Predicting the Role of Temozolomide Drug in Glioma By integrating Available Genomic Databases and Computational Methods

Aya ALHAJJ[1], Zehra BAYINDIR[1], Sleiman ALHAJJ[3], Lama ALHAJJ[3], Kashfia SAILUNAZ[4], Mehmet KAYA[5] and Reda ALHAJJ[4,6,7]

[1]Biomedical Engineering, Istanbul Medipol University, Istanbul, Turkey
[2]Molecular Medicine and Biotechnology, Istanbul Medipol University, Istanbul, Turkey
[3]International School of Medicine, Istanbul Medipol University, Istanbul, Turkey
[4]Department of Computer Science, University of Calgary, Alberta, Canada
[5]Department of Computer Engineering, Firat University, Elazığ, Turkey
[6]Department of Computer Engineering, Istanbul Medipol University, Istanbul, Turkey
[7]Department of Health Informatics, University of Southern Denmark, Odense, Denmark

Abstract. Cancer is one of the most serious and harmful diseases that threatens humanity. Currently there is no robust treatment which leads to guaranteed cure from cancer. Thus, researchers from various domains are still working hard to identify molecules such as genes and proteins which could be handled and targeted as cancer biomarkers. Various methods have been developed and the research spans wide range of techniques from wet lab testing by biologists to computational methods by computer scientists. The latter research is promising because it greatly reduces the number of molecules as potential biomarkers. This project investigated existing literature data by integrating text mining, as well as gene-gene interactions. Different genes are highlighted in relationship to Glioma and temozolomide.

Keywords: Temozolomide, Glioma, Cancer, Biomarkers, Molecular Databases, Machine Learning, Data Mining

1 Introduction

Glioblastoma is a deadly kind of brain tumor for which the mainstay of care is complete surgical resection followed by chemoradiotherapy, if possible [37]. In such conditions, age plays an essential role since it is linked to shorter survival and a higher risk of treatment-related toxicity. Malignant gliomas are the most common type of primary brain tumors and represents one of the most aggressive and lethal malignancies in adults. Malignant gliomas occur more frequently than other types of primary central nervous system tumors, having a combined incidence of 5-8/100,000 population [38]. Even with aggressive treatment using surgery, radiation, and chemotherapy, median

reported survival is less than 1 year. Temozolomide, a new drug, has shown promise in treating malignant gliomas and other difficult-to-treat tumors [39].

Temozolomide, a p.o. imidazotetrazine second-generation alkylating agent, is the leading compound in a new class of chemotherapeutic agents that enter the cerebrospinal fluid and do not require hepatic metabolism for activation [1, 33]. Temozolomide anti-cancer prodrug of temodar is an imidazotetrazine that has increased the prognosis of highly aggressive glioma [34]. This drug has been shown to cause some side effects including nausea, vomiting, loss of appetite, changes in taste, constipations, tiredness, dizziness, or headaches [40]. Although Temozolomide is used to treat cancer, it may rarely increase your risk of getting other cancers. In addition, Temozolomide decreases bone marrow function, an effect that may lead to a low number of blood cells such as red blood cells, white blood cells, and platelets. This effect can cause anemia, decrease your body's ability to fight an infection, or cause easy bruising/bleeding [2].

The therapeutic index of Temozolomide is dependent on molecular markers like MGMT (methylguanine mthyltransferase) and lack of DNA repair in glioma [36]. In addition to their molecular markers, isocitrate dehyfrogenase type 1 and type 2 (IDH1/2) mutations metabolically reprogram the cell to form neo-metabolite 2- hydroxyglutarate and epigenetic status controlling cytosine- phosphate-guanine (CpG). U98MG which is a completely sequenced [3] and commonly studied grade IV Glioma cell line [4] is wildtype for IDH1/2 is sensitive to Temozolomide [3].

Interest in Temozolomide as an antitumor agent derives from its broad-spectrum antitumor activity in tumor models in mice [5]. Invitro, Temozolomide has demonstrated schedule- dependent antitumor activity against a variety of malignancies including glioma, metastatic melanoma, and other difficult-to-treat cancers [5,6,7]. In preclinical studies, Temozolomide demonstrated distribution to all tissues, including penetration into the CNS; relatively low toxicity compared with its parent compound, mitozolomide and antitumor activity against a broad range of tumor types, including glioma, melanoma, mesothelioma, sarcoma, and lymphoma [5,6,7,8,9,10]. Its demonstrated ability to cross the blood-brain barrier is of special interest with respect to its activity in CNS tumors [35].

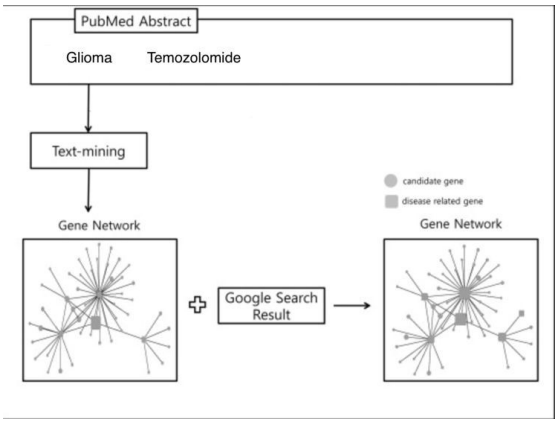

Fig. 1. Workflow of the developed approach.

2 Methods

2.1 Text Mining

The number of articles in PubMed about brain tumor Glioma are gradually increasing, therefore using text mining, the act of looking for hidden patterns in previously published text, and it will be very effective for researchers to restrict their list of target potential genes to research. We will use text mining using PubMiner [19] in this research. PubMiner will help us retrieve data from databases and web services such as PubMed [20], Uniprot [21], and HUGO [22].

The query for 'Glioma' between the years 1992-2022 to PubMed can retrieve over 47073 abstracts, and the query for 'Temozolomide' between the years 2000-2022 to PubMed can retrieve 8524 abstracts.

2.2 GeneMania

GeneMania [23] accounts for a few different types of interactions between genes such as co-expression, physical interaction, genetic interaction, shared protein domains, co-localization, pathway, as well as predicting relationships using orthological functional data from other organisms. Co- expressed genes are genes which had the same expression levels over the same conditions in a published study, where most of the gene expression data came from the Gene Expression Omnibus (GEO) database [24]. Another interaction in GeneMania is physical interaction, which means if two genes code for proteins that have a physical interaction, then the two genes have a connection. Two genes partake in the shared protein domain interaction if their proteins have the same protein domain. Two genes have co-localization interaction if their proteins are found in the same body tissue. Two genes share in the pathway interaction if they participate in the same reaction in a pathway.

The set of genes used for this study are the genes that are mentioned the most in the literature relevant to Glioma. These genes are EGFR, PTEN, TP53, ARAF, ATM, ATR, BARD1, BRAF, BRCA1, BRCA2, CDK12, CHEK1, CRKL, FANCA, FANCB, FANCE, FANCG, FANCL, HGF, IDH1, IDH2, MAP2K4, MAP3K1, MAPK1, NBN, NRAS, PALB2, RAF1, SLX4, SRC, TACC1, and TACC3. A gene-gene interaction was constructed by using these genes as inputs for our system.

2.3 String For DNA repair system

We used STRING database to find the targets for TMZ. The biological effect of TMZ depends on 3 DNA repair systems- MGMT, BER, and MMR. MGMT is a DNA 'suicide' repair enzyme. It repairs damaged guanine nucleotides by transferring the methyl at O6 site of the guanine to its cysteine residues thus avoiding gene mutation, cell death and tumorigenesis cause by alkylating agents [10]. BER is base excision repair that corrects DNA damage from oxidation, deamination, and alkylation, and it is initiated

by a DNA glycosylase that recognizes and removes the damage base, leaving a basic site that is further processed by short patch repair or long patch repair that largely uses different proteins to complete BER [11]. MMR is a system for recognizing and repairing erroneous insertion, deletion, and misincorporation of bases that can arise during DNA replication and recombination as well as repairing some forms of DNA damage [12]. Using the data from STRING we constructed a drug-target protein interaction network of Temozolomide.

2.4 Drug-Drug interaction

We explored the Stanford Network Analysis Project (SNAP) data repository [25] which has a public Biomedical Network Dataset Collection [26] compromising many datasets with information on relationships between entities. We used ChCh-Miner dataset [27, 28, 29] for this project which is a set of raw data stored as a CSV file of 48464 rows and 2 columns where each row represents an interaction between a particular pair of drugs. We specifically highlighted Temozolomide DrugBank code DB00853. Since the dataset is given in DrugBank form DB code, we matched each code with actual drug name from DrugBank database (Table 1).

Table 1. Representation of the interaction between a particular pair of drugs.

	drug_1_code	drug_1_name	drug_2_code	drug_2_name
0	DB00313	Valproic acid	DB00853	Temozolomide
1	DB00363	Clozapine	DB00853	Temozolomide
2	DB00331	Metformin	DB00853	Temozolomide
3	DB00853	Temozolomide	DB06688	Sipuleucel-T
4	DB00853	Temozolomide	DB04817	Metamizole
5	DB00853	Temozolomide	DB08895	Tofacitinib
6	DB01008	Natalizuman	DB00853	Temozolomide
7	DB00337	Pimecrolimus	DB00853	Temozolomide
8	DB00072	Trastuzumab	DB00853	Temozolomide
9	DB00853	Temozolomide	DB01097	Leflunomide
10	DB00853	Temozolomide	DB00864	Tacrolimus
11	DB00853	Temozolomide	DB06643	Denosumab

2.5 Cytoscape

U87 cells and U251 cells are glioblastoma cell lines and many in vitro studies have been made on these cells' lines. Some studies treated these cell lines with increasing dose of TMZ starting from 50 uM (micromole) to 600 uM over a period of 10 months. Results showed a decrease in the percentage of growth rate. The resistant Glioma cells showed similar proliferation and doubling time compared to their respective parental cell lines (U87P and U251P). It was tested to check whether preconditioning of TMZ resistant cells with metformin will be able to reverse the drug resistance U87 and U251

cells, if they were treated with metformin for 2 weeks and then they were exposed to TMZ for 3 days. They measured the cell survival rate, and it showed high survival rate after 3 days of TMZ treatment, and the survival rate dropped significantly for the cells that were pretreated with metformin. This experiment indicated that metformin could partially restore TMZ sensitivity in TMZ-resistant glioblastoma cells. It has been reported that metformin acts synergistically with Temozolomide to inhibit proliferation of Glioma cells [13]. To explore the potential regulatory networks among the reversed genes, we retrieved protein-protein interactions of these genes from the string database and reconstructed the protein-protein interaction network using Cytoscape software [31].

2.6 K-means

K-means clustering was performed using the "kmeans" function in 'STRING' [30]. We used 'STRING' dataset for this project to cluster the genes that Temozolomide interacts with. We used k-means clustering to cluster the genes to assist us in selecting new pathways for Temozolomide to check if it gets resistance to one gene. We clustered the genes into 3 groups by which cluster 1 has 16 genes, cluster 2 has 11 genes and the cluster 3 has 2 genes.

3 Results

3.1 Text Mining

According to the search query result between the years of 1992-2022, we found that the year that mentioned Glioma the most in the abstract is 2021with a count of 4662 and the year that mentioned Glioma the least was 2022(Fig. 2. (A)).

The search query results for Temozolomide, we found that the most year that mentioned Temozolomide was year 2021 and the least year was 1992 (Fig. 2. (B)).

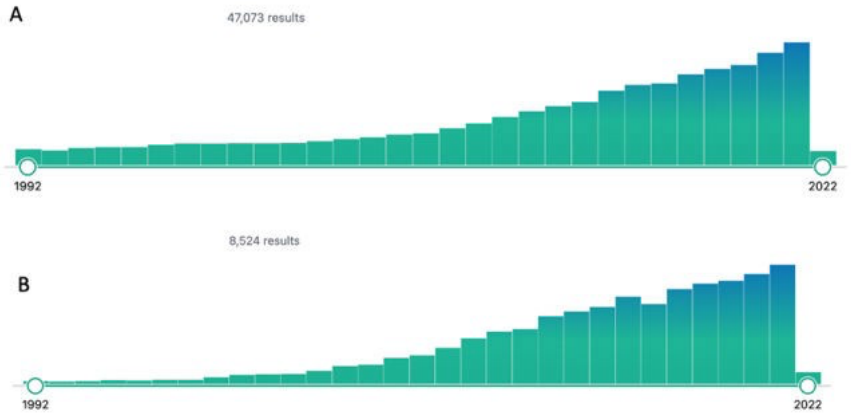

Fig. 2. Text mining result of (A) Glioma and (B) Temozolomide.

3.2 GeneMania

Results of GeneMania based on published studies showed that 39.90% of the genes in different publications had physical interactions, 18.71% of the genes showed co-expression, 8.91% have a shared pathway, 8.43% share protein domains, and 2.77% are co-localized (Fig. 3).

According to different publications, most genes in Glioma have gene-gene interactions (Fig. 4). Not all genes that interact with one another have shared protein domains. Fig. 5 shows the genes that have shared protein domains. For example, BRAF and EGFR have gene-gene interaction and share protein domains. 23 genes are co-localized (Fig. 8). All the present genes have physical interaction (Fig. 6) except TDRD7 and TACC3. All the genes have co- expression (Fig. 7) except ID3B.

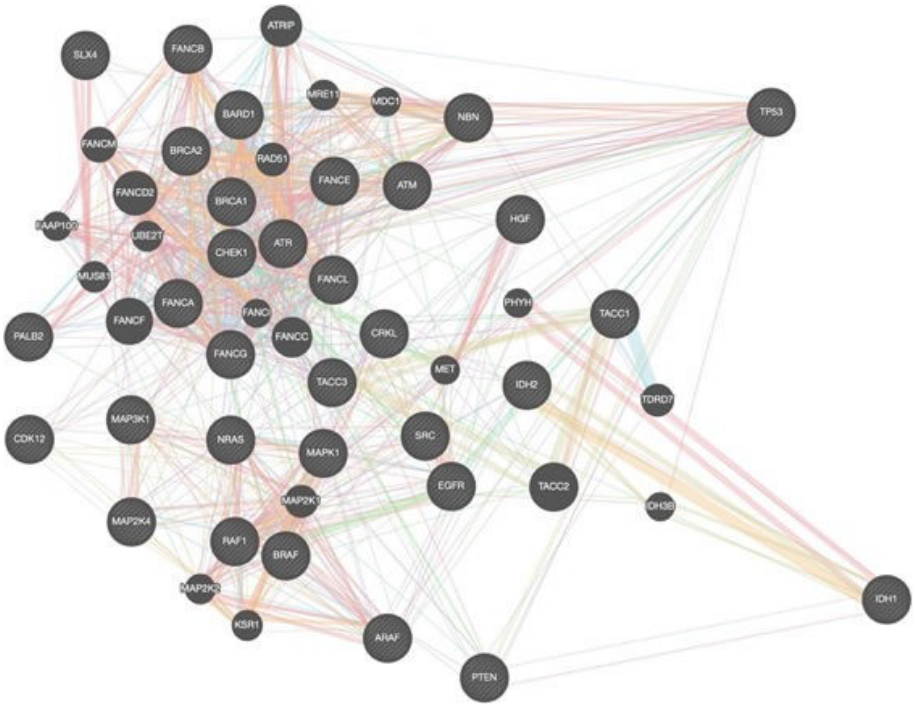

Fig. 3. GeneMania Results based on published studies

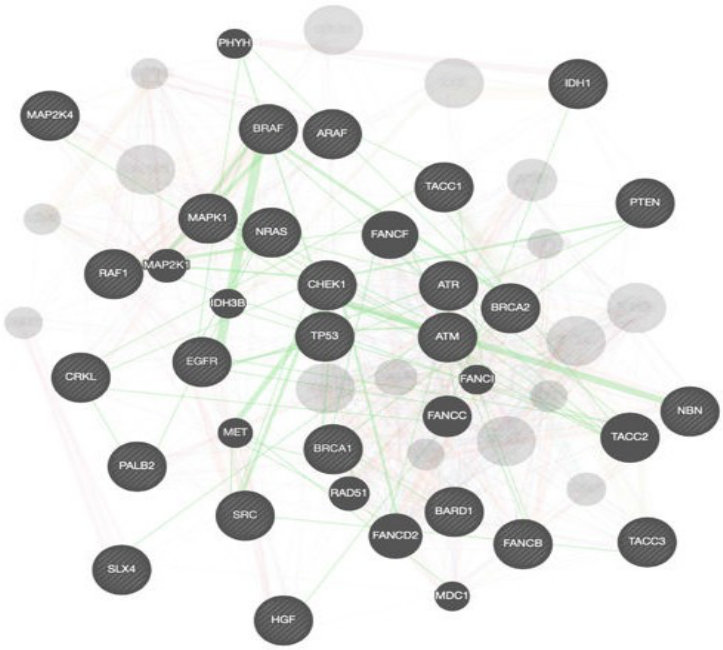

Fig. 4. Gene-Gene Interactions.

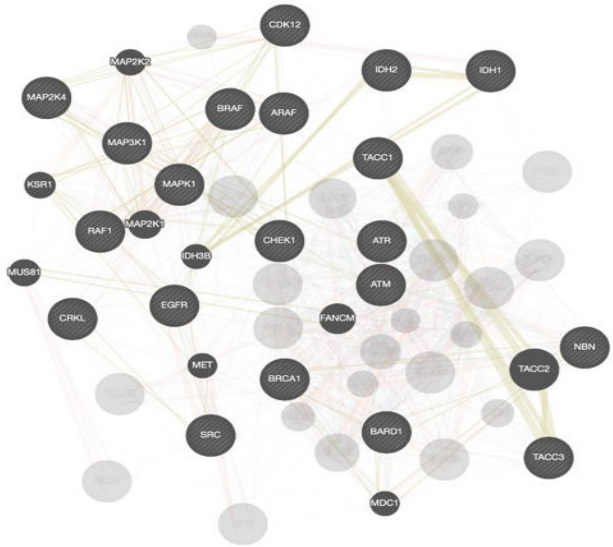

Fig. 5. Shared Protein Domains.

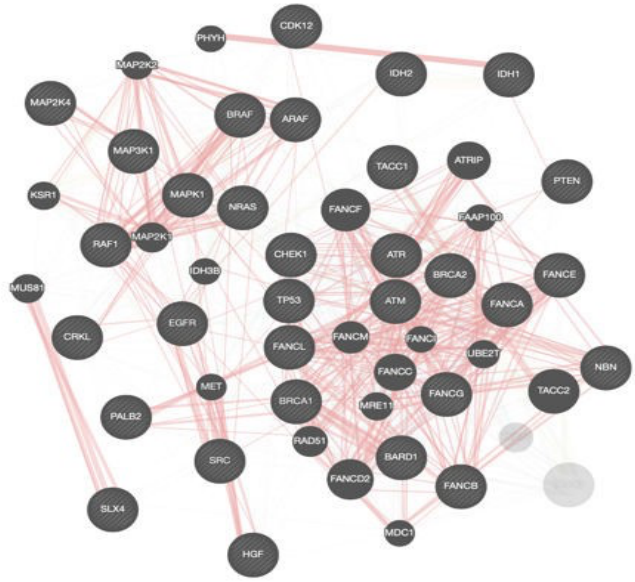

Fig. 6. Physical Interactions.

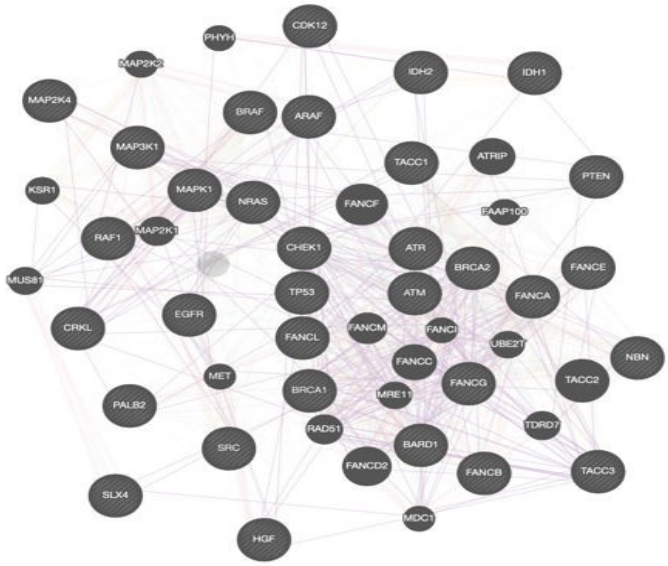

Fig. 7. Co-Expression.

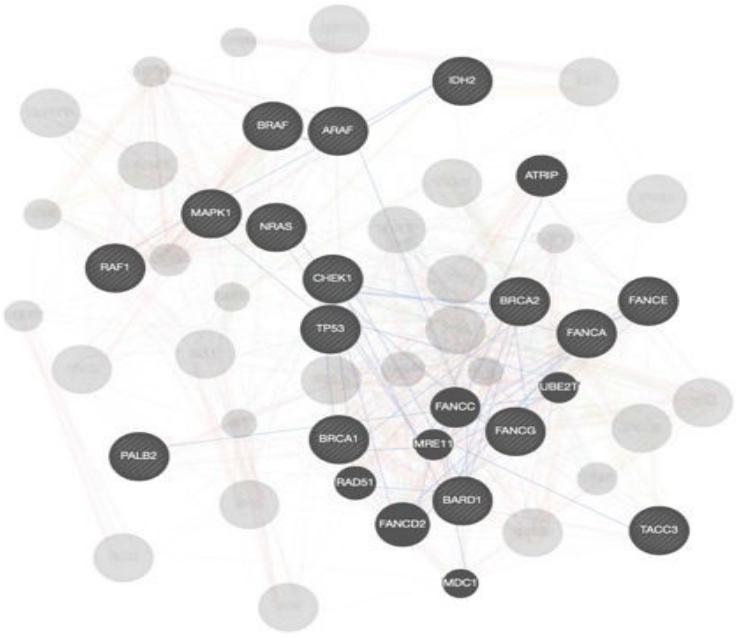

Fig. 8. Co-Localization.

3.3 String

Results showed that for MGMT systems, there were 51 nodes, 562 edges, the average node degree was 22, average local clustering coefficient was 0.778, and the PPI enrichment p-value was <1.0e-16 (Fig. 9). For BER system, there were 51 nodes, 555 edges, average node degree was 21.8, average local clustering coefficient was 21.8, and PPI enrichment p-value was <1.0e-16 (Fig. 10). For MMR system, there were 51 nodes, 908 edges, average node degree was 35.6, average local clustering coefficient was 0.848, and PPI enrichment p-value was <1.0e-16 (Fig. 11). Temozolomide targets these 3 systems and based on these 3 systems found 5 common genes which are XPA, MSH6, TP53, MSH2, and ERCC1.

Notably, TMZ has been reported to induce apoptosis in melanoma cells, and the inactivation of MGMT results in a high level of resistance to TMZ and impairs the expression of MSH2/MSH6 through the over expression of P53 [16]. Therefore, these findings suggested the hypothesis that TMZ might exert anti-tumor effects through MSH2/MSH6/TOP2B/XRCC3 in Glioma patients via a regulated interaction with the TP53 signaling axis.

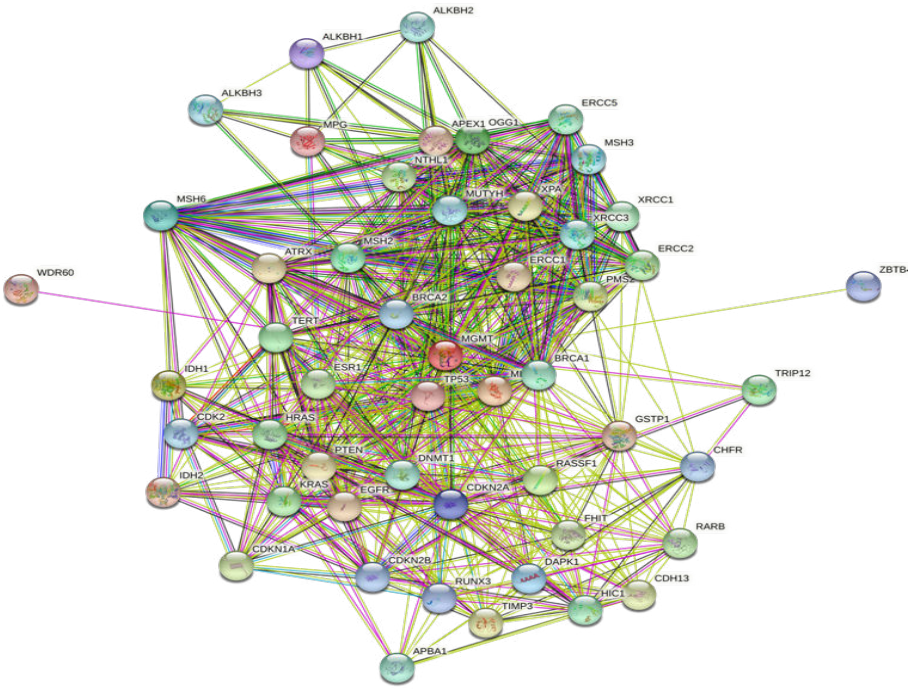

Fig. 9. MGMT System.

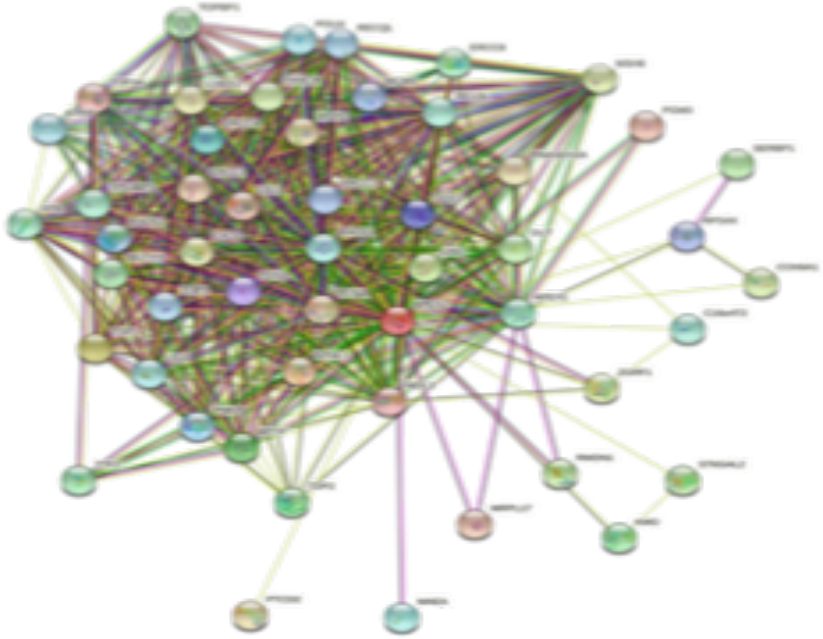

Fig. 10. BER System.

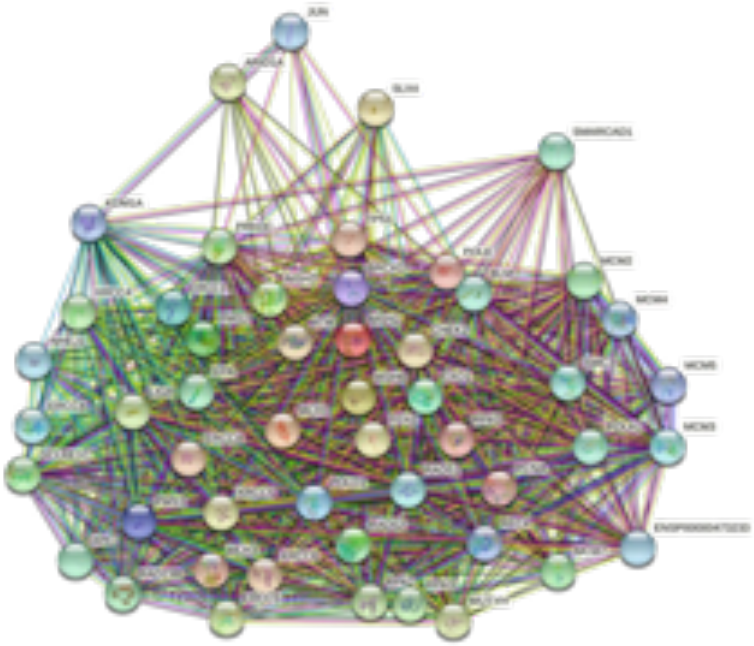

Fig. 11. MMR System.

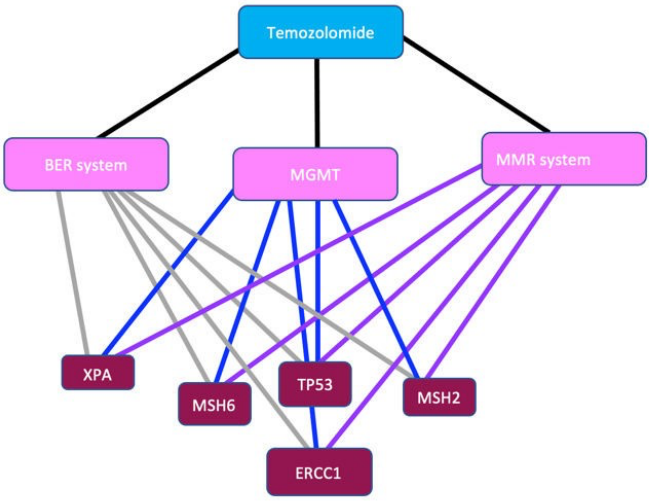

Fig. 12. Drug-Target protein interaction of Temozolomide.

3.4 K-Means

The genes are clustered into 3 clusters- red, blue and green (Fig. 13-Table 2). All the clustered genes have 36 nodes, 233 edges, with an average node degree of 12.9 and average local clustering coefficient of 0.87. Observing each cluster separately gives us some valuable insights about the genes. Cluster one, colored red (Fig. 13), has 13 nodes, 38 edges, with average node degree of 5.85 and average local clustering coefficient of 0.798. On the other hand, the green cluster (cluster 2) (Fig. 14) has 17 nodes, 98 edges with average node degree of 11.5 and average local lustering coefficient of 0.905. And finally, the blue cluster (cluster 3) (Fig. 15) has 6 nodes, 12 edges, with an average node degree of 4 and average local clustering coefficient of 0.9.

If we look at the molecular function of the clusters recording in the 'Gene Ontology' database, cluster 1 has isocitrate dehydrogenase activity, Mitogen-activated protein kinase binding, MAP kinase activity, and Phosphoprotein binding. Cluster 2 has Damaged DNA binding and DNA binding. Cluster 3 has Exodeoxyribonuclease iii activity. These results suggest that if Temozolomide is resistant by target genes, some other drugs can be used to activate it again or some chemical structure changes for Temozolomide can be made.

Table 2. Clustering of the genes.

Cluster 1	Cluster 2	Cluster 3
ARAF	TACC3	MRC1
BRAF	ATM	RAD9A
CRKL	ATR	CHEK1
HGF	BARD1	FANCB
IDH1	BRCA1	HUS1
IDH2	BRCA2	RAD1
MAP2K4	CDK12	
MAP3K1	FANCA	
MAPK1	FANCE	
NRAS	FANCG	
RAF1	FANCL	
SRC	MSH3	
TACC1	MSH6	
	PALB2	
	NBN	

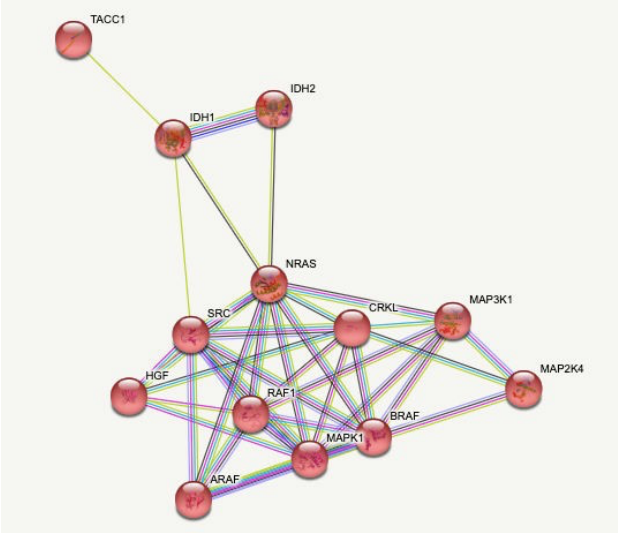

Fig. 13. Cluster 1 - Red Nodes.

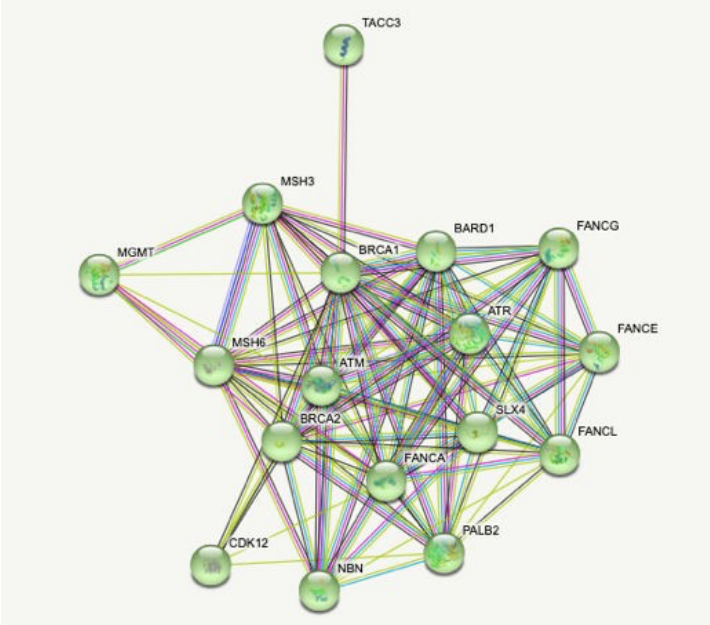

Fig. 14. Cluster 2 - Green Nodes.

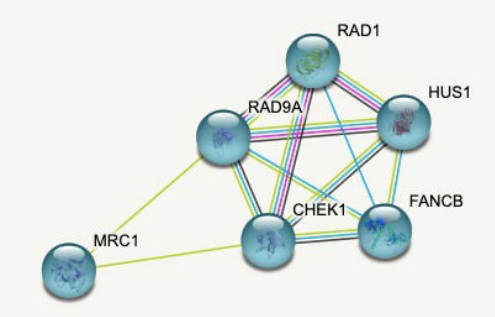

Fig. 15. Cluster 3 - Blue Nodes.

4 Conclusion

Glioma is one of the most aggressive adult primary brain tumors that can be life threating if not diagnosed and treated in time. The alkylating agent Temozolomide, an imidazotetrazine class, is used to treat glioblastoma multiforme and it is the leading compound in a new class of chemotherapeutic agents. However, resistance to Temozolomide is a barrier to effective therapy, decreasing medication effectiveness and life expectancy [18]. In this study, we covered the genes related to Glioma and its treatment with TMZ. We used some web-based tools to show how we can overcome the TMZ resistant for effective treatment of glioma. We showed which genes interact with which one, what their interaction means, and some drugs that can be used as working with TMZ in synergistic. In the future, we would like to explore the research topic more from various perspectives, specifically the cell lines U87 and U251 to find out what role they truly have in Glioma, which genes favor it, and which do not.

References

1. Friedman HS, Kerby T, Calvert H. Temozolomide and treatment of malignant glioma. Clinical cancer research. 2000 Jul 1;6(7):2585-97.
2. HealthLinkBC, Topotecan – Injection, https://www.healthlinkbc.ca/medications/topotecan-injection.
3. Clark MJ, Homer N, O'Connor BD, Chen Z, Eskin A, Lee H, Merriman B, Nelson SF. U87MG decoded: the genomic sequence of a cytogenetically aberrant human cancer cell line. PLoS genetics. 2010 Jan 29;6(1):e1000832.
4. Pei J, Moon KS, Pan S, Lee KH, Ryu HH, Jung TY, Kim IY, Jang WY, Jung CH, Jung S. Proteomic analysis between U87MG and U343MG-A cell lines: searching for candidate proteins for glioma invasion. Brain Tumor Research and Treatment. 2014 Apr 1;2(1):22-8.
5. Lee SY. Temozolomide resistance in glioblastoma multiforme. Genes & diseases. 2016 Sep 1;3(3):198-210.
6. Stevens MF, Hickman JA, Langdon SP, Chubb D, Vickers L, Stone R, Baig G, Goddard C, Gibson NW, Slack JA, Newton C. Antitumor activity and pharmacokinetics in mice of 8-carbamoyl-3-methyl-imidazo [5, 1-d]-1, 2, 3, 5-tetrazin-4 (3 H)-one (CCRG 81045; M & B

39831), a novel drug with potential as an alternative to dacarbazine. Cancer research. 1987 Nov 15;47(22):5846-52.
7. Friedman HS, Dolan ME, Pegg AE, Marcelli S, Keir S, Catino JJ, Bigner DD, Schold Jr SC. Activity of temozolomide in the treatment of central nervous system tumor xenografts. Cancer research. 1995 Jul 1;55(13):2853-7.
8. Plowman J, Waud WR, Koutsoukos AD, Rubinstein LV, Moore TD, Grever MR. Preclinical antitumor activity of temozolomide in mice: efficacy against human brain tumor xenografts and synergism with 1, 3-bis (2-chloroethyl)-1-nitrosourea. Cancer research. 1994 Jul 15;54(14):3793-9.
9. Pera MF, Köberle B, Masters JR. Exceptional sensitivity of testicular germ cell tumour cell lines to the new anti-cancer agent, temozolomide. British journal of cancer. 1995 May;71(5):904-6.
10. Carter CA, Waud WR, Plowman J. Responses of human melanoma, ovarian, and colon tumor xenografts in nude mice to oral temozolomide. InProc Am Assoc Cancer Res 1994 (Vol. 35, p. 297).
11. Wedge SR, Porteous JK, Newlands ES. Effect of single and multiple administration of an O 6-benzylguanine/temozolomide combination: an evaluation in a human melanoma xenograft model. Cancer chemotherapy and pharmacology. 1997 Jun;40:266-72.
12. Yu W, Zhang L, Wei Q, Shao A. O6-methylguanine-DNA methyltransferase (MGMT): challenges and new opportunities in glioma chemotherapy. Frontiers in oncology. 2020 Jan 17;9:1547.
13. Krokan HE, Bjørås M. Base excision repair. Cold Spring Harbor perspectives in biology. 2013 Apr 1;5(4):a012583.
14. Wikipedia, DNA mismatch repair, https://en.wikipedia.org/w/index.php?title=DNA_mismatch_repair&o ldid=1040787236.
15. Lee JE, Lim JH, Hong YK, Yang SH. High-dose metformin plus temozolomide shows increased anti-tumor effects in glioblastoma in vitro and in vivo compared with monotherapy. Cancer research and treatment: official journal of Korean Cancer Association. 2018 Oct 15;50(4):1331-42.
16. Naumann SC, Roos WP, Jöst E, Belohlavek C, Lennerz V, Schmidt CW, Christmann M, Kaina B. Temozolomide-and fotemustine-induced apoptosis in human malignant melanoma cells: response related to MGMT, MMR, DSBs, and p53. British journal of cancer. 2009 Jan;100(2):322-33.
17. Shi J, Dong B, Zhou P, Guan W, Peng Y. Functional network analysis of gene-phenotype connectivity associated with temozolomide. Oncotarget. 2017 Oct 10;8(50):87554.
18. Zhang J, Stevens MF, Laughton CA, Madhusudan S, Bradshaw TD. Acquired resistance to temozolomide in glioma cell lines: molecular mechanisms and potential translational applications. Oncology. 2010 Mar 31;78(2):103-14.
19. Eom JH, Zhang BT. PubMiner: machine learning-based text mining for biomedical information analysis. Genomics & Informatics. 2004;2(2):99-106.
20. National Library of Medicine, PubMed, https://pubmed.ncbi.nlm.nih.gov/.
21. Uniprot, UniProt, https://www.uniprot.org/.
22. Human Genome Organization, HUGO, https://www.hugo-international.org/.
23. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome biology. 2008 Jun;9:1-5.
24. Gene Expression Omnibus, NCBI, https://www.ncbi.nlm.nih.gov/geo/.
25. Leskovec J, Stanford Network Analysis Project, https://snap.stanford.edu/index.html.

26. Leskovec J, Stanford Biomedical Network Dataset Collection, https://snap.stanford.edu/biodata/index.html.
27. Leskovec J, Drug-drug interaction network, https://snap.stanford.edu/biodata/datasets/10001/10001-ChCh-Miner.html.
28. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic acids research. 2018 Jan 4;46(D1):D1074-82.
29. Stanford SNAP Group. Miner: Gigascale multimodal biological network. GitHub Repository. 2017.
30. STRING, STRING Consortium, https://string-db.org/.
31. National Resource for Network Biology, Cytoscape, https://cytoscape.org/.
32. Karve AS, Desai JM, Gadgil SN, Dave N, Wise-Draper TM, Gudelsky GA, Phoenix TN, DasGupta B, Yogendran L, Sengupta S, Plas DR. A review of approaches to potentiate the activity of temozolomide against glioblastoma to overcome resistance. International Journal of Molecular Sciences. 2024 Mar 12;25(6):3217.
33. Smerdi D, Moutafi M, Kotsantis I, Stavrinou LC, Psyrri A. Overcoming Resistance to Temozolomide in Glioblastoma: A Scoping Review of Preclinical and Clinical Data. Life. 2024 May 24;14(6):673.
34. Ntafoulis I, Kleijn A, Ju J, Jimenez-Cowell K, Fabro F, Klein M, Chi Yen RT, Balvers RK, Li Y, Stubbs AP, Kers TV. Ex vivo drug sensitivity screening predicts response to temozolomide in glioblastoma patients and identifies candidate biomarkers. British Journal of Cancer. 2023 Oct 12;129(8):1327-38.
35. Iturrioz-Rodríguez N, Sampron N, Matheu A. Current advances in temozolomide encapsulation for the enhancement of glioblastoma treatment. Theranostics. 2023;13(9):2734.
36. Han X, Abdallah MO, Breuer P, Stahl F, Bakhit Y, Potthoff AL, Pregler BE, Schneider M, Waha A, Wüllner U, Evert BO. Downregulation of MGMT expression by targeted editing of DNA methylation enhances temozolomide sensitivity in glioblastoma. Neoplasia. 2023 Oct 1;44:100929.
37. Obrador E, Moreno-Murciano P, Oriol-Caballo M, López-Blanch R, Pineda B, Gutiérrez-Arroyo JL, Loras A, Gonzalez-Bonet LG, Martinez-Cadenas C, Estrela JM, Marqués-Torrejón MÁ. Glioblastoma therapy: past, present and future. International journal of molecular sciences. 2024 Feb 21;25(5):2529.
38. Wang LM, Englander ZK, Miller ML, Bruce JN. Malignant glioma. Human Brain and Spinal Cord Tumors: From Bench to Bedside. Volume 2: The Path to Bedside Management. 2023 Jul 16:1-30.
39. Kaina B. Temozolomide, procarbazine and nitrosoureas in the therapy of malignant gliomas: update of mechanisms, drug resistance and therapeutic implications. Journal of clinical medicine. 2023 Nov 30;12(23):7442.
40. Jezierzański M, Nafalska N, Stopyra M, Furgoł T, Miciak M, Kabut J, Gisterek-Grocholska I. Temozolomide (TMZ) in the Treatment of Glioblastoma Multiforme—A Literature Review and Clinical Outcomes. Current Oncology. 2024 Jul 12;31(7):3994-4002.