RESEARCH ARTICLE

Editorial Process: Submission:05/19/2024 Acceptance:12/14/2024

The G Protein-Coupled Receptor-Related Gene Signatures for Diagnosis and Prognosis in Glioblastoma: A Deep Learning Model Using RNA-Seq Data

Ghazaleh Khalili-Tanha[1], Nima Khalili-Tanha[2], Masoumeh Farahani[3], Mostafa Rezaei-Tavirani[4]*, Elham Nazari[4]*

Abstract

Background: Glioblastoma (GBM) is the most aggressive cancer in the central nervous system in glial cells. Finding novel biomarkers in GBM offers numerous advantages that can contribute to early detection, personalized treatment, improved patient outcomes, and advancements in cancer research and drug development. Integrating machine learning with RNAseq data in medicine holds significant potential for identifying novel biomarkers in various diseases, including cancer. **Methods:** Gene expression raw data was used to detect differentially expressed genes (DEGs) within a cohort of 532 GBM patients. The molecular pathway analysis, disease ontology, and protein-protein interactions of DEGs were assessed. Machine learning methods were performed to identify candidate genes. Survival curves were estimated using the Kaplan–Meier method and Cox proportional hazard to find prognostic biomarkers. **Results:** The molecular pathway analysis revealed that key dysregulated genes are in GPCRs, class A rhodopsin−like, MAPK signaling pathway, and calcium regulation in cardiac cells. Additionally, survival analysis showed that ten downregulated genes, including *CPLX3, GPR162, LCNL1, SLC5A5, GPR61, GPR68, IL1RL2, HCRTR1, AIPL1*, and *SYTL1*, and also ten upregulated genes, including *C1orf92, CATSPER1, CCDC19, EPS8L1, FAIM3, FAM70B, FCN3, GPR157, IGFBP1*, and *MYBPH* decreased the overall survival in GBM patients. Furthermore, the machine learning detected twenty genes, among which LRRTM2 and OPRL1 were candidates with high correlation coefficients. **Conclusion:** Our data suggest that genes belonging to G Protein-Coupled Receptors play a critical role in various aspects of glioblastoma progression and pathogenesis. Four members of GPCRs, including *GPR162, GPR61, GPR68*, and *GPR157*, can be considered prognostic biomarkers. Additionally, the combination of A2BP1 and GPR157 was reported as a diagnostic marker.

Keywords: Glioblastoma- RNA-Seq- Machine Learning- Deep Learning- Biomarkers

Asian Pac J Cancer Prev, **25 (12)**, 4201-4210

Introduction

Glioblastoma multiforme (GBM) is the most common and aggressive form of brain cancer. It is a type of glioma originating from the brain's glial cells. GBM typically affects adults between 45 and 70 [1, 2]. According to the American Cancer Society, the incidence rate of glioblastoma in the United States is estimated at approximately 3.2 cases per 100,000 people per year. According to the Central Brain Tumor Registry of the United States (CBTRUS), the 5-year survival rate for patients diagnosed with glioblastoma is around 5-10% or 12-15 months [3-5]. GBM is a multifactorial disease, and it is associated with genetic mutations and chromosomal abnormalities. Environmental factors for developing GBM are such as radiation exposure, specifically inherited syndromes, and a family history of brain tumors. GBM can cause various symptoms to vary depending on the location of cancer, but commonly include headaches, seizures, cognitive changes, weakness, and changes in vision or speech. Current treatment for GBM involves a combination of surgery, radiation therapy, and chemotherapy [6]. Biomarkers commonly studied in glioblastoma consist of mutations and dysregulation in various genes, including *IDH1, Tp53, EGFR, Ki-67*, and methylation of MGMT [7].

The discovery of new biomarkers can improve the study and understanding of the mechanisms underlying cancer. This has many advantages, including (1) Emerging new diagnostic tools and early detection. (2) Personalized

*[1]Department of Medical Genetics and Molecular Medicine, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. [2]Department of Small Animal Clinical Sciences, Western College of Veterinary Medicine, University of Saskatchewan, 52Campus Drive, Saskatoon, SK S7N 5B4, Canada. [3]Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [4]Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. *For Correspondence: Tavirani@sbmu.ac.ir*

medicine can involve selecting the most effective treatment options and avoiding unnecessary or ineffective treatments, thus improving treatment response rates. (3) Prognostic information provides important prognostic information, helping to predict the likelihood of disease progression and response to treatment. (4) Monitoring treatment response can allow for real-time monitoring of treatment response. (5) Cost-effectiveness, in which healthcare providers can focus on targeted treatments, avoiding unnecessary or ineffective therapies. This can reduce costs associated with treatment, minimize side effects, and improve patient quality of life [8-10].

Machine learning (ML) is a subset of artificial intelligence (AI) that uses algorithms and statistical models to enable computers to learn and make predictions or decisions without being explicitly programmed [11]. These days, by leveraging ML, researchers can train algorithms on large datasets of RNA-seq data from healthy individuals and patients with a specific condition. These algorithms can then learn patterns and relationships between gene expression levels and disease outcomes to identify potential biomarkers. Furthermore, machine learning algorithms can also be used to develop predictive models that incorporate RNA-seq data and other clinical and demographic variables to make accurate predictions about disease prognosis or treatment response. These predictive models can aid in personalized medicine by enabling doctors to identify patients most likely to benefit from a particular treatment or identify potential adverse reactions to certain medications based on RNA-seq data [12, 13].

The main aim of the current study was to develop a Machine Learning-based integrated bioinformatics approach for the identification of novel biomarkers in glioblastoma. Previous research has predominantly relied on bioinformatics for biomarker discovery, a method that, while valuable, typically necessitates manual data interpretation, rendering it time-consuming and less effective for managing large datasets. In contrast, this study integrates bioinformatics with machine learning techniques to enhance the process of biomarker discovery. The application of machine learning algorithms enables a more efficient analysis of complex datasets, reveals patterns that traditional methods might overlook, and significantly improves the accuracy and reliability of the findings [14].

In the current study, we identified biomarkers of GBM by a Machine Learning-based integrated bioinformatics approach. The survival analysis used the Kaplan–Meier method to predict prognostic biomarkers. Furthermore, we evaluated disease ontology, molecular pathways, protein-protein network, and correlations between candidate genes and glioblastoma with clinical data. Additionally, the diagnostic biomarker was detected by the machine learning technology and ROC curve (Figure 1A).

Materials and Methods

Data collection

RNA-seq data and associated clinicopathological data from patients diagnosed with glioblastoma (GBM) were downloaded from the GDAC database (http://gdac.broadinstitute.org/). In this study, we included 537 samples consisting of 532 tumor tissues and five normal samples and analyzed a dataset of 20,531 genes.

Data preprocessing and identification of differentially expressed genes

Raw RNA-seq data were normalized using the R software DESeq2 to remove samples with null expression and duplicates. Data preprocessing was used to show differentially expressed genes (DEG) based on specific criteria with a P value of <0.05 and $|\log2\text{-fold change (FC)}| > 1.5$ were considered a significant threshold.

Genome, disease ontology, and pathway enrichment analysis

Enrichment analysis and discovery of significant DEG pathways based on Gene Ontology (GO), GSEA, Reactome, and Human Disease Ontology (DO). Data were visualized with the respective packages in R with a p-value < 0.05.

Machine Learning Algorithm

A machine learning technique using deep learning algorithms was performed to analyze DEG values and identify prognostic markers of glioblastoma. Deep learning is a specialized field of machine learning that uses artificial neural networks to model and understand complex data patterns and extract advanced features from input data. A characteristic selection weight corresponding to the correlation is used when building the model. The correlation matrix shows the correlation coefficients between several variables. It helps to understand the relationships between variables and identify data patterns. The correlation coefficient, usually expressed as "randquot; to quantify, indicates the strength and direction of the linear relationship between two variables. It ranges from -1 to 1, indicating negative and positive associations, respectively. The closer the correlation coefficient is to -1 or 1, the stronger the relationship.

Computer workflow

The model was run in Python 3.7 with a learning rate of 0.01, the activation function chosen was Rectifier, and the model was trained for ten epochs. Generally, in deep learning, it is customary to partition the data into two categories, namely training and testing, as part of the standard workflow. The process involves modeling the training set, assessing performance on the test set, and repeating the model and data processing techniques to enhance overall performance and generalization capabilities. The dataset was split into a training set comprising 70% of the data and a testing set containing 30% of the data. In the machine learning workflow, the models were trained on the training set by utilizing fixed optimal hyperparameter values for AUC (area under the curve), accuracy, MSE (mean squared error), and R2 (R-squared). Accuracy can be calculated by determining the percentage of correctly predicted values from the total number of predictions made. This aids in obtaining a comprehensive model evaluation and ensures accurate

data classification. The calculation parameters for accuracy are TP (true positive), TN (true negative), FP (false positive), and FN (false negative). To calculate the accuracy, we can use the formula: Accuracy = (True Positives + True Negatives) / (True Positives + True Negatives + False Positives + False Negatives).

The Mean Squared Error (MSE) is a calculation that determines the average squared difference between the predicted values and the actual values. The degree of concordance between the model and predictions is displayed, wherein lower values signify superior performance. R2, also known as R-squared, represents a statistical measurement that reveals the extent to which the independent variables can elucidate the variation in the dependent variable. The numeric scale spans from 0 to 1, wherein higher values signify a more optimal alignment between the model and the collected data. The AUC (Area Under the Curve) is a metric used to determine the effectiveness of a model in discerning between positive and negative classes in a binary classification task. In other words, The probability of a randomly selected positive case being ranked higher than a negative one is quantified. The model's performance improves as the AUC value increases.

Protein-Protein Interaction (PPI)

Researchers in biology widely use the STRING database, providing valuable insights into the complex interactions that occur within living organisms. Accessible through its website (https://string-db.com), the STRING database is a valuable tool for studying protein-protein interactions. We analyzed the protein-protein interaction network of DEGs to uncover their biological relationship. Functional genomes and cellular pathways rely heavily on these networks for a comprehensive understanding. These interactions greatly influence the understanding of cellular pathways and functional genomes. The minimum threshold we set for the effective inclusion score is considered more significant than 0.4.

Identification of prognostic biomarkers

Survival analysis methods, including Cox proportional hazards and Kaplan-Meier plot, were performed to obtain the best DEG values using R4.2.1 software to evaluate independent prognostic biomarkers. All data were screened based on criteria consisting of HR and borderline: 1 and $P < 0.05$.

ROC curve

A pooled receiver operating characteristic (ROC) curve was performed to identify diagnostic biomarkers. The ROC curve allows visualization of the trade-off between sensitivity and specificity. The overall performance of the test can be determined by calculating the area under the ROC curve (AUC). AUC 1 indicates a perfect test that can completely distinguish disease from health. ROC curve analysis provides a comprehensive assessment of the diagnostic test's accuracy and helps determine its clinical utility in distinguishing between diseased and healthy cases.

Validation by other Datasets

To further confirm the candidate genes, two microarray datasets (GSE4290 and GSE68848) from the gene expression omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) were evaluated. The GEO2R tool was used to identify DEGs between human GBM tissues and control samples.

Results

Identification of DEGs and Pathway Enrichment Analysis

The data consists of 532 GBM samples and five controls. As shown in Table 1, the average age of patients was 58.13 years, most of whom were male. The data was downloaded and normalized, and then DEGs were obtained. 4940 DEGs were identified, of which 2302 were upregulated, and 2638 were downregulated genes, respectively. Heat map and principal component analysis (PCA) provide additional information that can help better understand differences or similarities between patients and controls regarding traits and potential genes (Figures 1B and C). Enrichment analysis indicated that DEGs were remarkably enriched in GPCRs, class A rhodopsin−like, MAPK signaling pathway, and Calcium regulation in cardiac cells (Figures 2A, B, and C). Disease Ontology (DO) results showed that DEGs are associated with other diseases, such as unipolar depression, memory impairment, mental disorders, and inflammation (Figure 2D). The protein-protein interaction network illustrated that hub genes related to candidate genes *GPR162, GPR61, GPR68*, and *GPR157* (Figure 3).

Machine Learning Data Analysis

The results of the machine learning analysis are shown in Table 2. The deep learning method achieved an AUC of 1, an accuracy of 98.24%, and an R^2 0f of 0.999 in predicting cervical cancer. Twenty genes were candidates for further analysis to identify diagnostic and prognostic biomarkers (Table 2).

Table 1. The Clinicopathological Characteristics of Glioblastoma Patients

Clinicopathological Variables	No. of patients (%) /mean ± SD
Patients	532
Mean age (Years, mean ± SD)	58.13 ± 14.07
Gender	
Male	328 (61.7)
Female	204 (38.3)
Vital status	
Alive	148 (27.8)
Dead	384 (72.2)
Race	
Asian	11 (2.1)
White	455 (85.5)
Black	46 (8.6)

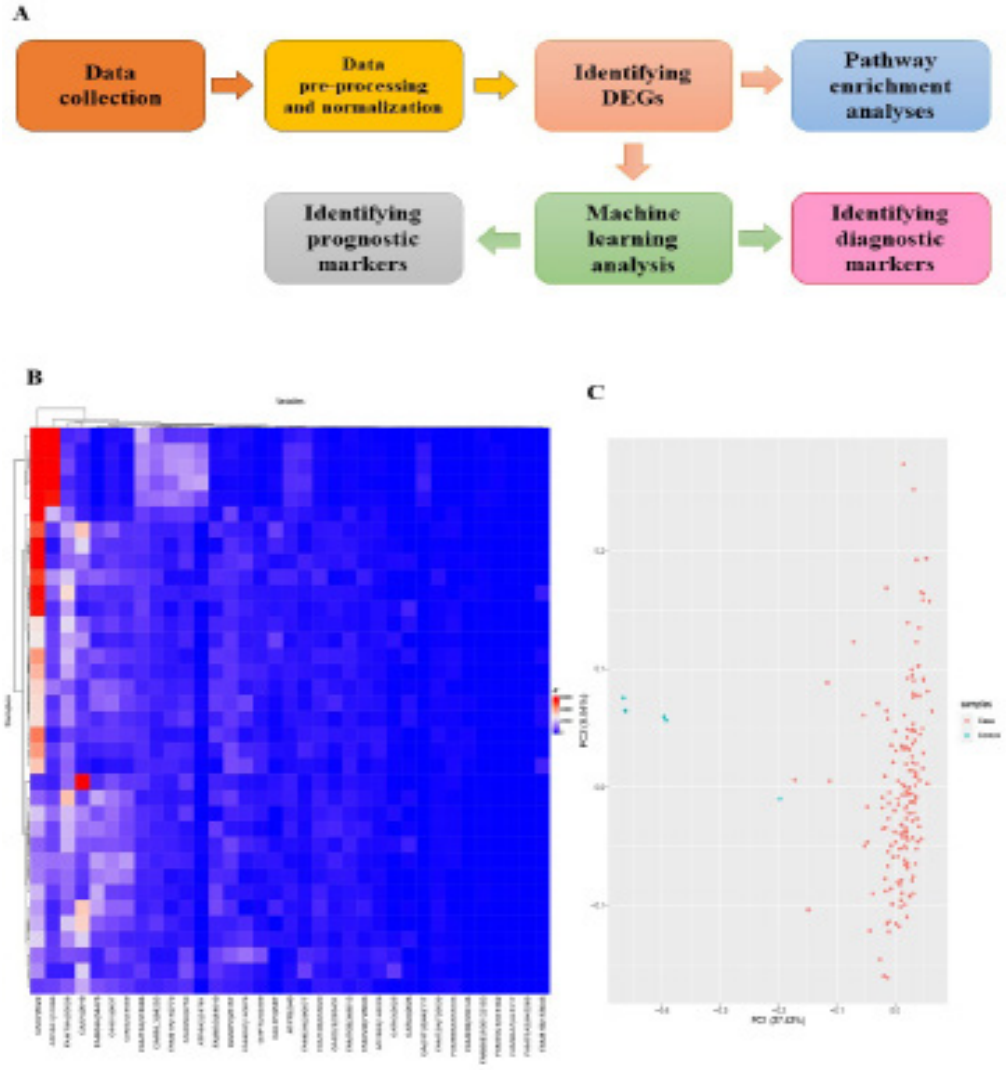

Figure 1. (A) The overall workflow, (B) Then heatmap of DEGs of GBM was drawn by R software, (C) Principal component analysis (PCA) plot.

Identify prognostic markers

Kaplan-Meier survival was used to assess the impact of prognostic candidate genes in cervical cancer. The result revealed that three downregulated genes, *CPLX3, GPR162,* and *LCNL1*, reduced overall survival (OS) (Figures 3A, C, E, and G). Our survival analysis showed that the downregulation of *GPR162* and *GPR61* and upregulation of *GPR68* and *GPR157* are prognostic markers for GBM and decreased the OS of the patients. In addition, univariate Cox proportional hazards regression was performed for DEGs to evaluate prognostic markers. The data showed that ten up-regulated genes and six down-regulated genes impaired the OS of the patients (Table 3).

Identification of diagnostic markers with the ROC curve

Our study showed that A2BP1 reached the highest value with an AUC of 0.99 (95% CI, sensitivity 0.93 and specificity 1.00). In addition, the combination of *A2BP1* with other genes, including *ABR, AIPL1, C1orf92, CPLX3, EPS8L1, GPR157, HCRTR1, IL1RL2, KCNAB1, OPRL1, PPIAL4C,* and *SLC5A5* showed potential as a diagnostic marker with AUC, sensitivity, and specificity value of 0.98, 1.00 and 0.99, respectively (Figure 4A).

Investigation of the correlations of Clinical/ Demographic with glioblastoma

The information showed a correlation between age and race (p-value= = 0.003). The results showed the average age of Asian patients was lower than that of white and African-American patients (Figure 4B).

External validation

The expression levels of candidate genes in GBM patients were verified using two GEO datasets (GSE4290 including 77 GBM samples and 23 control samples based on GPL570 and GSE68848 including 228 GBM samples and 28 control samples based on GPL570) Figure 4C presents the heatmap of expression changes.

Discussion

Glioblastoma is the most aggressive cancer in the central nervous system in glial cells. Finding novel biomarkers in GBM offers numerous advantages that

Table 2. The Candidate Gene Detected by Machine Learning Analysis

Candidate genes	Coefficient	Performance of Deep Learning approach
LRRTM2	1	
OPRL1	0.989824	
PPIAL4C	0.975116	
DKFZp434J0226	0.96836	MSE: 4.324857E-8
FOXD4L2	0.966007	
PPM1E	0.964098	RMSE: 2.079629E-4
RAPGEF5	0.963645	
DIRAS1	0.96363	R^2: 0.99999845
SYNGAP1	0.962822	
LCNL1	0.961324	AUC: 1.0
ABR	0.671373	
LOXHD1	0.669941	Accuracy: 98.24%
KIF5A	0.669211	
CPLX3	0.667917	
ADRBK2	0.667077	
GPR162	0.665177	
KCNAB1	0.662677	
CMTM4	0.653963	
OR13J1	0.651024	
TP53TG5	0.650316	

Table 3. The Prognostic Biomarkers Identified by Kaplan–Meier Survival Analysis and Cox Proportional Hazards Regression

	P-value	log2FoldChange
Upregulated Genes		
MYBPH	0.011	4.732408
C1orf92	0.041	3.884442
CATSPER1	0.028	3.703642
FCN3	0.022	3.431704
IGFBP1	0.011	2.491103
GPR157	0.013	2.242312
CCDC19	0.006	2.893859
FAIM3	0.0012	2.69397
FAM70B	0.0012	1.981646
EPS8L1	6e−04	1.629814
Downregulated Genes	P-value	log2FoldChange
CPLX3	0.0038	-4.76083
SLC5A5	0.028	-4.61565
GPR61	0.0019	-4.75389
GPR68	0.0093	-3.40969
IL1RL2	0.027	-3.25222
LCNL1	0.024	-2.91717
HCRTR1	0.0081	-2.54194
AIPL1	0.0046	-2.38213
SYTL1	0.0046	-1.90601
GPR162	0.049	-1.84766

can contribute to early detection, personalized treatment, improved patient outcomes, and advancements in cancer research and drug development [9]. Recently, novel approaches, including machine learning, a branch of artificial intelligence, have been widely used in medicine to process and analyze large amounts of data, such as RNA sequencing (RNAseq). RNAseq is a technology used to measure gene expression levels by sequencing the RNA molecules in a biological sample. Integrating machine learning with RNAseq data in medicine holds significant potential for identifying novel biomarkers in various diseases, including cancer [15-18].

The DEGs' ontology and pathways analysis results revealed that they were significantly enriched in GPCRs, class A rhodopsin−like, MAPK signaling pathway, and Calcium regulation in cardiac cells. Class A Rhodopsin-like GPCRs are a specific G-protein coupled receptors (GPCRs) subclass. Class A Rhodopsin-like GPCRs have a similar structure to rhodopsin. The large family of cell surface receptors plays pivotal roles in various biological processes, including regulating neurotransmission, hormone signaling, and sensory perception. The previous studies indicated that the dysregulation of GPCR family members is associated with central nerve system tumors, and targeting GPCRs is considered a promising strategy for cancer treatment [19]. A bioinformatic analysis of GEO datasets on glioma transcriptomic data, including GSE43289, GSE4290, and GSE19728, showed that fourteen GPCR genes, nine upregulated and four downregulated, are related to the high severity of glioma [20]. Feve et al. [21] evaluated the GPCRs expression in mRNA and protein levels by TaqMan Low-Density Arrays and mass spectrometry in different glioblastoma cell lines, including U-87, TG1, OB1, HA, and f-NSCs. The results showed that *LPHN2, GPR56, F2R, FZD7, FZD3, GPRC5B, FZD1*, and *BDKRB2* genes overexpressed in cells. GnRH, Gonadotropin-releasing hormone receptor, a rhodopsin-like G-protein coupled receptor overexpressed in GBM cell line LN229. The data indicated that treating cells with GnRH agonists decreased cell proliferation. This is due to the interaction of GnRH and KNG1which, a cell proliferation regulation gene [22]. Our survival analysis showed that the downregulation of *GPR162* and *GPR61* and upregulation of GPR68 and GPR157 are prognostic markers for GBM and decreased the OS of the patients. GPR162 is a member of the rhodopsin-like G protein-coupled receptor (GPCR) family, which plays a critical role in DNA damage. GPR162 interacts with STING, a regulator of beta-interferon, to activate the STING-signaling pathway, resulting in increases in the expression level of different chemokines. Long et al. reported that the overexpression of GPR162 increases radiosensitivity in mice models for lung cancer and causes tumor shrinkage. Furthermore, the results of in-vitro indicated that the high expression level significantly decreases cell proliferation, invasion, and migration [23]. A study reported four members of GPCRs, including *GPR62, GPR77, GPR61*, and *GPR63* in different parts of the brain. They showed that GPR61 expression in

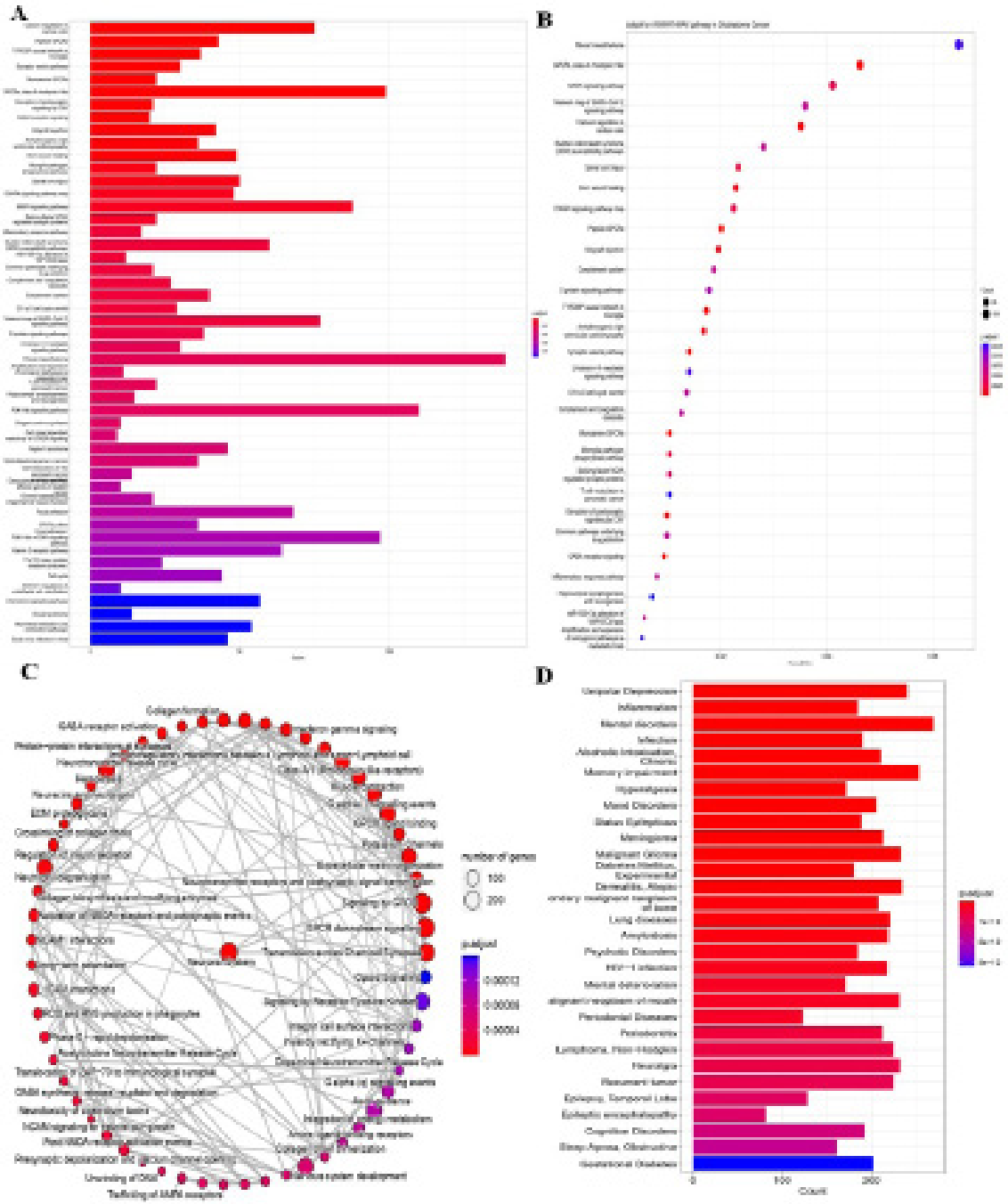

Figure 2. GO Functional Annotation (A, B, and C) Bar Plot and Dot Plot of Molecular Pathway in GBM, and (D) Disease Ontology (DO). The P-value is less than 0.05 and is shown by the color.

mRNA level in the thalamus, putamen, and caudate of humans, as well as, in the hippocampus, hypothalamus, thalamus, and cortex of rat brain [24]. Previous studies align with our results and reported that GPR68 and GPR157 are highly expressed in tumorigenesis; therefore, targeting them is a potential therapeutic strategy in cancer treatment [25-27]. An RNA-seq analysis showed that the GPR68 and GPRC5A overexpressed ten-fold higher in pancreatic cancer cells and Pancreatic Cancer-Associated Fibroblasts (CAFs)[25]. Ahmad et al. determined that O-6-Methylguanine-DNA Methyltransferase (MGMT) can cause Temozolomide resistance. Inhibiting GPR68 decreases the MGMT expression and increases the sensitivity of glioblastoma cells to Temozolomide by NF-kB Pathway [28]. Another study showed that GPR68 inhibition by siRNA and CRISPR in pancreatin cell lines (Panc02 and A549) activates the Warburg effect and induces ferroptosis and radiosensitivity [29]. *GPR157* has been identified as a modulator for neuronal differentiation of radial glial progenitor cells (RGPs) located at the primary cilium. Activation of *GPR157* by CSF-derived signals promotes the differentiation of RGPs into mature neurons. Understanding the role of *GPR157* in neural development may provide insights into the mechanisms underlying neurogenesis and potentially lead to therapeutic strategies for neurodevelopmental disorders

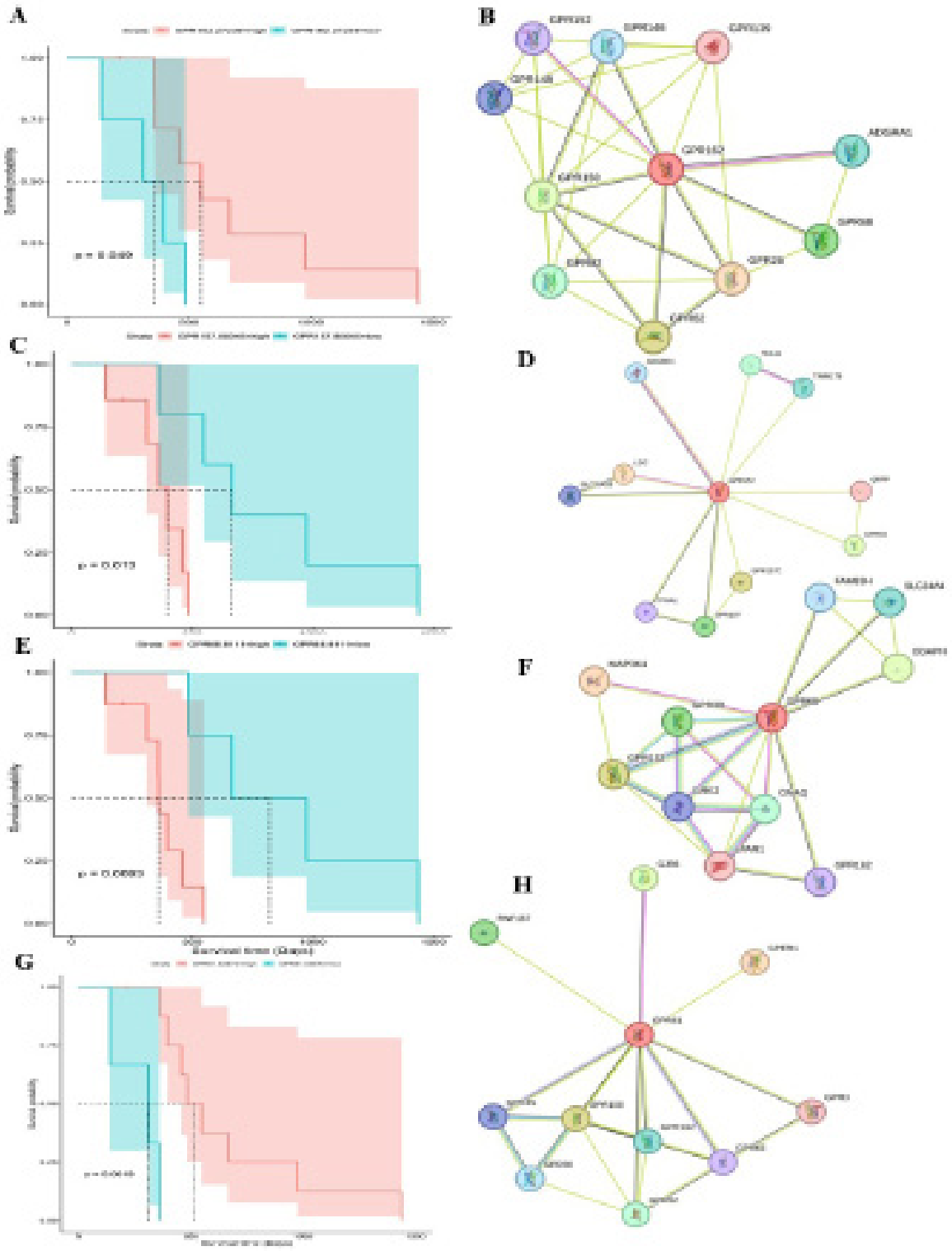

Figure 3. (A and B) Kaplan–Meier and Protein–Protein Interaction (PPI) Network of *GPR162*, (C and D) Kaplan–Meier and Protein–protein interaction (PPI) network of *GPR157*, and (E and F) Kaplan–Meier and Protein–protein interaction (PPI) network of *GPR162*, (G and H) Kaplan–Meier and Protein–protein interaction (PPI) network of GPR61 p-value< 0.05.

[30]. The ROC curve analysis revealed that A2BP1 alone and combined with other genes, including *GPR157*, are diagnostic biomarkers in glioblastoma. Our results showed that the downregulation of A2BP1 is associated with developing glioblastoma. In agreement with our results, Dai et al. indicated that low expression level of A2BP1 in advanced glioma is related to poor OS in patients. They reported that the loss of A2BP1 causes tumor growth of GBM via neutralization terminal differentiation [31]. An investigation showed that using anti-A2BP1 and detecting it by Western blotting and IHC. Their results revealed A2BP1 is highly expressed in the cerebellum and

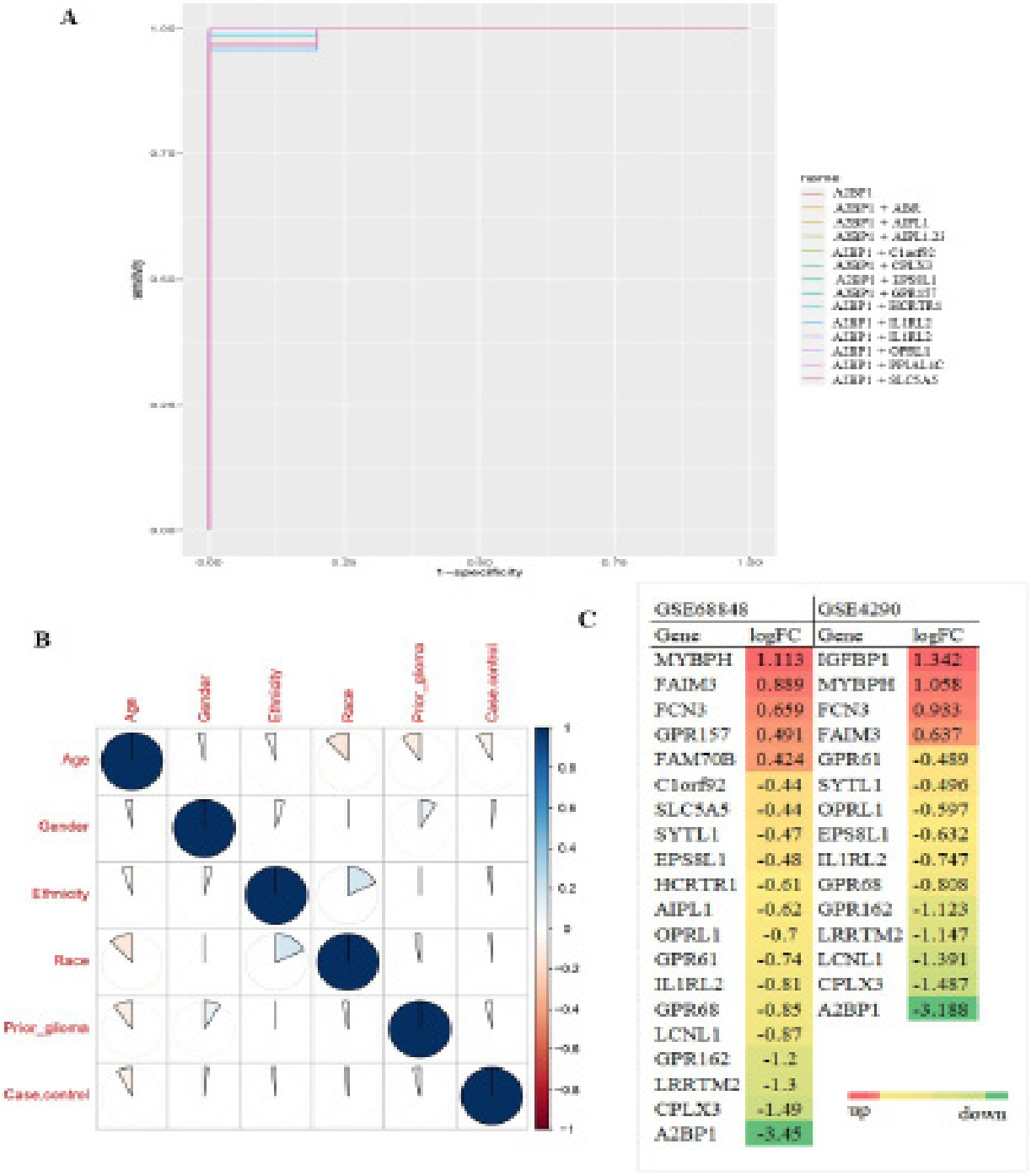

Figure 4. (A) The combineROC curve of A2BP1, and the combination of A2BP1 with with other genes, including *ABR, AIPL1, C1orf92, CPLX3, EPS8L1, GPR157, HCRTR1, IL1RL2, KCNAB1, OPRL1, PPIAL4C*, and *SLC5A5,* (B) The correlation matrix shows a significant co-relationship between clinical/demographic features in GBM, blue and red circles present positive and negative correlations, respectively. The size of the circle and color intensity are the sign of correlation coefficients. (C) Expression changes of candidate genes using GEO datasets with adjusted p-value<0.05.

neocortex of rats [32]. The dysregulation was reported in different neurological disorders, such as autism spectrum disorders and schizophrenia [33, 34].

The machine learning analysis showed twenty genes with higher correlation coefficients, including *LRRTM2* and *OPRL1*. *LRRTM2*, also known as Leucine-Rich Repeat Transmembrane Neuronal 2, is a gene that codes for a protein involved in synaptic adhesion and neuronal development. It is primarily expressed in the brain and plays a role in the formation and function of synapses, which are the connections between neurons. Studies have shown that dysregulation of Lrrtm2 protein can affect synaptic connectivity and neurotransmission, leading to abnormal brain development and neurological phenotypes [35, 36]. De et al. [37] reported that *LRRTM2* is located on the hippocampal neurons and interacts with Neurexin1 to regulate excitatory synapses. Moreover, they indicated that the knockdown of *LRRTM2* decreases presynaptic

differentiation. The *OPRM1* gene, also known as the mu-opioid receptor gene, encodes the mu-opioid receptor protein. *OPRM1* receptor plays a vital role in response to opioids, including pain relief and euphoria. A Genome-wide association study (GWAS) reported different susceptibility novel markers for glioma, including *OPRL1, GALNT6, HAR1A, PHLBD1, JAK1,* etc [38]. Xu et al. revealed that methylation of two genes, *OPRL1* and *OPRM1*, is a diagnostic biomarker in Alzheimer's and elevates disease risk [39]. Previous evidence showed *OPRL1* as a marker in various types of cancers, including cervical cancer [40], oral cancer [41], and colorectal cancer [42].

In conclusion, our data suggest that genes belonging to G Protein-Coupled Receptors play a critical role in various aspects of glioblastoma progression and pathogenesis. Four members of GPCRs, including *GPR162, GPR68, GPR61,* and *GPR157*, can be considered prognostic biomarkers. Furthermore, the results showed that the combination of two genes, *A2BP1* and *GPR157*, is a diagnostic marker for glioblastoma. Furthermore, investigating genetic variation in the G protein-coupled receptor-related gene and its expression across various cancer types may offer valuable insights for cancer therapy. Therefore, further investigations and clinical evaluation of these biomarkers are necessary to develop more accurate diagnostic and prognostic tools, refine treatment strategies, and improve patient care and outcomes.

Author Contribution Statement

Ghazaleh Khalili-Tanha, Nima Khalili-Tanha, Masoumeh Farahani and Elham Nazari have gathered study data and written the manuscript. Elham Nazari has contributed to the study design and approved the final version of the manuscript. All the authors read and approved the final version of the manuscript.

Acknowledgements

Funding information

This study was supported by a grant from the Shahid Beheshti University of Medical Science to Elham Nazari.

Ethics approval

This research involving bioinformatic analysis of online data has received approval from the appropriate ethics committee to ensure compliance with ethical standards and regulations.

Data Availability

The data was downloaded from the TCGA portal (https:// tcga- data. nci. nih. gov/). The datasets generated and/or analyzed during this study are available from the corresponding author upon reasonable request.

Conflict of interest

The authors have no conflict of interest to disclose.

References

1. Grochans S, Cybulska AM, Simińska D, Korbecki J, Kojder K, Chlubek D, et al. Epidemiology of glioblastoma multiforme–literature review. Cancers. 2022;14(10):2412. https://doi.org/10.3390/cancers14102412.
2. Schaff LR, Mellinghoff IK. Glioblastoma and other primary brain malignancies in adults: A review. Jama. 2023;329(7):574-87. https://doi.org/10.1001/jama.2023.0023.
3. Fabbro-Peray P, Zouaoui S, Darlix A, Fabbro M, Pallud J, Rigau V, et al. Association of patterns of care, prognostic factors, and use of radiotherapy–temozolomide therapy with survival in patients with newly diagnosed glioblastoma: A french national population-based study. J Neurooncol. 2019;142(1):91-101. https://doi.org/10.1007/s11060-018-03065-z.
4. Ostrom QT, Patil N, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS. Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2013–2017. Neuro Oncol. 2020;22(Supplement_1):iv1-iv96. https://doi.org/10.1093/neuonc/noaa200.
5. Alcedo-Guardia R, Labat E, Blas-Boria D, Vivas-Mejia PE. Diagnosis and new treatment modalities for glioblastoma: Do they improve patient survival? Curr Mol Med. 2016;16(5):447-64.
6. McKinnon C, Nandhabalan M, Murray SA, Plaha P. Glioblastoma: Clinical presentation, diagnosis, and management. BMJ. 2021;374:n1560. https://doi.org/10.1136/bmj.n1560.
7. Feyissa AM, Worrell GA, Tatum WO, Chaichana KL, Jentoft ME, Cazares HG, et al. Potential influence of idh1 mutation and mgmt gene promoter methylation on glioma-related preoperative seizures and postoperative seizure control. Seizure. 2019;69:283-9. https://doi.org/10.1016/j.seizure.2019.05.018.
8. Bodaghi A, Fattahi N, Ramazani A. Biomarkers: Promising and valuable tools towards diagnosis, prognosis and treatment of covid-19 and other diseases. Heliyon. 2023;9(2):e13323. https://doi.org/10.1016/j.heliyon.2023.e13323.
9. Sarhadi VK, Armengol G. Molecular biomarkers in cancer. Biomolecules. 2022;12(8):1021. https://doi.org/10.3390/biom12081021.
10. Huangfu S, Pan J. Novel biomarkers for predicting response to cancer immunotherapy. Front Immunol. 2023;14:1179913. https://doi.org/10.3389/fimmu.2023.1179913.
11. Mandair D, Reis-Filho JS, Ashworth A. Biological insights and novel biomarker discovery through deep learning approaches in breast cancer histopathology. NPJ Breast Cancer. 2023;9(1):21. https://doi.org/10.1038/s41523-023-00518-1.
12. Nazari E, Biviji R, Roshandel D, Pour R, Shahriari MH, Mehrabian A, et al. Decision fusion in healthcare and medicine: A narrative review. Mhealth. 2022;8:8. https://doi.org/10.21037/mhealth-21-15.
13. Nazari E, Chang H-CH, Deldar K, Pour R, Avan A, Tara M, et al. A comprehensive overview of decision fusion technique in healthcare: A systematic scoping review. Iran Red Crescent Med J 2020;22(10). https://doi.org/10.32592/ircmj.2020.22.10.30.
14. Atas Guvenilir H, Doğan T. How to approach machine learning-based prediction of drug/compound–target interactions. J Cheminform. 2023;15(1):16. https://doi.org/10.1186/s13321-023-00689-w.
15. Nazari E, Pourali G, Khazaei M, Asadnia A, Dashtiahangar M, Mohit R, et al. Identification of potential biomarkers

in stomach adenocarcinoma using machine learning approaches. Curr Bioinform. 2023;18(4):320-33. https://doi.org/10.2174/1574893618666230227103427.
16. Khojasteh-Leylakoohi F, Mohit R, Khalili-Tanha N, Asadnia A, Naderi H, Pourali G, et al. Down regulation of cathepsin w is associated with poor prognosis in pancreatic cancer. Sci Rep. 2023;13(1):16678. https://doi.org/10.1038/s41598-023-42928-y.
17. Asadnia A, Nazari E, Goshayeshi L, Zafari N, Moetamani-Ahmadi M, Goshayeshi L, et al. The prognostic value of asphd1 and zbtb12 in colorectal cancer: A machine learning-based integrated bioinformatics approach. Cancers. 2023;15(17):4300. https://doi.org/10.3390/cancers15174300.
18. Khalili-Tanha G, Mohit R, Asadnia A, Khazaei M, Dashtiahangar M, Maftooh M, et al. Identification of zmynd19 as a novel biomarker of colorectal cancer: Rna-sequencing and machine learning analysis. J Cell Commun Signal. 2023:1-17. https://doi.org/10.1007/s12079-023-00779-2.
19. Stephan G, Ravn-Boess N, Placantonakis DG. Adhesion g protein-coupled receptors in glioblastoma. Neurooncol Adv. 2021;3(1):vdab046. https://doi.org/10.1093/noajnl/vdab046.
20. Ko E-A, Zhou T. Gpcr genes as a predictor of glioma severity and clinical outcome. J Int Med Res. 2022;50(7):03000605221113911. https://doi.org/10.1177/03000605221113911.
21. Feve M, Saliou J-M, Zeniou M, Lennon S, Carapito C, Dong J, et al. Comparative expression study of the endo–g protein coupled receptor (gpcr) repertoire in human glioblastoma cancer stem-like cells, u87-mg cells and non malignant cells of neural origin unveils new potential therapeutic targets. PLoS One. 2014;9(3):e91519. https://doi.org/10.1371/journal.pone.0091519.
22. Tripathi PH, Akhtar J, Arora J, Saran RK, Mishra N, Polisetty RV, et al. Quantitative proteomic analysis of gnrh agonist treated gbm cell line ln229 revealed regulatory proteins inhibiting cancer cell proliferation. BMC cancer. 2022;22(1):133. https://doi.org/10.1186/s12885-022-09218-8.
23. Long Y, Guo J, Chen J, Sun J, Wang H, Peng X, et al. Gpr162 activates sting dependent DNA damage pathway as a novel tumor suppressor and radiation sensitizer. Signal Transduct Target Ther. 2023;8(1):48. https://doi.org/10.1038/s41392-022-01224-3.
24. Lee DK, George SR, Cheng R, Nguyen T, Liu Y, Brown M, et al. Identification of four novel human g protein-coupled receptors expressed in the brain. Mol Brain Res. 2001;86(1-2):13-22. https://doi.org/10.1016/s0169-328x(00)00242-4.
25. Insel PA, Sriram K, Wiley SZ, Wilderman A, Katakia T, McCann T, et al. Gpcromics: Gpcr expression in cancer cells and tumors identifies new, potential biomarkers and therapeutic targets. Front pharmacol. 2018;9:431. https://doi.org/10.3389/fphar.2018.00431.
26. Wiley SZ, Sriram K, Salmerón C, Insel PA. Gpr68: An emerging drug target in cancer. Int J Mol Sci. 2019;20(3):559. https://doi.org/10.3390/ijms20030559.
27. Wiley SZ, Sriram K, Liang W, Chang SE, French R, McCann T, et al. Gpr68, a proton-sensing gpcr, mediates interaction of cancer-associated fibroblasts and cancer cells. FASEB J. 2018;32(3):1170. https://doi.org/10.1096/fj.201700834R.
28. Ahmad JD. Inhibition of gpr68 sensitizes gbm to temozolomide treatment via the nf-kb pathway: University of Maryland, Baltimore; 2021.
29. Neitzel LR, Lee HG, Shukla H, Mahmood J, Williams CH, Hong CC. Inhibition of gpr68 signaling activated downstream of the warburg effect induces cell death and enhances radiosensitivity in diverse cancer cell types. Cancer Res. 2023;83(7Supplement):444.
30. Takeo Y, Kurabayashi N, Nguyen MD, Sanada K. The g protein-coupled receptor gpr157 regulates neuronal differentiation of radial glial progenitors through the gq-ip3 pathway. Sci Rep. 2016;6(1):25180. https://doi.org/10.1038/srep25180.
31. Dai J, Bing Z, Zhang Y, Li Q, Niu L, Liang W, et al. Integrated mrnaseq and micrornaseq data analysis for grade iii gliomas. Mol Med Rep. 2017;16(5):7468-78. https://doi.org/10.3892/mmr.2017.7545.
32. Hamada N, Ito H, Iwamoto I, Mizuno M, Morishita R, Inaguma Y, et al. Biochemical and morphological characterization of a2bp1 in neuronal tissue. J Neurosci Res. 2013;91(10):1303-11. https://doi.org/10.1002/jnr.23266.
33. Davis L, Maltman N, Mosconi M, Macmillan C, Schmitt L, Moore K, et al. Rare inherited a2bp1 deletion in a proband with autism and developmental hemiparesis. Am J Med Genet A. 2012;158(7):1654-61. https://doi.org/10.1002/ajmg.a.35396.
34. Hammock EA, Levitt P. Developmental expression mapping of a gene implicated in multiple neurodevelopmental disorders, a2bp1 (fox1). Dev Neurosci. 2011;33(1):64-74. https://doi.org/10.1159/000323732.
35. Ramsey AM, Tang A-H, LeGates TA, Gou X-Z, Carbone BE, Thompson SM, et al. Subsynaptic positioning of ampars by lrrtm2 controls synaptic strength. Sci Adv. 2021;7(34):eabf3126. https://doi.org/10.1126/sciadv.abf3126.
36. Hayer SN. Neuronal activity-dependent transcriptional regulation of the synaptic cell adhesion molecules lrrtm1 and lrrtm2; 2014.
37. De Wit J, Sylwestrak E, O'Sullivan ML, Otto S, Tiglio K, Savas JN, et al. Lrrtm2 interacts with neurexin1 and regulates excitatory synapse formation. Neuron. 2009;64(6):799-806. https://doi.org/10.1016/j.neuron.2009.12.019.
38. Atkins I, Kinnersley B, Ostrom QT, Labreche K, Il'yasova D, Armstrong GN, et al. Transcriptome-wide association study identifies new candidate susceptibility genes for glioma. Cancer Res. 2019;79(8):2065-71. https://doi.org/10.1158/0008-5472.CAN-18-2888.
39. Xu C, Liu G, Ji H, Chen W, Dai D, Chen Z, et al. Elevated methylation of oprm1 and oprl1 genes in alzheimer's disease. Mol Med Rep. 2018;18(5):4297-302. https://doi.org/10.3892/mmr.2018.9424.
40. Le TM, Nguyen HDT, Lee E, Lee D, Choi YS, Cho J, et al. Transcriptomic immune profiles can represent the tumor immune microenvironment related to the tumor budding histology in uterine cervical cancer. Genes. 2022;13(8):1405. https://doi.org/10.3390/genes13081405.
41. Ishikawa R, Imai A, Mima M, Yamada S, Takeuchi K, Mochizuki D, et al. Novel prognostic value and potential utility of opioid receptor gene methylation in liquid biopsy for oral cavity cancer. Curr Probl Cancer. 2022;46(2):100834. https://doi.org/10.1016/j.currproblcancer.2021.100834.
42. Aridi T, El-Kurdi A, Khoueiry P, Fakhruddin N, Mahfouz R, Zaatari G. Detection of genetic markers involved in metastatic colorectal carcinoma to lymph nodes. 2022. https://doi.org/10.21203/rs.3.rs-1988504/v1.