

**COMPUTATIONAL IDENTIFICATION OF BIOMARKER AND
THERAPEUTIC TARGETS IN NEURODEGENERATIVE AND BRAIN
DISORDERS**

A Dissertation

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF TECHNOLOGY

in

BIOTECHNOLOGY

by

Smita Yadav

(24/IBT/13)

Under the supervision of

Prof. Pravir Kumar



**Department of Biotechnology
DELHI TECHNOLOGICAL UNIVERSITY**

(Formerly Delhi College of Engineering)

Shahbad Daultapur, Bawana Road, Delhi -110042, INDIA

May, 2026

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all those who have contributed to the completion of this thesis. Firstly, I would like to thank my supervisor, Prof. Pravir Kumar, for his guidance, support, and invaluable feedback throughout this research journey. His expertise and encouragement have been instrumental in shaping this work. I am deeply grateful to him for having faith in my abilities since day one of my M.Tech journey. I express my heartiest and most sincere thanks to all the PhD scholars, Neetu ma'am, Shrutikirti ma'am, Aastha ma'am, Apurva ma'am, Nishita ma'am, and my M. Tech fellow Sanskriti for giving the necessary pieces of advice and working without any stress. They took time out from their research work and helped me learn new things. I sincerely thank all of them for treating me like their sibling and helping me learn even the complicated things. I will always remember their words of motivation throughout my life. In the end, I would like to thank my mother and sisters, who stood as a pillar behind me and constantly motivated me. I perceive this project opportunity as a big milestone in my career development. Additionally, I would like to thank my teachers and colleagues for their encouragement, support, and meaningful discussions that have contributed to the development of my ideas. Finally, I am thankful to the Department of Biotechnology for providing me with the resources and opportunities necessary for conducting this research. Thank you to everyone who has been a part of this journey and has helped me reach this milestone.

Smita Yadav
24/IBT/13



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daultapur, Main Bawana Road, Delhi-42

DECLARATION

I Smita Yadav 24/IBT/13 hereby certify that the work which is being presented in the thesis entitled "Computational Identification of biomarker and therapeutic targets in neurodegenerative and brain disorder " in partial fulfillment of the requirements for the award of the Degree of Master of Technology, submitted in the Department of Biotechnology, Delhi Technological University is an authentic record of my own work carried out during the period from 2024 to 2026 under the supervision of Prof. Pravir Kumar.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **Smita Yadav** (24/IBT/13) has carried out her research work presented in this thesis entitled "**Computational Identification of biomarker and therapeutic targets in neurodegenerative and brain disorder**" for the award of Master of Technology from the Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself, and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Date:

Prof. Pravir Kumar

Supervisor

Department of Biotechnology

Delhi Technological University

Prof. Yasha Hasija

Head of Department

Department of Biotechnology

Delhi Technological University

COMPUTATIONAL IDENTIFICATION OF BIOMARKER AND THERAPEUTIC TARGETS IN NEURODEGENERATIVE AND BRAIN DISORDERS

Smita Yadav
ABSTRACT

Aim: The objective of the current study was to identify candidate sets of differentially expressed and hub genes related to Multiple Sclerosis (MS) using bioinformatics methods, and to identify potential phytochemical inhibitors targeting the BUB1B protein in Glioblastoma (GBM) through molecular docking analysis.

Background: Neurological diseases like Multiple Sclerosis (MS) and Glioblastoma (GBM) are highly complex diseases, characterised by impairment of the nervous system and few therapeutic choices. Multiple Sclerosis involves immune-mediated demyelination and degeneration of the nervous system, while Glioblastoma is a rapidly growing, aggressive malignant brain tumour that is resistant to treatment. The identification of molecular Therapeutic targets and biomarkers is key to the advancement of diagnosis and therapeutic approaches to the disease. Increased knowledge about disease mechanisms and potential therapeutics using recent bioinformatics and computational drug discovery advances. Effective tools for understanding disease mechanisms and identifying potential therapeutics using bioinformatics and computational advances in drug discovery.

Methodology: Gene expression data (GSE21942) were downloaded from the Gene Expression Omnibus (GEO) database and analysed with GEO2R to generate a list of differentially expressed genes. The STRING database was used to conduct functional enrichment analysis, Gene Ontology (GO) pathway analysis, KEGG pathway analysis and protein–protein interaction (PPI) network construction. Degree centrality analysis was used to determine hub genes. For molecular docking studies, the three-dimensional structure of the BUB1B protein was obtained from the Data Bank on Proteins. Phytochemicals were selected and screened using EasyDock Vina. Protein–ligand interaction analysis was performed using LigPlot+, and the structural visualisation was performed using PyMOL.

Results: A total of 1276 genes were identified as being DE (697 up-regulated and 579 down-regulated). Multiple Sclerosis was found to be regulated by hub genes: RPS6, RPS28, RPS18, CD79A, and CD79B. Functional enrichment analysis showed ribosomal activity, protein translation, inflammatory pathways, and immune response were enriched pathways. Based on the molecular docking study, α -peltatin had the highest binding affinity towards the BUB1 B protein with a docking score of — 10.6 kcal/mol, indicating that it has therapeutic potential against Glioblastoma

LIST OF PUBLICATION

1. Conference paper:

Smita Yadav and P. Kumar, “Computational Identification of BUB1B-Targeting Phytochemicals for GBM Therapy Using Molecular Docking,” *2026 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2026, IEEE, doi: 10.1109/IATMSI68868.2026.11465606.

Keywords: Glioblastoma (GBM); BUB1B; Molecular Docking; Phytochemicals; EasyDock Vina; LigPlot+; PyMOL; Binding Affinity; Protein–Ligand Interaction; Computational Biology; Structure-Based Drug Discovery; Bioinformatics; Therapeutic Targeting; Cancer Therapy.

2. Poster:

Smita Yadav¹, Pravir Kumar¹, “Identification of Novel ALDH1A3 Inhibitors in Glioblastoma Using Molecular Docking Machine Learning Based Virtual Screening”

Presented at: SNCI, Jamia Hamdard, New Delhi

TABLE OF CONTENTS

Acknowledgement	2
Candidate's Declaration	3
Certificate	4
Abstract	5
List of Publications	6
Table of Contents	7-8
List of Tables, Figures and Abbreviations	9-10
CHAPTER-1 INTRODUCTION	11-14
CHAPTER-2 LITERATURE REVIEW	15
2.1 Multiple Sclerosis and Bioinformatics Approaches	15
2.2 Hub Gene Identification and Protein–Protein Interaction Networks	15
2.3 Functional Enrichment and Pathway Analysis	16
2.4 Glioblastoma and BUB1B as a Therapeutic Target	16
2.5 Molecular Docking and Phytochemical-Based Drug Discovery	17
2.6 Integration of Bioinformatics and Drug Discovery Approaches	18
2.7 Study Rationale and Significance	18
CHAPTER-3 METHODOLOGY	19
3.1 Research Methodology Overview	19
3.2 Bioinformatics Analysis for Multiple Sclerosis	19
• 3.2.1 Dataset Selection	19
• 3.2.2 Differential Gene Expression Analysis	19
• 3.2.3 Data Normalisation	20
• 3.2.4 Building Protein–Protein Interaction (PPI) Networks	20
• 3.2.5 Analysis of Functional Enrichment	20
• 3.2.6 Hub Gene Identification	20
3.3 Molecular Docking for Glioblastoma	20
• 3.3.1 Preparing the Target Protein	20

• 3.3.2 Ligand Selection and Preparation	21
• 3.3.3 Molecular Docking Procedure	21
• 3.3.4 Selection of Top Candidate Compounds	21
• 3.3.5 Protein–Ligand Interaction Analysis	22
• 3.3.6 Structural Visualisation	22
3.4 Integrated Computational Framework	22
CHAPTER-4 RESULTS	24
4.1 Determination of Genes with Differential Expression	24
4.2 Data Normalization	24
4.3 Functional Enrichment Analysis	25
4.4 Hub Gene Identification and the PPI Network	28
4.5 Molecular Docking Results	30
4.6 Binding Affinity Analysis	30
4.7 Protein–Ligand Interaction Analysis	32
4.8 Structural Visualisation	33
CHAPTER 5: DISCUSSION AND CONCLUSION	35-37
REFERENCES	38-42
List of Publications and Certificates	43-47
Plagiarism Report	48
Curriculum Vitae	49

LIST OF FIGURES

S.No.	Fig.no	Name of the figure	Page No.
1	Fig 3.1	Integrated computational workflow for bioinformatics analysis and molecular docking	23
2	Fig 4.1	Volcano plot of differentially expressed genes in Multiple Sclerosis	24
3	Fig 4.2	Box plot showing normalisation of gene expression data	25
4	Fig 4.3a	GO biological process enrichment analysis of upregulated DEGs	26
5	Fig 4.3b	GO molecular function enrichment analysis of upregulated DEGs	26
6	Fig 4.3c	GO cellular component enrichment analysis of upregulated DEGs	26
7	Fig 4.3d	KEGG pathway enrichment analysis of upregulated DEGs	26
8	Fig 4.3e	Monarch phenotype enrichment analysis of upregulated DEGs	26
9	Fig 4.3f	GO biological process enrichment analysis of downregulated DEGs	26
10	Fig 4.3g	GO molecular function enrichment analysis of downregulated DEGs	27
11	Fig 4.3h	GO cellular component enrichment analysis of downregulated DEGs	27
12	Fig 4.3i	KEGG pathway enrichment analysis of downregulated DEGs	27
13	Fig 4.3j	Monarch phenotype enrichment analysis of downregulated DEGs	27
14	Fig 4.4a	Protein–protein interaction network of upregulated DEGs	29
15	Fig 4.4b	Protein–protein interaction network of downregulated DEGs	29
16	Fig 4.7	LigPlot+ interaction diagram of α -peltatin with BUB1B protein	32-33
17	Fig 4.8	Three-dimensional visualisation of α -peltatin docked with the BUB1 B protein	34

LIST OF TABLES

S. No	Name of Table	Page No.
1	Docking scores and key interacting residues of top-ranked phytochemicals with the BUB1 B protein	31

LIST OF ABBREVIATIONS

ADMET	Absorption, Distribution, Metabolism, Excretion and Toxicity
ATP	Adenosine Triphosphate
BBB	Blood–Brain Barrier
BCR	B-Cell Receptor
BP	Biological Process
BUB1B	BUB1 Mitotic Checkpoint Serine/Threonine Kinase B
CDK	Cyclin-Dependent Kinase
CNS	Central Nervous System
DEG	Differentially Expressed Gene
DNA	Deoxyribonucleic Acid
GBM	Glioblastoma
GEO	Gene Expression Omnibus
GO	Gene Ontology
H-bond	Hydrogen Bond
KEGG	Kyoto Encyclopedia of Genes and Genomes
kcal/mol	Kilocalories per Mole
log2FC	Log2 Fold Change
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
mRNA	Messenger Ribonucleic Acid
PDB	Protein Data Bank
PDBQT	Protein Data Bank, Partial Charge (Q), and Atom Type (T) Format
PPI	Protein–Protein Interaction
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
RMSD	Root Mean Square Deviation

CHAPTER 1

INTRODUCTION

Neurological disorders are a major problem on the planet, as they impact millions of people across the world and lead to disability in the long term, low quality of life, and mortality [11,12]. The disorders include a broad spectrum of central nervous system diseases such as neurodegenerative diseases, autoimmune diseases, and malignant diseases [40,41]. Two of them, Multiple Sclerosis (MS) and Glioblastoma (GBM), are imperative and widely researched conditions because of their complex molecular pathways, heterogeneous nature, and limited treatment efficacy [11,19]. Even with the innovations in medical science of the time, both diseases pose challenges for early identification and tracking, as well as for effective treatment; hence, the need to develop new methods to understand their underlying biology and potential treatment modalities [13,20]. MS is a chronic autoimmune disease that is defined by the inflammation, demyelination and progressive neurodegeneration of the central nervous system [11,15]. It is a leading cause of non-traumatic neurological disability in the world and is mostly seen in young adults [12]. The disease is characterised by a defective immune response where autoreactive T cells and B cells destroy the myelin sheath around neurons, resulting in a loss of nerve signal conduction and neuro- pathology [15]. There is a broad spectrum of clinical manifestations in MS patients, such as motor defects, sensory disorders, visual dysfunction, and cognitive issues [11]. This heterogeneity is challenging to diagnostic like magnetic resonance imaging (MRI) and cerebrospinal fluid examination frequently cannot reflect the entire range of molecular changes related to disease onset and progression [14]. Molecular mechanisms of MS are multifaceted interactions between immune dysregulation, inflammatory signalling, and neuronal damage [15,41]. The effects of these interactions include dramatic changes in gene expression patterns, which can be studied either by high-throughput transcriptomic methods, such as microarray and RNA sequencing [4,8]. The discovery of variously expressed genes (DEGs) in diseased and healthy states can give valuable information about the mechanism of diseases and aid in finding possible biomarkers for diagnosis and progression [3,49]. Larger-scale analysis of gene expression has been made easier through curated datasets in publicly available databases, including the Gene Expression Omnibus (GEO) [1,2]. Computational analysis tools such as GEO2R can help the researcher to compare statistics and detect DEGs effectively [3].

Nonetheless, gene expression analysis does not always give all the information about the disease biology. Genes do not act alone but rather in complex ways; the protein-protein interaction network analysis has become an effective method to learn more about the functional connections between proteins [16,42]. The topological properties of PPI networks can be used to identify hub genes that are core to disease-related pathways by building PPI networks [17]. These hub genes usually play a central role in regulating biological functions, and they may be potential biomarkers or therapeutic targets. Unless otherwise specified, the data presented in

this table represents a subset of functional enrichment data, e.g. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO), which are available for further analysis pathways, can be used to give an insight into the biological significance of the genes by providing links to particular cellular functions and pathways [7,8]. Unlike MS, Glioblastoma is one of the most severe and malignant types of primary tumours in the brain [19]. GBM is defined by the high invasiveness, the rapid cell growth, and the resistance to the traditional treatment regimens [21]. GBM patients have a median survival rate of about 12- 15 months, which demonstrates the necessity to develop better treatment potions [20]. Tumour heterogeneity is one of the biggest problems in GBM treatment, which leads to different cell populations in the tumour's microenvironment [22]. Moreover, the existence of cells leads to the recurrence and resistance of the tumor to therapy [23]. The molecular pathways that result in GBM are the deregulation of important signalling pathways and cell cycle controls, which facilitate unchecked cell division and tumor development [23]. One of these, BUB1B, which is a mitotic checkpoint kinase, has become a key protein important in ensuring the stability of the chromosome during cell division [24,25]. BUB1B overexpression has been linked to tumor growth, genomic instability and chemotherapy and radiotherapy resistance [26]. Because of its crucial contribution to the survival and progression of the tumors BUB1B has become an attractive therapeutic target in the treatment of GBM [25]. The ability to target these proteins could help to interfere with vital cellular functions in cancerous cells with minimal impact on normal tissues, thus enhancing treatment outcomes [45]. The screening of large libraries of chemical compounds is time-consuming and resource-intensive by traditional experimental methods to uncover effective therapeutic agents that target particular proteins [44]. In this respect, computational drug discovery techniques, especially molecular docking, have turned out to be useful towards fast tracking the identification of drug candidates [29,30]. Molecular docking, which enables the researcher to predict the binding affinity and stability of the interaction and modes of binding [29]. This method will allow a quick screen of compound libraries and finding potentially promising candidates to be further validated in an experiment [30]. Phytochemical which are bioactive plant-based products, have attracted a lot of interest in recent years as potential therapeutic agents against cancer [35,36]. Such natural compounds have varied pharmacological effects [37]. Notably, numerous phytochemicals are capable of bypassing the blood-brain barrier, which means that they are good candidates in brain tumour treatment, including GBM [36]. Molecular docking combined with phytochemical screening is an effective approach to finding natural compounds that have the potential to act against particular disease-related proteins [29,35]. Computational biology has advanced at a very fast pace, which has made it possible to combine various methods in the study of complex diseases [42,43]. Large-scale biological data analysis using bioinformatics techniques allows identifying disease-related genes, pathways, and in silico drug discovery methods allow identifying therapeutic compounds against particular proteins [3,42]. Together, these methods offer a holistic way old analyzing disease pathology and working on specific treatment. Bioinformatics analysis can be used in the case of MS to identify molecular biomarkers related to immune dysregulation and neurodegeneration [15]. Molecular docking in GBM has the potential to discover compounds that can inhibit the activity of crucial proteins like BUB1B [29]. Development has been made, but there are still several problems in the area of computational biology [44]. Most of the studies are based on single datasets, which can restrict

the externalisation of the results [2]. Also, computational predictions are frequently not tested experimentally, which is required to establish biological relevance [45]. Additionally, biomarker discovery and drug development are not well integrated, as these fields are often studied separately [43]. To fill this gap, an integrated computational framework that integrates various approaches must be developed to give a comprehensive insight into the biology of a disease [42]. Here, the current research seeks to employ combined computational methods to predict biomarkers and therapeutic targets in neurological diseases. In particular, the study will analyse the result of gene expression to determine differentially expressed genes and hub genes related to multiple sclerosis and the use of molecular docking tools to determine phytochemical inhibitors of BUB1B in Glioblastoma. This study aims to add to the knowledge of the disease and the development of treatment by integrating bioinformatics and in silico drug discovery. In addition, advances in transcriptomic technologies and computational biology have greatly enhanced the knowledge of molecular mechanisms behind neurological diseases [3,42]. The high-throughput sequencing technologies like RNA sequencing and microarray analysis yield large-scale biological data, which may be analysed by bioinformatics tools to discover disease-associated molecular signatures [48,49]. These technologies have allowed researchers to study global changes in gene expression that have helped to better identify genes and pathways associated with disease progression [3]. Several MS studies have shown that there are abnormalities in genes related to inflammatory response, activation of immune cells, oxidative stress and neuronal degeneration [15,41]. Immune-related pathways come into play in a dysfunctional way and play a significant role in Axonal injury and demyelination in the central nervous system. More specifically, disease severity and progression have been associated with changes in T cell, B cell, cytokine, and inflammatory mediator activity [11,15]. These molecular changes may thus serve as useful biomarkers for early diagnosis and tracking of disease progression. The knowledge of functional relationships among the disease-associated genes has been further improved through the application of 4D+ of the protein–protein interaction (PPI) network analysis [16]. The identification of highly interconnected genes called hub genes through PPI networks is useful for identifying genes that play crucial roles in biological pathways and cellular regulation [17]. Hub genes have been identified as central nodes in biological networks and are potential therapeutic targets and markers [18]. These genes can also be annotated to specific pathways and processes using tools of functional enrichment analysis, such as Gene Ontology (GO) and KEGG pathway analysis, which help to provide a biological context for these genes [7,8]. Likewise, cell cycle regulation, apoptosis, the DNA repair pathway, and signalling pathways of mitosis are all critical processes altered in the development and progression of Glioblastoma [19,23]. One of the critical mitotic checkpoint kinases, BUB1B, is a component of chromosome segregation and maintenance of genomic stability during cell division [24,25]. The overexpression of BUB1B has been linked to uncontrolled tumour growth, chromosomal instability and chemoradiotherapy resistance in the glioblastoma patient [26]. Thus, the inhibition of BUB1B activity becomes an attractive therapeutic approach to combat tumour proliferation and to enhance the therapeutic outcome [25]. Traditional drug discovery methods are time-consuming, costly and involve huge experimental screening [44]. On the other hand, computational drug discovery techniques like molecular docking can be used to quickly and cheaply identify potential therapeutic compounds [29,30]. Molecular docking determines the interaction between a drug and a target

protein using binding affinity, orientation, and stability [29]. Compounds with more binding interactions with target proteins are regarded as candidates for further experimental testing for drug development [30]. Phytochemicals have been gaining more attention in recent years because of their wide range of pharmacological activities and reduced toxicity when compared to synthetic drugs [35,36]. Phytochemicals have anti-inflammatory, anti-oxidative, and anti-cancer qualities that may help treat brain tumors and neurological conditions [37]. More than a few phytochemicals have demonstrated the ability to penetrate the blood-brain barrier, thus having therapeutic potential for central nervous system diseases [36]. Therefore, the combination of phytochemical screening and molecular docking can expedite the discovery of new drugs against proteins that are associated with a disease like BUB1B [29,35]. Recently, the use of integrated computational methods that integrate bioinformatics analysis and molecular docking has gained great significance in biomedical research [42, 43]. These approaches help identify disease biomarkers and can also aid in finding therapeutic compounds that are specific to a disease. The combination of transcriptomic analysis, network biology, and structure-based drug discovery gives a whole picture of disease mechanisms and therapeutic possibilities [42]. Both of these strategies are consistent with the principles of precision medicine that aim to create personalised therapeutic strategies based on the molecular characteristics of the diseases [43]. While there has been significant advances in computational biology, there are a number of drawbacks in existing studies [44]. The majority of the studies are based on one data set and have not been validated enough to ensure reliability and repeatability [2]. Besides, the computational prediction derived from molecular docking studies is frequently followed by molecular dynamics simulations, pharmacokinetic analysis and experimental studies [38]. Thus, there is an unmet need for a computational model that integrates and validates biomarker discovery with therapeutic target identification [42]. For this reason, the present study aimed to overcome these limitations by combining transcriptomic analysis based on bioinformatics with molecular docking approaches. Finding molecular biomarkers linked to multiple sclerosis and screening phytochemical substances for the BUB1B protein in glioblastoma are the main goals of the research effort. It is anticipated that this integrated computational approach will aid in the development of specific treatment approaches for neurological illnesses as well as an understanding of the underlying mechanisms of disease [42, 43].

CHAPTER 2

LITERATURE REVIEW

2.1 Multiple Sclerosis and Bioinformatics Approaches

Inflammation, demyelination, and neurodegeneration are the hallmarks of multiple sclerosis (MS), a chronic inflammatory disease of the central nervous system [11,15]. It is an affliction of young adults and is among the most typical reasons for neurologically impaired individuals in the world [12]. An abnormal immune response pathophysiologically mediates the disease, whereby autoreactive T cells and B cells attack the myelin sheath of the neurons, leading to impaired neuronal communication and, consequently, neurological dysfunction and cognitive impairment [15]. The clinical symptom heterogeneity and disease progression are significant challenges to diagnosis and treatment [14]. The understanding of MS on a molecular level has greatly improved in the past few years with the development of computational biology [42]. Bioinformatics methods allow analysing a large amount of biological data, which provides information on how genes can change their expression under disease conditions [3]. The microarray and RNA sequencing technologies enable researchers to analyse the expression of thousands of genes at the same time, which helps in identifying disease-related molecular signatures [48]. Investigation into large-scale computational analysis using curated datasets of gene expression is made accessible in publicly available repositories like the Gene Expression Omnibus (GEO) [1,2]. Applications such as GEO2R involve statistical procedures to reveal the differentially expressed genes (DEGs) between the disease and control groups [3]. DEGs are genes that have tremendous variations in the expression level and are regarded as significant signs of disease-related molecular changes [49]. Studies on bioinformatics have indicated that MS is linked with deregulation of immune response, inflammation, neuronal signalling pathways, and genes [15]. The discovery of these genes is essential to comprehend the mechanism of the disease as well as to create potential biomarkers for diagnosis and prognosis [3]. Nevertheless, the analysis of gene expression does not present the full picture of the analysis of disease since biological systems are prone to complex interactions of molecules[42].

2.2 Hub Gene Identification and Protein–Protein Interaction Networks

Network-based methods such as protein-protein interaction (PPI) analysis have become commonly used in systems biology to overcome the limitations of single-gene analysis [16,42]. PPI networks enable a global perspective of the interaction between proteins in a cellular context and are used to identify important regulatory elements that are involved in disease processes [16]. STRING database is typically utilised to form PPI networks since it combines known and predicted protein interactions from various sources [5]. Using the mapping of DEGs onto PPI networks, researchers can identify functional relationships between proteins and can

detect highly interconnected nodes called hub genes [17]. These hub genes are major contributors of network stability and are usually key regulators of biological functions [18]. Using PPI network analysis in the context of Multiple Sclerosis, a few key hub genes related to disease progression have been identified. Ribosomal proteins (RPS6, RPS28 and RPS18) have been reported to be among the key upregulated genes, indicating the increased protein synthesis and the cellular responses to stress. Alternatively, immune-related genes like CD79A and CD79B are almost always down-regulated, which points to impaired B-cell signalling and immune dysregulation [15]. It is specifically important that the hub genes are identified, as these genes can be used as potential biomarkers and therapeutic targets [18]. Targeting hub genes can have a significant effect on disease pathways since these key nodes in biological networks are essential to the development of illness [16]. As a result, PPI network analysis provides a useful approach for understanding disease mechanisms at the systems level [42].

2.3 Functional Enrichment and Pathway Analysis

To determine the biological meaning of differentially expressed genes, there is an important step of functional enrichment analysis [46]. This method is accomplished by mapping genes to established biological terms like Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [7,8]. Gene Ontology analysis divides genes into three broad categories: cellular components, molecular functions and biological processes [7]. Enrichment analysis in MS has demonstrated that the upregulated genes are commonly related to ribosomal activity, translation of proteins and stress response of the cell. Conversely, the downregulated genes are associated with the immune processes enriched, such as B cell receptor signalling and antigen processing [15]. KEGG pathways analysis also gives information about signalling pathways related to disease [8]. Research has established that MS has a link with immune response, inflammation and neurodegeneration pathways [41]. The result underscores the multifactorial interrelation of immune dysregulation and neurodegenerative damage in MS pathogenesis [15]. Besides GO and KEGG analyses, other phenotype-based methods, such as Monarch analysis, correlated changes in gene expression with the apparent disease phenotype. This assists in the comprehension of how molecular changes can be converted to clinical manifestations. In general, functional enhancement analysis improves the analysis of transcriptomic data and gives a more in-depth insight into the disease processes [46]. Nevertheless, the majority of studies concentrate on identifying the pathways but do not integrate them with therapeutic approaches, limiting their translational opportunities [43].

2.4 Glioblastoma and BUB1B as a Therapeutic Target

The most common and malignant type of primary brain tumor is glioblastoma (GBM). which is highly invasive, exhibits quick growth, and is resistant to therapy [19]. Despite the improvement in the surgical procedures, radiotherapy and chemotherapy, GBM patients are very unfortunate because the median survival time is still quite short, about 12-15 months [20]. Tumor heterogeneity is seen as one of the greatest challenges in GBM treatment as a result of the existence of different populations of cells in the tumor microenvironment [22]. Moreover, cancer stem-like cells also play a role in cancer recurrence and resistance to treatment, which

makes GBM especially challenging to treat [23]. On a molecular level, GBM development is caused by the imbalance in the activity of key signalling pathways and cell cycle regulators [23]. One of these, BUB1B (BUBR1), a mitotic checkpoint kinase, is a key protein in the process of ensuring correct segregation of chromosomes during cell division [24,25]. BUB1B has been linked to tumor progression, genomic instability, and resistance to therapy because of its overexpression [26]. Research has shown that BUB1B inhibition can disrupt mitotic mechanisms, cause cell cycle arrest and decrease tumor growth [24]. This renders BUB1B a promising therapeutic agent to treat GBM [25]. Nevertheless, finding effective inhibitors with the ability to selectively target BUB1B and with a low level of toxicity is an outstanding challenge [44].

2.5 Molecular Docking and Phytochemical-Based Drug Discovery

Experimental screening techniques, which are resource-intensive and time-consuming, are traditionally used to discover new therapeutic agents [44]. The computational methods, especially the molecular docking, have brought revolution to the discovery of drugs since they facilitate the rapid screening of large compound libraries [29,30]. Molecular docking is a form of simulating the interaction between a target protein and small molecules, which enables researchers to predict the binding affinity and stability of the interaction [29]. The values of the binding energy of the docking studies give an estimate of how strongly a ligand and a protein bind, with a lower value representing a stronger interaction [30]. The phytochemicals, which are naturally found bioactive compounds of plants, have been of great interest as a possible therapeutic agent [35,36]. These compounds have pharmacological activities such as antioxidant, anti-inflammatory, and a wide and anti-cancer properties [37]. Notably, in most cases, phytochemicals can cross the blood-brain barrier, thus they are good candidates to treat brain diseases like GBM [36]. The recent computational investigations have revealed a number of phytochemicals that have a great binding affinity to BUB1B [29]. Promising interaction profiles, such as stable hydrogen bonding and hydrophobic interaction with the important residues in the ATP-binding pocket of BUB1B, have been observed with compounds like α -peltatin, 2'-peltatin, and demethyl dehydropodophyllotoxin. These results indicate that phytochemicals could be used as possible BUB1B inhibitors and could also help in the creation of specific therapies against GBM [35]. Nonetheless, the majority of docking research is reduced to affinity prediction and does not provide additional validation of molecular dynamics simulations, pharmacokinetics, and experimental research [38]. Hence, more detailed methods that involve computational predictions and biological validation are required.

2.6 Integration of Bioinformatics and Drug Discovery Approaches

Despite the remarkable advances in the fields of biomarker discovery and drug development, the two are usually looked into separately [43]. Bioinformatics approaches primarily focus on the discovery of disease-related genes and pathways, whereas computational drug discovery focuses on identifying therapeutic compounds that target specific proteins [42]. A combined methodology that incorporates these techniques can give more insight into the mechanisms of disease [42]. Bioinformatics analysis enables the discovery of molecular biomarkers associated

with disease progression in Multiple Sclerosis [15]. Molecular docking aids in the discovery of compounds that can interact with important proteins like BUB1B in Glioblastoma [29]. The gap between disease diagnosis and treatment can be filled by connecting gene-level understanding with target-based drug discovery [42]. The proposed integrated framework is consistent with the ideas of precision medicine, which tries to create customised therapeutic approaches to the individual molecular profiles [43].

2.7 Study Rationale and Significance

Although much has been done to understand biology, there are still a few challenges that may be encountered in understanding and treating neurological disorders. Most studies are based on individual datasets, which might not be sufficient to reflect the heterogeneity of diseases [2]. Also, the majority of the research activities are either geared towards the discovery of biomarkers or drug discovery, and a lack of integration between these two research directions is observed [43]. Various studies have found differentially expressed genes and hub genes in Multiple Sclerosis, but there are still few studies that translate the results into therapeutic approaches [15]. In a similar case, in glioblastoma, though BUB1B has been suggested as a possible therapeutic target, limited research has been done on phytochemical-based inhibitors using a thorough computational analysis [24]. In addition, the molecular docking studies are usually limited to the binding affinity prediction and are not validated by the advanced techniques that include the molecular dynamics simulations and pharmacokinetic analysis [38]. The above limitations emphasise the need to have integrated computational platforms involving transcriptomic analysis and drug discovery [42]. To deal with these issues, the current study uses an integrated approach of computations that integrates bioinformatics-based gene expression analysis to identify biomarkers in Multiple Sclerosis and molecular docking-based screening of phytochemicals targeting BUB1B in Glioblastoma. This method will offer a detailed insight into the pathophysiology of diseases and, at the same time, identify a source of potential therapeutic compounds. The study, by connecting the biomarker discovery with the target-based drug screening, leads to the design of more efficient diagnostic and therapeutic approaches to neurological disorders.

CHAPTER 3

METHODOLOGY

3.1 Research Methodology Overview

The current research uses an integrated computational platform to bring together bioinformatics analysis and structure-based drug discovery to examine neurological disorders [42,43]. The methodology is aimed at identifying molecular biomarkers associated with Multiple Sclerosis (MS) and potential therapeutic compounds targeting Glioblastoma (GBM) [11,19]. The study has two significant parts. The former is the transcriptomic analysis of the data of gene expression to find hub genes and differentially expressed genes (DEGs) associated with MS [3]. These involve dataset retrieval, differential expression, protein-protein interaction (PPI) network and functional enrichment analysis [16,46]. The second component focuses on structure-based drug discovery, in which molecular docking is used to identify phytochemicals that target the BUB1B protein in Glioblastoma [29,30]. This includes the preparation of proteins, ligand selection, docking analysis, and interaction evaluation [27]. The study, by combining these methods, makes a linkage between disease-related molecular processes and the identification of therapeutic targets [42].

3.2 Bioinformatics Analysis for Multiple Sclerosis

3.2.1 Dataset Selection

The publicly accessible gene expression information was obtained from the Gene Expression Omnibus (GEO) database. for multiple sclerosis [1,2]. The dataset GSE21942 was used for this study because it contains the gene expression profiles of both MS patients and healthy control individuals. The 29 samples in the collection were produced using microarray technology. Because it offers clearly defined sickness and control groups, this dataset is suitable for a differential gene expression study.

3.2.2 Differential Gene Expression Analysis

The GEO2R online program, which is based on the limma statistical package, was used to conduct differential gene expression analysis [3]. MS and control were the two groups into which the samples were split. For each gene, log₂ fold change (log₂FC) values and matching p-values were computed using GEO2R. Filtering criteria were used to identify statistically significant genes. Genes were deemed significant if their p-value was less than 0.05. Among these, genes were classified as upregulated if their log₂ fold change was larger than +1, and as downregulated if their log₂ fold change was smaller than -1. Genes that didn't meet these requirements were excluded from the analysis.

3.2.3 Data Normalisation

Log2 transformation was used to normalise gene expression data so that samples could be compared [48]. This stage guarantees that variations in gene expression have biological significance while reducing technical variances. To ensure that the dataset was properly normalised, boxplot analysis was utilised.

3.2.4 Building Protein–Protein Interaction (PPI) Networks

Protein–protein interaction networks were constructed using the STRING database in order to examine the functional connections between the discovered genes [5]. *Homo sapiens* was chosen as the organism, and a high confidence interaction score of 0.900 was used to guarantee trustworthy interactions. For gene sets that were upregulated and downregulated, distinct PPI networks were created. To increase network clarity and concentrate on biologically significant relationships, isolated nodes were eliminated.

3.2.5 Analysis of Functional Enrichment

To ascertain the biological importance of the discovered genes, STRING was used for functional enrichment analysis [46]. Genes were categorised into cellular components, molecular functions, and biological processes using Gene Ontology (GO) analysis [7]. To find pathways connected to multiple sclerosis, KEGG pathway analysis was performed [8]. Phenotype analysis was also used to link alterations in gene expression to features of the disease. These investigations revealed networks pertaining to immune response, inflammation, and cellular control and shed information on the molecular mechanisms driving multiple sclerosis [15].

3.2.6 Hub Gene Identification

Using degree centrality, hub genes were found based on the PPI network's topological characteristics [17]. Because they are essential to biological networks and the development of disease, the genes with the greatest number of connections were referred to as hub genes [18]. These hub genes were chosen as possible multiple sclerosis biomarkers.

3.3 Molecular Docking for Glioblastoma

3.3.1 Preparing the Target Protein

The Protein Data Bank provided the target protein BUB1B's three-dimensional structure (PDB ID: 4R8Q). To guarantee compatibility with docking simulations, the protein was produced using AutoDock Tools [28]. Water molecules, co-crystallised ligands, and unnecessary heteroatoms were eliminated during preparation. Kollman charges were assigned, and polar hydrogen atoms were added. The PDBQT format was used to store the prepared protein structure.

3.3.2 Ligand Selection and Preparation

Based on their documented anticancer effects, 80 phytochemicals were selected [35,36]. Open Babel software was used to obtain these chemicals' structures and transform them into a three-dimensional format. To enable conformational flexibility, rotatable bonds were designed, and Gasteiger charges were introduced. PDBQT format was used to store the prepared ligands.

3.3.3 Molecular Docking Procedure

EasyDock Vina, which is an engine that is based on AutoDock Vina, was used to perform the molecular docking [26]. To achieve good binding of the ligand with the BUB1B protein, a grid box was delimited around the active site of the BUB1B protein. The grid parameters were set as follows:

- Centre: X = -28.379
- Y = 19.179
- Z = -12.680
- Size: 25 × 25 × 25 Å.

The ligand's lowest binding energy was selected for further study. The binding affinity of each ligand was determined as a binding energy (kcal/mol); the smaller the value, the stronger the binding [27].

3.3.4 Selection of Top Candidate Compounds

The binding affinity values from molecular docking analysis were used as the basic criterion for ranking the docked ligands. The binding energy (kcal/mol) is a measure of the strength and stability of how the substance and the target protein interact, and each compound was assessed based on this value. The compounds that showed more negative values for the binding energy were believed to possess a stronger binding affinity and higher interaction stability. The largest number of compounds were grouped into the top candidates based on this ranking, which is based on the lowest binding energy values for all compounds screened as phytochemicals. These compounds were found to be potential inhibitors of the BUB1 B protein as they form a stable interaction within the active site of the protein. Choosing the best candidates is significant to advanced analysis since compounds that bind strongly to the protein are likely to display strong inhibitory activity for the target protein. Therefore, the compounds that were shortlisted were regarded as potential drug candidates and were subjected to more analysis of interaction and validation [29,30].

3.3.5 Protein–Ligand Interaction Analysis

Protein–ligand interaction was analysed using the two-dimensional representation between ligands and protein residues, LigPlot+ [32]. It will allow one to clearly see hydrogen bonds and hydrophobic interactions between the ligand and the amino acid residues of the active site, which occur in the target protein. A few significant amino acid residues were identified in the analysis, which are important for the binding of the ligand. The hydrogen bonding interactions

and the hydrophobic interactions are important types of interactions that contribute to the protein-ligand complex's stability and selectivity and to the binding affinity and structural stability, respectively. These interactions are crucial for proper orientation of the ligand in the active site as well as for effective inhibition of the target protein. This analysis has the advantage of identifying the specific residues that participate in binding and can be used to guide the choice of compounds that hold promise for further studies [30,32].

3.3.6 Structural Visualisation

Protein ligand interactions were visualised in three dimensions using PyMOL [33]. The software provides a detailed analysis of the spatial structure, delivering the ligand to the active region of the target protein, making it easy to view the orientation, geometric structure and conformational positioning of the ligand. This step was crucial for the validation of the results of the docking, as it allowed us to see the fit of the ligand in the binding pocket of the BUB1B protein visually. The analysis also contributed to the understanding of the stability of the protein–ligand complex by analysing the position of important interacting residues, and the overall structural compatibility of the ligand in the active site. Moreover, three-dimensional visualisation helps to understand the atomic-level molecular interactions, which helps to confirm the accuracy of docking prediction and helps with understanding for further drug design and optimisation studies [29,33].

3.4 Integrated Computational Framework

This study's methodology combines molecular docking with bioinformatics analysis to offer a thorough understanding of neurological illnesses [42]. While the molecular docking technique makes it easier to find possible therapeutic drugs that target BUB1B in glioblastoma, the bioinformatics approach makes it possible to identify disease-associated genes and biomarkers in multiple sclerosis [3]. This integrated methodology connects target-based pharmaceutical development with gene expression data, bridging the gap between drug discovery and illness knowledge [43]. This strategy is appropriate for cutting-edge biomedical research since the application of computational tools guarantees effectiveness, affordability, and scalability [44].

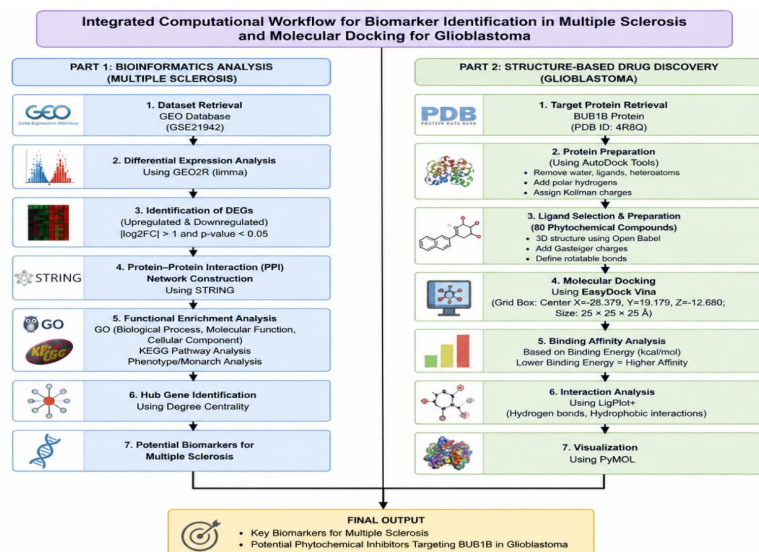


Figure 3.1: Integrated computational workflow illustrating bioinformatics analysis for identification of differentially expressed and hub genes in Multiple Sclerosis, and molecular docking-based screening of phytochemical compounds targeting BUB1B protein in Glioblastoma.

CHAPTER 4

RESULTS

4.1 Determination of Genes with Differential Expression

GEO2R was used to conduct a differential gene expression study between samples from multiple sclerosis and healthy controls [3]. There were 1276 differentially expressed genes (DEGs) found, comprising 579 downregulated and 697 upregulated genes.

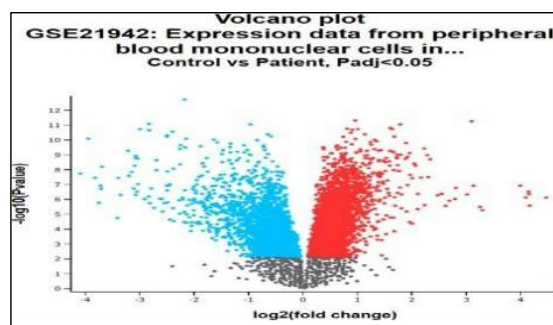


Fig 4.1 Volcano plot of the GSE21942 dataset to visualise genes whose expression is up or down in MS patients compared to healthy controls. Grey colour means the genes that are not significant, blue colour means genes that have been down-regulated ($\log_2FC < -1$, $p < 0.05$), and red represents genes that have been up-regulated ($\log_2FC > 1$, $p < 0.05$).

4.2 Data Normalization

In order to reduce technical variance and ensure comparability of gene expression between all samples, standardisation of data was performed [48]. To ensure accurate comparisons between samples, variance was stabilised by log2 transformation of the expression values during normalization step [48]. A box plot of the distribution of each sample's gene expression values is displayed both before and after normalization in Figure 4.2. All the median expression levels are in the same range after normalization, showing that the data is distributed evenly. This indicates that the data is suitably normalized and is suitable for further analysis, including network analysis and differential gene expression. Therefore, normalization of gene expression data is a crucial step in transcriptomic analysis as it can reduce the variations that might occur during experimental and technical processes involved in sample processing and data acquisition [48,49]. The consistency of the normalization process also increases confidence in downstream analyses of the differences in gene expression. Normalization of expression values

yielded a uniform distribution of expression levels, further supporting the quality and consistency of the dataset, and improving confidence in the identification of differentially expressed genes and related biological pathways [3]. Further, by normalizing the data, it can serve as a basis for subsequent functional enrichment analysis, building networks of interactions between proteins and identifying hub genes in advanced bioinformatics analyses [16,46].

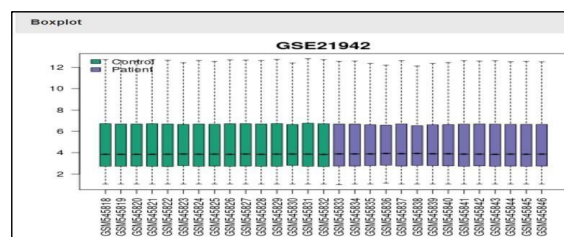


Fig 4.2 Data normalisation and group comparability are confirmed by a box plot of normalised gene expression values for MS and control samples.

4.3 Functional Enrichment Analysis

The STRING database [5] was used for a functional enrichment study of the DEGs, or differentially expressed genes. To find biological processes and pathways that were noticeably enriched, Gene Ontology (GO) and KEGG pathway studies were performed [7,8]. Results indicated that the genes that were upregulated were mainly associated with the processes of ribosomal activity, protein translation, and suggested a higher metabolic activity of the cells. The downregulated genes, by comparison, showed significant associations with immune response, B-cell receptor signalling and inflammatory pathways, which suggests that there is an interference with the immune regulation in Multiple Sclerosis [15]. The enrichment results of the genes that are upregulated are shown in Figure 4.3 (a-e), and the enrichment results of the genes that are downregulated are shown in Figure 4.3 (f-j)



Fig 4.3a Gene Ontology (GO) biological process enrichment analysis showing the significantly enriched biological pathways associated with upregulated differentially expressed

genes (DEGs) in Multiple Sclerosis. Fig 4.3b Gene Ontology (GO) molecular function enrichment analysis illustrating the significantly enriched molecular functions associated with upregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

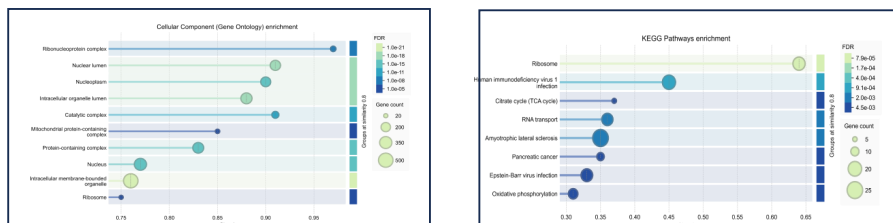


Fig 4.3c Gene Ontology (GO) cellular component enrichment analysis showing the significantly enriched cellular components associated with upregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

Fig 4.3d KEGG pathway enrichment analysis illustrating the significantly enriched biological pathways associated with upregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

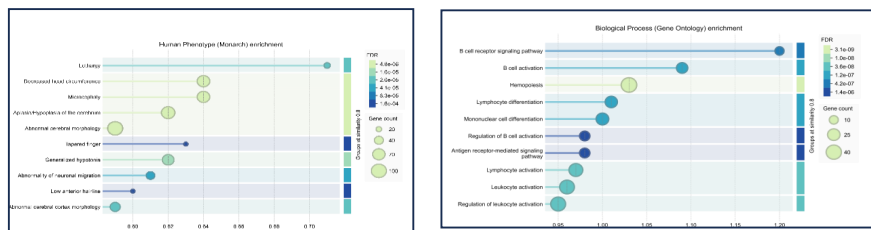


Fig 4.3e Monarch phenotype enrichment analysis showing the significantly associated phenotypic characteristics linked with upregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

Fig 4.3f Gene Ontology (GO) biological process enrichment analysis showing the significantly enriched biological pathways associated with downregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

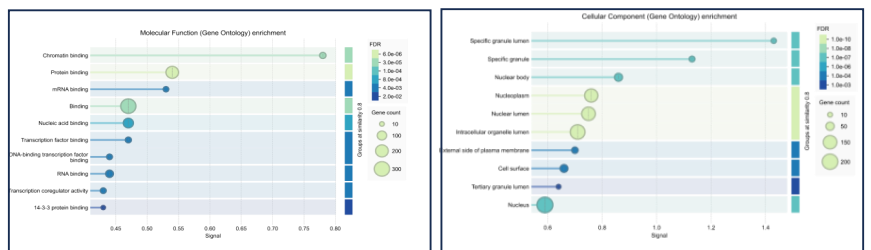


Fig 4.3g Gene Ontology (GO) molecular function enrichment analysis illustrating the significantly enriched molecular functions associated with downregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

Fig 4.3h Gene Ontology (GO) cellular component enrichment analysis showing the significantly enriched cellular components associated with downregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

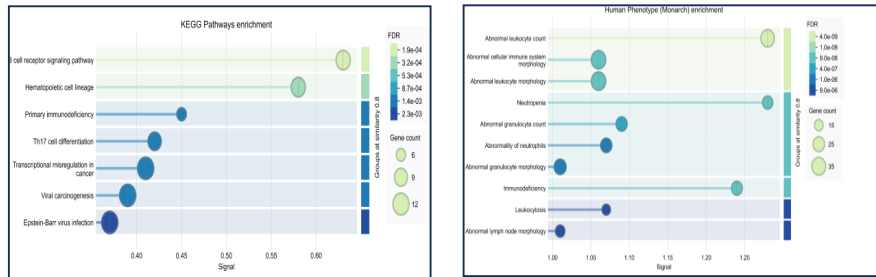


Fig 4.3i KEGG pathway enrichment analysis illustrating the significantly enriched biological pathways associated with downregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

Fig 4.3j Monarch phenotype enrichment analysis showing the significantly associated phenotypic characteristics linked with downregulated differentially expressed genes (DEGs) in Multiple Sclerosis.

4.4 Hub Gene Identification and the PPI Network

The protein-protein interaction (PPI) network was built based on the STRING database to investigate the functional relationship among the identified differentially expressed genes [5]. The network analysis revealed that there was a significant interaction between the genes, suggesting that they are linked in some way to help each other function. The PPI networks of both upregulated and downregulated genes are shown in Figure 4.4 (a–b). Such networks consist of nodes (proteins) and the interactions between them (edges). Hub genes were nodes with a high degree centrality [17] that were determined as hub genes. Hub gene analysis showed that the genes that were upregulated included RPS6, RPS28 and RPS18, indicating the genes were related to the activity of the ribosomes and protein synthesis. On the other hand, the two genes CD79A and CD79B were identified as the primary hub genes that were down-regulated, suggesting their role in immune response and B-cell signalling pathways [15]. These data indicate the relevance of hub genes in the molecular pathways of Multiple Sclerosis. Moreover, hub genes identified by PPI network analysis give valuable clues on molecular mechanisms

related to Multiple Sclerosis disease [16,42]. Hub genes are found in the centre of the gene interaction network and play an important role in the regulation of many cellular pathways [17] and are therefore biologically significant. Highly connected ribosomal proteins (RPS6, RPS28, RPS18) indicate greater translational activity and/or metabolic changes in disease states. The downregulation of the immune-related hub genes CD79A and CD79B also helps to confirm a role for immune dysregulation and impaired B cell signalling in MS pathogenesis [15]. Beyond this, PPI network analysis can be used to obtain an understanding of the functional coordination between disease-associated proteins and can identify potential biomarkers and therapeutic targets [18]. Those genes that are better-connected within a network are more likely to be critical for maintaining biological stability and disease progression [16]. Hence, the discovered hub genes could be useful molecular markers in the diagnosis and prognosis of MS and as targets for therapeutic interventions [42].

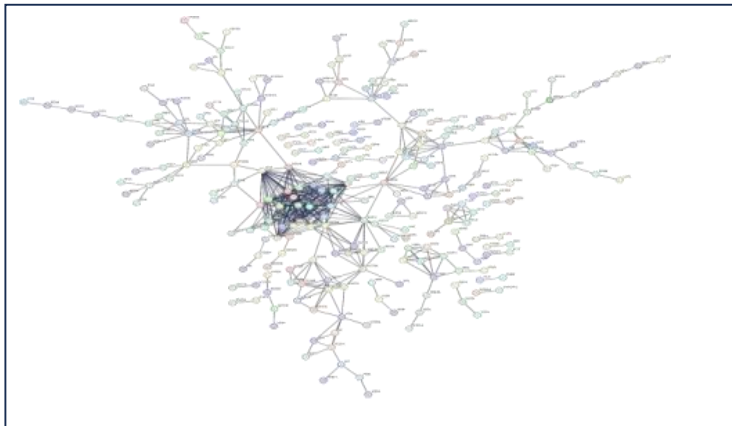


Fig 4.4 a PPI network of upregulated DEGs in MS generated using STRING (confidence score > 0.900). Nodes stand in for the proteins, and the interactions are represented by edges. The highly connected nodes denote hub genes, which include RPS6, RPS28 and RPS18.

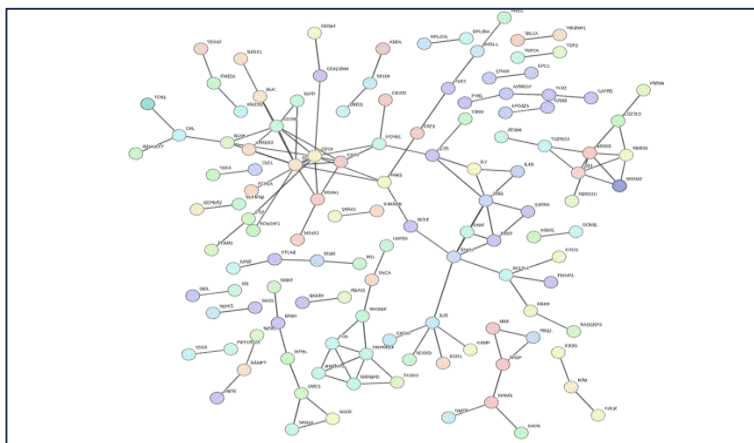


Fig 4.4 b PPI of downregulated DEGs in MS generated by use of STRING (confidence score >0.900). The proteins are represented by the nodes and the interactions between them by the edges. The important hub genes are CD79A and CD79B.

4.5 Molecular Docking Results

To determine the binding affinity of the chosen phytochemical compounds to the BUB1B protein, molecular docking was carried out [29,30]. The docking analysis has been conducted using EasyDock Vina, which generated several binding conformations of each ligand in the active site of the protein [27]. The docking was used to give binding energy values of each compound, which are used to determine the strength of the ligand's interaction with the target protein [29]. The smaller the binding energy values, the stronger and more the interaction is [30]. Moreover, molecular docking analysis gave detailed information regarding the interaction between the selected phytochemical compounds and the active site of the BUB1B protein [29]. Multiple binding conformations generated during docking revealed that the compounds were able to fit effectively within the ATP-binding pocket of the target protein. The docking poses generated were used for determining the most stable docking pose for each ligand with the lowest binding energy values [30]. In addition, the docking study showed that some compounds showed stable interactions with the essential active site residues of the BUB1B protein, suggesting that it has the potential to inhibit the protein [29,35]. The binding affinity values obtained indicate that the selected phytochemicals have good interaction properties and can disrupt the normal function of the target protein. These interactions are crucial in the process of structure-based drug discovery as a tighter interaction between the ligand and the protein is linked to increased inhibitory power [30]. Also, the docking results offer preliminary indications of the therapeutic potential of phytochemical compounds against Glioblastoma by targeting the protein BUB1B, which is part of the mitotic checkpoint [24,25]. Computational

methods have been useful in speeding up drug discovery, and compounds with high binding affinity has been identified to support the usefulness of these methods in drug discovery and screening of bioactive molecules [44]. Based on the above facts, the molecular docking analysis carried out in the present study is an important step in screening for possible lead molecules for experimental validation and therapeutic development [29,38].

4.6 Binding Affinity Analysis

Eight phytochemical compounds were screened against the BUB1B protein, and the compounds were ranked by their affinity to the protein [29]. The result revealed that α -peltatin (POHX0006) had the best binding strength using a docking score of -10.6 kcal/mol, showing a strong interaction with the target protein. Other compounds like 2 peltatin and POHX0054 were also shown to have significant binding affinities, implying their potential as inhibitors of the BUB1 B protein [35,36]. The binding values of the top-ranked compounds are summarised in the Table. Additionally, binding affinity analysis was used to gain information about the binding strength and stability of the selected phytochemical compounds in the selected BUB1B protein [29]. Molecular docking-derived binding affinity scores are regarded as an important parameter to assess the inhibitory activity of a compound against a protein. Strong and stable protein–ligand interactions are typically correlated with compounds having lower binding energy values [30]. The docking scores in this study suggest that the phytochemicals screened in the present study have significant binding potential towards the active site of the BUB1B protein. The most binding compound among all the screened was α -peltatin, which may be more stable of interaction and therefore useful as a therapeutic drug for Glioblastoma [35]. In addition, the existence of good docking scores for other compounds, like β -peltatin and POHX0054, further highlights the importance of phytochemicals for use as inhibitors against the mitotic checkpoint protein [36,37]. Structural differences in the compounds (conformation, functional groups, interaction ability) with the protein binding pocket might cause the variation of the docking scores [29]. The results indicate that compounds with higher binding affinity might have more inhibitory activity and possibly be used for further pharmacological investigation. Hence, the binding affinity analysis carried out in this study will be used as a valuable foundation for the identification and prioritisation of potential lead compounds for further drug discovery studies against Glioblastoma [44].

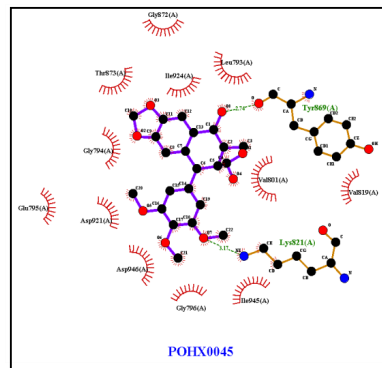
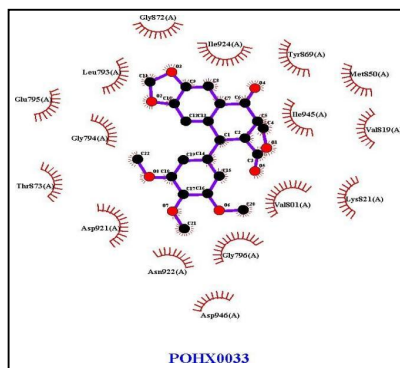
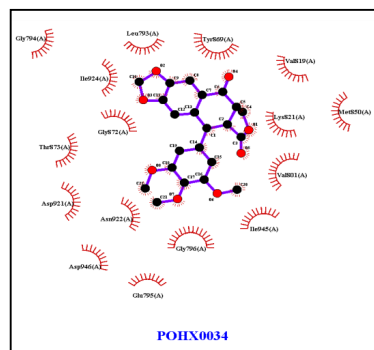
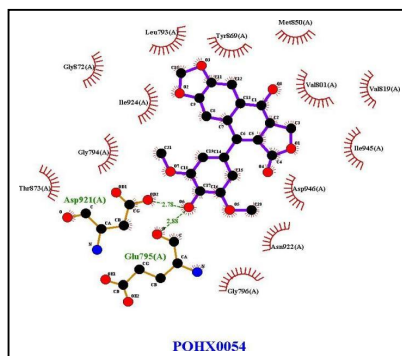
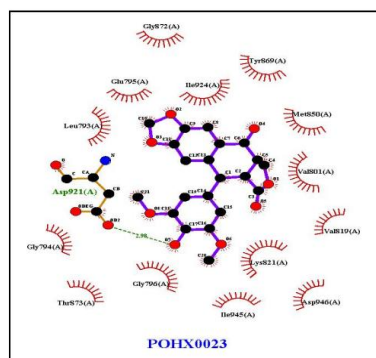
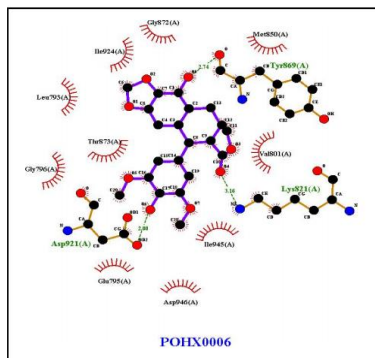
DOCKING SCORES AND KEY INTERACTING RESIDUES OF TOP RANKED PHYTOCHEMICALS WITH BUB1B

ID	Phytochemical	Score	Key interacting residues
POHX0006	α -Peltatin	-10.6	Glu795, Gly796, Val801, Lys821, Asp921,

			Ile945
POHX0023	β -Peltatin	-10.4	Glu795, Gly796, Val801, Ile924, Lys821, Asp921, Ile945
POHX0054	4'-O-Demethyl-dehydropodophyllotoxin	-10.2	Gly794, Gly796, Glu795, Val801, Ile921, Ile945
POHX0034	Picropodophyllin	-9.8	Gly794, Gly796, Glu795, Val801, Ile924, Lys821, Asp921
POHX0033	Isopicropodophyllotoxin	-9.7	Glu795, Gly796, Val801, Ile924, Lys821, Asp921, Ile945
POHX0045	Ellagic acid	-9.7	Glu795, Gly796, Val801, Ile924, Lys821, Asp921, Ile945
POHX0047	Podophyllotoxin	-9.6	Glu795, Gly796, Val801, Ile924, Lys821, Asp921, Ile945
POHX0017	Podophyllin	-9.6	Glu795, Gly796, Val801, Ile924, Lys821, Asp921, Ile945
POHX0011	4-Demethyl-desoxypodophyllotoxin	-9.5	Val801, Gly796, Glu795, Ile945
POHX0036	Ophyllotoxin	-9.5	Glu795, Gly796, Val801, Ile924, Lys821, Asp921, Ile945

4.7 Protein–Ligand Interaction Analysis

LigPlot+ was used to analyse the protein-ligand interaction between the selected compounds and the BUB1B protein. It was observed that the most effective compounds were stable key amino acid residues in the active site through hydrophobic interactions and hydrogen bonding. Specifically, residues like Glu795, Gly796 and Val801 were discovered as playing a significant role in ligand binding. The protein–ligand complex is stabilised and becomes more complicated by these specific interactions. Figure 4.7 shows the pattern of interaction of the top compound. Image LigPlot+ (2D interaction diagram of α -peltatin)



4.8 Structural Visualization

The protein-ligand complexes of the selected compounds were visualised in three dimensions using PyMOL to confirm the binding orientation and stability of the chosen substances in the BUB1B active site protein. The results of the visualisation revealed that α -peltatin was properly fitted in the ATP-binding pocket of the protein, forming a stable complex. The docking findings are further confirmed by the proper orientation and interaction of the ligand in the binding site.



Fig 4.8 Three-dimensional visualisation of the BUB1B- alpha- peltatin complex, where the binding orientation of the ligand in the active site of the protein is shown.

CHAPTER 5

DISCUSSION AND CONCLUSION

This paper used an integrated approach of calculating using bioinformatics analysis and molecular docking to examine molecular pathways and potential pharmacological targets in neurological diseases, namely Multiple Sclerosis (MS) and Glioblastoma (GBM) [3,29]. Transcriptomic analysis combined with structure-based drug discovery was an all-encompassing tool to comprehend the disease progression and identify possible therapeutic approaches [42,43]. The bioinformatics analysis results showed massively different patterns of gene expression in MS, and this indicates a great molecular change is related to the pathology of the disease [3,48]. The increased expression of genes related to ribosomal functions and protein synthesis, which is enriched, suggests that the cellular metabolic processes are increased, and that this might be related to cellular stress responses and disease progression [15]. These findings are consistent with earlier research, such point to the role of the altered protein synthesis in neurodegenerative and autoimmune diseases [40,41]. By contrast, down-regulated genes were the most common ones in relation to the immune response and the B-cell signalling pathways, highlighting the important role of immune dysregulation in the pathogenesis of MS [15]. This immune imbalance could be one of the factors leading to the gradual degenerative changes in the central nervous system [11]. Subsequent protein-protein interaction (PPI) network analysis allowed identification of key hub genes, including those of the upregulated genes, RPS6, RPS28 and RPS18, and those of the downregulated genes, CD79A and CD79B [16,18]. These hub genes are very interconnected, and they are very important in ensuring that cells are maintained in a good state of homeostasis and biological functions [17]. The fact that the ribosomal proteins are hub genes implies they are involved in increased translational activity, whereas immune-related genes indicate that there is impaired immune signalling. These hub genes can be potential biomarkers used in early diagnosis, monitoring of diseases, and targeting of therapies in Multiple Sclerosis, as these hub genes are centrally located in the network [18]. Molecular docking analysis of the BUB1B protein, which is one of the key regulators of the mitotic checkpoint signalling, identified several phytochemical compounds with a high binding affinity to this protein in the Glioblastoma component of the study [24,29]. Out of the screened compounds, α -peltatin was found to have the highest binding affinity, which means that it can be a promising therapeutic candidate. Other compounds, including 2-peltatin and POHX0054, also had a strong interaction with the target protein and therefore are of possible use as alternative inhibitors [35,36].

The protein-ligand interaction analysis further showed that these compounds develop stable Hydrophobic interactions and hydrogen bonding with important amino acid residues in the active site of the BUB1B protein, such as Glu795, Gly796 and Val801 [31,32]. Such interactions are critical in ensuring binding stability and specificity, which is paramount in effective inhibition of protein functioning [30]. In addition, the structural visualization has

verified the presence of the ligands in the right place in the ATP-binding pocket of the protein, which supports the reliability and accuracy of the docking results [33]. These results suggest that phytochemicals that have been selected can disrupt the process of cell cycle regulation in tumor cells and consequently inhibit the growth of tumor cells [24]. The combination of transcriptomic analysis and molecular docking in the current study is a comprehensive and effective plan to connect the disease-related biomarkers with the identification of therapeutic targets [42]. Although bioinformatics analysis can be used to identify the key genes and pathways involved in the disease processes, molecular docking can be used to screen and evaluate a large number of compounds that target specific proteins in a very short period of time [29,30]. Such an integrated methodology can contribute to a better comprehension of the biology of diseases and help to design specific and tailored therapeutic solutions [43]. Along with such promising results, this study has some limitations. The analysis was done on a single gene expression dataset, which is not a complete representation of Multiple Sclerosis complexity and heterogeneity [2]. Also, molecular docking outcomes are computational predictions and are not necessarily accurate in vivo [30]. Consequently, the therapeutic potential of the identified compounds needs to be verified further through the use of molecular dynamics simulations, pharmacokinetic analysis, and experimental studies [38]. Overall, the present research has been able to highlight the important molecular biomarkers related to Multiple Sclerosis and the potential phytochemical inhibitors of the BUB1B protein in Glioblastoma. The findings illustrate the usefulness of combined computational methods in biomedical research and show their potential in eliminating the gap between disease knowledge and drug discovery [42,44]. The results of the current study will form a very solid foundation for further research in the future, to create new diagnostic tools and therapeutic approaches that are specific to neurological disorders.

Moreover, results from this study highlight the need for considering integrated computational biology techniques in understanding the complex molecular mechanisms involved in neurological disorders [42,43]. The identification of a collection of genes with varying expression levels, and a set of hub genes, gives insights into the immune dysregulation, the inflammation and the neurodegeneration pathways of Multiple Sclerosis [15,41]. These molecular markers can play a major role in the identification of new diagnostic markers and individualised therapeutic approaches in disease management [43]. The significance of these functional enrichment analysis results showed that the genes identified are involved in significant cellular and molecular pathways, such as ribosomal activity, immune response and inflammatory signalling pathways [7,8]. Ribosomal pathways and protein processing were enriched in the upregulated genes, indicating enhanced cellular metabolism and protein synthesis in disease conditions [15]. The downregulation of immune-associated pathways, on the other hand, suggests immune regulatory abnormalities and B-cell signalling abnormalities, both of which are believed to play an important role in the disease course of Multiple Sclerosis [11,15]. These observations are in line with studies that have already been published on autoimmune and neurodegenerative diseases [40,41]. Further, the results of the identification of hub genes through PPI network analysis are helpful to provide information on the central regulators that are central to the progression of disease [16,17]. Hub genes can have multiple functions and be important for the maintenance of multiple biological pathways in the cell

simultaneously, because they are highly connected in biological networks [18]. Hence, the genes RPS6, RPS28, RPS18, CD79A and CD79B could be useful as molecular markers in early diagnosis, prognosis and therapeutic targets in Multiple Sclerosis [15,18]. Furthermore, the present molecular docking study revealed the applicability of structural based drug discovery strategies in the identification of therapeutic molecules for the disease related protein targets [29,30]. The compounds selected for phytochemicals showed good binding ability to the BUB1B protein, which suggests their potential inhibitory effect on Glioblastoma [35,36]. The interactions between α -peltatin and important amino acid residues within the ATP binding pocket of the BUB1B protein indicate the potential mechanism of its interference with the regulation of the mitotic checkpoint and proliferation of tumour cells [24,25]. The results of this study are in line with earlier reports on the therapeutic value of natural products for cancer therapy [35,37]. The other major implications of the study are that phytochemicals can be utilized as therapeutic agents because they have relatively lower toxicity and high biological activity [36,37]. The pharmacological activities of these compounds present in plants are also very diverse, including antioxidant, anti-inflammatory, and anticancer properties, and could be responsible for the therapeutic effect in neurological diseases and brain tumors [37]. Moreover, a number of phytochemicals have been shown to be able to cross the blood–brain barrier, making them more promising for the treatment of diseases affecting the central nervous system, including Glioblastoma [36]. Thus, the combination of transcriptomic analysis and molecular docking offers a holistic approach for associating disease-related biomarkers with the identification of new therapeutic targets [42]. These multifaceted strategies could greatly speed up drug discovery and enhance clarity of disease biology at the molecular level [43]. Traditional lab methods can also be accelerated, money can be saved, and the number of experiments conducted can be reduced through the use of computational tools [44]. As a result, computational biology has become an integral part of today's biomedical research and precision medicine approaches [42,43]. Even though the results from this study were encouraging, additional studies with molecular dynamics simulation, ADMET analysis, and experimental testing would be beneficial to ensure the therapeutic potential and biological stability of the compounds identified [38,44]. The results of the study can be more reliable and clinically applicable in future studies with a greater quantity of samples and verification of experiments. Furthermore, introducing A.I. and machine learning technologies to transcriptomic and docking analysis could further boost the prediction accuracy and identification of other novel therapeutic targets in neurological diseases [43,44]. Overall, the present study has proven that the integrated computational approaches, which include bioinformatics analysis, network biology and molecular docking, can yield valuable insights into the mechanisms of disease and into therapeutic target discovery in neurological diseases [42]. This study's results can help in the future development of diagnostic markers and targeted therapeutic approaches for Multiple Sclerosis and Glioblastoma, to aid precision medicine and computational drug discovery.

REFERENCES

- [1] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridisation array data repository. *Nucleic Acids Res.* 2002;30(1):207–210.
- [2] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 2013;41(Database issue):D991–D995.
- [3] Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York: Springer; 2005. p. 397–420.
- [4] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57–63.
- [5] Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein–protein association networks with increased coverage. *Nucleic Acids Res.* 2019;47(D1):D607–D613.
- [6] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498–2504.
- [7] Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. *Nat Genet.* 2000;25(1):25–29.
- [8] Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
- [9] Kanehisa M, Furumichi M, Sato Y, et al. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 2021;49(D1):D545–D551.
- [10] Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology. *Bioinformatics.* 2009;25(8):1091–1093.
- [11] Compston A, Coles A. Multiple sclerosis. *Lancet.* 2008;372(9648):1502–1517.

- [12] Reich DS, Lucchinetti CF, Calabresi PA. Multiple sclerosis. *N Engl J Med*. 2018;378(2):169–180.
- [13] Filippi M, Bar-Or A, Piehl F, et al. Multiple sclerosis. *Nat Rev Dis Primers*. 2018;4(1):43.
- [14] Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17(2):162–173.
- [15] Dendrou CA, Fugger L, Friese MA. Immunopathology of multiple sclerosis. *Nat Rev Immunol*. 2015;15(9):545–558.
- [16] Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet*. 2011;12(1):56–68.
- [17] Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature*. 2001;411(6833):41–42.
- [18] Yu H, Kim PM, Sprecher E, et al. The importance of bottlenecks in protein networks. *PLoS Comput Biol*. 2007;3(4):e59.
- [19] Ostrom QT, Cioffi G, Waite K, et al. CBTRUS statistical report: primary brain and other CNS tumors diagnosed in the United States. *Neuro Oncol*. 2021;23(Suppl 3):iii1–iii105.
- [20] Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med*. 2005;352(10):987–996.
- [21] Tan AC, Ashley DM, López GY, et al. Management of glioblastoma: state of the art and future directions. *CA Cancer J Clin*. 2020;70(4):299–312.
- [22] Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in glioblastoma. *Science*. 2014;344(6190):1396–1401.
- [23] Brennan CW, Verhaak RGW, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell*. 2013;155(2):462–477.
- [24] Ricke RM, Jeganathan KB, van Deursen JM. Bub1 overexpression induces aneuploidy and tumor formation. *J Cell Biol*. 2011;193(6):1049–1064.

- [25] Wang Z, Katsaros D, Shen Y, et al. Biological and clinical significance of MAD2L1 and BUB1B expression in ovarian cancer. *Cancer Res.* 2007;67(14):6906–6913.
- [26] Cahill DP, Lengauer C, Yu J, et al. Mutations of mitotic checkpoint genes in human cancers. *Nature.* 1998;392(6673):300–303.
- [27] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking. *J Comput Chem.* 2010;31(2):455–461.
- [28] Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4. *J Comput Chem.* 2009;30(16):2785–2791.
- [29] Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des.* 2011;7(2):146–157.
- [30] Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening. *Nat Rev Drug Discov.* 2004;3(11):935–949.
- [31] Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams. *J Chem Inf Model.* 2011;51(10):2778–2786.
- [32] Wallace AC, Laskowski RA, Thornton JM. Ligplot: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* 1995;8(2):127–134.
- [33] DeLano WL. The PyMOL molecular graphics system. San Carlos, CA: DeLano Scientific; 2002.
- [34] Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—a visualization system for exploratory research. *J Comput Chem.* 2004;25(13):1605–1612.
- [35] Newman DJ, Cragg GM. Natural products as sources of new drugs. *J Nat Prod.* 2020;83(3):770–803.
- [36] Cragg GM, Pezzuto JM. Natural products as a vital source for the discovery of cancer chemotherapeutic agents. *Med Princ Pract.* 2016;25(Suppl 2):41–59.
- [37] Atanasov AG, Zotchev SB, Dirsch VM, Supuran CT. Natural products in drug discovery.

Nat Rev Drug Discov. 2021;20(3):200–216.

[38] Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron.* 2018;99(6):1129–1143.

[39] Lipinski CA. Lead- and drug-like compounds. *Drug Discov Today Technol.* 2004;1(4):337–341.

[40] Glass CK, Saijo K, Winner B, et al. Mechanisms underlying inflammation in neurodegeneration. *Cell.* 2010;140(6):918–934.

[41] Lassmann H. Multiple sclerosis pathology. *Cold Spring Harb Perspect Med.* 2018;8(3):a028936.

[42] Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004;5(2):101–113.

[43] Hood L, Friend SH. Predictive, personalized, preventive, participatory medicine. *Nat Rev Clin Oncol.* 2011;8(3):184–187.

[44] Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity. *Nat Rev Drug Discov.* 2010;9(3):203–214.

[45] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646–674.

[46] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis. *Proc Natl Acad Sci USA.* 2005;102(43):15545–15550.

[47] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses. *Nucleic Acids Res.* 2015;43(7):e47.

[48] Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods. *Bioinformatics.* 2003;19(2):185–193.

[49] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

[50] Peng J, Wang Y, Guan J, et al. An end-to-end heterogeneous graph representation learning-based framework for drug-target interaction prediction. *Nat Mach Intell.* 2021;3(5):393–403.

[51] Dendrou CA, Fugger L, Friesse MA. Immunopathology of multiple sclerosis. *Nat Rev Immunol.* 2015;15(9):545–558. (Reference used for Figure 2.1: Immunopathogenesis of Multiple Sclerosis)

[52] Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12(1):56–68. (Reference used for Figure 2.2: Systems biology and disease module identification)

[53] Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer.* 2009;9(3):153–166. (Reference used for Figure 2.3: Cell cycle regulation and checkpoint control related to BUB1B)

LIST OF PUBLICATION

1. Conference paper:

Smita Yadav and P. Kumar, “Computational Identification of BUB1B-Targeting Phytochemicals for GBM Therapy Using Molecular Docking,” *2026 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2026, IEEE, doi: 10.1109/IATMSI68868.2026.11465606.

Keywords: Glioblastoma (GBM); BUB1B; Molecular Docking; Phytochemicals; EasyDock Vina; LigPlot+; PyMOL; Binding Affinity; Protein–Ligand Interaction; Computational Biology; Structure-Based Drug Discovery; Bioinformatics; Therapeutic Targeting; Cancer Therapy.

2. Poster:

Smita Yadav¹, Pravir Kumar¹, “Identification of Novel ALDH1A3 Inhibitors in Glioblastoma Using Molecular Docking Machine Learning Based Virtual Screening”

Presented at: SNCI, Jamia Hamdard, New Delhi



IEEE Xplore®

[Browse](#)
[My Settings](#)
[Help](#)

Institutional Sign In

IEEE

All

ADVANCED SEARCH

Conferences > 2026 IEEE International Confe...

Computational Identification of BUB1B-Targeting Phytochemicals for GBM Therapy Using Molecular Docking

Publisher: IEEE

[Cite This](#)
[PDF](#)

Smita Yadav · Pravir Kumar · All Authors

2

Full

Text Views

Abstract

Authors

Figures

References

Keywords

Metrics

Abstract:

Glioblastoma is a very aggressive, rapidly growing cancerous brain tumor that is rebellious against standard treatment. At the centre of its development is BUB1B, a checkpoint kinase that is routinely upregulated in the spindle assembly. In glioblastoma cells, BUB1B is essential for tumor survival and genomic fidelity, which is a strong focus of treatment. In this investigation, we used computational methods to find phytochemicals that inhibit the BUB1B kinase domain. We selected a phytochemical collection of 80 phytochemicals, performed molecular docking through EasyDock Vina to determine affinities, and then visualised the interaction with LigPlot+ and PyMOL. The experiment revealed three of the most outstanding compounds: O-POHX0006 (peltatin alpha), POHX0023 (peltatin beta), and demethyl dehydropodophyllotoxin POHX0054 (high) binding affinities, and alpha peltatin is -10.6kcal/mol . Further validation using molecular dynamics simulations and ADMET profiling in vitro assays is necessary to confirm this finding. This study provides the foundation for the phytochemical treatment of glioblastoma.

Published in:

2026 IEEE International Conference on Interdisciplinary Approaches In Technology and Management for Social Innovation (IATMSI)

Date of Conference:

12-14 March 2026

DOI:

10.1109/IATMSI68868.2026.11465006

Date Added to IEEE Xplore:

07 April 2026

Publisher:

IEEE

Conference Location:

Gwalior, India

ISBN Information:

Recommended for You (Beta)

1.1 AI Era Innovation Matrix

Revolutionize 3D-Chip Design With...

IEEE International Conference on...

Learn More

I. Introduction

The most common and aggressive primary brain malignancy is glioblastoma. Brain malignancy has a median survival of 15 months [1], [2]. Despite advances in surgery, radiotherapy, and temozolomide therapy, most patients experience tumor relapse [2]. This negative progress is caused by the high invasiveness of glioblastoma, high heterogeneity of tumors, and resistance of the tumors to therapy due to the presence of the heterogeneous stem-like cells leading to tumor relapse [3]. Many of the chemotherapeutic drugs are also limited by the blood brain barrier, which explains the urgent need to develop new molecular targets and improved treatment methods [8]. BUB1B is emerging as an important competitor, an essential protein that regulates mitotic chromosome positioning, which, in case of its dysfunction in tumors, leads to uncontrolled development and chromosomal abnormalities [4], [5]. Studies indicate that there is a

Back to Results

Need Full-Text

access to IEEE Xplore for your organization?

CONTACT IEEE TO SUBSCRIBE >

More Like This

Renewable energy cost-benefit analysis under Malaysian feed-in-tariff

2012 IEEE Student Conference on Research and Development (SCoReD)

Published: 2012

SWOT Analysis of Feed-In Tariff Policy in Indonesia

2022 International Seminar on Intelligent Technology and Its Applications (ISITIA)

Published: 2022

Show More

DIGITAL TRANSFORMATION: Moving Toward a Digital Society

eLEARNING COURSE PROGRAM

LEARN MORE

IEEE

Feedback

44

Provisional Acceptance Notification – ICETEA 2026

External Inbox x



Microsoft CMT <noreply@msr-cmt.org>

Mar 28, 2026, 4:43 PM



to me ▼

Dear Author,

We extend warm greetings on behalf of Team ICETEA-2026.

Paper ID: 1217: Integrated Bioinformatics Analysis of Differentially Expressed Genes in Multiple Sclerosis

Thank you for submitting your manuscript to the 2nd IEEE International Conference on Emerging Technologies in Engineering Applications (ICETEA-2026).

It is with great pleasure to inform you that your paper has been provisionally accepted for presentation at the IEEE International Conference on Emerging Technologies in Engineering Applications (ICETEA-2026), Sri Manakula Vinayagar Engineering College, Puducherry, India





DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis: “**Computational Identification of biomarker and therapeutic targets in neurodegenerative and brain disorders**” Total Pages: 50 Name of the Scholar **Smita Yadav**
(24/IBT/13).

Supervisor

Prof. Pravir Kumar Department of Biotechnology

This is to report that the above thesis was scanned for similarity detection. Process and outcome are given below:

Software used: **Turnitin**, Similarity Index: **3%**, Total Word Count: **11,592**

Date:

Candidate's Signature

Signature of Supervisor

Pravir Kumar
NEW THESIS

Nishita Singh

Document Details

Submission ID
trncoid::27535:138630813

Submission Date
May 12, 2026, 11:46 AM GMT+5:30

Download Date
May 12, 2026, 11:59 AM GMT+5:30

File Name
NEW THESIS.docx

File Size
4.3 MB

32 Pages
9,214 Words
55,382 Characters



Page 1 of 36 - Cover Page

Submission ID trncoid::27535:138630813



Page 2 of 36 - Integrity Overview

Submission ID trncoid::27535:138630813

3% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text
- Small Matches (less than 10 words)

Match Groups

- 26 Not Cited or Quoted 3%
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%
Matches that are still very similar to source material
- 0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 2% Internet sources
- 1% Publications
- 2% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.



Page 2 of 36 - Integrity Overview

Submission ID trncoid::27535:138630813

Smita Yadav

+91-8429258291

| callsmitayadav@gmail.com

| [LinkedIn](#)

ACADEMIC DETAILS

Degree/Qualification	University	Year	CGPA
Master in Industrial Biotechnology	Delhi Technological University, New Delhi	2024-26	7.55
B.Tech (Biotechnology)	KMCLU, Lucknow	2020-24	8.45
Higher Secondary Education	Umanath Singh Higher Secondary School	2018-20	7.0
Secondary Education	Prasad International School, Jaunpur	2017-18	7.5

EXPERIENCE

- **Hands-on Industrial Training in Specialized Biotech Techniques**, MRD Life Sciences *July 2025*
 - Gained intensive training in molecular biology, microbiology, and analytical biotechnology techniques.
 - Executed core experiments: DNA/RNA isolation, cloning, PCR/Real-Time PCR, SDS-PAGE, and Western blotting for genetic and protein analysis.
 - Operated advanced instruments (ELISA reader, gel documentation system, HPLC) and performed downstream processing including chromatography, dialysis, and salt precipitation.
- **Food Quality Assurance & Laboratory Testing**, RFRAC, Lucknow *June – August 2023*
 - Performed physicochemical testing of food samples (moisture, pH, fat, ash).
 - Conducted microbial analysis and bakery product quality checks.
 - Operated lab instruments and supported QA workflows.
- **Advanced Biotechnology Instrumentation**, Cytogene R&D, Lucknow *Oct – Nov 2022*
 - Practiced microbiology techniques: culturing, staining, sterilization, and aseptic handling.
 - Performed molecular biology techniques: DNA isolation, PCR, and agarose gel electrophoresis.
 - Learned to operate and handle advanced laboratory instruments for molecular research.

PROJECTS

- **ESKAPE Pathogens – Antibiotic Resistance Mechanisms**, CSIR-IMTECH Chandigarh *Jan – June 2024*
 - Conducted a six-month research project analyzing resistance mechanisms in clinically significant ES-KAPE pathogens.
 - Performed bacterial isolation, spotting, spreading, colony picking, and screening.
 - Designed and executed spotting assays using *Staphylococcus aureus* resistant isolates.
 - Contributed to understanding bacterial resistance phenotypes for potential antimicrobial strategies.

PUBLICATIONS

- **Computational Identification of BUB1B Targeting Phytochemical for GBM Therapy Using Molecular Docking** *IEEE IATMSI, 2026*
Molecular Docking, Computational Drug Discovery
- **Integrated Bioinformatics Analysis of Differentially Expressed Genes in Multiple Sclerosis** *IEEE ICETEA, 2026*
Bioinformatics, Differential Gene Expression Analysis

SKILLS

- **Laboratory & Microbiology:** Aseptic techniques, bacterial culture & isolation, media preparation, staining, sterilization.
- **Molecular Biology:** DNA/RNA extraction, PCR, Real-Time PCR, cloning, restriction digestion, transformation.
- **Protein Analysis:** SDS-PAGE, Western blotting, ELISA.

- **Analytical Techniques:** HPLC, chromatography, ELISA reader, gel documentation system.
- **Other Competencies:** Laboratory record keeping, data interpretation, report preparation, teamwork in research environments.

CERTIFICATIONS & WORKSHOPS

- Six-month dissertation at CSIR IMTECH (Chandigarh).
- One-month training in Advanced Biotechnology Instrumentation at Cytogene R&D.
- Two-day workshop on Vaccine Technology.
- National Conference on Environmental and Industrial Biotechnology (NCEIB-2020).
- 45-day training in Food Standards & Quality Control at RFRAC Lucknow.

Undertaking

I am taking full responsibility for this project work and the thesis written for my dissertation. This thesis is neither copied from any place nor written in an AI/ChatGPT-assisted manner. In case of any discrepancies or falsification of information, I understand that my degree may be cancelled.

I will take responsibility for the dissertation. If something is found wrong, take an appropriate action.

Name of Student:

Roll No.

Signature: