MASTER OF SCIENCE (BIOTECHNOLOGY) [SHALLU CHAUHAN] 2023

PREDICTION OF PECILOCIN AS A POTENTIAL THERAPEUTIC REGIME IN COUNTERING GLIOBLASTOMA USING COMPUTATIONAL APPROACH

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE AWARD OF THE DEGREE

OF

MASTER OF SCIENCE

IN

BIOTECHNOLOGY

Submitted By:

SHALLU CHAUHAN

(2K21/MSCBIO/43)

Under the guidance of:

PROF. PRAVIR KUMAR

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road – 110042

MAY, 2023

CANDIDATE'S DECLARATION

I**, Shallu Chauhan,** (Roll No.: 2K21/MSCBIO/43) of **M.Sc. Biotechnology,** declare that this work which is presented in this Major Project titled "**Prediction of Pecilocin as A Potential Therapeutic Regime in Countering Glioblastoma Using Computational Approach**" submitted to the Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirements for the award of the degree of Master of Science, is novel and my own, carried out during a period from 2th January, 2023 to 20th May, 2023, under the supervision of **Prof. Pravir Kumar**.

I also state that this work has not previously formed the basis for the award of any Degree or other similar title or recognition.

This work has been accepted in an IEEE conference with Scopus indexed proceedings.

The details of which are as follows:

Title of Paper: "Pecilocin as Potential Drug Candidate in GBM Suppression: Bioinformatic Approach to Identify Gene Expression Signature Using ADMET Analysis and Molecular Docking Studies"
Names of Authors: Shallu Chauhan and Pravir Kumar

Name of the Conference: IEEE Bangalore Humanitarian Technology Conference (IEEE B-HTC 2023)
Conference date with Venue: 24th March, 2023, JSS Hospital, Mysuru

Registration for the conference: Completed

Status of the Paper (Accepted/Published/Communicated): Accepted

Date of Paper Communication: 18th January, 2023

Date of Paper Acceptance: 20th February, 2023

Place: Delhi

Date: 2/06/2023

Shallu
2
2/06/23

SHALLU CHAUHAN

(2K21/MSCBIO/43)

SUPERVISOR CERTIFICATE

To the best of my awareness, the project named **"Prediction of Pecilocin as A Potential Therapeutic Regime in Countering Glioblastoma Using Computational Approach"** has never been submitted anywhere else, in whole or in part, for any degree or diploma at this university or anywhere. I further certify that the student's publication and indexing information is accurate.

Place: Delhi

Date: 02/06/2023

02/06/2023

Prof. Pravir Kumar

Professor, HOD, and Guide

Department of Biotechnology

Delhi Technological University

PROOF OF PRESENTATION

Title of the conference Paper: "Pecilocin as Potential Drug Candidate in GBM Suppression: Bioinformatic Approach to Identify Gene Expression Signature Using ADMET Analysis and Molecular Docking Studies"

Name of the Authors: Shallu Chauhan and Pravir Kumar

JSS ACADEMY OF HIGHER EDUCATION AND RESEARCH
JSS HOSPITAL, MYSURU

Certificate

to certify that

Shallu Chauhan

has presented paper entitled

Pecilocin as Potential Drug Candidate in GBM Suppression: Bioinformatic Approach to Identify Gene Expression Signature Using ADMET Analysis and Molecular Docking Studies

in **IEEE Bangalore Humanitarian Technology Conference (IEEE B-HTC 2023)** organized by JSS Academy of Higher Education and Research, JSS Hospital, Mysuru, during 24th and 25th March 2023

Dr. Sudarshan Patil Kulkarni
Chair - IEEE Mysore Subsection
Professor, Dept. of ECE,
SJCE (JSSSTU), Mysuru

Dr. Prashant M Vishwanath
Dean (Research)
JSS AHER, Mysuru -15

JSS ACADEMY OF HIGHER EDUCATION AND RESEARCH
JSS HOSPITAL, MYSURU

Certificate

to certify that

Prof. Pravir Kumar

has presented paper entitled

Pecilocin as Potential Drug Candidate in GBM Suppression: Bioinformatic Approach to Identify Gene Expression Signature Using ADMET Analysis and Molecular Docking Studies

in **IEEE Bangalore Humanitarian Technology Conference (IEEE B-HTC 2023)** organized by JSS Academy of Higher Education and Research, JSS Hospital, Mysuru, during 24th and 25th March 2023

Dr. Sudarshan Patil Kulkarni
Chair - IEEE Mysore Subsection
Professor, Dept. of ECE,
SJCE (JSSSTU), Mysuru

Dr. Prashant M Vishwanath
Dean (Research)
JSS AHER, Mysuru -15

ACKNOWLEDGEMENT

I would like to take this prospect to express my sincere appreciation and debt of gratefulness to my mentor and supervisor, **Prof. Pravir Kumar**, of the Department of Biotechnology at Delhi Technological University, for all of his invaluable assistance, constant support, and motivating suggestions throughout my research period, without which the dissertation would not have been finished within the allotted time frame.

I want to thank the entire faculty of the department of biotechnology for their unwavering assistance. I also want to express gratefulness for the assistance and guidance that I received from my lab seniors, **Dr. Rohan Gupta, Ms. Mehar Sahu, and Mr. Sudhanshu Sharma**, throughout the duration of the work that was accomplished.

On a personal note, I want to thank my family and friends for their love, encouragement, and well wishes.

Shallu Chauhan

2K21/MSCBIO/43

PREDICTION OF PECILOCIN AS A POTENTIAL THERAPEUTIC REGIME IN COUNTERING GLIOBLASTOMA USING COMPUTATIONAL APPROACH

ABSTRACT

The utmost detrimental form of brain cancer is glioblastoma. Current GBM survival rates are around two years because of fast cellular migration and genetically programmed therapeutic escape mechanisms. GBM has been considered to originate in central nervous system (CNS) neuroglia cells. Glioma stem cells (GSCs) have also been found in several studies, a small subset of tumor-initiating and tumor-propagating cells with features similar to neural stem cells. This study aims to identify genetic markers associated with GBM survival by analyzing publicly available gene expression and clinical data from the Gene Expression Omnibus (GEO) datasets.

The differential expression genes (DEGs) among mutant and normal groups were identified, and their survival rates in patients were evaluated. Additionally, enrichment analysis using Enrichr and protein-protein interaction (PPI) were executed to determine significance of these genes. Through the analysis of the GDC Data portal, seven genes (GRIN2A, BCL11A, CAMTA1, ERBB3, WIF1, HLF, and CHN1) were identified as having a substantial impact on GBM development. Furthermore, docking research showed the interaction among natural substances and CAMTA1 protein, revealing a strong affinity between them. The ADME/T study highlighted the probable of Pecilocin as a glioblastoma therapy option.

The dissertation's outline includes an introduction that provides background information on GBM and its current treatment challenges, followed by a literature review that covers various aspects of GBM. The materials and methods section

describes the datasets used to identify DEGs and the workflow followed for data analysis of GBM patient. The results and discussion section presents the findings of sequence similarity analysis, target receptor structures, and the analysis of receptor-ligand interactions. In conclusion, this study identifies potential genetic markers associated with GBM survival and highlights Pecilocin as a promising therapeutic option. The research provides insights into the molecular mechanisms underlying GBM and opens avenues for further investigation in the field. The dissertation emphasizes the importance of understanding GBM biology to develop effective treatments and improve patient outcomes.

Keywords: *Glioblastoma, Datasets, Differential expression genes, Biomarkers, GBM survival, CAMTA1, Pecilocin*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

S.No	Abbreviation	Full Form
1	CNS	Central Nervous System
2	GSCs	Glioma stem cells
3	GEO	Gene Expression Omnibus
4	DEGs	Differential expression genes
5	PPI	Protein-protein interaction
6	GBM	Glioblastoma Multiforme
7	GDC	Genomic Data Commons
8	CAMTA1	Calmodulin Binding Transcription Activator 1
9	ADME/T	Chemical absorption, distribution, metabolism, excretion, and toxicity
10	NCBI	National Center for Biotechnology Information Gene Expression Omnibus
11	GO	Gene Ontology
12	KEGG	Kyoto Encyclopedia of Genes and Genomes
13	VEGF	Vascular endothelial growth factor
14	LOH	Loss of heterozygozity
15	PTEN	Phosphatase and tensin homolog
16	EGFR	Epidermal growth factor receptor
17	IDH1	Isocitrate dehydrogenase 1
18	IDH2	Isocitrate dehydrogenase 2
19	PPI	Protein-protein interaction
20	GESA	Gene set enrichment analysis
21	COCONUT	Collection of Open Natural Products Database
22	SDF	Structure-Data File
23	PDB	Protein Data Bank

CHAPTER– 1

INTRODUCTION

1.1 BACKGROUND

About 60% of adult brain tumors are caused by the prevalent and severe primary malignant brain tumor glioblastoma multiforme (GBM) [1]. GBM is the utmost prevalent and devastating form of tumor, presenting significant challenges in treatment. Despite advancements in medical science, the prognosis for GBM patients remains poor, emphasizing the need for improved therapeutic options. GBM is primarily driven by genetic mutations, which are primarily responsible for tumor formation and progression. Understanding the genetic markers responsible for GBM survival is essential for identifying potential targets for therapy. Treatment for GBMs is extremely difficult due to their extreme invasiveness and tendency to metastasize to other parts of the brain. Radiation therapy, chemotherapy, and surgical removal are the current accepted treatments for GBM [2]. Unfortunately, these treatments tend to have short-term effects, and most GBM patients eventually succumb away from the illness. It is believed that environmental as well as genetic factors contribute to the development of GBM [3].

Genetic modifications, for instance those in the TP53, EGFR, and IDH1 genes, are linked to GBM [4]. Furthermore, the development of GBM has been linked to a number of environmental variables, including exposure to ionising radiation, particular chemicals, and particular viruses [5]. The goal of recent research has been to find novel genetic targets for GBM treatment. Gene-directed therapy, which entails focusing on particular genes responsible for GBM cell growth and dissemination, is one potential

strategy. Although gene-directed therapy has showed promise in clinical studies, further study is required to create more powerful cures [6].

In addition to gene-directed therapy, other promising approaches for GBM treatment include immunotherapy and targeted therapy. Utilising the body's own immune system, immunotherapy contests cancer cells, while targeted therapy employs drugs that specifically target molecules present on cancer cells. Both immunotherapy and targeted therapy have demonstrated some success in clinical trials, but additional research is essential to improve their efficacy [7]. The pursuit of novel treatments for GBM remains an active area of research. Through ongoing investigations, it is hoped that new therapies will be developed to enhance the survival rates of GBM patients while providing improved quality of life.

In current trends, the advent of high output technologies has enabled the collection of large-scale gene expression datasets, providing valuable resources to study the prognosis and molecular features of patient. The NCBI-GEO is widely used repository that offers publicly accessible data on GBM patients. By analyzing these datasets, researchers can identify DEGs and investigate their potential role in GBM.

1.2 Problem Statement

Despite numerous efforts, treating GBM remains a significant challenge due to its aggressive nature and limited treatment options Standard-of-care treatment usually entails surgical excision of the tumour, proceeded by radiation therapy and temozolomide chemotherapy [8]. However, the development of resistance and tumor recurrence are common, leading to poor patient outcomes. As a result, it is imperative to find new therapeutic targets and GBM strategies for treatment.

1.3 Objectives of the Study

The primary impartial of this finding is to utilize a bioinformatic approach to investigate gene expression signatures and potential targets for GBM. The objectives are:

[1] To collect microarray expressed datasets for GBM from the NCBI-GEO.

[2] To identify DEGs by comparing gene expression profiles of GBM tumors with normal brain tissues.

[3] To perform enrichment analysis of the DEGs using KEGG pathway and GO analysis, providing understandings for the biological processes and molecular pathways interlinked with GBM.

[4] To analyze protein-protein interactions (PPI) among the DEGs to identify potential protein interactions and networks involved in GBM.

[5] To evaluate the survival significance of the DEGs by analyzing their association with patient survival using the GDC Data Portal.

[6] To investigate the potential of natural substance Pecilocin as a glioblastoma therapy option through molecular docking studies, assessing its affinity for the CAMTA1 protein.

[7] To conduct ADMET analysis to assess drug-like properties of potential candidates, including Pecilocin.

By achieving these objectives, this finding aims to contribute to the understanding of GBM at molecular level, identify potential therapeutic targets, and explore the efficacy of Pecilocin as a novel treatment option.

In conclusion, this chapter provided an introduction to the background of GBM, highlighting the challenges in its treatment and the importance of identifying genetic markers for GBM survival. The problem statement emphasized the need for improved therapeutic options, and the objectives of the study were outlined to address these challenges. In the following chapters, the methodology, results, discussion, and conclusion of the study will be presented in detail, shedding light on potential advancements in GBM treatment using a bioinformatic approach.

CHAPTER- 2

LITERATURE REVIEW

2.1 GBM Disease - Outline to Prevalence, Causes, and Pathophysiology

2.1.1 Prevalence and Causes

GBM is exceedingly prevalent and destructive brain tumor, accounting for the common of malignant brain tumors in adults. The occurrence of GBM is believed to be roughly 3 per 100,000 people, with an older age group having a higher prevalence [9]. The precise origins of GBM are unknown, however various risk factors have been discovered. These include exposure to ionizing radiation, certain genetic syndromes, and prior history of brain tumors or neurosurgery [10]. Additionally, certain genetic mutations and alterations subsidize to the development and progression of GBM.

2.1.2 Genetic Basis of GBM

Genetic alterations show a crucial role in the pathogenesis of GBM. The most prevalent genetic mutation seen in GBM is chromosome 10 loss of heterozygozity (LOH), which results in the inactivation of the tumour suppressor gene PTEN (phosphatase and tensin homolog) [11]. PTEN is tangled in cell growth and survival, and its inactivation leads to uncontrolled cell proliferation and tumor formation. Another frequently observed genetic alteration is the extension of the EGFR gene,

resulting in over secretion of the EGFR protein. EGFR over expression promotes cell growth, contributing to the aggressive nature of GBM [12].

In addition to PTEN and EGFR, other genetic alterations have also been identified in GBM. These include alterations in TP53, which codes for the tumor suppressor protein p53, controls the cell cycle and repairs DNA [13]. The loss of p53 function allows for the accumulation of further genetic alterations and promotes tumor growth. Other genetic mutations commonly found in GBM involve the genes encoding IDH1 and IDH2. Alterations in these genes result in the abnormal production of the metabolite 2-hydroxyglutarate, leading to epigenetic changes and altered cellular metabolism [14].

Understanding the genetic basis of GBM is critical to uncovering new therapeutic targets. Targeted therapies aimed at specific genetic alterations have shown promise in preclinical and clinical studies, highlighting the importance of personalized medicine approaches for GBM treatment.

2.2 Overview of Previous Studies on GBM Biomarkers and Therapeutic Targets

Numerous studies have been conducted to identify biomarkers and therapeutic targets for GBM. The availability of large-scale gene expression datasets has aided in the discovery of differentially expressed genes (DEGs) linked to GBM. Through bioinformatic analysis, these studies have acknowledged DEGs that play perilous roles in GBM development, and patient survival.

Functional enrichment analysis of DEGs has provided understandings for biological processes and molecular pathways interlinked with GBM. KEGG pathway analysis has revealed the involvement of various pathways, including the PI3K-Akt signaling

pathway, focal adhesion, and cell cycle regulation, in GBM pathogenesis [15]. Additionally, Gene Ontology (GO) analysis has highlighted the significance of biological processes such as cell proliferation, angiogenesis, and immune response in GBM.

Protein-protein interaction (PPI) network analysis has further expanded our understanding of the molecular interactions and networks involved in GBM. By examining the physical and functional associations between proteins, PPI analysis has identified key hubs and modules within the network, providing potential targets for therapeutic intervention [16].

Moreover, survival analysis has been executed to evaluate the forecasting of DEGs in GBM. These studies have identified genes whose expression levels are associated with patient survival, providing potential biomarkers for predicting prognosis and response to treatment.

2.3 Potential Therapeutic Targets and Treatment Strategies

The identification of potential therapeutic targets is crucial for developing effective treatment strategies for GBM. Several targets have been proposed based on previous studies.

The EGFR pathway, which is typically affected in GBM, is one expected target. Preclinical studies have showed potential for EGFR inhibitors like erlotinib and gefitinib [17]. Additionally, inhibitors targeting downstream signaling molecules, such as PI3K and mTOR, have demonstrated efficacy in inhibiting GBM cell growth and survival [18].

An additional potential target is VEGF pathway, which is decisive for the angiogenesis of tumors. Anti-VEGF agents, including bevacizumab, have been investigated in clinical trials and have shown some clinical benefit, particularly in terms of reducing tumor size and edema [19].

Immunotherapy approaches have also gained attention in GBM treatment. Strategies that target immune checkpoint molecules, like CTLA-4 and PD-1, have demonstrated efficacy in treating other cancer types and are currently being investigated for use in GBM [20]. Additionally, adoptive cell-based therapies are being researched as potential GBM treatments, with chimeric antigen receptor (CAR) T-cell therapy [21]

Additionally, molecular targeted therapies aimed at specific genetic alterations, such as IDH1 mutations, are being explored. Small molecule inhibitors targeting mutant IDH1 have shown promise in pre-clinical models and are used in clinical trials [22].

2.4 Limitations and Gaps in the Literature

Despite the progress made in identifying potential biomarkers and therapeutic targets for GBM, several challenges and gaps in the literature remain. One limitation is the heterogeneity of GBM, both at the genomic and cellular levels [23]. This heterogeneity contributes to treatment resistance and the varying responses observed in patients. Therefore, personalized medicine approaches that consider the individual genetic profile and tumor characteristics are needed.

Another limitation is the lack of effective treatment options for recurrent GBM. While initial treatments, such as surgery, radiation, and chemotherapy, may provide temporary relief, the development of resistance and tumor recurrence is common.

Therefore, new therapeutic approaches that focus exclusively on recurrent GBM need to be explored [24].

Moreover, the BBB poses a significant challenge in the delivery of therapeutics to GBM [25]. The BBB bounds the diffusion of medicines in brain, making it challenging to achieve therapeutic concentrations within the tumor. Developing strategies to overcome the BBB and enhance drug delivery to GBM is an area of active research.

Furthermore, the identification of biomarkers for forecast and treatment response in GBM, is still an ongoing challenge. While several potential biomarkers have been identified, further validation and standardization are required to establish their clinical utility [26].

In conclusion, this chapter provided a comprehensive literature review on GBM, including its prevalence, genetic basis, previous studies on biomarkers and therapeutic targets, and current treatment strategies. The limitations and gaps in the literature were highlighted, emphasizing the need for personalized medicine approaches, novel therapeutic strategies for recurrent GBM, strategies to dazed the BBB, and reliable biomarkers for prognosis and treatment response. The subsequent chapters will present the methodology, results, discussion, and conclusion of the study, addressing these challenges and contributing to the advancement of GBM research and treatment.

CHAPTER- 3

METHODOLOGY OF STUDY

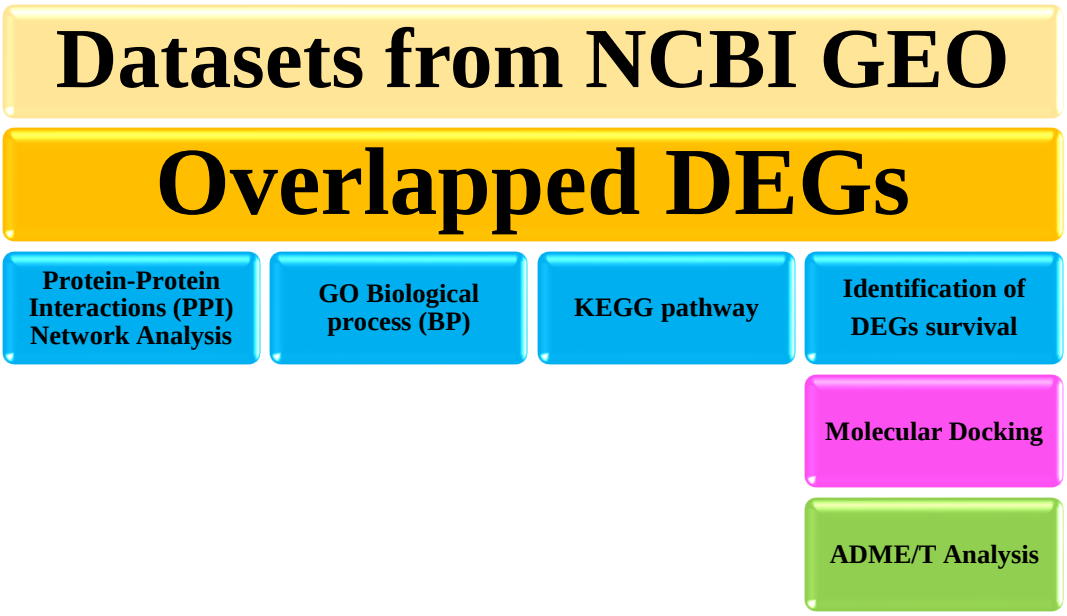

Figure 3.1: Block schematic of the multi-stage methodology to study GBM

3.1 Data Set Collection

In this chapter, we describe the methodology employed in our study to investigate potential therapeutic targets for glioblastoma multiforme (GBM). The following sections outline the steps involved in data set collection, detection of DEGs, enrichment analysis, PPI network analysis, data collection for molecular docking, preparation of receptors and ligands, molecular docking using SwissDock, ADMET analysis, and visualization and processing of results using Chimera.

The first step in our study was to collect microarray gene expression datasets for GBM. We obtained the datasets from the NCBI GEO database, a repository of publicly available gene expression data. We selected two datasets, namely GSE4290 and GSE12657, which contained gene expression profiles of GBM samples and corresponding normal brain samples. These datasets were chosen based on their

sample size, availability of clinical information, and relevance to our research objectives.

3.2 Identification of Differential Expression Genes (DEGs)

To find out DEGs among GBM patients and normal brain samples, we employed GEO2R web tool. GEO2R allows for the analysis of gene expression data by comparing samples from different experimental groups [27]. We inputted the raw expression data from the selected datasets and applied appropriate filters to normalize the data and remove outliers.

Next, we established criteria to select DEGs based on fold change and statistical significance. Genes with a fold change of at least 2 and a p-value below a predetermined threshold (e.g., 0.05) were considered differentially expressed [28]. These criteria helped us identify genes that exhibited significant changes in expression levels between GBM and normal brain samples.

3.3 Functional Enrichment Analysis of DEGs

For gaining insights into biological processes and molecular pathways associated with the find out DEGs, we performed functional enrichment analysis. We utilized gene set enrichment analysis (GSEA) through the EnrichR web tool [29]. GSEA compares the expression of a predefined gene set against the remaining genes in the dataset, allowing us to determine whether the genes in our DEG list were enriched in specific biological pathways.

We applied the KEGG pathway analysis as well as GO findings using EnrichR. KEGG pathway identified the molecular pathways in which the DEGs were involved, providing information on the underlying biological mechanisms. GO analysis provided insights into the gene ontology terms associated with the DEGs, together with biological processes, cellular components, and molecular functions [29].

3.4 Protein-Protein Interactions (PPI) Network Analysis

To understand the interactions and functional relationships between the DEGs, we constructed a protein-protein interactions (PPI) network [30]. We utilized the STRING database, which consolidates known and predicted protein-protein interactions from various sources.

Using the DEG list as input, we retrieved the corresponding protein IDs and submitted them to the STRING database. The resulting PPI network provided a visual representation of the interactions between the proteins [31]. We analyzed the network to identify key hubs and modules, which are indicative of proteins with critical roles in GBM development and progression.

3.5 Data Collection for Molecular Docking

In addition to exploring gene expression data, we aimed to investigate potential natural compounds as therapeutic agents for GBM. To achieve this, we collected data on natural compounds with potential anticancer properties.

We employed COCONUT (Collection of Open Natural Products Database) to analyze natural products. COCONUT provides a comprehensive collection of natural

compounds and their associated biological activities. Based on the literature and COCONUT database search, we selected a set of natural compounds with reported anticancer properties for further analysis [32].

We retrieved the chemical structures of the selected natural compounds in Structure-Data File (SDF) format from the PubChem database, a comprehensive resource for chemical information.

Additionally, we chose the CAMTA1 gene as a target for our molecular docking analysis. CAMTA1 has been implicated in GBM progression and might serve as a probable therapeutic target [33]. We selected the CAMTA1 protein structure from the PDB, which offers experimentally determined protein structures.

3.6 Preparation of Receptors and Ligands

To prepare the CAMTA1 protein structure for docking analysis, we performed necessary pre-processing steps. This included removing any water molecules, ions, and other heteroatoms from the protein structure and assigning proper bond orders and protonation states [34].

For the ligands (natural compounds), we converted the SDF structures obtained from PubChem into mol2 format using Discovery Studio software. Mol2 format is commonly used for molecular docking and is compatible with the SwissDock tool we utilized [34].

3.7 Molecular Docking using SwissDock

We employed the SwissDock web tool for molecular docking analysis. SwissDock utilizes a protein-ligand docking algorithm to predict the binding affinity and binding mode between proteins and small molecules [35].

The ligands (natural compounds) and the prepared CAMTA1 protein structure were uploaded to the SwissDock server. The docking simulation was performed to predict the most favorable binding conformations and orientations of ligands within the binding site of CAMTA1 gene.

The docking results were analyzed based on various parameters, including binding energy, which serves as an indicator of the strength of the protein-ligand interaction. Lower binding energies suggest stronger binding affinity between the ligand and the protein.

3.8 ADMET Analysis

To estimate the pharmacokinetic properties and drug-likeness of the selected ligands, we employed the SwissADME web tool. SwissADME predicts various ADMET parameters for small molecules [36].

We analyzed the physicochemical properties of the ligands, including molecular weight, water solubility, lipophilicity (logP), and polar surface area. We also evaluated drug-likeness using Lipinski's rule, which assesses properties related to oral bioavailability and drug development [36].

3.9 Visualization and Processing of Results Using Chimera

For visualization and study of molecular structures and interactions, we utilized the Chimera software. Chimera is a powerful tool for visualizing protein-ligand complexes, generating molecular graphics, and performing structural analysis [37], [38].

We imported the docking results, including the protein-ligand complexes with their predicted binding modes, into Chimera. This allowed us to visualize the interactions between the ligands and the CAMTA1 protein and analyze the binding conformation and key residues involved in the interaction.

In conclusion, this chapter described the methodology employed in our study to investigate potential therapeutic targets for GBM. We explained the data set collection process, identification of DEGs, functional enrichment analysis, PPI network analysis, data collection for molecular docking, preparation of receptors and ligands, molecular docking using SwissDock, ADMET analysis, and visualization and processing of results using Chimera [36]. These methods were essential in addressing our research objectives and will provide valuable insights into potential therapeutic strategies for GBM.

CHAPTER- 4

RESULTS

In this chapter, we present the findings from each stage of the study, including the identification of DEGs, enrichment analysis, protein-protein interactions (PPI) network analysis, DEGs survival analysis, molecular docking results, and ADMET analysis. These results provide insights into hypothetical therapeutic targets for GBM.

4.1 Identification of Differential Expression Genes (DEGs)

Utilizing the analysis of gene expression data from the selected datasets (GSE4290 and GSE12657), we identified a set of DEGs that exhibited significant differences in expression between GBM and normal brain samples. We applied a fold change threshold of at least 2 and statistical significance threshold (p-value < 0.05) to determine DEGs.

We evaluated the common DEGs and found 285 overlapping genes. The DEGs included genes involved in various biological processes and molecular functions relevant to GBM, such as cell cycle instruction, apoptosis, cell proliferation, and DNA repair. These findings suggest dysregulation of these processes in GBM and provide potential targets for therapeutic interventions.

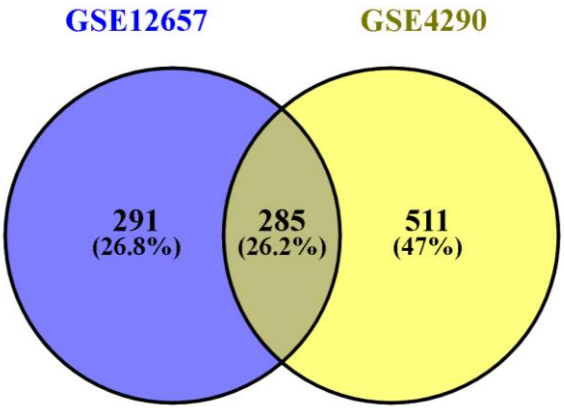

Figure 4.1: The Venn diagram of 285 overlapping up-regulated genes identified from two datasets GSE4290 and GSE12657 of GBM

4.2 Functional Enrichment Analysis

Functional enrichment analysis was executed for 285 overlapping DEGs to study the biological processes and molecular pathways associated with the DEGs. Through KEGG pathway and GO analysis, we identified several enriched pathways and gene ontology terms.

The KEGG pathway study revealed the involvement of DEGs in critical signaling pathways. These pathways play significant roles in cancer development and progression.

The GO analysis provided information about the biological processes, cellular components, and functions related to DEGs. We observed enrichment in processes such as cell division, cell adhesion, and response to DNA damage, indicating their relevance to GBM pathogenesis.

TABLE I. **KEGG PATHWAY HUMAN ASSOCIATED WITH GLIOBLASTOMA**

Name	**P-value**	**Adjusted p-value**	**Odds Ratio**	**Combined score**
GABAergic synapse	1.002e-18	2.044e-16	21.49	890.59
Morphine addiction	3.313e-17	3.379e-15	19.49	739.46
Nicotine addiction	1.714e-10	4.371e-9	23.86	536.55
Insulin secretion	1.235e-12	4.200e-11	15.37	421.47
Glutamatergic synapse	5.500e-13	2.244e-11	12.83	362.15
Retrograde endocannabinoid signaling	3.730e-13	1.902e-11	10.84	310.35
Calcium signaling pathway	8.986e-15	6.110e-13	8.72	282.06
Long-term potentiation	2.570e-9	4.370e-8	14.09	278.76
Circadian entrainment	1.218e-9	2.258e-8	11.17	229.27
Oxytocin signaling pathway	6.958e-10	1.420e-8	8.44	177.92

TABLE II. **GO BIOLOGICAL FUNCTION ASSOCIATED WITH GLIOBLASTOMA**

Name	**P-value**	**Adjusted p-value**	**Odds Ratio**	**Combined score**
presynaptic active zone organization (GO:1990709)	5.922e-7	0.00002584	140.31	2011.92
anterograde trans-synaptic signaling (GO:0098916)	3.452e-33	2.861e-30	16.70	1247.98
chemical synaptic transmission (GO:0007268)	6.222e-36	1.032e-32	15.27	1238.18
protein localization to axon (GO:0099612)	0.00002803	0.0007151	104.86	1099.12
inhibitory synapse assembly (GO:1904862)	2.443e-7	0.00001366	58.66	893.05
positive regulation of non-membrane protein tyrosine kinase activity (GO:1903997)	0.00005547	0.001332	69.90	685.00
Neurotransmitter secretion (GO:0007269)	2.050e-11	4.856e-9	23.94	589.27
regulation of neurotransmitter receptor activity (GO:0099601)	1.353e-12	5.606e-10	20.89	570.94
regulation of postsynaptic neurotransmitter receptor activity (GO:0098962)	0.000007924	0.0002628	46.76	549.21

regulation of nervous system process (GO:0031644)	1.844e-7	0.00001132	32.59	505.38

4.3 Protein-Protein Interactions (PPI) Network Analysis

The PPI network system of the 285 overlapping DEGs was constructed that allowed us to explore the interactions and functional relationships between the DEGs. The network revealed several key hub proteins that exhibited high connectivity and served as potential central players in the network.

By identifying modules within the network, we uncovered groups of DEGs that were closely interconnected and likely involved in common biological processes. These modules included proteins associated with cell cycle regulation, DNA repair, and cellular signaling.

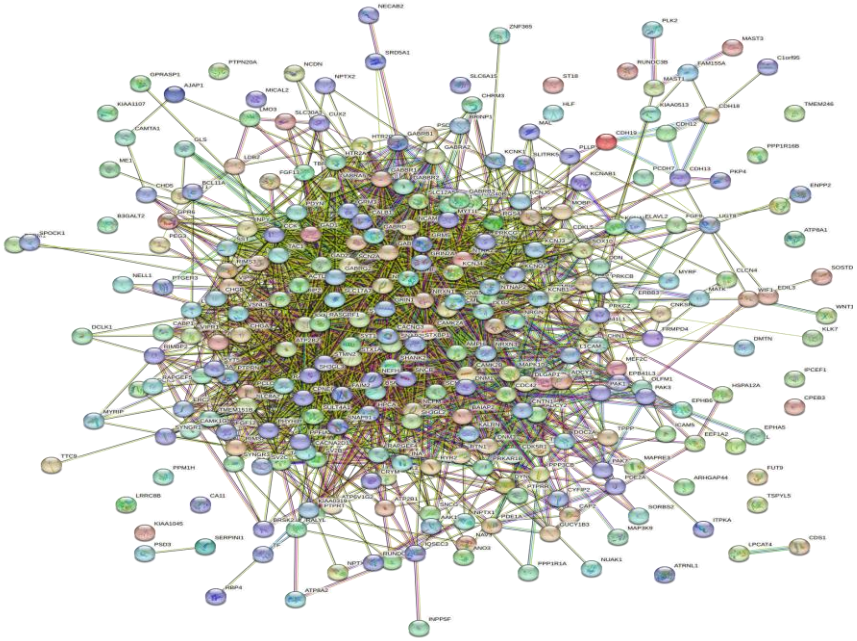

Figure 4.2: The Protein-protein interaction network of 285 overlapped genes which were common in two datasets

4.4 DEGs Survival Analysis

To investigate the prognostic implication of the DEGs, we accomplished survival analysis using available clinical data from the GBM patients in the selected datasets. For the overlapping 285 DEGs among the GBM datasets, using GDC Data Portal, seven genes (GRIN2A, BCL11A, CAMTA1, ERBB3, WIF1, HLF, and CHN1) were found noteworthy in the patients' survival having $P < 0.05$.

The survival study revealed that certain DEGs were associated with patient survival outcomes. Some genes showed a significant correlation with poor prognosis, while others were associated with improved survival. These findings highlight the probable DEGs for the prognostic biomarkers for GBM.

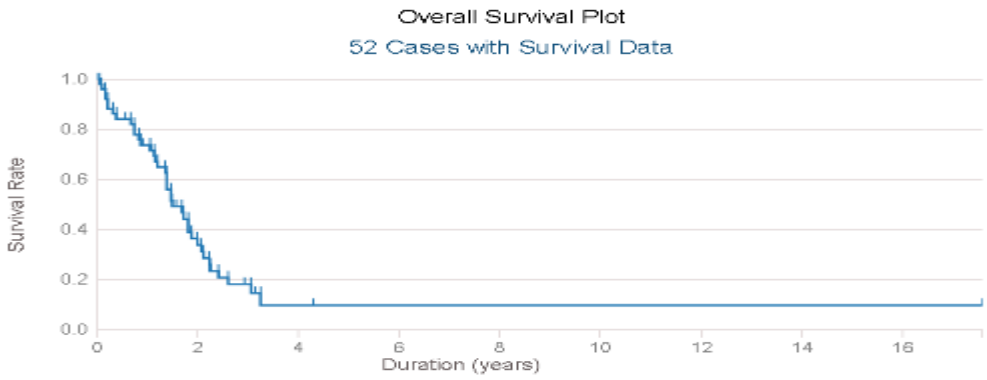

Figure 4.3: Overall survival plot of patients having mutated genes

4.5 Molecular Docking Results

In the molecular docking analysis, we evaluated the binding affinities and interactions between the selected natural compounds and the CAMTA1 protein, a potential therapeutic target for GBM.

The docking results provided insights into the binding modes and orientations of the ligands within the binding site of CAMTA. We observed favorable binding interactions between specific ligands and the protein, indicating potential therapeutic efficacy.

Out of 50 natural compounds docked, 9 compounds met the binding energy less than -7.13 Kcal/mol (energy of standard) as shown in TABLE III. Out of these ten, Pecilocin has the furthermost negative binding energy which suggests it has the best binding affinity with the gene (Fig.4.4).

Based on binding energy values, we identified ligands with higher affinity towards the CAMTA1 protein. These ligands have the potential to inhibit the protein's function and may serve as promising candidates for further experimental validation.

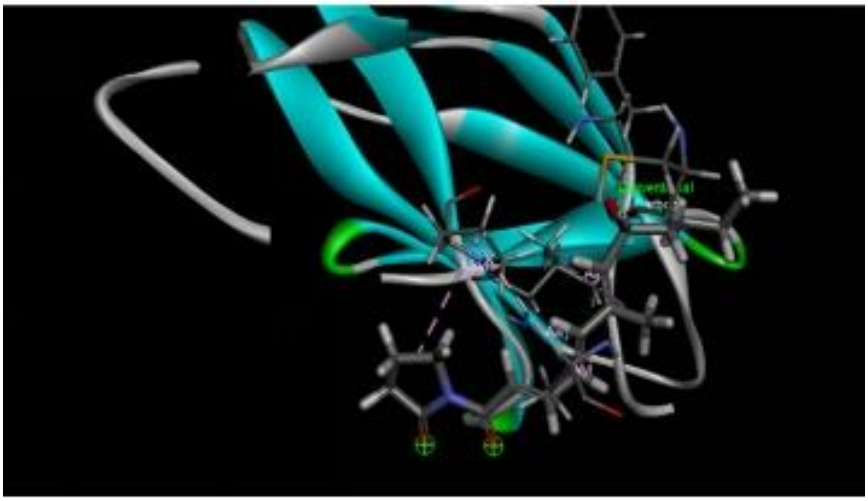

Figure 4.4: Pecilocin binds at CAMTA1 active site

TABLE III. MOLECULAR DOCKING AND ADME/T ANALYSIS RESULT

Name	Cluster	Element	Estimated ΔG (kcal/mol)	Lipinski	BBB permeant
Chrysin (Standard)	1	0	-7.13	Yes; 0 violation	Yes
Supral/ Pecilocin	1	5	-8.37	Yes; 0 violation	Yes
Nigerapyrone E	9	8	-7.23	Yes; 0 violation	Yes
Rankinidine	7	0	-7.28	Yes; 0 violation	Yes
Kuguacin M	0	0	-7.56	Yes; 0 violation	Yes
Anthrakunthone	1	4	-7.21	Yes; 0 violation	Yes
Laurequinone	1	1	-7.1	Yes; 0 violation	Yes
Atractylenolide I	1	1	-7.31	Yes; 0 violation	Yes
Melophlin D	0	1	-8.32	Yes; 0 violation	Yes
Glaucocalyxin A	0	0	-7.51	Yes; 0 violation	yes

4.6 ADMET Analysis

The ADMET analysis provided information on pharmacokinetic properties and drug-likeness of selected ligands. We considered various parameters, including molecular weight, lipophilicity (logP), water solubility, and drug-likeness according to Lipinski's rule.

The analysis revealed that the selected ligands exhibited favorable physicochemical properties and drug-likeness profiles, indicating their potential for further development as therapeutic agents. These findings support the feasibility of utilizing natural compounds as potential GBM therapeutics.

In conclusion, this chapter presented the results obtained from each stage of the study. We identified DEGs associated with GBM, performed functional enrichment analysis to uncover relevant pathways and processes, analyzed PPI networks to explore protein interactions, conducted survival analysis to evaluate the prognosis of DEGs, evaluated molecular docking result for potential therapeutic ligands, and performed ADMET analysis to assess drug-likeness. These findings contribute to a comprehensive understanding of potential therapeutic targets and compounds for GBM treatment.

CHAPTER- 5

DISCUSSION & CONCLUSION

In this chapter, we provide an in-depth interpretation and analysis of the results obtained from the study. We discuss the implications and significance of the findings, highlight the drawbacks of the study, and provide suggestions for additional research in the field of glioblastoma multiforme (GBM) treatment.

5.1 Interpretation and Analysis of Results

We begin the discussion by summarizing the key findings from each stage of the study. The identification of differential expression genes (DEGs) revealed dysregulated biological processes and molecular functions in GBM, such as cycle regulation, cell proliferation, and DNA repair. These findings align with the known characteristics of GBM, highlighting the importance of these processes in GBM pathogenesis.

The functional enrichment analysis provided insights into biological pathways and GO linked with DEGs. Additionally, the GO analysis highlighted the relevance of processes such as cell division, cell adhesion, and response to DNA damage in GBM development and progression.

The PPI network scrutiny revealed pivot proteins and modules associated with key biological processes in GBM. This network-based approach allowed us to identify potential central players and their interactions within the DEG network, providing a more comprehensive understanding of the molecular mechanisms underlying GBM.

The survival analysis demonstrated the prognostic significance of certain DEGs in GBM patients. These findings have implications for patient stratification and personalized treatment approaches. Genes associated with poor prognosis could serve as potential targets for more aggressive therapeutic interventions, while those associated with improved survival may guide the development of novel prognostic biomarkers.

The molecular docking analysis identified natural compounds with favorable binding affinities and interactions with the CAMTA1 protein. These findings suggest the potential of these compounds as therapeutic agents for GBM. However, further experimental validation is necessary to confirm their efficacy and safety.

The ADMET analysis provided important insights into the pharmacokinetic properties and drug-likeness of the selected ligands. Ligands with favorable physicochemical properties and drug-likeness profiles may have better chances of success in further preclinical and clinical development.

5.2 Implications and Significance of Findings

The findings from this study have several inferences for treatment of GBM. The identification of DEGs associated with critical biological processes and pathways provides potential targets for therapeutic interventions. Modulating these dysregulated processes could potentially disrupt GBM growth and progression.

The survival analysis revealed DEGs that are correlated with patient prognosis. These genes may serve as prognostic biomarkers, and the advancement of personalized

treatment approaches. Identifying patients with poor prognosis at an early stage can guide the implementation of more aggressive therapies and improve patient outcomes. The molecular docking results suggest the potential of natural compounds as therapeutic agents for GBM. These compounds can be further investigated and optimized to enhance their efficacy and specificity. The identification of ligands with favorable binding affinities towards the CAMTA1 protein provides a starting point for future drug discovery efforts.

5.3 Limitations of the Study

It is important to acknowledge the drawbacks of the study. Firstly, the study was based on publicly available gene expression datasets, which may have inherent biases and variations. The use of additional datasets or experimental validation is warranted to confirm the robustness of the findings.

Secondly, the molecular docking analysis provides insights into the potential binding interactions between ligands and the CAMTA1 protein. However, it is crucial to validate these findings through *in vitro* as well as *in vivo* analysis to measure the functional effects of ligand-protein interactions and their impact on GBM cells.

Furthermore, the ADMET analysis provides initial insights into the pharmacokinetic properties of the ligands. However, more extensive *in vitro* experiments are necessary to fully judge the safety, bioavailability, and metabolism of these compounds.

5.4 Suggestions for Further Research

Based on the findings and limitations of this study, several recommendations for further study can be made. Primarily, conducting experimental validation, such as cell-based assays and animal models, can provide functional evidence for the identified DEGs, potential therapeutic targets, and ligand-protein interactions.

Additionally, investigating the combination therapy of natural compounds with existing GBM treatments, such as chemotherapy and radiation therapy, could be explored. Synergistic effects of these compounds with standard therapies may enhance treatment outcomes and overcome drug resistance.

Furthermore, exploring the character of other molecular players and pathways in GBM, beyond DEGs discovered in this finding, could offer more comprehensive thoughtful of the disease. Integrated multi-omics approaches, such as genomics, proteomics, and epigenomics, can uncover additional layers of complexity and potential therapeutic targets.

Finally, conducting clinical trials to estimate the safety and efficacy of identified compounds for GBM survival is crucial for their translation into clinical practice. Investigating the therapeutic potential of these compounds in combination with standard treatments can provide valuable insights into their clinical utility.

In conclusion, this chapter discussed the understanding and analysis of the outcomes obtained from the study. The verdicts have important implications for the

understanding and treatment of GBM, providing potential therapeutic targets, prognostic biomarkers, and natural compounds with therapeutic potential. However, it is important to acknowledge the drawbacks of the study and suggest additional research to validate and expand upon these findings, ultimately improving GBM treatment strategies.

REFERENCES

[1] G. F. Buchanan and G. B. Richerson, "A dietary supplement for SUDEP prevention?," *Nat Rev Neurol*, vol. 12, no. 9, pp. 495–496, Sep. 2016, doi: 10.1038/nrneurol.2016.114.

[2] A. Iasonos and J. O'Quigley, "Early phase clinical trials—are dose expansion cohorts needed?," *Nat Rev Clin Oncol*, vol. 12, no. 11, pp. 626–628, Nov. 2015, doi: 10.1038/nrclinonc.2015.174.

[3] N. McCarthy, "A growing concern," *Nat Rev Cancer*, vol. 14, no. 3, pp. 154–154, Mar. 2014, doi: 10.1038/nrc3689.

[4] S. Zheng, M. G. Chheda, and R. G. W. Verhaak, "Studying a Complex Tumor," *The Cancer Journal*, vol. 18, no. 1, pp. 107–114, Jan. 2012, doi: 10.1097/PPO.0b013e3182431c57.

[5] M. L. Bondy *et al.*, "Brain tumor epidemiology: Consensus from the Brain Tumor Epidemiology Consortium," *Cancer*, vol. 113, no. S7, pp. 1953–1968, Oct. 2008, doi: 10.1002/cncr.23741.

[6] A. Kwiatkowska, M. Nandhu, P. Behera, E. Chiocca, and M. Viapiano, "Strategies in Gene Therapy for Glioblastoma," *Cancers (Basel)*, vol. 5, no. 4, pp. 1271–1305, Oct. 2013, doi: 10.3390/cancers5041271.

[7] M. Touat, A. Idbaih, M. Sanson, and K. L. Ligon, "Glioblastoma targeted therapy: updated approaches from recent biological insights," *Annals of Oncology*, vol. 28, no. 7, pp. 1457–1472, Jul. 2017, doi: 10.1093/annonc/mdx106.

[8] A. Narayana *et al.*, "A clinical trial of bevacizumab, temozolomide, and radiation for newly diagnosed glioblastoma," *J Neurosurg*, vol. 116, no. 2, pp. 341–345, Feb. 2012, doi: 10.3171/2011.9.JNS11656.

[9] Y. Kong *et al.*, "Identification of Immune-Related Genes Contributing to the Development of Glioblastoma Using Weighted Gene Co-expression Network Analysis," *Front Immunol*, vol. 11, Jul. 2020, doi: 10.3389/fimmu.2020.01281.

[10] N. A. Butowski, "Epidemiology and Diagnosis of Brain Tumors," *CONTINUUM: Lifelong Learning in Neurology*, vol. 21, pp. 301–313, Apr. 2015, doi: 10.1212/01.CON.0000464171.50638.fa.

[11] D. Matias *et al.*, "Dual treatment with shikonin and temozolomide reduces glioblastoma tumor growth, migration and glial-to-mesenchymal transition," *Cellular Oncology*, vol. 40, no. 3, pp. 247–261, Jun. 2017, doi: 10.1007/s13402-017-0320-1.

[12] J. P. Fisher and D. C. Adamson, "Current FDA-Approved Therapies for High-Grade Malignant Gliomas," *Biomedicines*, vol. 9, no. 3, p. 324, Mar. 2021, doi: 10.3390/biomedicines9030324.

[13] Z. Z. Choudhury, U. N. Chowdhury, S. Ahmad, M. B. Islam, J. M. W. Quinn, and M. A. Moni, "Machine Learning and Bioinformatics Models to Identify Gene Expression Patterns of Glioblastoma Associated with Disease Progression and Mortality," in *2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*, IEEE, Dec. 2021, pp. 1–6. doi: 10.1109/IC4ME253898.2021.9768525.

[14] Y. Zhang *et al.*, "Bioinformatics analysis of potential core genes for glioblastoma," *Biosci Rep*, vol. 40, no. 7, Jul. 2020, doi: 10.1042/BSR20201625.

[15] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: new perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Res*, vol. 45, no. D1, pp. D353–D361, Jan. 2017, doi: 10.1093/nar/gkw1092.

[16] D. Szklarczyk *et al.*, "STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Res*, vol. 47, no. D1, pp. D607–D613, Jan. 2019, doi: 10.1093/nar/gky1131.

[17] M.-E. Halatsch, U. Schmidt, J. Behnke-Mursch, A. Unterberg, and C. R. Wirtz, "Epidermal growth factor receptor inhibition for the treatment of glioblastoma multiforme and other malignant brain tumours," *Cancer Treat Rev*, vol. 32, no. 2, pp. 74–89, Apr. 2006, doi: 10.1016/j.ctrv.2006.01.003.

[18] Q.-W. Fan and W. A. Weiss, "Targeting the RTK-PI3K-mTOR Axis in Malignant Glioma: Overcoming Resistance," 2010, pp. 279–296. doi: 10.1007/82_2010_67.

[19] L. M. Ellis and D. J. Hicklin, "Pathways Mediating Resistance to Vascular Endothelial Growth Factor–Targeted Therapy," *Clinical Cancer Research*, vol. 14, no. 20, pp. 6371–6375, Oct. 2008, doi: 10.1158/1078-0432.CCR-07-5287.

[20] H. Y. Chan, J. Choi, C. Jackson, and M. Lim, "Combination immunotherapy strategies for glioblastoma," *J Neurooncol*, vol. 151, no. 3, pp. 375–391, Feb. 2021, doi: 10.1007/s11060-020-03481-0.

[21] A. Rodriguez, C. Brown, and B. Badie, "Chimeric antigen receptor T-cell therapy for glioblastoma," *Translational Research*, vol. 187, pp. 93–102, Sep. 2017, doi: 10.1016/j.trsl.2017.07.003.

[22] M. Touat, A. Idbaih, M. Sanson, and K. L. Ligon, "Glioblastoma targeted therapy: updated approaches from recent biological insights," *Annals of Oncology*, vol. 28, no. 7, pp. 1457–1472, Jul. 2017, doi: 10.1093/annonc/mdx106.

[23] M. A. Qazi *et al.*, "Intratumoral heterogeneity: pathways to treatment resistance and relapse in human glioblastoma," *Annals of Oncology*, vol. 28, no. 7, pp. 1448–1456, Jul. 2017, doi: 10.1093/annonc/mdx169.

[24] T. I. Janjua *et al.*, "Frontiers in the treatment of glioblastoma: Past, present and emerging," *Adv Drug Deliv Rev*, vol. 171, pp. 108–138, Apr. 2021, doi: 10.1016/j.addr.2021.01.012.

[25] Z. Deng, Z. Sheng, and F. Yan, "Ultrasound-Induced Blood-Brain-Barrier Opening Enhances Anticancer Efficacy in the Treatment of Glioblastoma: Current Status and Future Prospects," *J Oncol*, vol. 2019, pp. 1–9, Nov. 2019, doi: 10.1155/2019/2345203.

[26] A. Shergalis, A. Bankhead, U. Luesakul, N. Muangsin, and N. Neamati, "Current Challenges and Opportunities in Treating Glioblastoma," *Pharmacol Rev*, vol. 70, no. 3, pp. 412–445, Jul. 2018, doi: 10.1124/pr.117.014944.

[27] F. Guo *et al.*, "Screening and Identification of Key Biomarkers in Lower Grade Glioma via Bioinformatical Analysis," *Appl Bionics Biomech*, vol. 2022, pp. 1–12, Jan. 2022, doi: 10.1155/2022/6959237.

[28] C. Song and G. C. Tseng, "Hypothesis setting and order statistic for robust genomic meta-analysis," *Ann Appl Stat*, vol. 8, no. 2, Jun. 2014, doi: 10.1214/13-AOAS683.

[29] M. V. Kuleshov *et al.*, "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update," *Nucleic Acids Res*, vol. 44, no. W1, pp. W90–W97, Jul. 2016, doi: 10.1093/nar/gkw377.

[30] L. Zhou *et al.*, "Bioinformatics analyses of significant genes, related pathways and candidate prognostic biomarkers in glioblastoma," *Mol Med Rep*, Aug. 2018, doi: 10.3892/mmr.2018.9411.
[31] F. Cao, Y. Fan, Y. Yu, G. Yang, and H. Zhong, "Dissecting Prognosis Modules and Biomarkers in Glioblastoma Based on Weighted Gene Co-Expression Network Analysis," *Cancer Manag Res*, vol. Volume 13, pp. 5477–5489, Jul. 2021, doi: 10.2147/CMAR.S310346.
[32] M. Sorokina, P. Merseburger, K. Rajan, M. A. Yirik, and C. Steinbeck, "COCONUT online: Collection of Open Natural Products database," *J Cheminform*, vol. 13, no. 1, p. 2, Dec. 2021, doi: 10.1186/s13321-020-00478-9.
[33] D. Schraivogel *et al.*, "CAMTA1 is a novel tumour suppressor regulated by miR-9/9 * in glioblastoma stem cells," *EMBO J*, vol. 30, no. 20, pp. 4309–4322, Oct. 2011, doi: 10.1038/emboj.2011.301.
[34] A. E. Lohning, S. M. Levonis, B. Williams-Noonan, and S. S. Schweiker, "A Practical Guide to Molecular Docking and Homology Modelling for Medicinal Chemists," *Curr Top Med Chem*, vol. 17, no. 18, Jun. 2017, doi: 10.2174/1568026617666170130110827.
[35] Y. Liu, M. Grimm, W. Dai, M. Hou, Z.-X. Xiao, and Y. Cao, "CB-Dock: a web server for cavity detection-guided protein–ligand blind docking," *Acta Pharmacol Sin*, vol. 41, no. 1, pp. 138–144, Jan. 2020, doi: 10.1038/s41401-019-0228-6.
[36] A. Daina, O. Michielin, and V. Zoete, "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," *Sci Rep*, vol. 7, no. 1, p. 42717, Mar. 2017, doi: 10.1038/srep42717.
[37] A. Grosdidier, V. Zoete, and O. Michielin, "SwissDock, a protein-small molecule docking web service based on EADock DSS," *Nucleic Acids Res*, vol. 39, no. suppl, pp. W270–W277, Jul. 2011, doi: 10.1093/nar/gkr366.
[38] G. Bitencourt-Ferreira and W. F. de Azevedo, "Docking with SwissDock," 2019, pp. 189–202. doi: 10.1007/978-1-4939-9752-7_12.

turnitin

Similarity Report ID: oid:27535:36440980

● 8% Overall Similarity

Top sources found in the following databases:

- 6% Internet database
- Crossref database
- 4% Submitted Works database
- 3% Publications database
- Crossref Posted Content database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	dspace.dtu.ac.in:8080 Internet	3%
2	Sanya Madaan, Pravir Kumar. "Stilbostemin C as a Potential Candidate ... Crossref	<1%
3	res.mdpi.com Internet	<1%
4	ir.uitm.edu.my Internet	<1%
5	journal.waocp.org Internet	<1%
6	Qi-Lin Zhang, Han-Yue Qiu, Ming-Zhong Liang, Bang Luo, Xiu-Qiang Wa... Crossref	<1%
7	researchsquare.com Internet	<1%
8	Queen Mary and Westfield College on 2012-03-26 Submitted works	<1%

Sources overview

9	Farzana Kabir, Mohd Nazmul Hasan Apu. "Multi-omics analysis predict... Crossref	<1%
10	UC, San Diego on 2022-02-14 Submitted works	<1%
11	aei.pitt.edu Internet	<1%
12	inpuma.um.edu.my Internet	<1%
13	pubmed.ncbi.nlm.nih.gov Internet	<1%
14	Kumaraswamy Naidu Chitrala, Prakash Nagarkatti, Mitzi Nagarkatti. "P... Crossref	<1%
15	The Hong Kong Polytechnic University on 2006-04-20 Submitted works	<1%
16	Anjali Sinha, Bharmjeet, Asmita Das. "An In-Silico Comparison Of Gins... Crossref	<1%
17	Feng Zhong, Hui-Ping Lu, Gang Chen, Yi-Wu Dang et al. "The clinical si... Crossref	<1%
18	Imperial College of Science, Technology and Medicine on 2017-09-04 Submitted works	<1%
19	Zakia Zinat Choudhury, Utpala Nanda Chowdhury, Shamim Ahmad, M. ... Crossref	<1%
20	bio-protocol.org Internet	<1%

Sources overview

Pecilocin As Potential Drug Candidate in GBM Suppression: Bioinformatic Approach to Identify Gene Expression Signature Using ADMET Analysis and Molecular Docking Studies

Shallu Chauhan
Molecular Neuroscience and Functional Genomics Laboratory,
Dept. of Biotechnology
Delhi Technological University
Delhi – 110042, India
shallu22chauhan@gmail.com

Prof. Pravir Kumar*
Molecular Neuroscience and Functional Genomics Laboratory,
Dept. of Biotechnology
Delhi Technological University
Delhi – 110042, India
pravirkumar@dtu.ac.in

***Abstract*— Despite significant medical improvements, the treatment of the most common and dangerous brain tumor, glioblastoma multiforme (GBM), remains a difficult challenge. Genetic mutation is the major contributing factor to this lethal condition**[1]**. It is crucial to recognize the genetic markers responsible for GBM survival. The GEO dataset provides publicly available gene expression and clinical data on GBM patients to study the prognosis of patient death rate and thereby identify GBM biomarkers.**

We performed functional enrichment analysis and protein-protein interaction (PPI). In this study, we used the GDC data portal to analyze linked genes with clinical data to determine the most significant genes of GBM. To distinguish the genes with different expression levels between the mutant and normal groups, we identified differential expression genes (DEGs) and evaluated their survival rate in patients. Furthermore, we used the GDC Data portal to estimate their significance in GBM survival.

A natural substance called Pecilocin might be evaluated as a glioblastoma therapy option. For a better understanding of the interactions between the bioactive substances and CAMTA1 protein, molecular docking research was performed. According to the molecular docking analysis, Pecilocin has a stronger affinity for the CAMTA1. Pecilocin's potential for further research was also revealed by the ADME/T study.

Keywords*— *Glioblastoma, Datasets, Differential expression genes, Biomarkers, GBM survival CAMTA1, Pecilocin,

Introduction

GBM has been considered to originate in the central nervous system (CNS) glial cells[2]. However, recent research suggests that GBM is caused by the neural stem in the brain's subventricular zone[3]. The standard-of-care treatment consists of surgical deletion of the tumor, along with radiation, and temozolomide chemotherapy[4], [5]; but, as the low survival rate indicates, these therapies have not been helpful in preventing tumor progression.

The most serious and frequent form of primary astrocytoma is glioblastoma multiforme (GBM). Since genetic mutation is the primary cause of this fatal condition, it is crucial to determine the genetic markers involved in GBM survival.

The Gene expression omnibus (GEO) dataset provides sufficient publicly available gene expression and clinical data on GBM patients to study the prognosis of patient death rate and thereby identify GBM biomarkers. In this study, bioinformatics tools are used to identify the mutant genes associated with the GBM compared with normal mRNA expression data from the brain tissues[6]. We used datasets GSE4290 and GSE12657 for the relative study of gene expression data between Glioblastoma tumors and normal brain tissues. Gene set functional enrichment analysis, gene ontology (GO), KEGG pathways, and protein-protein interaction (PPI), also revealed their significance[7]. We identified seven genes (GRIN2A, BCL11A, CAMTA1, ERBB3, WIF1, HLF, and CHN1) by using the GDC Data portal which has a substantial effect on GBM development. We performed Molecular docking to check gene and drug interaction. ADME/T analysis showed water-solubility, lipophilicity, pharmacokinetics, and drug-likeness parameters of the potential drug.

Method and Material

Data Set Collection

The microarray gene expression datasets for Glioblastoma used in this study were collected from the National Center for Biotechnology Information Gene Expression Omnibus (NCBI-GEO) (https://www.ncbi.nlm.nih.gov/gds). The datasets are GSE4290 and GSE12657. GSE4290 and GSE12657 datasets are used for the relative study of gene expression data between Glioblastoma tumors and normal brain tissues.

Identification of Differential Expression Genes (DEGs)

GEO2R web tool was used to analyze and identify differentially expressed genes for GBM and normal brain samples were compared with GBM patients[8]

(https://www.ncbi.nlm.nih.gov/geo/geo2r/). The characteristics of choosing DEGs were |log(foldchange)| > 2 and P < 0.05 [9]. The overlapping portions between two dataset genes were identified using the Venn web tool.

Functional Enrichment Analysis of DEGs

All of the DEGs that were identified by the microarray GBM datasets were subjected to the KEGG pathway[10] and GO analysis using the gene set enrichment analysis (GSEA) tool EnrichR (https://maayanlab.cloud/Enrichr/). The biological process Gene Ontology (GO) was employed as an annotation source in this research, and its limit was set at P 0.05.

Protein-Protein Interactions (PPI) Network Analysis

Using the web-based STRING database, a protein-protein interactions (PPI) network of the 285 overlapped DEGs was created (https://string-db.org/). A database that offers both physical and functional protein associations is called the PPI network[11]. The protein interaction network of differentially expressed genes, which can be used to assess potential protein interactions, was obtained for the 285 DEGs that were identified in the study.

Data Collection for Molecular Docking

COCONUT (COlleCtion of Open Natural ProdUcTs) is a web-based open-source project for Natural Products storage, search, and analysis (https://coconut.naturalproducts.net/). For the identification of natural products as a drug for Glioblastoma, 50 natural compounds showing antipsychotic properties and one already known natural compound Chrysin[12] were taken as standard for reference. PubChem was employed to download the SDF structures of these natural compounds. Out of the seven most frequently mutated genes, CAMTA1 gene was selected for molecular docking and structure downloaded from PDB. CAMTA1 is an antitumor gene that regulates proliferation and the cell cycle in glioma cells by inhibiting AKT phosphorylation[13]. CAMTA1 was prepared for carrying out docking using Discovery Studio. 50 ligands were converted into mol2 format using Discovery studio. After the ligands and protein preparation, docking was performed using SwissDock (http://www.swissdock.ch/docking).

ADMET Analysis

SwissADME which is a web-based tool allowed us to predict ADME parameters for drug discovery (http://www.swissadme.ch/). The Canonical SMILES were taken from PubChem and the virtual physicochemical properties of all fifty ligands are analyzed using the SwissADME tools, including water-solubility, lipophilicity, pharmacokinetics, and druglikeness parameters[14].

Results

Differentially Expressed Genes (DEGs) Identification

The two selected datasets were analyzed by comparing normal samples with GBM cases. We identified 285 overlapping DEGs among two datasets (Fig.1.). We evaluated the common DEGs and found 285 overlapping genes.

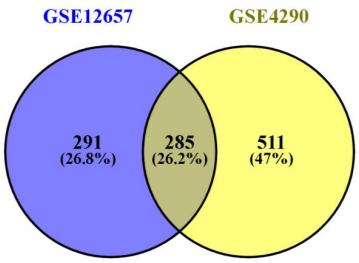

Fig. 1. The Venn diagram of 285 overlapping up-regulated genes identified from two datasets GSE4290 and GSE12657 of GBM

Functional Enrichment Analysis

The Functional enrichment study of the GO Biological process (BP) and KEGG pathway was carried out to illustrate the functions and molecular pathways associated with GBM using 285 overlapped DEGs. TABLE 1. shows KEGG pathways associated with GBM and TABLE II. shows GO Biological process associated with Glioblastoma.

Analysis of the PPI network

The PPI network for the 285 overlapping DEGs was constructed using the String tool. The network where 280 nodes each represent a protein and 2120 edges each indicating an interaction between each node pair (Fig.2.).

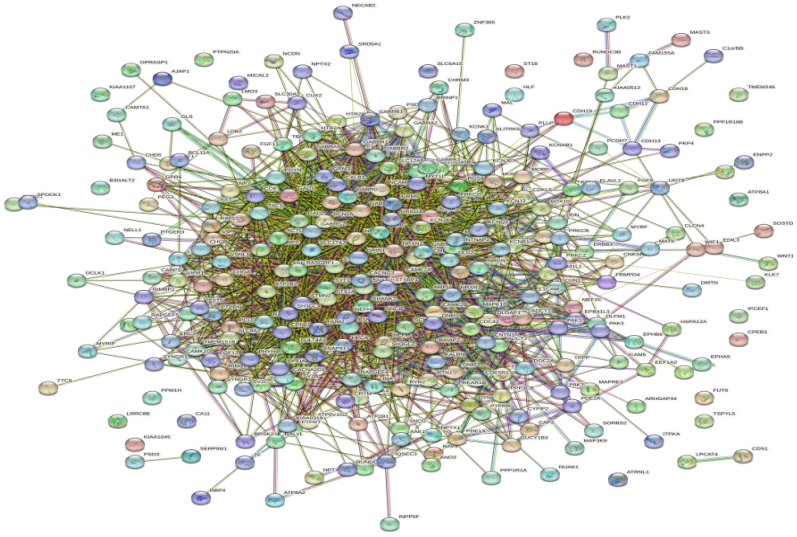

Fig. 2. The Protein-protein interaction network of 285 genes which were common in two datasets

TABLE III. KEGG PATHWAY HUMAN ASSOCIATED WITH GLIOBLASTOMA

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	GABAergic synapse	1.002e-18	2.044e-16	21.49	890.59
2	Morphine addiction	3.313e-17	3.379e-15	19.49	739.46
3	Nicotine addiction	1.714e-10	4.371e-9	23.86	536.55
4	Insulin secretion	1.235e-12	4.200e-11	15.37	421.47
5	Glutamatergic synapse	5.500e-13	2.244e-11	12.83	362.15
6	Retrograde endocannabinoid signaling	3.730e-13	1.902e-11	10.84	310.35
7	Calcium signaling pathway	8.986e-15	6.110e-13	8.72	282.06
8	Long-term potentiation	2.570e-9	4.370e-8	14.09	278.76
9	Circadian entrainment	1.218e-9	2.258e-8	11.17	229.27
10	Oxytocin signaling pathway	6.958e-10	1.420e-8	8.44	177.92

TABLE IV. GO BIOLOGICAL FUNCTION ASSOCIATED WITH GLIOBLASTOMA

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	presynaptic active zone organization (GO:1990709)	5.922e-7	0.00002584	140.31	2011.92
2	anterograde trans-synaptic signaling (GO:0098916)	3.452e-33	2.861e-30	16.70	1247.98
3	chemical synaptic transmission (GO:0007268)	6.222e-36	1.032e-32	15.27	1238.18
4	protein localization to axon (GO:0099612)	0.00002803	0.0007151	104.86	1099.12
5	inhibitory synapse assembly (GO:1904862)	2.443e-7	0.00001366	58.66	893.05
6	positive regulation of non-membrane spanning protein tyrosine kinase activity (GO:1903997)	0.00005547	0.001332	69.90	685.00
7	Neurotransmitter secretion (GO:0007269)	2.050e-11	4.856e-9	23.94	589.27
8	regulation of neurotransmitter receptor activity (GO:0099601)	1.353e-12	5.606e-10	20.89	570.94
9	regulation of postsynaptic neurotransmitter receptor activity (GO:0098962)	0.000007924	0.0002628	46.76	549.21
10	regulation of nervous system process (GO:0031644)	1.844e-7	0.00001132	32.59	505.38

Identification of DEGs survival

For the survival analysis, we considered the overlapping 285 DEGs among the GBM datasets. Using GDC Data Portal, seven genes (GRIN2A, BCL11A, CAMTA1, ERBB3, WIF1, HLF, and CHN1) were found significant in the patients' survival having $P < 0.05$. The survival patterns of the mutated and normal groups of these seven genes are shown in Fig.3.

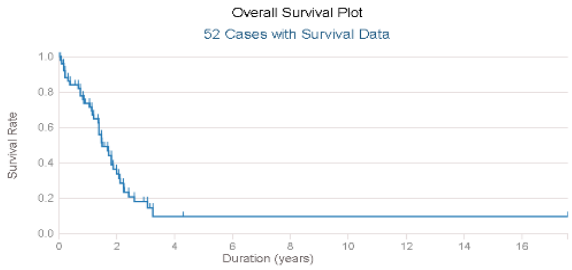

Fig. 3. Overall survival plot of patients having mutated genes

Molecular Docking Result

SwissDock was used in studying the natural compounds with interactions of glioblastoma protein CAMTA1. Out of 49 natural compounds docked, 9 compounds met the required parameters binding energy less than -7.13 Kcal/mol (energy of standard compound) as shown in TABLE III. Out of these ten, Pecilocin has the most negative binding energy which suggests it has the best binding affinity with the protein (Fig.4.).

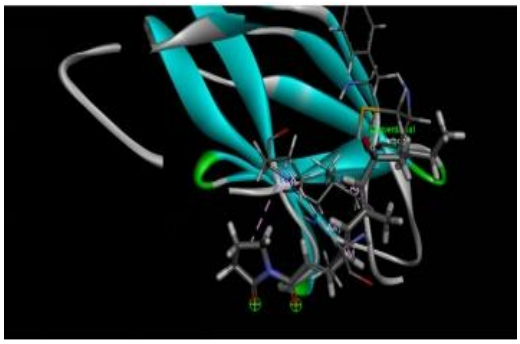

Fig. 4. Pecilocin binds at CAMTA1 active site

ADMET Analysis

ADMET analysis was performed using SwissADME and Pecilocin was found to have high GI absorption capacity, and BBB permeant. Lipinski rule is one of the most popular parameters used to evaluate the drug-likeness of small molecules. With a bioavailability score of 0.55 and 0 violations, Pecilocin satisfies with Lipinski rule's drug-likeness standards for bioavailable drugs.

Discussion

In recent decades, significant efforts have been undertaken to improve the health efficiency of GBM patients. In this work, we discovered 285 DEGs to be overlapped in two datasets. To determine the molecular function of the typical DEGs, we conducted GO biological process and KEGG pathway analysis. Using the overlapping DEGs found, a topological analysis of the PPI network was performed to evaluate probable protein interactions. Using GDC Data Portal, seven genes (GRIN2A, BCL11A, CAMTA1, ERBB3, WIF1, HLF, and CHN1) were found significant in the patient's survival. The standard-of-care treatment consists of surgical deletion of the tumor, along with radiation, and temozolomide chemotherapy[15]. In recent years, natural compounds

TABLE V. MOLECULAR DOCKING AND ADME/T ANALYSIS RESULT

Name	Cluster	Element	FullFitness (kcal/mol)	Estimated ΔG (kcal/mol)	GI absorption	Lipinski	BBB permeant
Chrysin (Standard)	1	0	-1622.82	-7.13	High	Yes; 0 violation	Yes
Supral/ Pecilocin	1	5	-1672.16	-8.37	High	Yes; 0 violation	Yes
Nigerapyrone E	9	8	-1617.32	-7.23	High	Yes; 0 violation	Yes
Rankinidine	7	0	-1576.32	-7.28	High	Yes; 0 violation	Yes
Kuguacin M	0	0	-1605.56	-7.56	High	Yes; 0 violation	Yes
Anthrakunthone	1	4	-1603.09	-7.21	High	Yes; 0 violation	Yes
Laurequinone	1	1	-1385.28	-7.1	High	Yes; 0 violation	Yes
Atractylenolide I	1	1	-1635.19	-7.31	High	Yes; 0 violation	Yes
Melophlin D	0	1	-1667.3	-8.32	High	Yes; 0 violation	Yes
Glaucocalyxin A	0	0	-1596.91	-7.51	High	Yes; 0 violation	yes

have been used as a drug using in silico techniques such as molecular docking which helped to facilitate the drug development procedure. Using this technique, we have predicted Pecilocin as a potent inhibitor of CAMTA1 gene in Glioblastoma. It has shown promising effects as a better inhibitor with binding energy -8.37 in comparison to other inhibitors such as Chrysin with the binding energy of -7.13 and found to be BBB permeant and have good drug-likeness characteristics.

Conclusion & Future Scope

In this study, survival analysis of 20 GBM patients strengthened that most of the considered 285 common DEGs are only associated with GBM. Out of 285, seven genes which are GRIN2A, BCL11A, CAMTA1, ERBB3, WIF1, HLF, and CHN1, were found to be associated with GBM survival. Final survival analysis revealed that CAMTA1 mutation was associated with prognosis in GBM patients. Overall, CAMTA1 could be a suitable prognostic marker for GBM. Molecular docking of CAMTA1 genes with 50 natural compounds revealed that Pecilocin could be the potent drug for GBM. ADMET analysis of Pecilocin depicts BBB permeant, following the Lipinski rule, with 0 violations, and has better drug-likeness. The actual wet-lab work will support in choosing the Pecilocin molecule as a potential GBM target. We can also use seven genes for clinical therapies and for precision medications which will be beneficial for GBM survival.

Acknowledgment

We express our gratitude to the senior administration of Delhi Technological University for their continuous support and assistance.

References

[1] A. H. Thorne, C. Zanca, and F. Furnari, "Epidermal growth factor receptor targeting and challenges in glioblastoma," *Neuro Oncol*, vol. 18, no. 7, 2016, doi: 10.1093/neuonc/nov319.

[2] F. He and Y. E. Sun, "Glial cells more than support cells?," *International Journal of Biochemistry and Cell Biology*, vol. 39, no. 4. 2007. doi: 10.1016/j.biocel.2006.10.022.

[3] C. Neftel *et al.*, "An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma," *Cell*, vol. 178, no. 4, 2019, doi: 10.1016/j.cell.2019.06.024.

[4] D. Matias *et al.*, "Dual treatment with shikonin and temozolomide reduces glioblastoma tumor growth, migration and glial-to-mesenchymal transition," *Cellular Oncology*, vol. 40, no. 3, 2017, doi: 10.1007/s13402-017-0320-1.

[5] J. P. Fisher and D. C. Adamson, "Current FDA-Approved Therapies for High-Grade Malignant Gliomas.," *Biomedicines*, vol. 9, no. 3, Mar. 2021, doi: 10.3390/biomedicines9030324.

[6] Z. Z. Choudhury, U. N. Chowdhury, S. Ahmad, M. B. Islam, J. M. W. Quinn, and M. A. Moni, "Machine Learning and Bioinformatics Models to Identify Gene Expression Patterns of Glioblastoma Associated with Disease Progression and Mortality," in *2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*, Dec. 2021, pp. 1–6. doi: 10.1109/IC4ME253898.2021.9768525.

[7] Y. Zhang *et al.*, "Bioinformatics analysis of potential core genes for glioblastoma," *Biosci Rep*, vol. 40, no. 7, 2020, doi: 10.1042/BSR20201625.

[8] H. H. F. Loong *et al.*, "Patient-derived tumor organoid predicts drugs response in glioblastoma: A step forward in personalized cancer therapy?," *Journal of Clinical Neuroscience*, vol. 78, 2020, doi: 10.1016/j.jocn.2020.04.107.

[9] M. E. Ritchie *et al.*, "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Res*, vol. 43, no. 7, 2015, doi: 10.1093/nar/gkv007.

[10] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: New perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Res*, vol. 45, no. D1, 2017, doi: 10.1093/nar/gkw1092.

[11] D. Szklarczyk *et al.*, "STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Res*, vol. 47, no. D1, 2019, doi: 10.1093/nar/gky1131.

[12] J. Wang *et al.*, "Chrysin suppresses proliferation, migration, and invasion in glioblastoma cell lines via mediating the ERK/Nrf2 signaling pathway," *Drug Des Devel Ther*, vol. 12, 2018, doi: 10.2147/DDDT.S160020.

[13] Z. He *et al.*, "CAMTA1, a novel antitumor gene, regulates proliferation and the cell cycle in glioma by inhibiting AKT phosphorylation.," *Cell Signal*, vol. 79, p. 109882, Mar. 2021, doi: 10.1016/j.cellsig.2020.109882.

[14] A. Daina, O. Michielin, and V. Zoete, "SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," *Sci Rep*, vol. 7, 2017, doi: 10.1038/srep42717.

[15] V. A. Cruz Flores *et al.*, "Combined Therapy With Avastin, a PAF Receptor Antagonist and a Lipid Mediator Inhibited Glioblastoma Tumor Growth," *Front Pharmacol*, vol. 12, Sep. 2021, doi: 10.3389/fphar.2021.746470.