



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: IV Month of publication: April 2023

DOI: <https://doi.org/10.22214/ijraset.2023.50885>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

NGS and Sequence Analysis with Biopython for Prospective Brain Cancer Therapeutic Studies

Uma Kumari¹, Shruti Gupta²

^{1, 2}Bioinformatics Project and Research Institute, Noida - 201301, India

Abstract: Malignant primary brain tumour in adults and in analysis driven by various genomic alterations. Next generation sequencing (NGS) provides about the genetic landscape of tumours and might detect targetable mutations. Genomic DNA was extracted from blood samples or tumor tissues number of biological pathways in cancer involve the products of genes that are not mutated, but epigenetically when we found regulated, for example by altered transcription factor availability, repressor activity or gene methylation receptor-type protein tyrosine phosphatases (RTPs) comprises PTPRZ and PTPRG, primary human glioblastomas suggested a close association between PTPRZ1 (human PTPRZ) expression and cancer stemness. The functional roles of PTPRZ activity in glioma that sphere-forming cells in human U251 glioblastoma cell lines showed high expression levels of PTPRZ-B. Potential PTPRZ1-MET fusion, that is caused when PTPRZ1 and the MET oncogene are fused in 8% to 16% of gliomas and is linked to a less favorable outcome nevertheless is also a promising biomarker for kinase inhibitor therapy. In accordance to the studies that are currently available, receptor-type tyrosine-protein phosphatase zeta (PTPRZ1) is found in various tumour tissues and plays a role in migration, cell proliferation, the EMT and adhesions, as well as treatment resistance by binding with a variety of substances.

Keywords: Biopython; Somatic mutations; Structural analysis; Sequence alignment; Glioblastoma;

I. INTRODUCTION

A brain tumour is an uncontrolled mass of cell development in or near the brain. If a brain tumour becomes large enough to press against nearby nerves, blood vessels, or tissues, it may impair brain function. Two to three percent of malignant tumours are brain tumours, which account for 85 to 90 percent of all primary CNS tumours and up to 35 percent of malignant tumours with a five-year survival rate of about 90% for benign tumours (Neugut AI et al., 2019). Adults with brain tumours are diagnosed between the ages of 40 and 70 (Fox BD et al., 2011). According to the International Association of Cancer Registries (IARC), over 28,000 cases of brain tumours are recorded in India every year, and over 24,000 people are estimated to die from brain tumours. There will be 24,810 diagnoses of malignant brain tumours (14,280 in men and 10,530 in women). In this paper we mainly focus onto the glioblastoma which is the grade IV tumor. The most prevalent and aggressive primary malignant tumour of the brain and central nervous system, glioblastoma multiforme (GBM), accounts for 14.5% of all central nervous system tumours and 48.6% of malignant central nervous system tumours (Koshy et al., 2012). According to the World Health Organization (WHO) classification, which classifies GBM as a grade-IV cancer, gliomas are identified by their nomenclature and are diagnosed using this system globally. Based on histological criteria that identify four different forms of this tumour (Louis et al., 2016).

- 1) Glioblastoma, IDH-mutant
- 2) Not-elsewhere-classified (NEC) Glioblastoma
- 3) Glioblastoma, isocitrate dehydrogenase (IDH) wildtype
- 4) Glioblastoma not otherwise specified (NOS)

Isocitrate dehydrogenase's affinity for the common substrate, isocitrate, is decreased by the mutation of IDH1 and IDH2, although its affinity for alpha-ketoglutarate is increased. As a result, alpha-ketoglutarate is transformed into the cometabolite 2-hydroxyglutarate, which is thought to prevent cell differentiation by competitively inhibiting the enzymes that metabolise alpha-ketoglutarate. As a result, DNA as well as histone demethylation actions of dioxygenases are blocked (Reiter-Brennan et al., 2018) henceforth causes hypoxia-inducible factor 1 to become active (HIF-1) (Fu et al., 2012). It is now understood that low-grade glioma development is primarily influenced by these epigenetic alterations (Reiter-Brennan et al., 2018). Expansion of the epidermal growth factor receptor (EGFR) and loss of function of both the phosphatase as well as tensin homolog (PTEN) are two additional genetic changes that are required for secondary GBM to develop from low-grade gliomas including an IDH alteration (Lai et al., 2011).

Moreover, the categorization lists the following as criteria for the diagnosis of GBM, IDH-wildtype: TERT promoter mutation, EGFR gene amplification, simultaneous acquisition of the complete chromosome 7 as well as deletion of the entire chromosome 10 (+7/-10). Hence, if an adult IDH-wildtype diffuse & astrocytic glioma exhibits microvascular overgrowth, TERT promoter mutations, necrosis, EGFR gene amplification, or +7/10 chromosome copy number variations, GBM, IDH-wildtype must be suspected. Nonetheless, a variety of pediatric-type gliomas ought to be investigated to diagnose paediatric patients with IDH-wildtype diffuse astrocytomas (Louis, D.N et al., 2021). The proangiogenic and inflammatory milieu that glioblastoma produces causes endothelial cells to express more adhesion molecules and have fewer tight junctions, which results in a highly porous blood-brain barrier (BBB). These modifications enable leukocytes to extravasate through the endothelial wall of the brain and enter the tumour mass after leaving the blood stream. Particularly, GBM exhibits a nuanced interaction between immune control, tumor-induced immunodeficiency, and cancer occurrence. Despite an immune infiltrate, the tumour microenvironment in GBM is extremely immunosuppressive, which promotes recurrence and a poor prognosis (Gieryng et al., 2017). The environment that surrounds cancer cells is known as the tumour microenvironment. This consists of extracellular matrix, stromal, vascular, plus immune cells as well as any substances that are secreted. In the immunological infiltrate, lymphocyte, macrophages, as well as microglia make up the majority of the cells. A significant portion of the tumour mass is made up of macrophages and microglia, whose phenotypes and activities have been profoundly altered by the tumour cells (Hambardzumyan et al., 2016). Because of their pro - angiogenic and immunosuppressive characteristics, tumor-associated macrophages are frequently thought to promote tumour growth. MDSCs in mice can be broadly categorised into monocytic as well as granulocytic subtypes and are cells that exhibit both CD11b & Gr1 surface markers. Granulocytic MDSCs are infrequently detected in GBM tumours (Chen Z et al., 2017). Many immune cell subtypes have been seen in glioblastoma tumours, although immunosuppressive cells predominate. Current research has refined this idea further by identifying different Treg and dendritic cell subsets as well as subpopulations inside the immunoregulatory macrophage sector of malignancies. In orthotopic animal models, GBM-secreted IL-6 produces immunosuppressive peripheral myeloid cells that exhibit elevated amounts of PD L1 (Lamano et al., 2019). This result is consistent with peripheral blood CyTOF (Cytometry by Time of Flight) fingerprinting of patients, which shows a large rise of myeloid-derived suppressor cells (MDSC), but also not Treg, in circulation (Alban et al., 2018). As a result, it's possible that the myeloid cells that invade GBM tumours have already been trained to promote tumour growth. The discovery that the presence of MDSC in mice GBM tumours is linked to a decrease in the number of tumor-infiltrating lymphocytes ((Raychaudhuri et al., 2015) further supports the significance of MDSC, both of the monocytic or granulocytic subtype (Movahedi et al., 2008). The existence of Treg has been extensively discussed in a variety of cancer types, although its usefulness in predicting the course of the disease in glioblastoma is in question. Independent researchers have noted a rise in the quantity of FOXP3+ Treg in glioblastoma and other higher grade brain tumours (Jacobs et al., 2010). These Treg may be drawn to the tumour microenvironment through a number of pathways, such as IDO activity (Wainwright et al., 2012) as well as CCL2 synthesis, that reacts with CCR4 over the membrane of Treg (Chang et al., 2016). It's interesting to note that CCR8 is a different chemokine receptor that has recently been discovered as an indicator that is clearly expressed in at least a sub - set of tumor-infiltrating Treg, and not Treg in the periphery, as aspect of a tumor-infiltrating Treg transcriptional signature which is conserved throughout species and tumour forms (Magnuson et al., 2018). In GBM immune checkpoints, Negative regulatory mechanisms known as immune checkpoints prevent the activation and growth of T cells. On T cells, inhibitory functions are frequently carried out by the cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4), programmed cell death-1 receptor (PD-1), and T-cell inhibitory receptor (TIM-3) via interaction with the respective ligands. PD-1 inhibits T cell proliferation and effector functions by interacting with its ligand PD-L1, which is produced by stromal cells like APCs. This inhibition can be reversed by employing antibodies that are specific for PD-L1. Moreover, the expression of PD-1 here on membrane of these T cells is a sign indicating functional exhaustion (Barber DL et al., 2006). By coercing the microenvironment's cells, such as TAMs, to express large levels of PD-L1, tumours have developed to hijack this mechanism for their own gain. Recent research has shown that GBM has much more PD-1+ tumor-infiltrating lymphocytes and PD-L1 expression. This supports the idea that immune checkpoint inhibition, which prevents the PD-1/PD-L1 axis from working, could be used as a viable treatment for GBM (Preusser M et al., 2015). Recent research has shown that nanoparticles have the ability to transport therapeutic payload to the mouse brain. After a low-dose irradiation to promote brain absorption, solid lipid nanoparticles can deliver short interfering RNAs to the mouse glioblastoma microenvironment (Erel-Akbaba G et al., 2019). As a GBM targeting ligand for systemically delivered nanoparticles, LinTT1 (AKRGARSTA), a new tumour penetrating peptide that binds cell surface p32, has also been reported (Säälük et al., 2019). Moreover, nanoscale immunoconjugates (NICs) with covalently bonded anti-CTLA-4 or anti-PD-1 were created on a natural biopolymer scaffold called poly(-L-malic acid) enabling systemic distribution throughout the BBB enabling stimulation of local brain anti-tumor immune responses (Galstyan et al., 2019).

(Ha Young Woo et.al, 2020), postulates that diffuse gliomas have been discovered to include oncogenic gene fusions, which could be used as therapeutic targets. Therefore, we evaluated the clinical, pathological, as well as genetic characteristics of gene fusion in 356 diffuse gliomas obtained from 2017 to 2019 utilise next generation sequencing analysis (Illumina TruSightTumor 170 panel). The relevant oncogenic gene fusions were discovered in 53 cases of glioblastomas: MET (n = 18), FGFR (n = 12), AKT3 (n = 1), NTRK (n = 5), EGFR (n = 14), RET (n=2), and PDGFRA fusions (n = 1). Both IDH-wildtype but also IDH-mutant glioblastomas consistently exhibited gene fusions (8.8% and 9.4%, $p = 1.000$). The only fusion with genetic characteristics similar to secondary glioblastomas (i.e., highest incidence of IDH mutation, ATRX absence, TP53 mutation, as well as loss of EGFR) was PTPRZ1-MET fusion while primary glioblastomas remained similar to certain other gene fusion kinds (i.e., highest amount of IDH-wildtype, PTEN mutation, TERT mutation and EGFR augmentation). Multivariable analysis indicates that perhaps the PTPRZ1-MET fusion significantly linked to a poor progression-free survival among individuals with IDH-wildtype glioblastoma. (Veronica Vellani et al., 2023), illustrated that the purpose of this study was to describe the genetic makeup of glioma patients and to talk about how next-generation sequencing affects glioma detection and treatment. In a bid to better comprehend the existence of a significant relation among non-morphological information on brain MRI acquired with diffusion as well as perfusion methods and molecular data, the research sets out to characterise the molecular profiles of patients with gliomas. It also aimed to establish an entirely novel concept for molecular diagnostics relying on NGS analysis. Molecular evaluation was carried out utilising the Oncomine Focus Assay and an NGS methodology (Thermo Fisher Scientific). The above targeted sequencing test enables the detection of hotspots, copy number variants (CNVs), single nucleotide variations (SNVs), insertions and deletions, and gene fusions in 52 genes having known significance in solid tumours. The assay enables simultaneous evaluation of DNA and RNA together in single procedure. Furthermore, they gathered information on the demographics and clinical characteristics of the patients, as well as information on prior treatments and their effects on overall survival (OS) as well as progression-free survival (PFS), or time to progression out from NGS analysis (TTP). Through automatically uploading required data out from Torrent Suite™ Software, sequences were matched to a hg19 reference genome as well as variant calling was carried out by utilizing Ion Reporter version 5.4. Oncomine™ Focus—520—w2.4—DNA plus Fusions—Single Sample has been the workflow analysis. This approach identifies and annotate gene fusions using targeted RNA libraries or low-frequency somatic variations (SNPs, Insertions and deletions, and CNVs) using targeted DNA libraries for the Oncomine™ Focus Test. When we compared the prevalence of SNVs according to the kind of tumour, we discovered that GBM patients had much less of them (31.6%) than other glioma patients (68.4%). In contrast to GBM, other gliomas have a higher prevalence of the IDH1 mutation. Several low-frequency variants, like IDH2, BRAF, as well as PIK3A, were more prevalent in GBM patients although small EGFR variants were found among patients having other gliomas. In comparison to other gliomas (17.1%), the percentage of CNVs considerably greater in GBM (82.9% among all CNVs). In instance, individuals with GBM had more EGFR (22/26) as well as CDK4 (6/8) alterations compared to individuals who had other gliomas, but PDGFRA CNVs have only been found in people with GBM (3/3). Pralsetinib (BLU-667), a strongly RET inhibitor being researched for the treatment of various solid cancers, was administered to one person with RET fusion in a stage 1/2 study (NCT03037385).

II. METHODOLOGY

We have used a variety of databases and technologies in this study to better understand my glioblastoma sample (5H08). We employed NGS TOOLS and technologies on the sample because our paper is focused on NGS (Next Generation Sequencing) and biopython. Both databases and technologies are utilised to identify the sample's phenotypic and genotypic characteristics. Moreover, tools are utilised to visualise the sample's graphical representation in terms of similarity and other factors. For instance, Protein Data Bank (PDB), Molecular Modelling Database (MMDB), Basic Local Search Alignment Tool (BLAST), Constraint Based Multiple Alignment Tool (COBALT), PDBsum and Biopython is utilised to analyse the sample receptor type tyrosine protein phosphatase zeta 1 (PTPRZ1) (5H08).

- 1) *Protein Data Bank (PDB)* - The Protein Data Bank (PDB) is a repository for the three-dimensional structures of nucleic acids, proteins as well as other biological macromolecules that have been experimentally identified. In order to preserve experimental data generated by the emerging field of macromolecular crystallography, which was beginning to disclose three-dimensional structures, the PDB was created in 1971. Nowadays, the PDB also stores atomic coordinates and associated experimental data from electron microscopy and NMR spectroscopy. It is crucial for comprehending important branches of research, including as basic biology, energy, medicine, drug development, and education. The US PDB Data Center has existed since 1999 and is known as the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB). More than 6,200 novel atomic level biomolecular structure, experimental data, or metadata from PDB Data Depositors throughout the Americas and Oceania were processed in 2017 by RCSB PDB.

- 2) *Molecular Modelling Database (MMDB)* -MMDB, commonly known as the NCBI structure database, is a collection of empirically determined three-dimensional biological structures. It is a subset of three-dimensional structures derived from the RCSB protein data repository and is additionally known as the Entrez Structure database. MMDB, is also known as micromolecular database. Each MMDB record includes a cross-reference to the PDB record that served as its original source.
- 3) *Basic Local Alignment Search Tool (BLAST)* - The Basic Local Alignment Search Tool (BLAST) identifies areas of sequence similarity. The software compares protein or nucleotide sequences and determines the statistical importance of matches. In addition to assisting in the identification of gene family members, BLAST can be used to infer functional and evolutionary links between sequences. BLAST is a computer algorithm that is available at NCBI for online use. Since BLAST is a paradigmatic of local alignment, end-to-end alignment is not possible. There are various BLAST search types. Four basic search types are available via NCBI's WebBLAST: Blastn, Blastp, Blastx, tBlastn. Using a heuristic approach that resembles the Smith-Waterman algorithm, BLAST looks for higher scoring sequence alignments between both the query sequence as well as the available sequences in the database.
- 4) *Constraint Based Multiple Alignment tool (COBALT)*- Using RPS-BLAST, BLASTP, and PHI-BLAST, the multiple sequence alignment programme COBALT discovers a set of pairwise constraints generated from conserved domain databases, protein motif databases, and sequence similarity. An incremental multiple alignment is then added pairwise constraints. By grouping together sequences that have a lot of words in common and finding conserved domains or motif similarities for only single sequence for each cluster, computation time gets decreased.
- 5) *PDBsum*- Laskowski and colleagues at the European Bioinformatics Institute manage PDBsum in Janet Thornton's lab (EBI). A visual or pictorial database called PDBsum gives a quick overview of the information contained in each 3D structure that has been deposited in the Protein Data Bank (PDB). PDBsum is just a web service that offers structural details on Protein Data Bank entries (PDB). Either the two search boxes on the PDBsum home page, or the entry's 4-character PDB code, can be used to reach it. The investigations, which are mostly image-based, include a wide range of topics, such as protein secondary structure, interactions with protein ligands and protein DNA, structural quality PROCHECK assessments (Ramachandran plot analysis) and more.
- 6) *Biopython* - A multinational team of developers created the collection of free tools known as Biopython for biological computation. The development of Python libraries and apps to meet the demands of present and upcoming bioinformatics work is a worldwide collaborative effort. It could read and write to a wide range of file formats, and it has categories to describe biological sequences plus sequence annotations. Additionally, it enables programmatic access to internet biological data databases such as those at NCBI. Sequence alignment, population genetics, phylogenetics, protein structure, sequence motifs, and machine learning are added as separate modules to Biopython's repertoire. Google offers collaborative research laboratories where biopython studies can be carried out.

III. RESULT

- 1) *Protein Data Bank Analysis* - We retrieved the 3D structure of our sample from the PDB and determined the ligand that was involved in it. The ligand 3-[2-Ethoxy-5-(trifluoromethyl)benzyl]sulfanyl-N-(phenylsulfonyl)thiophene-2-carboxamide is found in 5H08 and has the identification number 7WL.

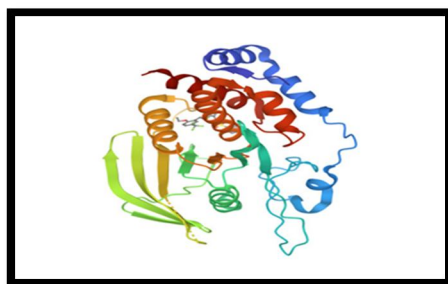


Fig 1: showing the 3D structure of human Receptor-type tyrosine-protein phosphatase zeta(5H08)

- 2) *Molecular Modelling Database (MMDB) Analysis* - Using the MMDB, we have turned a four-letter code into a sequence, or we may consider that doing so transforms a structure into a sequence. and we identify the region of the amino acid sequence that functions as the active site and the catalytic site.

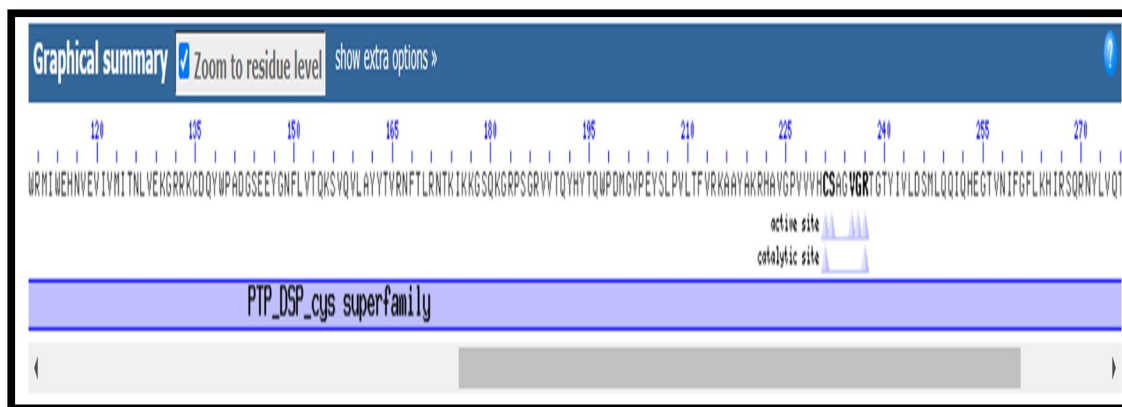


Fig 2: representing the active and catalytic site present in the sequence. active site, on conserved domain R-PTPc-G-1. 5 of 5 of the residues that compose this conserved feature have been mapped to the query sequence and catalytic site, on conserved domain R-PTPc-G-1. 2 of 2 of the residues that compose this conserved feature have been mapped to the query sequence.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

>pdb|5H08|A Chain A, Receptor-type tyrosine-protein phosphatase zeta

GPAIPKHFPHKADLHASSGFTEEFEEVQSCTVDLGITADSSNHPDNKHKNRYINIVAYDHSRVKLAQL
AEKDGKLTIDYINANYVDGYNRPKAYIAAQGLKSTAEDEFWRMIWEHNVEVIVMITNLVEKGRRKCDQYWP
ADGSEEYGNFLVTQKSVQVLAAYTVRNFTLRNTKIKKGSQKGRPSGRVVTQYHYTQWPDMPGVPEYSLPVL
TFVRKAAYAKRHA VGPVVVHCSAGVGRGTGTIVLDSMLQQIQHEGTVNIFGFLKHRSQRNYLVQTEEQY
VFIHDTLVEAILSKETEV

3) *Basic Local Alignment Search Tool (BLAST) Analysis* - I utilize BLAST to determine sequence similarity and even to obtain an alignment of my query sequence with certain other species. Here, I utilised BLAST to generate multiple sequence alignment of the sample as well as a graphical representation of my data in comparison to other species.

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	protein tyrosine phosphatase receptor type Z1 [Homo sapiens]	Homo sapiens	633	745	99%	0.0	100.00%	2347	KAI4015577.1
<input checked="" type="checkbox"/>	receptor-type tyrosine-protein phosphatase zeta isoform X2 [Nomascus leucogenys]	Nomascus leuco...	633	745	99%	0.0	100.00%	2308	XP_030681373.1
<input checked="" type="checkbox"/>	protein tyrosine phosphatase receptor type Z1 [Homo sapiens]	Homo sapiens	633	745	99%	0.0	100.00%	2347	KAI2547646.1
<input checked="" type="checkbox"/>	receptor-type tyrosine-protein phosphatase zeta isoform X2 [Hylobates moloch]	Hylobates moloch	633	745	99%	0.0	100.00%	2307	XP_032617472.1
<input checked="" type="checkbox"/>	receptor-type tyrosine-protein phosphatase zeta isoform 4 precursor [Homo sapiens]	Homo sapiens	633	744	99%	0.0	100.00%	2308	NP_001356324.1
<input checked="" type="checkbox"/>	receptor-type tyrosine-protein phosphatase zeta isoform X2 [Pan troglodytes]	Pan troglodytes	633	744	99%	0.0	100.00%	2308	XP_009452364.1

Fig 3: representing the sequence similarity in terms of statistical theory.

4) *Constraint Based Multiple Alignment Tool (COBALT) Analysis* - Cobalt is utilised to depict the sequence of our interest in various colours, where each colour stands for a distinct identity.

We obtain two outcomes from the COBALT in the following ways:

Rasmol and shapely to determine the graphical view of the sample protein tyrosine phosphatase receptor type zeta1 (PTPR Z1) in homo sapiens.

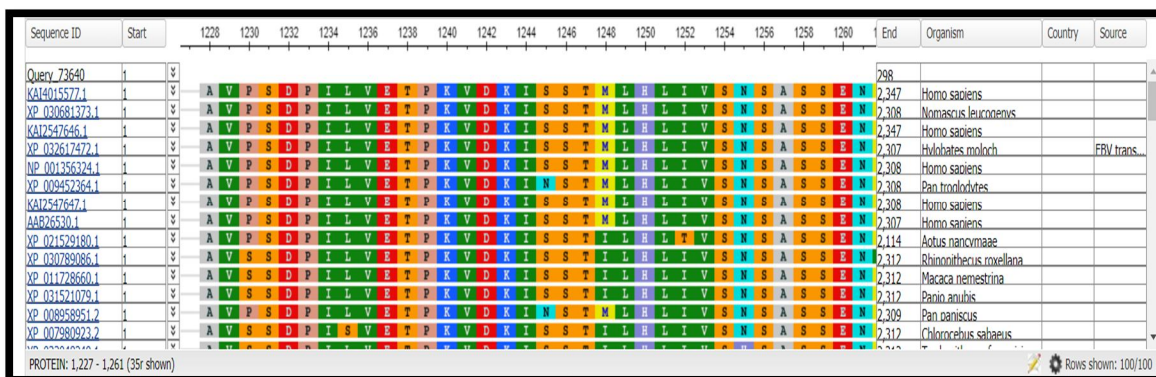


Fig 4: Rasmol displays the colors unique to each amino acid. For instance, alanine is in grey colour, valine is in dark green colour, asparagine is in sky blue colour and methionine is in yellow colour.

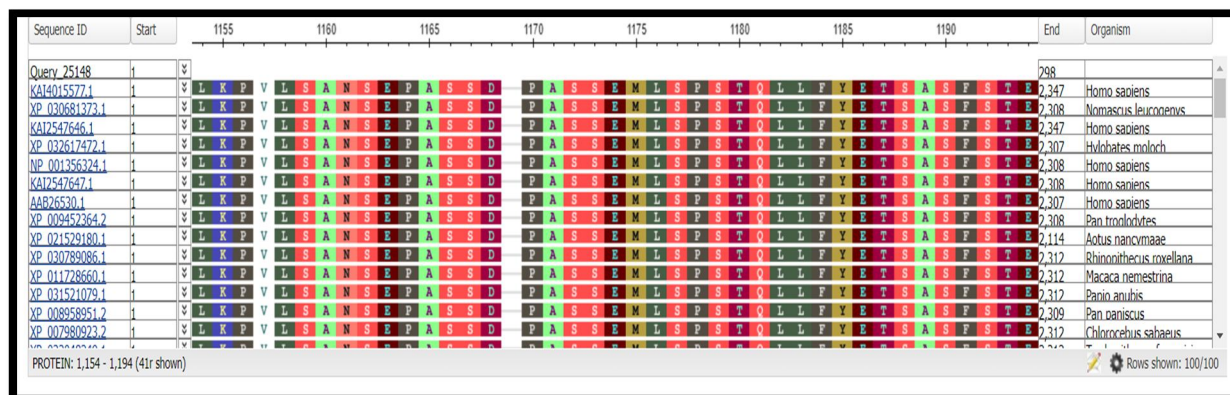


Fig 5: Shapely utilizes multiple colour schemes for amino acids based on the periodic table. Here, grey line represents gap. Shapely represents alanine in light green colour, valine is in white colour, asparagine is in orange green colour and methionine is in light brown colour.

5) *PDBsum Analysis* – We utilised PDBsum to determine the PROCHECK analysis of our sample 5H08.

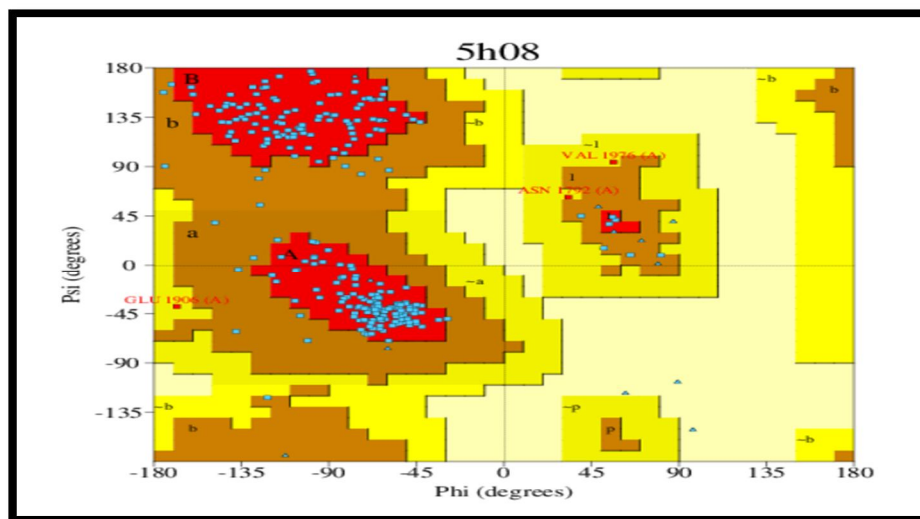


Fig 6: PROCHECK analysis (Ramachandran plot of PTPRZ1 sample in humans): here, red indicates most favourable region, brown indicates additional allowed region whereas yellow represent generously allowed region as per the values of PROCHECK statistics.

PROCHECK statistics		
1. Ramachandran Plot statistics		
	No. of residues	%-tage
Most favoured regions [A,B,L]	220	87.3%*
Additional allowed regions [a,b,l,p]	29	11.5%
Generously allowed regions [~a,~b,~l,~p]	3	1.2%
Disallowed regions [XX]	0	0.0%
Non-glycine and non-proline residues	252	100.0%
End-residues (excl. Gly and Pro)	3	
Glycine residues	16	
Proline residues	11	
Total number of residues	282	

Fig 7: representing the PROCHECK statistics of PTPRZ1(5H08) sample.

6) *Biopython Analysis* - To grasp our sample in a coding approach, we utilized biopython.

Our sample Human sapiens microRNA 3928 (MIR3928), microRNA (NR 037496.1), has undergone multiple alignment, analysis, and bio import.

```
from Bio.Seq import Seq

my_dna = Seq("GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGGAGGAACCTTGGAGCTTCGGC")

print(my_dna)

GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGGAGGAACCTTGGAGCTTCGGC

print(my_dna.reverse_complement())

print(my_dna)

GCCGAAGCTCCAAGGTTCTCCGCTTCTGCCCCAGGCGGAACCTTAGAGCTTCAGC
GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGGAGGAACCTTGGAGCTTCGGC
```

```
from Bio.Seq import Seq

messenger_rna = Seq("GCUGAAGCUUAAGGUUCCGCGGGCAGGAAGCGGAGGAACCUUGGAGCUUCGGC")

messenger_rna.translate()

/usr/local/lib/python3.9/dist-packages/Bio/Seq.py:2804: BiopythonWarning: Partial codon, len(sequence) not a multiple of three. Explicitly trim the sequence or add trailing N before tr
warnings.warn(
Seq('AEALRFRLRAGSGGTLELR')
```

Fig 8: above results representing the bio import sequence analysis of Human sapiens microRNA 3928 (MIR3928) sample.


```
from Bio.Seq import Seq
from Bio.SeqRecord import SeqRecord
from Bio.Align import MultipleSeqAlignment
from Bio.codonalign import build

seq1 = SeqRecord(Seq('GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGAGGAACCTTGGAGCTTCGGC'), id='pro1')
seq2 = SeqRecord(Seq('GCTGAAGCTCTGTTCCGCTGCGGGCAGGAAGCGAGGAACCGAGCGGC'), id='pro2')
pro1 = SeqRecord(Seq('AEALRFRLRAGSGGTLELR'), id='pro1')
pro2 = SeqRecord(Seq('AEAL-FRLRAGRG-T-E-R'), id='pro2')
aln = MultipleSeqAlignment([pro1, pro2])
print(seq1, pro1, aln)
print("=====")
print(seq2, pro2, aln)
print("=====")
codon_aln = build(aln, [seq1, seq2])
print(codon_aln)
```

Fig 9: The aforementioned code is used to align the sample's multiple sequences. We used two sequences of human glioblastoma to align the nucleotide and protein sequences in this.

```
ID: pro1
Name: <unknown name>
Description: <unknown description>
Number of features: 0
Seq('GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGAGGAACCTTGGAGCTTCGGC') ID: pro1
Name: <unknown name>
Description: <unknown description>
Number of features: 0
Seq('AEALRFRLRAGSGGTLELR') Alignment with 2 rows and 19 columns
AEALRFRLRAGSGGTLELR pro1
AEAL-FRLRAGRG-T-E-R pro2
*==*==*==*==*==*==*==*==*==*
ID: pro2
Name: <unknown name>
Description: <unknown description>
Number of features: 0
Seq('GCTGAAGCTCTGTTCCGCTGCGGGCAGGAAGCGAGGAACCGAGCGGC') ID: pro2
Name: <unknown name>
Description: <unknown description>
Number of features: 0
Seq('AEAL-FRLRAGRG-T-E-R') Alignment with 2 rows and 19 columns
AEALRFRLRAGSGGTLELR pro1
AEAL-FRLRAGRG-T-E-R pro2
*==*==*==*==*==*==*==*==*==*
CodonAlignment with 2 rows and 57 columns (19 codons)
GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGAGGAACCTTGGAGCTTCGG pro1
GCTGAAGCTCTG---TTCCGCTGCGGGCAGGAAGAGGA---ACC---GAG---CGG pro2
```

Fig 10: This is the outcome we receive when we use the prior code.

```
from Bio.Seq import Seq
list_of_seqs = [Seq("GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGAGGAACCTTGGAGCTTCGGC"), Seq("GCTGAAGCTCTAAGGTTCCGCGGAGGAACCTTGGAGCTTCGGC")]
concatenated = Seq("")
for s in list_of_seqs:
    concatenated += s

print(s)

GCTGAAGCTCTAAGGTTCCGCGGAGGAACCTTGGAGCTTCGGC

print(list_of_seqs)

[Seq('GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGAGGAACCTTGGAGCTTCGGC'), Seq('GCTGAAGCTCTAAGGTTCCGCGGAGGAACCTTGGAGCTTCGGC')]

print(concatenated)

GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGAGGAACCTTGGAGCTTCGGCGCTGAAGCTCTAAGGTTCCGCGGAGGAACCTTGGAGCTTCGGC
```

Fig 11: Here the concatenated code is utilised for the human glioblastoma microRNA sample.

3326


```

description: RecName: Full=Transmembrane glycoprotein NMB; AltName: Full=Hematopoietic growth factor inducible neurokinin-1 type {ECO:0000303|PubMed:126097}
authors: Weterman M.A.J., Ajubi N., van Dinter I.M.R., Degen W.G.J., van Muijen G.N.P., Ruiter D.J., Bloemers H.P.J.
title: NMB, a novel gene, is expressed in low-metastatic human melanoma cell lines and xenografts.
authors: Bandari P.S., Qian J., Yehia G., Joshi D.D., Maloof P.B., Potian J., Oh H.S., Gascon P., Harrison J.S., Rameshwar P.
title: Hematopoietic growth factor inducible neurokinin-1 type: a transmembrane protein that is similar to neurokinin 1 interacts with substance P.
authors: Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J.S., Grimaldi C.,
title: The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.
authors: Scherer S.W., Cheung J., MacDonald J.R., Osborne L.R., Nakabayashi K., Herbrick J.-A., Carson A.R., Parker-Katiraei L., Skaug J., Khaja R., Zhang J., Hudek A.K., Li M., Haddad
title: Human chromosome 7: DNA sequence and biology.
authors: Mural R.J., Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fa
title:
authors: Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H., Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R., Wyllie K., Sekhon M., Becker M.C., O'Laughlin M.
title: The DNA sequence of human chromosome 7.
authors: The MGC Project Team
title: The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC).
authors: Basur V., Yang F., Kushimoto T., Higashimoto Y., Yasumoto K., Valencia J., Muller J., Vieira W.D., Watabe H., Shabanowitz J., Hearing V.J., Hunt D.F., Appella E.
title: Proteomic analysis of early melanosomes: identification of novel melanosomal proteins.
authors: Tse K.F., Jeffers M., Pollack V.A., McCabe D.A., Shadish M.L., Khramtsov N.V., Hackett C.S., Shenoy S.G., Kuang B., Boldog F.L., MacDougall J.R., Rastelli L., Herrmann J., Gal
title: CR011, a fully human monoclonal antibody-aurostatin E conjugate, for the treatment of melanoma.
authors: Kuan C.T., Wakiya K., Dowell J.M., Herndon J.E. II, Reardon D.A., Graner M.W., Riggins G.J., Wikstrand C.J., Bigner D.D.
title: Glycoprotein nonmetastatic melanoma protein B, a potential molecular therapeutic target in patients with glioblastoma multiforme.
authors: Chi A., Valencia J.C., Hu Z.-Z., Watabe H., Yamaguchi H., Mangini N.J., Huang H., Canfield V.A., Cheng K.C., Yang F., Abe R., Yamagishi S., Shabanowitz J., Hearing V.J., Wu C.
title: Proteomic and bioinformatic characterization of the biogenesis and function of melanosomes.
authors: Chen R., Jiang X., Sun D., Han G., Wang F., Ye M., Wang L., Zou H.
title: Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry.
authors: Vaca Jacome A.S., Rabilloud T., Schaeffer-Reiss C., Rompals M., Ayoub D., Lane L., Bairoch A., Van Dorsselaer A., Carapito C.
title: N-terminome analysis of the human mitochondrial proteome.
authors: Sjoebloom T., Jones S., Wood L.D., Parsons D.W., Lin J., Barber T.D., Mandelker D., Leary R.J., Ptak J., Silliman N., Szabo S., Buckhaults P., Farrell C., Meeh P., Markowitz S.
title: The consensus coding sequences of human breast and colorectal cancers.
authors: Yang C.F., Lin S.P., Chiang C.P., Wu Y.H., H'ng W.S., Chang C.P., Chen Y.T., Wu J.Y.
title: Loss of GPNMB Causes Autosomal-Recessive Amyloidosis Cutis Dyschromica in Humans.
classification: ['Eukaryota', 'Metazoa', 'Chordata', 'Craniata', 'Vertebrata', 'Euteleostomi', 'Mammalia', 'Eutheria', 'Euarchontoglires', 'Primates', 'Haplorrhini', 'Catarrhini', 'Hom

```

Fig 14: Figure showing the comprehensive examination of our sample of human glioblastoma to obtain information on the description, author, title, and classification relative to other organisms.

```

from Bio import motifs
from Bio.Seq import Seq

DNA_motif = [ Seq("GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGGAGAACCT"),

Seq("GCTGAAGCTCTGTTCCGCTGCGGGCAGGAAGAGGAGAACCGGCGC"),

]

seq = motifs.create(DNA_motif)

print(seq)

print(seq.counts)

GCTGAAGCTCTAAGGTTCCGCTGCGGGCAGGAAGCGGAGAACCT
GCTGAAGCTCTGTTCCGCTGCGGGCAGGAAGAGGAGAACCGGCGC

      0      1      2      3      4      5      6      7      8      9      10     11     12     13     14     15     16     17     18     19     20     21     22     23     24     25
A:  0.00  0.00  0.00  0.00  2.00  2.00  0.00  0.00  0.00  0.00  0.00  1.00  1.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
C:  0.00  2.00  0.00  0.00  0.00  0.00  0.00  2.00  0.00  2.00  0.00  0.00  0.00  0.00  1.00  1.00  0.00  2.00  2.00  0.00  1.00  2.00  0.00  0.00  1.00  1.00
G:  2.00  0.00  0.00  2.00  0.00  0.00  2.00  0.00  0.00  0.00  0.00  1.00  1.00  0.00  1.00  0.00  1.00  0.00  0.00  1.00  1.00  0.00  1.00  2.00  1.00  1.00
T:  0.00  0.00  2.00  0.00  0.00  0.00  0.00  0.00  2.00  0.00  2.00  0.00  1.00  1.00  0.00  1.00  1.00  0.00  0.00  1.00  0.00  0.00  1.00  0.00  0.00  0.00

```

Fig 15: Representing the DNA motif sequence code for the sample of homo sapiens microRNA.

IV. CONCLUSION

We obtained the sample via MRI, biopsy or NCBI OR PDB and then further sample analysis began with the computation. Secondly, several tools have been used to extract features. For instance, graphical representation shows the catalytic or active site, sequence similarity and also shows the amino acid residues of the sample via various tools like MMDB, BLAST etc. Additionally, association of genomics have performed to check the translations and mutations. Furthermore, data analysis have been performed by NGS and biopython.

REFERENCES

- [1] Alban TJ, Alvarado AG, Sorensen MD, Bayik D, Volovetz J, Serbinowski E, Mulkearns-Hubert EE, Sinyuk M, Hale JS, Onzi GR, McGraw M, Huang P, Grabowski MM, Wathen CA, Ahluwalia MS, Radivoyevitch T, Kornblum HI, Kristensen BW, Vogelbaum MA, Lathia JD. Global immune fingerprinting in glioblastoma patient peripheral blood reveals immune-suppression signatures associated with prognosis. *JCI Insight*. 2018;3:122264. doi: 10.1172/jci.insight.122264.
- [2] Barber DL, Wherry EJ, Masopust D, Zhu B, Allison JP, Sharpe AH, et al. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature* (2006) 439(7077):682–7. 10.1038/nature04444.
- [3] Chang AL, Miska J, Wainwright DA, Dey M, Rivetta CV, Yu D, Kanojia D, Pituch KC, Qiao J, Pytel P, Han Y, Wu M, Zhang L, Horbinski CM, Ahmed AU, Lesniak MS. CCL2 produced by the glioma microenvironment is essential for the recruitment of regulatory T cells and Myeloid-Derived suppressor cells. *Cancer Research*. 2016;76:5671–5682. doi: 10.1158/0008-5472.CAN-16-0144.
- [4] Chen Z, Feng X, Herting CJ, Garcia VA, Nie K, Pong WW, et al. Cellular and molecular identity of tumor-associated macrophages in glioblastoma. *Cancer Res* (2017) 77(9):2266–78. 10.1158/0008-5472.CAN-16-2310.
- [5] Erel-Akbaba G, Carvalho LA, Tian T, Zinter M, Akbaba H, Obeid PJ, Chiocca EA, Weissleder R, Kantarci AG, Tannous BA. Radiation-Induced targeted Nanoparticle-Based gene delivery for brain tumor therapy. *ACS Nano*. 2019;13:4028–4040. doi: 10.1021/acsnano.8b08177.
- [6] Fox BD, Cheung VJ, Patel AJ, Suki D, Rao G (2011) Epidemiology of metastatic brain tumors. *Neurosurg Clin N Am* 22: 1-6.
- [7] Fu, Y.; Zheng, S.; Zheng, Y.; Huang, R.; An, N.; Liang, A.; Hu, C. Glioma derived isocitrate dehydrogenase-2 mutations induced up-regulation of HIF-1 α and β -catenin signaling: Possible impact on glioma cell metastasis and chemo-resistance. *Int. J. Biochem. Cell Biol.* 2012, 44, 770–775.
- [8] Galstyan A, Markman JL, Shatalova ES, Chiechi A, Korman AJ, Patil R, Klymyshyn D, Tourtellotte WG, Israel LL, Braubach O, Ljubimov VA, Mashouf LA, Ramesh A, Grodzinski ZB, Penichet ML, Black KL, Holler E, Sun T, Ding H, Ljubimov AV, Ljubimova JY. Blood-brain barrier permeable nano immunconjugates induce local immune responses for glioma therapy. *Nature Communications*. 2019;10:3850. doi: 10.1038/s41467-019-11719-3.
- [9] Giering, A.; Pszczolkowska, D.; Walentynowicz, K.A.; Rajan, W.D.; Kaminska, B. Immune microenvironment of gliomas. *Lab. Invest.* 2017, 97, 498–518.
- [10] Hambardzumyan, D.; Gutmann, D.H.; Kettenmann, H. The role of microglia and macrophages in glioma maintenance and progression. *Nat. Neurosci.* 2016, 19, 20–27.
- [11] Jacobs JF, Idema AJ, Bol KF, Grotenhuis JA, de Vries IJ, Wesseling P, Adema GJ. Prognostic significance and mechanism of treg infiltration in human brain tumors. *Journal of Neuroimmunology*. 2010;225:195–199. doi: 10.1016/j.jneuroim.2010.05.020.
- [12] Koshy, M.; Villano, J.L.; Dolecek, T.A.; Howard, A.; Mahmood, U.; Chmura, S.J.; Weichselbaum, R.R.; McCarthy, B.J. Improved survival time trends for glioblastoma using the SEER 17 population-based registries. *J. Neurooncol.* 2012, 107, 207–212.
- [13] Kumari Uma and Choudhary, Ashok Kumar, "Genome Sequence Analysis of Solanum Lycopersicum by Applying Sequence Alignment Method to Determine the Statistical Significance of an Alignment", *International Journal of Bio-Technology and Research (IJBT)*, 2249-6858;9-12, 2016.
- [14] Kumari Uma, & Choudhary, A. K, "Genome sequence analysis of solanum lycopersicum showing the phylogenetic relationship based on multiple sequence alignment and conserved domain proteins", *International journal of advanced biotechnology and research*, 7(4), 2012-2014, 2016.
- [15] Vinita Kukreja, Uma Kumari, "Genome Annotation of Brain Cancer and Structure Analysis by applying Drug Designing Technique", *International Journal of Emerging Technologies and Innovative Research*, 9(5):473-k479, May, 2022.
- [16] Lai, A.; Kharbanda, S.; Pope, W.B.; Tran, A.; Solis, O.E.; Peale, F.; Forrest, W.F.; Pujara, K.; Carrillo, J.A.; Pandita, A.; et al. Evidence for sequenced molecular evolution of IDH1 mutant glioblastoma from a distinct cell of origin. *J. Clin. Oncol.* 2011, 29, 4482–4490.
- [17] Lamano JB, Lamano JB, Li YD, DiDomenico JD, Choy W, Veliceasa D, Oyon DE, Fakurnejad S, Ampie L, Kesavabhotla K, Kaur R, Kaur G, Biyashev D, Unruh DJ, Horbinski CM, James CD, Parsa AT, Bloch O. Glioblastoma-Derived IL6 induces immunosuppressive peripheral myeloid cell PD-L1 and promotes tumor growth. *Clinical Cancer Research*. 2019;25:3643–3657. doi: 10.1158/1078-0432.CCR-18-2402.
- [18] Louis, D.N.; Perry, A.; Wesseling, P.; Brat, D.J.; Cree, I.A.; Figarella-Branger, D.; Hawkins, C.; Ng, H.K.; Pfister, S.M.; Reifenberger, G.; et al. The 2021 WHO Classification of Tumors of the Central Nervous System: A summary. *Neuro-Oncology* 2021, 23, 1231–1251.
- [19] Louis, D.N.; Perry, A.; Reifenberger, G.; von Deimling, A.; Figarella-Branger, D.; Cavenee, W.K.; Ohgaki, H.; Wiestler, O.D.; Kleihues, P.; Ellison, D.W. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A summary. *Acta Neuropathol.* 2016, 131, 803–820.
- [20] Magnuson AM, Kiner E, Ergun A, Park JS, Asinowski N, Ortiz-Lopez A, Kilcoyne A, Paoluzzi-Tomada E, Weissleder R, Mathis D, Benoist C. Identification and validation of a tumor-infiltrating treg transcriptional signature conserved across species and tumor types. *PNAS*. 2018;115:E10672–E10681. doi: 10.1073/pnas.1810580115.
- [21] Movahedi K, Williams M, Van den Bossche J, Van den Bergh R, Gysemans C, Beschin A, De Baetselier P, Van Ginderachter JA. Identification of discrete tumor-induced myeloid-derived suppressor cell subpopulations with distinct T cell-suppressive activity. *Blood*. 2008;111:4233–4244. doi: 10.1182/blood-2007-07-099226.
- [22] Neugut AI, Sackstein P, Hillyer GC, Jacobson JS, Bruce J, et al. (2019) Magnetic Resonance Imaging-Based Screening for Asymptomatic Brain Tumors: A Review. *Oncologist* 24.
- [23] NCT03037385. Phase 1/2 Study of the Highly-selective RET Inhibitor, Pralsetinib (BLU-667), in Participants With Thyroid Cancer, Non-Small Cell Lung Cancer, and Other Advanced Solid Tumors (ARROW).
- [24] Preusser M, Lim M, Hafler DA, Reardon DA, Sampson JH. Prospects of immune checkpoint modulators in the treatment of glioblastoma. *Nat Rev Neurol* (2015) 11(9):504–14. 10.1038/nrneuro.2015.139
- [25] Raychaudhuri B, Rayman P, Huang P, Grabowski M, Hambardzumyan D, Finke JH, Vogelbaum MA. Myeloid derived suppressor cell infiltration of murine and human gliomas is associated with reduction of tumor infiltrating lymphocytes. *Journal of Neuro-Oncology*. 2015;122:293–301. doi: 10.1007/s11060-015-1720-6.
- [26] Reiter-Brennan, C.; Semmler, L.; Klein, A. The effects of 2-hydroxyglutarate on the tumorigenesis of gliomas. *Contemp. Oncol.* 2018, 22, 215–222.
- [27] Säälik P, Lingasamy P, Toome K, Mastandrea I, Rousoo-Noori L, Tobi A, Simón-Gracia L, Hunt H, Paiste P, Kotamraju VR, Bergers G, Asser T, Rätsep T, Ruoslahti E, Bjerkvig R, Friedmann-Morvinski D, Teesalu T. Peptide-guided nanoparticles for glioblastoma targeting. *Journal of Controlled Release*. 2019;308:109–118. doi: 10.1016/j.jconrel.2019.06.018.



- [28] Wainwright DA, Balyasnikova IV, Chang AL, Ahmed AU, Moon K-S, Auffinger B, Tobias AL, Han Y, Lesniak MS. IDO expression in brain tumors increases the recruitment of regulatory T cells and negatively impacts survival. *Clinical Cancer Research*. 2012;18:6110–6121. doi: 10.1158/1078-0432.CCR-12-2130.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)