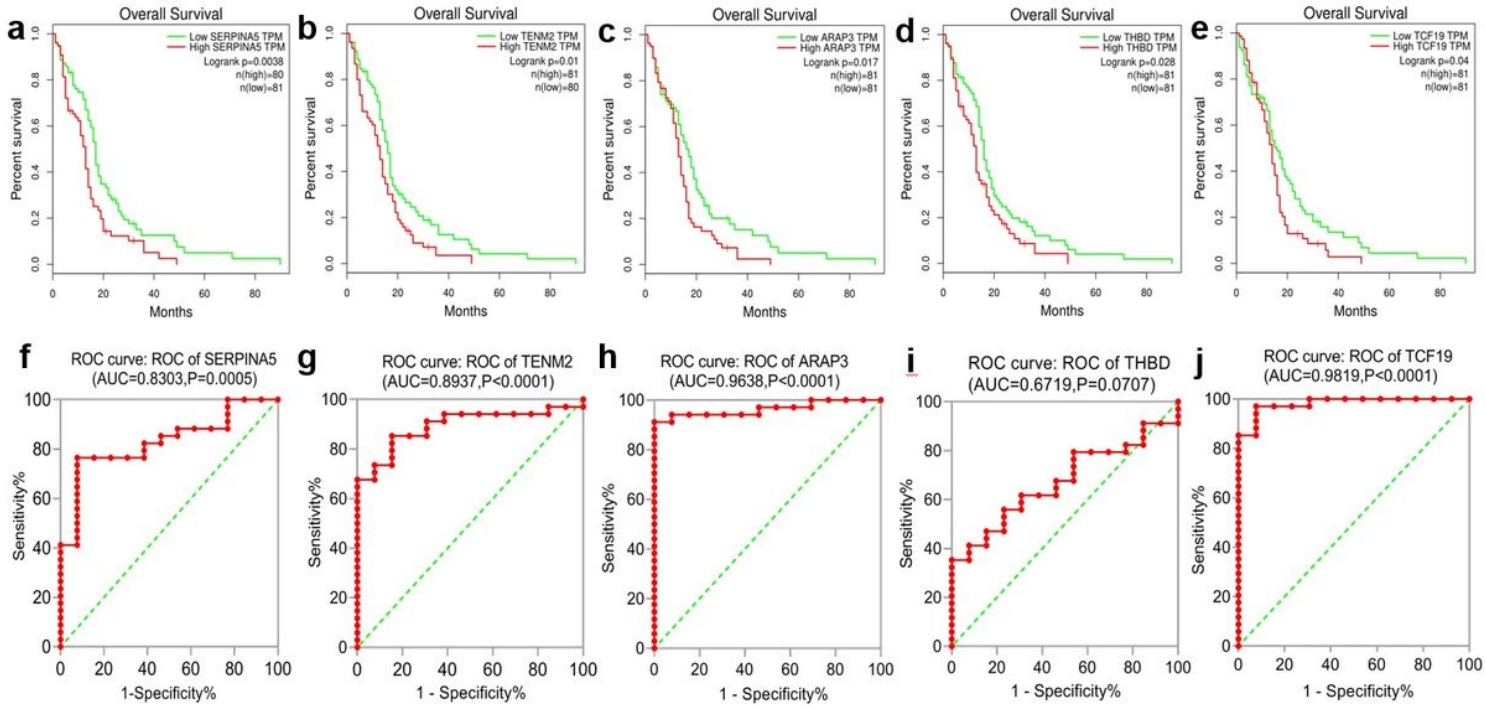

Figure 4

Prognostic survival analyses showed that SERPINA5, TENM2, ARAP3, THBD and TCF19 had significant overall survival (a-e); ROC analyses showed that SERPINA5, TENM2, ARAP3, TCF19 performed well for diagnosis prediction (f-h,j) but THBD did not (i).

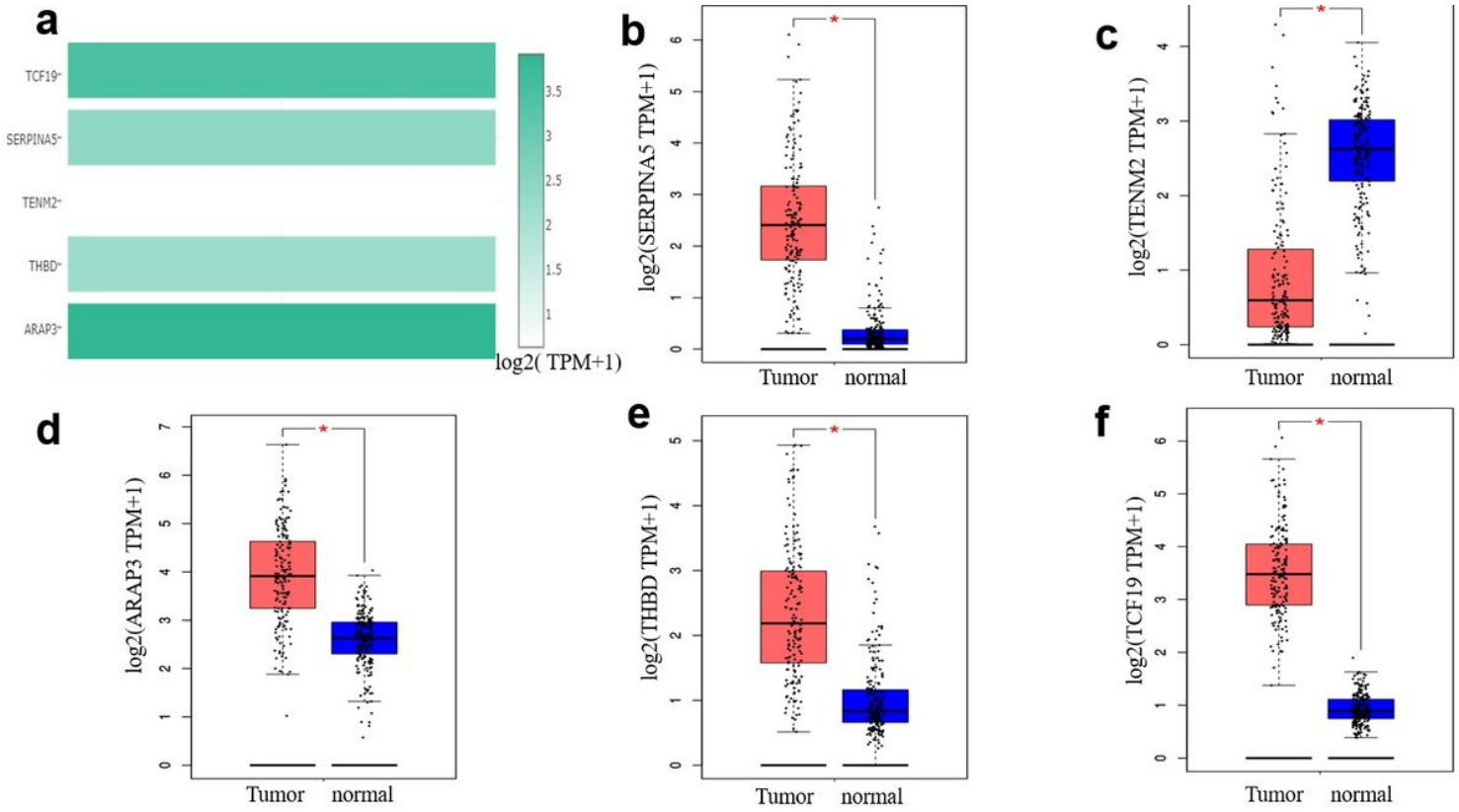

Figure 5

Comparation of expression levels of novel genes in GBM samples showed that ARAP3 had the highest expression value, TENM2 had the lowest expression value (a). Comparations of expression level of every novel gene in normal and GBM samples showed that all genes except TENM2 were highly expressed in GBM samples, TENM2 was highly expressed in normal samples (b-f).

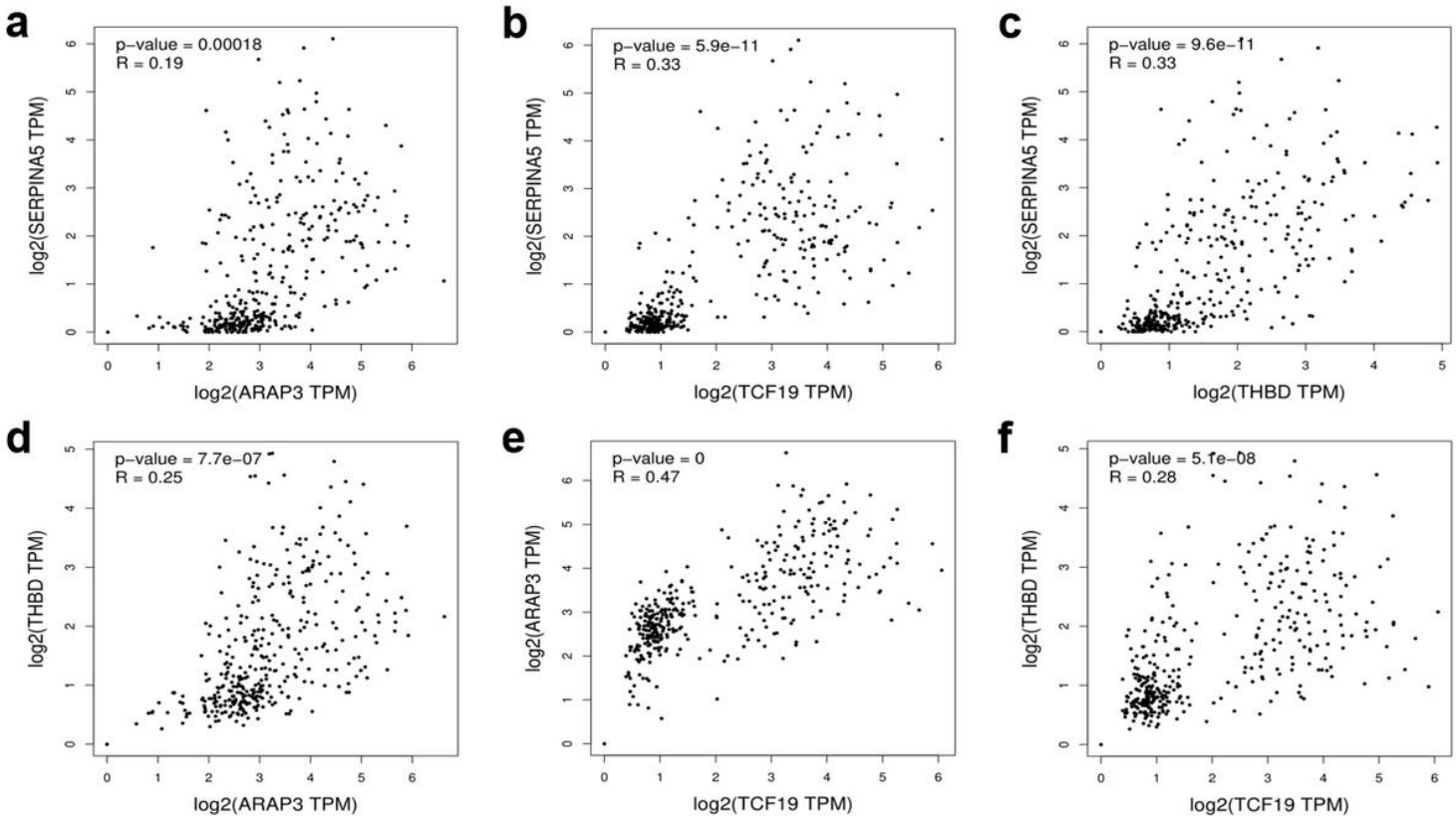

Figure 6

Four novel genes SERPINA5, ARAP3, THBD and TCF19 that were high expressed in GBM samples were positively correlated with each other and correlation between ARAP3 and TCF19 was the strongest.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xlsx
- Additionalfile2.xlsx
- Additionalfile3.xlsx