RESEARCH ARTICLE

**Editorial Process:** Submission:06/28/2024 Acceptance:11/09/2024

# Bioinformatics Examination of Glioblastoma Identifies a Potential Panel of Therapeutic Biomarkers

**Fatemeh Bandarian[1], Farideh Razi[2], Zahra Razzaghi[3], Mohammad Rostami Nejad[4], Babak Arjmand[5,6], Alireza Ahmadzadeh[7,8]***

## Abstract

**Objective:** Glioblastoma, previously recognized as glioblastoma multiform (GBM), is the most aggressive and most common type of cancer that originates in the brain and has a very poor prognosis for survival. Glioblastoma, as one of the lethal cancers of the brain, is important to be studied in terms of molecular exploration. **Methods:** Bioinformatics approaches could be a promising complementary study for identifying more robust biomarkers. This study evaluates the gene expression profile of normal brain endothelial cells versus glioblastoma tumor cells with positive CD3 in more depth by applying R Studio and Cytoscape and its plug-ins. **Results:** A network of differentially expressed genes (DEGs) introduced promising candidates comprised of *TP53*, *EGFR, FN1, JUN*, and *CDC42* and their related biological processes. Comprised of differentially expressed genes, this panel's dysregulation could significantly affect the stability of the protein-protein interaction (PPI) network. Moreover, previous studies have validated these genes' relevance to this cancer type. **Conclusion:** In conclusion, the molecular profile of glioblastoma aids in drug targeting following thorough validation assessments. Five key genes and their related biological processes are possible drug targets to control glioblastoma.

**Keywords:** Glioblastoma (GBM)- Endothelial cells- Gene expression- Biomarker panel

*Asian Pac J Cancer Prev,* **25 (11)**, 4035-4041

## Introduction

Glioblastoma is the most common heterogenous type of brain cancer, knowing as the malignant tumor of glioma [1]. The likelihood of developing this tumor varies with sex, being more prevalent in males. In addition, the risk of acquiring this tumor increases with age [2]. Taking immediate action to address the issue is crucial due to its lethal nature and the lack of effective treatment options. The treatment choices are surgical approaches, radiation therapy, and alkylating chemotherapy. This deadly cancer leads to fatalities within a year of diagnosis [3]. Molecular analysis could be helpful in terms of identifying prominent biomarkers of the underlying disease mechanism. Moreover, various molecular approaches, including genetic, genomic, and epigenetic factors, suggest the presence of heterogeneity in Glioblastoma. This necessitates the development of targeted therapies tailored to the specific subtypes of individuals rather than a one-size-fits-all approach [4]. However, the complexity of Glioblastoma made it challenging for drug targeting design. The grim prognosis associated with this cancer makes it a prominent research focus, particularly in the realm of biomarker discovery for early detection, prognosis assessment, follow-up care, targeted therapies, and treatment strategies, as previously mentioned [5, 6]. Currently, many studies are focusing on targeting biomarkers of Glioblastoma with different drug-targeting attempts. On of these is RNA-based treatment, which targets genes and ncRNAs (non-coding RNAs). This regulator has also been reported to treat other cancers [7]. Our study utilized a prior investigation, in which one of their key findings was the identification of integrin α3β1 as a potential target involved in the formation of

*[1]Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran. [2]Diabetes Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran. [3]Laser application in medical sciences research center, Shahid Beheshti University of medical sciences, Tehran, Iran. [4]Celiac Disease and Gluten Related Disorders Research Center, Research Institute for Gastroenterology and Liver Disease, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [5]Cell Therapy and Regenerative Medicine Research Center, Endocrinology and Metabolism Molecular-Cellular Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran. [6]Iranian Cancer Control Center (MACSA), Tehran, Iran. [7]Proteomics Research Center, System Biology Institute, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [8]Department of lab sciences, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. *For Correspondence: a.ahmadzadeh@sbmu.ac.ir*

vessels in glioblastoma [1]. Taken together, by applying post-data analysis in terms of protein-protein interaction network analysis, further knowledge can be derived from the nature of Glioblastoma.

This scale-free network shows how many genes are significant in the network stability and foundation strength. The limited number of genes with these features highlights their possible roles as candidate biomarkers in terms of network centrals. The reliability of the data obtained by the network could be subjected to validation tests of backbone network investigations by setting some scores to the central genes of the network [8]. These scores show the importance of these genes in the network structure and functions. Deletion or dysregulation of these genes can greatly impact network instability and, ultimately, disease development. Therefore, in this study we considered the glioblastoma PPI network for identification of potential robust biomarkers for the future clinical purposes.

## Materials and Methods

Data collection: Expression data from a study of glioblastoma comparison with normal adjacent cells was conducted in 2022. A total of 9 samples, which were two groups of three normal brain endothelial cells (the NECs samples) and six glioblastoma tumor cells with positive CD3 (the TECs individuals) were the subjects of the study. This research subsequently led to a research paper authored by Eunnyung Bae and collaborators [1]. This data was derived from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Data are recorded as GSE137900 in GEO database.

*Pre-evaluation analysis*

GEO2R is an online tool available in GEO that compares comprehensively. Following this, the R programming software comprehensively analyzed the entire dataset according to the specified statistical criteria. Codes were obtained from GEO2R and then imported into R Studio for further examination. The statistical analyses including t-tests, Log Fold Change calculations, ANOVA, and using the Limma R package for the R analysis were applied. Different methods, including Boxplot, Uniform Manifold Approximation and Projection (UMAP), Venn diagram, and Mean difference, presented data visualization for this study.

*PPI network analysis*

The number of identified DEGs by R were determined and then were used for Cytoscape analysis. In Cytoscape 3.10.2 (https://cytoscape.org/), the protein-protein interaction network of Glioblastoma of these genes was constructed by the application of STRING database (https://string-db.org/). A network of quired genes with the confidence score cutoff 0.5 was then constructed. The network underwent some processes before the topological analysis with "Network Analyzer".

*Here, we focused on two centrality parameters*

degree and betweenness centrality. Degree centrality represents the number of genes connected to the gene of interest, while betweenness centrality indicates the gene's position within the network, determining whether it acts as a bridge facilitating connections within the network. Essentially, a gene's value in betweenness centrality is determined by its significance as a connector between pairs of nodes in the shortest paths [9]. Nodes or genes with high degree values are hubs, and those with high betweenness centrality values are considered bottlenecks. Nodes possessing both characteristics are referred to as hub-bottlenecks. However, if a node lacks one or both of these traits, it may be classified as either a hub-non-bottleneck or a non-hub-bottleneck. All these node classifications are examined in our study. To test and validate hub-bottlenecks, a backbone network was also constructed.

*Gene ontology analysis*

Central genes were then analyzed in terms of gene ontology to gain a better understanding of their role in the pathogenicity of Glioblastoma. The examination was conducted by ClueGO+CluePedia Plug-ins.

*Statistical analysis*

The statistical analysis are as follow: The statistical analyses include Kappa Statistics and the adjustment of p-values using Bonferroni correction to mitigate false discovery rates [10].

## Results

A box plot analysis was performed in Figure 1 to determine the comparability of normal and Glioblastoma, because the gene expression profiles of samples are median center. In this boxplot analysis, green represents normal samples, while purple represents patients with Glioblastoma.

Another visualization method for determining biological variation is UMAP, as shown in Figure 2. Normal groups are assigned a green color, whereas Glioblastoma is assigned purple. This analysis displays the distribution of Normal and Glioblastoma samples, visually comparing the similarities between each group's samples.

Detection of DEGs in a visualization pattern was handled with a Venn diagram (see Figure 3). A total of 3484 DEGs are present across our groups. These genes are appropriate candidates for further analysis regarding the PPI interaction network and gene ontology identifications.

A mean difference plot (see Figure 4) could illustrate the visualization of the distribution of regulation differences in genes. Differentially expressed genes are visualized as colored spots in this analysis. Fold change equal and above 1.5 are the colored spots.

After gene query in Cytoscape, several 2093 genes and 9765 connections were resulted. The network is obtained by a score confidence cut-off of 0.5 (data is not shown). Additionally, the network was analyzed for topological features defined as hub and bottlenecks. For this purpose, we processed the constructed network did some haircuts, ultimately providing a subnetwork of 1,760 nodes. This network is now ready for centrality analysis by Network Analyzer as the scatter plot indicates the distribution of

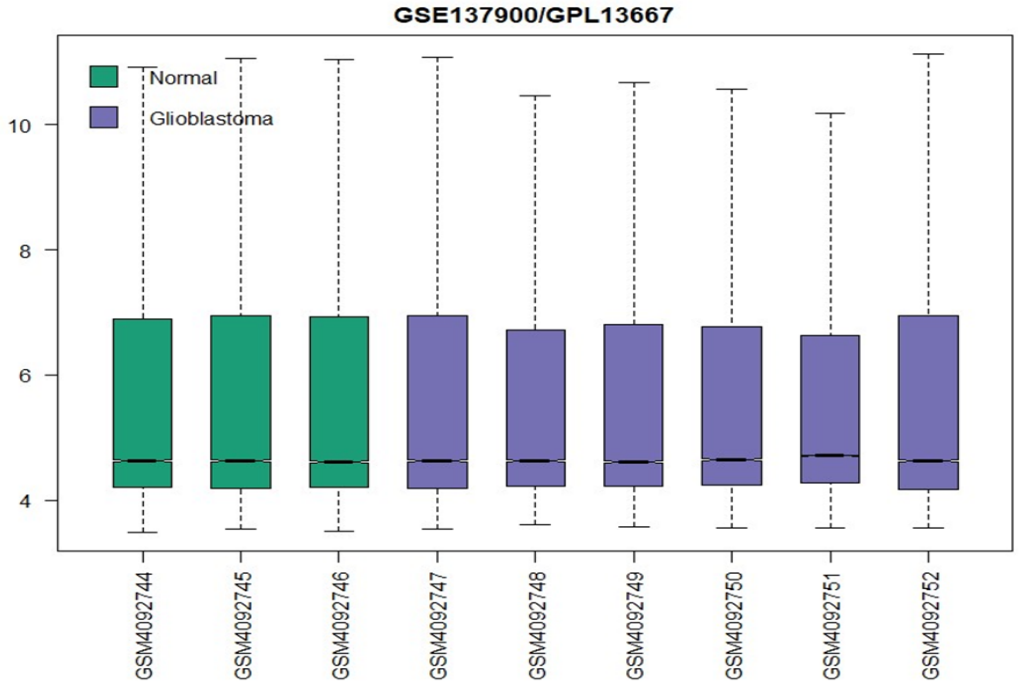


Figure 1. Boxplot View of Gene Expression Distribution Across Groups of Samples. Each sample is pointed out in the horizontal axis of box plot while distribution of normalized amounts of gene expression value for each gene is shown in the vertical axis.

centrality values (see Figure 5). This plot indicates a scale-free network with a limited number of nodes in the high centrality value. It is reported that top 10% of the nodes a PPI network based on degree value can be selected as hub nodes [11]. In the present study, large number of hubs were identified. Therefore, 1% of top nodes based on degree value were determined as the potent hubs. Table 1 lists the 1% highest-valued degree and betweenness centrality.

Table 1. The List of Hubs and Bottlenecks of the Glioblastoma Network

| Row | Display name | Degree | Betweenness Centrality |
|---|---|---|---|
| 1 | TP53 | 220 | 0.16 |
| 2 | EGFR | 209 | 0.12 |
| 3 | FN1 | 147 | 0.04 |
| 4 | JUN | 129 | 0.05 |
| 5 | CDC42 | 123 | 0.08 |
| 6 | CD44 | 118 | 0.02 |
| 7 | TGFB1 | 109 | 0.02 |
| 8 | HIF1A | 104 | 0.03 |
| 9 | COL1A1 | 94 | 0.01 |
| 10 | KDR | 93 | 0.02 |
| 11 | STAT1 | 92 | 0.03 |
| 12 | FGF2 | 92 | 0.02 |
| 13 | MMP2 | 89 | 0.01 |
| 14 | PPARG | 85 | 0.02 |
| 15 | NFKB1 | 84 | 0.02 |
| 16 | PECAM1 | 83 | 0.0 |
| 17 | CCND1 | 82 | 0.01 |
| 18 | APP | 80 | 0.03 |

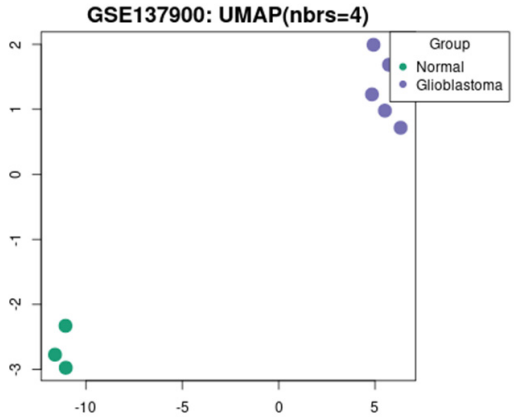


Figure 2. UMAP Plot of Normal-Glioblastoma Gene Expression Profiles Analysis. The horizontal and vertical axes (UMAP-1 and UMAP-2) determine separation of samples via analysis.

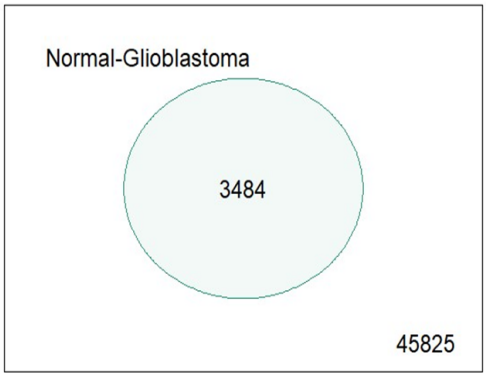


Figure 3. Venn Diagram View of DEGs Across Compared Groups of Samples. The Glioblastoma samples are differentiated from Normal individuals by 3484 significant DEGs.

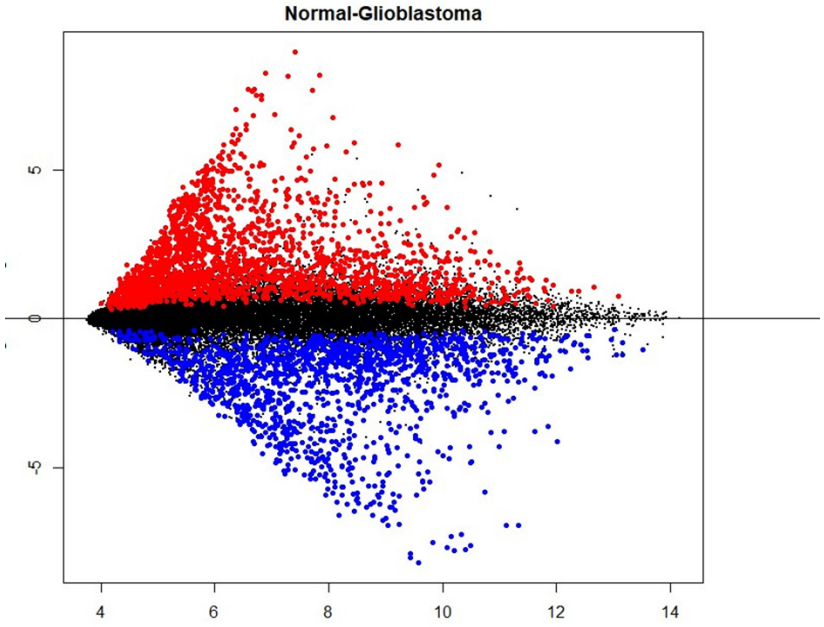


Figure 4. Mean Difference Depicts the Presence of Significant Down-Regulated and Up-Regulated Genes in Blue and Red Color.

Following the application of the top 1% highest-ranked hubs, we established a degree cutoff of 100 for our central genes, identifying hub bottlenecks and hub non-bottlenecks for further analysis. However, only hub bottlenecks are the focus of further analysis. To validate the identified Hub bottlenecks, a backbone network was created consisting of these five genes along with 50 neighboring nodes. The topological analysis of this backbone network indicates all the hub-bottleneck genes of the main network are present in the 14 top-ranked nodes of the backbone network. This finding indicates that the central nodes are valid and could have potential properties about centrality. The next step is identifying corresponding biological processes to the hub bottlenecks, providing more insight into our DEGs' behavior and underlying mechanisms (see Figure 6).

Using GO term fusion, the number of genes per group was set to 2, and the percentage of genes present in the groups was set to 3. This analysis also depicts the parents of nodes (biological processes) as smaller circles.

Only four of the five queried genes met the statistical criteria and cutoffs specified in the query. The biological analysis identified three major processes related to the Hub bottlenecks: positive regulation of miRNA maturation, regulation of fibroblast proliferation, and regulation of transcription from RNA polymerase II promoter in response to stress. *TP53* contributes to all derived biological processes terms, and this finding shows its

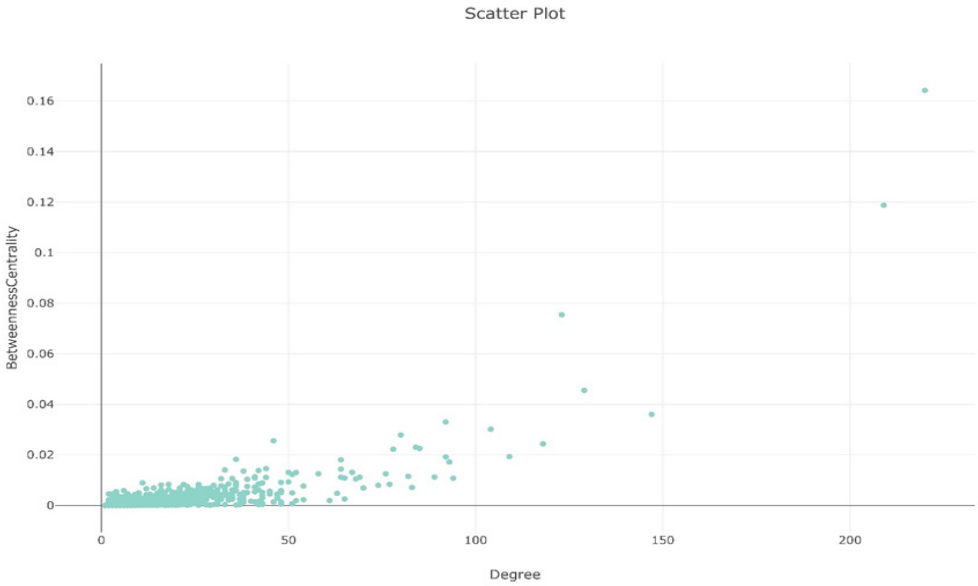


Figure 5. Distribution of Degree Versus Betweenness Pattern for the Subnetwork of Glioblastoma. The higher degree values are correlated to the larger values of betweenness centrality.

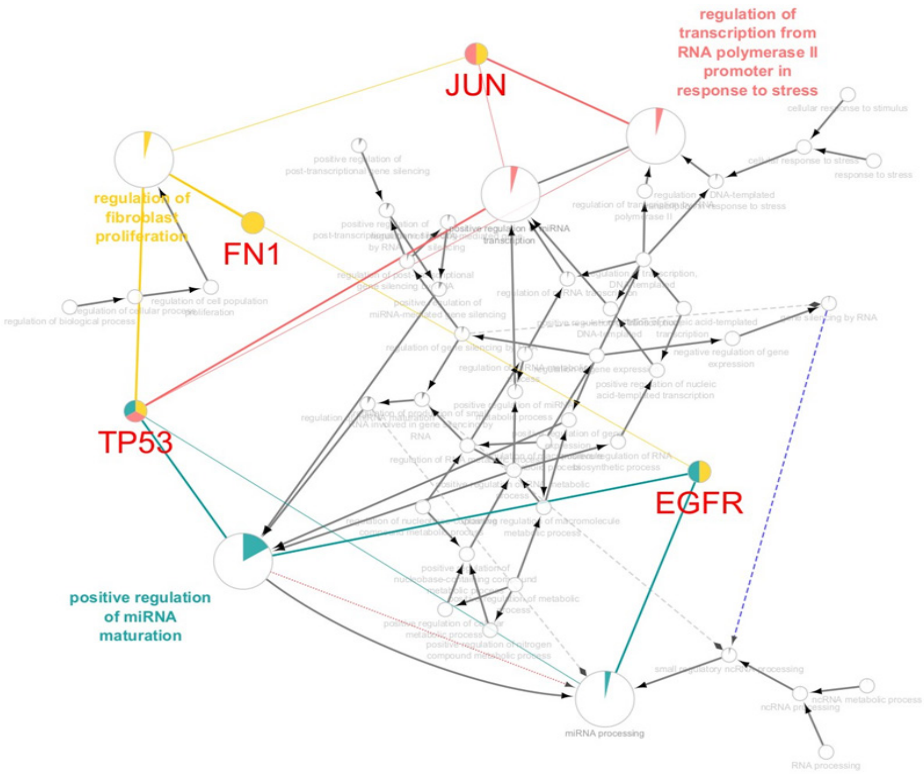


Figure 6. Percentage of Mapped Genes Per Term in a Backbone Network View. The associated genes are labeled in red. The group of biological processes are shown as larger circles.

significance in cancer.

## Discussion

The limited therapeutic options available for Glioblastoma (GBM) underscore the need for intensified molecular studies to advance research in this area [12]. One of the approaches is through computational analysis of expression data gathered by microarray [1]. This way, biomarkers and their corresponding biological processes identified by array study could be further analyzed and introduced as better candidates for future drug targeting approaches. The application for this purpose is PPI network analysis. In this complementary study, the physical interactions between the DEGs were assessed to better understand the role of potential biomarkers.

As depicted in Figures 1-4, the expression difference between the two samples was compared and visualized by applying boxplot analysis, UMAP plot, Venn diagram, and Meandiff plot. A considerable number of genes demonstrate both up-regulation and down-regulation in this sample.The total number of 3438 genes, almost all of which also indicate significant fold change, were then chosen for PPI network assessment. In the network construction, key factors such as statistical criteria were applied. First, for better and more robust analysis, a strong confidence score cut-off was set, and then the network was processed more till a subnetwork was obtained. This subnetwork is required for topological exploration considering hub and bottleneck parameters, which were

In addition, a scattered network showed a scale-free network since most of the network's elements have a low degree or betweenness [13]. A limited number of nodes have values of these parameters. These genes serve as hubs, bottlenecks, or both. A cut-off was then set to explore the most ranked genes in this analysis for reviewing the literature. Previous findings related to the linkage of these central genes to glioblastoma provide additional notable information for collecting the best biomarker panels. From this perspective, starting with *TP53*, a pivotal gene implicated in various types of cancer, can potentially expedite clinical interventions. This study identified tumor suppressor protein p53 (*TP53*) as an up-regulated gene in glioblastoma. This gene is involved in different malignancies, and GBM develops inflammation.

The mutation of this gene has been found to significantly correlate with the outcome of prognostic procedures and treatment approaches [14]. Due to its frequent association with glioblastoma in prior studies [15] and our own findings indicating its central PPI topological role in glioblastoma tumorigenesis, this gene emerges as a promising candidate for targeted drug interventions.

The next central hub-bottleneck, epidermal growth factor receptor (EGFR), showed resistance in glioblastoma treatment appears to be facilitated by epigenetic mechanisms, enabling cancer cells to adapt to drug exposure [16]. In the gene expression data, this gene is down-regulated. EGFR is also reported as a target for treatment approaches for different malignancies. Treatments for this gene include tyrosine kinase inhibitions and monoclonal antibodies since EGFR is a tyrosine kinase [17].

Fibronectin 1 (FN1) is also a hub bottleneck in the glioblastoma network that is significantly down-

regulated in this cancer. Previous studies suggested the prognostic role of this gene Previous studies suggested the prognostic role of this gene [18]. JUN is an up-regulated hub-bottleneck in Glioblastoma that limited tumors reported with its dysregulation. This gene is responsible for key functions in the cell including cell proliferation, cell death, and malignant conversion. The up-regulation of this c-Jun correlates with the progress of glioblastoma [19].

Cell division cycle 42 (CDC42) the last hub-bottleneck and, as a Rho-GTPase is the regulator of cell proliferation and the mechanism of apoptosis [20]. This gene is up-regulated in this cancer and is an important key for invasion [21]. Collectively, these hub-bottlenecks have been implicated in the development of Glioblastoma cancer, as evidenced by significant roles in previous studies. Furthermore, expression data analysis corroborates this finding. Moreover, our study reveals that these genes are important at various levels of molecular analysis, but they also form the backbone of the glioblastoma protein-protein interaction network. This positions them as crucial contributors to network stability and the foundation of the network itself. Targeting this panel comprising *TP53, EGFR, FN1, JUN*, and *CDC42* and their related biological processes could yield remarkable effects in treating glioblastoma followed by comprehensive complementary analysis.

In conclusion, it can be concluded that introducing potential hubs and bottlenecks in glioblastoma PPI network can minimize the side effects of traditional invasive therapies. It seems; *TP53*, *EGFR, FN1, JUN*, and *CDC42* and their related biological processes are suitable drug targets to control glioblastoma. It can be suggested that effect of expression modification of the mentioned genes (separately or in combination with each other) on controlling of glioblastoma be investigated. Nevertheless, additional analysis and validation analyses are encouraged in this regard.

## Author Contribution Statement

All authors contributed equally in this study.

## Acknowledgements

None.

*Conflict of interest*

There is no any conflict of interest

## References

1. Bae E, Huang P, Müller-Greven G, Hambardzumyan D, Sloan AE, Nowacki AS, et al. Integrin α3β1 promotes vessel formation of glioblastoma-associated endothelial cells through calcium-mediated macropinocytosis and lysosomal exocytosis. Nat Commun. 2022;13(1):4268. https://doi.org/10.1038/s41467-022-31981-2
2. Le Rhun E, Preusser M, Roth P, Reardon DA, van den Bent M, Wen P, et al. Molecular targeted therapy of glioblastoma. Cancer Treat Rev. 2019;80:101896. https://doi.org/10.1016/j.ctrv.2019.101896.
3. Jenkins MR, Drummond KJ. Car t-cell therapy for glioblastoma. N Engl J Med. 2024;390(14):1329-32. https://doi.org/10.1056/NEJMe2401307
4. Huse JT, Holland E, DeAngelis LM. Glioblastoma: Molecular analysis and clinical implications. Annu Rev Med. 2013;64:59-70. https://doi.org/10.1146/annurev-med-100711-143028.
5. Azodi MZ, Rezaei-Tavirani M, Nejad MR, Arjmand B, Razzaghi M. Uncovering systemic lupus erythematosus candidate proteomic biomarkers: A bioinformatics approach. J Cell Mol Anesth. 2022;7(1):40-8.
6. Abdullaeva S, Chubarev V, Valeeva A, Preferanskaya N, Neganova M, Smolyarchuk E, et al. Analysis of clinical success and molecular mechanisms of action of novel anti-glioblastoma drugs: A review. Curr Med Chem. 2024. https://doi.org/10.2174/0109298673281283240101053940.
7. Singh RR, Mondal I, Janjua T, Popat A, Kulshreshtha R. Engineered smart materials for rna based molecular therapy to treat glioblastoma. Bioact Mater. 2024;33:396-423. https://doi.org/10.1016/j.bioactmat.2023.11.007.
8. Naderi N, Zamanian Azodi M, Daskar Abkenar E, Shahidi Dadras M, Talaei R. Insulin dysregulation plays a critical role in colon inflammation: A bioinformatics approach. Gastroenterol Hepatol Bed Bench. 2018;11(Suppl 1):S85-S91.
9. Hosseinpour Z, Zamanian Azodi M, Jahani Sherafat S, Rezaei Tavirani M. Targeting colon cancer via antimicrobial rt2 peptide: A system biology study. Gastroenterol Hepatol Bed Bench. 2023;16(4):415-20. https://doi.org/10.22037/ghfbb.v16i4.2695.
10. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. Cluego: A cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25(8):1091-3. https://doi.org/10.1093/bioinformatics/btp101.
11. Rezaei-Tavirani M, Rezaei-Taviran S, Mansouri M, Rostami-Nejad M, Rezaei-Tavirani M. Protein-protein interaction network analysis for a biomarker panel related to human esophageal adenocarcinoma. Asian Pac J Cancer Prev. 2017;18(12):3357-63. https://doi.org/10.22034/APJCP.2017.18.12.3357.
12. Medikonda R, Dunn G, Rahman M, Fecci P, Lim M. A review of glioblastoma immunotherapy. J Neurooncol. 2021;151(1):41-53. https://doi.org/10.1007/s11060-020-03448-1.
13. Rezaei-Tavirani M, Tavirani SR. Gene screening of astrocytoma grade iii relative to grade ii via network analysis: A new molecular insight. International Journal of Cancer Management. 2018;11(11).
14. Ham SW, Jeon HY, Jin X, Kim EJ, Kim JK, Shin YJ, et al. Tp53 gain-of-function mutation promotes inflammation in glioblastoma. Cell Death Differ. 2019;26(3):409-25. https://doi.org/10.1038/s41418-018-0126-3.
15. Lee YJ, Seo HW, Baek JH, Lim SH, Hwang SG, Kim EH. Gene expression profiling of glioblastoma cell lines depending on tp53 status after tumor-treating fields (ttfields) treatment. Sci Rep. 2020;10(1):12272. https://doi.org/10.1038/s41598-020-68473-6.
16. Pan PC, Magge RS. Mechanisms of egfr resistance in glioblastoma. Int J Mol Sci. 2020;21(22):8471. https://doi.org/10.3390/ijms21228471.
17. Westphal M, Maire CL, Lamszus K. Egfr as a target for glioblastoma treatment: An unfulfilled promise. CNS drugs. 2017;31:723-35. https://doi.org/10.1007/s40263-017-0456-6
18. Wang J, Yan S, Chen X, Wang A, Han Z, Liu B, et al. Identification of prognostic biomarkers for glioblastoma based on transcriptome and proteome association analysis.

Technol Cancer Res Treat. 2022;21:15330338211035270. https://doi.org/10.1177/15330338211035270.
19. Blau L, Knirsh R, Ben-Dror I, Oren S, Kuphal S, Hau P, et al. Aberrant expression of c-jun in glioblastoma by internal ribosome entry site (ires)-mediated translational activation. Proc Natl Acad Sci U S A. 2012;109(42):E2875-84. https://doi.org/10.1073/pnas.1203659109.
20. Okura H, Golbourn BJ, Shahzad U, Agnihotri S, Sabha N, Krieger JR, et al. A role for activated cdc42 in glioblastoma multiforme invasion. Oncotarget. 2016;7(35):56958-75. https://doi.org/10.18632/oncotarget.10925.
21. Fortin SP, Ennis MJ, Schumacher CA, Zylstra-Diegel CR, Williams BO, Ross JT, et al. Cdc42 and the guanine nucleotide exchange factors ect2 and trio mediate fn14-induced migration and invasion of glioblastoma cells. Mol Cancer Res. 2012;10(7):958-68. https://doi.org/10.1158/1541-7786.MCR-11-0616.