ISSN No:-2456-2165

Data Analysis of Brain Cancer with Biopython

Vinita Kukreja[1]
[1]Trainee, Bioinformatics Project and Research Institute, Noida - 201301, India

Uma Kumari[2]*
*[2]Senior Bioinformatics Scientist, Bioinformatics Project and Research Institute, Noida - 201301, India

Abstract:- Biopython, an open source tools of Python for biological computation, was first published in 2000 by Brad Chapman and Jeff Chang. Biopython features consist of parsers for Bioinformatics file formats, access to online Bioinformatics databases, interfaces to common programs, a standard sequence class, etc. Glioblastoma is one of the most aggressive (grade IV) type of brain cancer, accounts for 15% of all brain tumors. Genetic alteration in Glioblastoma include EGFR and PDGFR amplification, TERT promoter mutation, alteration in TP53, NF1, PTEN and RB, loss of chromosome arm 10q and aberrations in RTK/Ras/PI3K signaling pathways. Pathway maps were used to understand a molecular interaction and reaction network in Glioma. Multiple sequence alignment tools helps us to analyze the area of similarity and evolutionary relationships between the sequences. Using Biopython tools we perform the analysis of the nucleotide sequences. This study introduces the application of a brain tumor detection algorithm using machine learning techniques.

Keywords:- *Brain Tumor, Computer Vision, Structure Analysis, Sequence Alignment, Deep Learning, Biopython.*

I. INTRODUCTION

In 1980, Guido van Rossum started working on Python and first published it in 1991 as Python 0.9.0.[1] Development of Biopython initiated in 1999 and it was first published in 2000 by Brad Chapman and Jeff Chang.[2] Python is a high-level programming language extensively used in commercial and academics, accessible to all the major operating system. It promotes basic syntax, object-oriented programming and a wide array of libraries.[3] Biopython is a member project of the Open Bioinformatics Foundation (OBF), which organises Biopython web site, source code repository, bug tracking database and email mailing lists. It also supports the related projects such as: BioPerl[4], BioJava[5], BioRuby and BioSQL.

Biopython is an open source compilation of Python tools for biological computation, created by an international team of developers.[2, 6] The main reason for development of Biopython is to make it easier for Python programming language user by creating high-quality, reusable modules and classes for the complex bioinformatics problems. Biopython consist of various features which include the ability to parse various Bioinformatics file formats (BLAST, Clustalw, FASTA, Genbank, PubMed, ExPASy, SCOP, KEGG, UniGene, and SwissProt), access to online Bioinformatics databases (NCBI and ExPASy), interfaces to common programs (Clustalw alignment program, Standalone Blast from NCBI, and command line tools from EMBOSS), a standard sequence class (dealing with sequences, sequence ids and sequence features), tools for performing common procedures on sequences (translation, transcription, and weight calculations), Bio.Motif module provide analysis of sequence motif (searching, comparing, and de novo learning),[2, 6] Bio.Phylo module used for the visualization of phylogenetic trees.[7]

An abnormal cell growth that have formed in the brain is a Brain tumor. Tumors can form in the brain or other parts of the central nervous system (CNS) (spine or cranial nerves). Brain controls most of the bodily functions which include awareness, movements, sensations, thoughts, speech, and memory. Tumors can affect these function and alters brain's ability to operate properly.[8, 9] There are more than 120 different types of brain tumor, based on the tissue they arise from. Brain tumors can be cancerous and non-cancerous or benign, but even non-cancerous tumors can be harmful due to its size and location.[10] Tumors that arises in the brain are called primary brain tumors and cancer that metastasize from other parts of the body to the brain are secondary brain tumors. Brain tumors can also classified as histological grading (I-IV) and molecular marker. Tumor diagnosis should be "layered" as histological classification, WHO grade, and molecular information and reported as "integrated diagnosis."[11]

Gliomas are one of the type of brain tumor that look like glial cells.[12] The most common type of malignant gliomas are Glioblastoma (grade IV), accounts for 15% of all brain tumors.[13] Glioblastoma is one of the most aggressive types of brain cancer because it arises from astrocytes cells that supports nerve cells and regulate the blood amount that reaches them, so having access to the large number of blood vessels helps cancer cells to grow and spread rapidly.[14] Another reason behind the aggressiveness of glioblastoma is their high recurrence rate. This is because tumor contains glioma stem cells (GSC), a type of self-regenerating cancer stem cell that controls the growth of tumors. In previous study, Subhas Mukherjee and his colleagues found high level of cyclin-dependent kinase 5 (CDK5) enzymes in GSC, the study shows that blocking this enzyme inhibits GSCs ability to self-regenerate.[15] The cause of some glioblastoma cases are unknown. Some uncommon risk factors include genetic disorders, previous radiation therapy[16, 17] and its association with viruses (SV40, [18] HHV-6,[19, 20] and cytomegalovirus[21]). Common genetic alteration in Glioblastomas include amplification of epidermal growth factor receptor (EGFR) and platelet-derived growth factor receptor (PDGFR); Telomerase reverse transcriptase (TERT) promoter mutation; alteration

ISSN No:-2456-2165

in tumor protein 53 (TP53), neurofibromin 1 (NF1), phosphate and tensin homologue (PTEN) and retinoblastoma (RB); loss of chromosome arm 10q, and aberrations in Receptor tyrosine kinase (RTK)/Ras/phosphoinositide 3-kinase (PI3K) signaling pathways.[22, 23] In this research journal, we study more about brain tumor with different Biopython (Bio.Seq import) and Bioinformatics tools. Bioinformatics tools and software propaganda is to develop a new biological technique and chemical database to help in the understanding of fundamental ADME processes that can control disease and health by the metabolic processes.[24] Bioinformatics strategies develop to significantly improve survival rates in patients and explore about how new models that allow us to bridge the gap between promising preclinical findings and identification of clinical translation.[25]

II. MATERIALS AND METHODS

To understand a network of molecular interaction and reaction in Glioma we used KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway maps. KEGG database is used to understand advanced functions and efficacies of the biological system from genomic and molecular-level information. According to developers, it is a computer representation of the biological system which include the integration of molecular building blocks of genes and proteins (genomic information) and chemical substances (chemical information) and wiring diagrams of molecular interaction and reaction networks (systems information), it also contains disease and drug information (health information). To study the alignment of more than two sequence we used Clustal Omega tool. Clustal Omega is a multiple sequence alignment tool that uses seeded guide trees and HMM (Hidden Markov Model) profile-profile techniques to generate alignments between three or more sequences. Alignment of multiple sequences emphasizes areas of similarity which perhaps associated with specific traits that are more highly conserved than other regions. Clustal Omega also used to analyze the evolutionary relationships between sequences through phylogenetic analysis. To study the gene prediction we used ORF Finder (Open Reading Frame Finder), to identify genomic DNA regions that encode genes including protein coding genes, RNA genes and regulatory genes. The ORF Finder is a graphical analysis tool which finds all open reading frames of a selectable sequence. This tool classify all open reading frames using the standard or alternative genetic codes. For the annotation of protein sequences, we study Conserved Domain Database (CDD), consist of analyzed multiple sequence alignment models for ancient domains and full-length proteins. CDD uses Reverse Position-Specific BLAST (RPS-BLAST) by comparing query sequence to the position-specific score matrices (PSSMs) of the conserved domain protein to find a several types of RPS-BLAST hits and domain model scope. CDD includes NCBI curated domain which uses information of 3D structure to define domain boundaries and provide understanding of sequence/structure/function relationships, it also imports data from external source databases. Analysis of nucleotide sequence were performed using Biopython tools. GC-content (guanine-cytosine content) is used to calculate the percentage of nitrogenous bases that are guanine (G) or cytosine (C) present in a DNA or RNA molecule and used to describe genomes. Biopython modules, Bio.Seq and Bio.SeqUtils are used to calculate the GC-content of a nucleotide sequence. Same as GC-content, Bio.Seq modules are used to measure the sequence length. Complementary sequence follows the lock-and-key principle, it shares the property between two DNA or RNA sequences. Complementary base pairing allows cells to copy information from one generation to another. Reverse Complement is formed by reversing the complementary sequence. Transcription is the synthesis of RNA molecule from DNA sequence with the help of RNA polymerase enzyme. To make a complementary RNA strand one of the DNA strands acts as a template. Translation is the process in which protein synthesize from mRNA template. At a time of translation, the sequence of nucleotides is translated into a sequence of amino acids, these amino acids form polypeptide chain which further bends and folds on itself to form a protein.

III. RESULT AND DISCUSSION

KEGG database utilizes genomic and molecular-level information to understand the functions and efficacies of the biological system. Figure 1, shows the molecular interaction network of Glioma's primary and secondary pathway. The highlighted route in these two pathway represent: Mutation-activated EGFR to RAS-ERK signaling pathway; EGFR-overexpression to PI3K signaling pathway; Amplified EGFR to PLCG-CAMK signaling pathway; Amplified PDGFR to PLCG-CAMK signaling pathway; Mutation-inactivated TP53 to transcription.

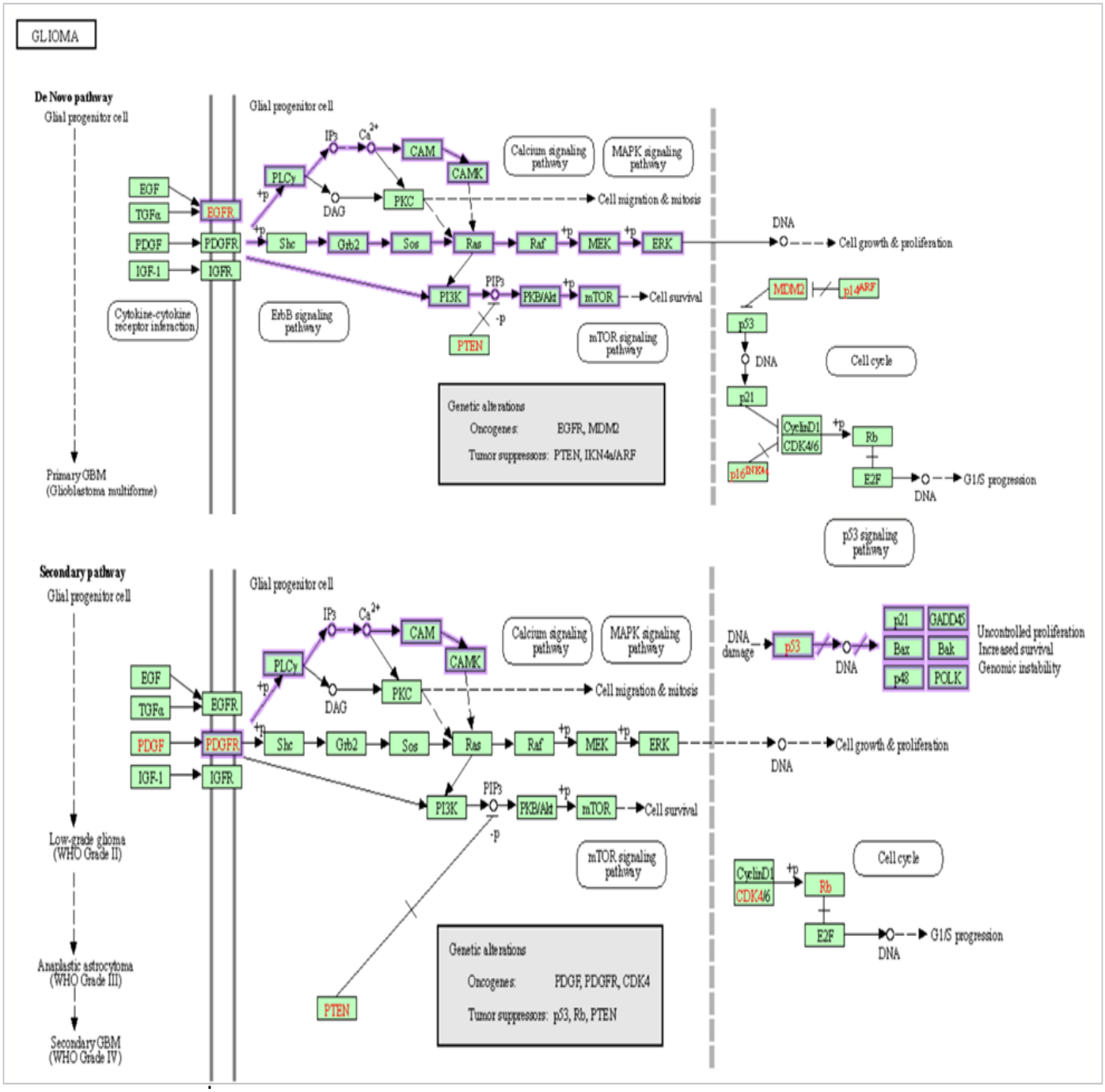


Fig 1 Molecular Interaction Network of Primary and Secondary Pathway of Glioma

Clustal Omega is a multiple sequence alignment tool which generate alignment between three or more sequence using seeded guide trees and HMM profile-profile techniques. In Figure 2, we observe a lot of variations in nucleotide sequences. The gap here represents the deletion in sequences and asterisk shows fully conserved alignment.

In Figure 3, the "length" of the branches represented by the values shown in the tree, indicating evolutionary distance between the sequences, i.e., the larger number represent the larger amount of genetic changes.

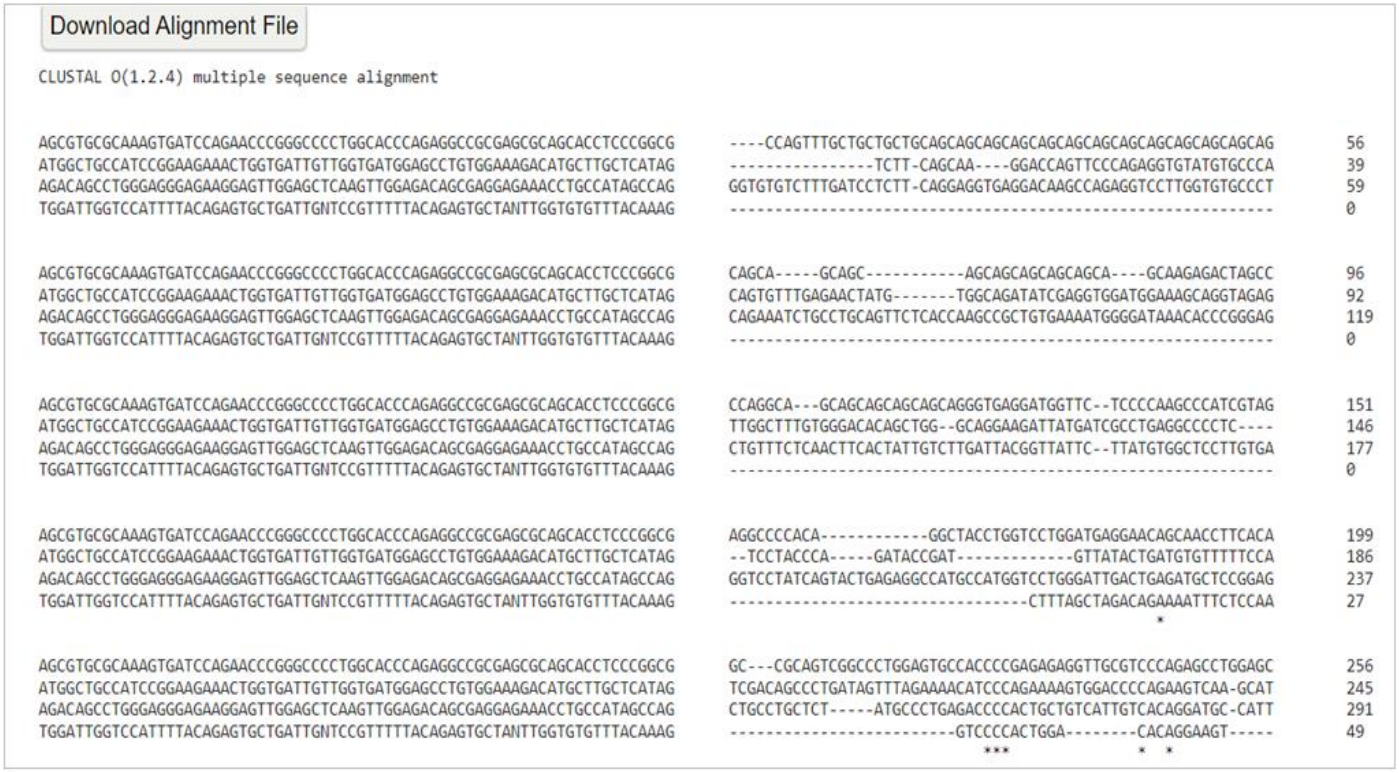


Fig 2 Multiple Sequence Alignment of Nucleotide Sequences

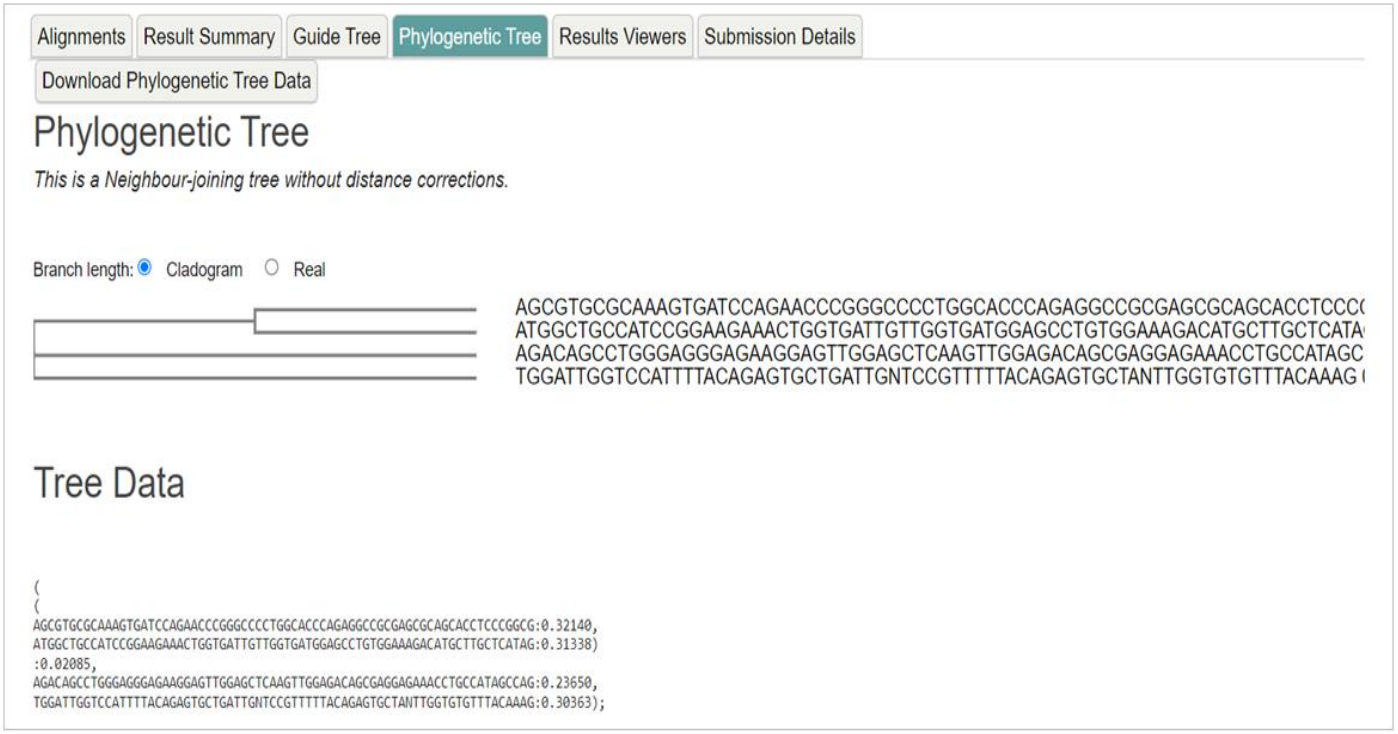


Fig 3 Cladogram of Nucleotide Sequences

Open Reading Frame (ORF) identifies all the possible protein coding region in the sequence. There would be 3 possible reading frames in each direction of the DNA sequence, i.e. there are total 6 possible reading frame (6 horizontal bars) in every DNA sequence. The 6 possible reading frames are +1, +2, +3 in the forward strand and -1, -2 and -3 in the reverse strand. Asterisk (*) represent Stop Codon whereas M codes for Start Codon. In figure 4 and 5, the result displays all the possible six reading frame present in the entered sequence query. The ORF is listed according to their size and the graphical representation of the sequence. The selected ORF is the ORF1, +1 reading frame in the forward strand. Nucleotide length of ORF1 is 96 and 31 is the protein length. For ORF1 start codon is placed at 169 while stop codon is placed at 264. The longest ORF among all is ORF6, -3 reading frame in the reverse strand. Nucleotide length and protein length of ORF6 is 360 and 119 respectively. For ORF6 start codon is placed at 360 while stop codon is placed at >1.

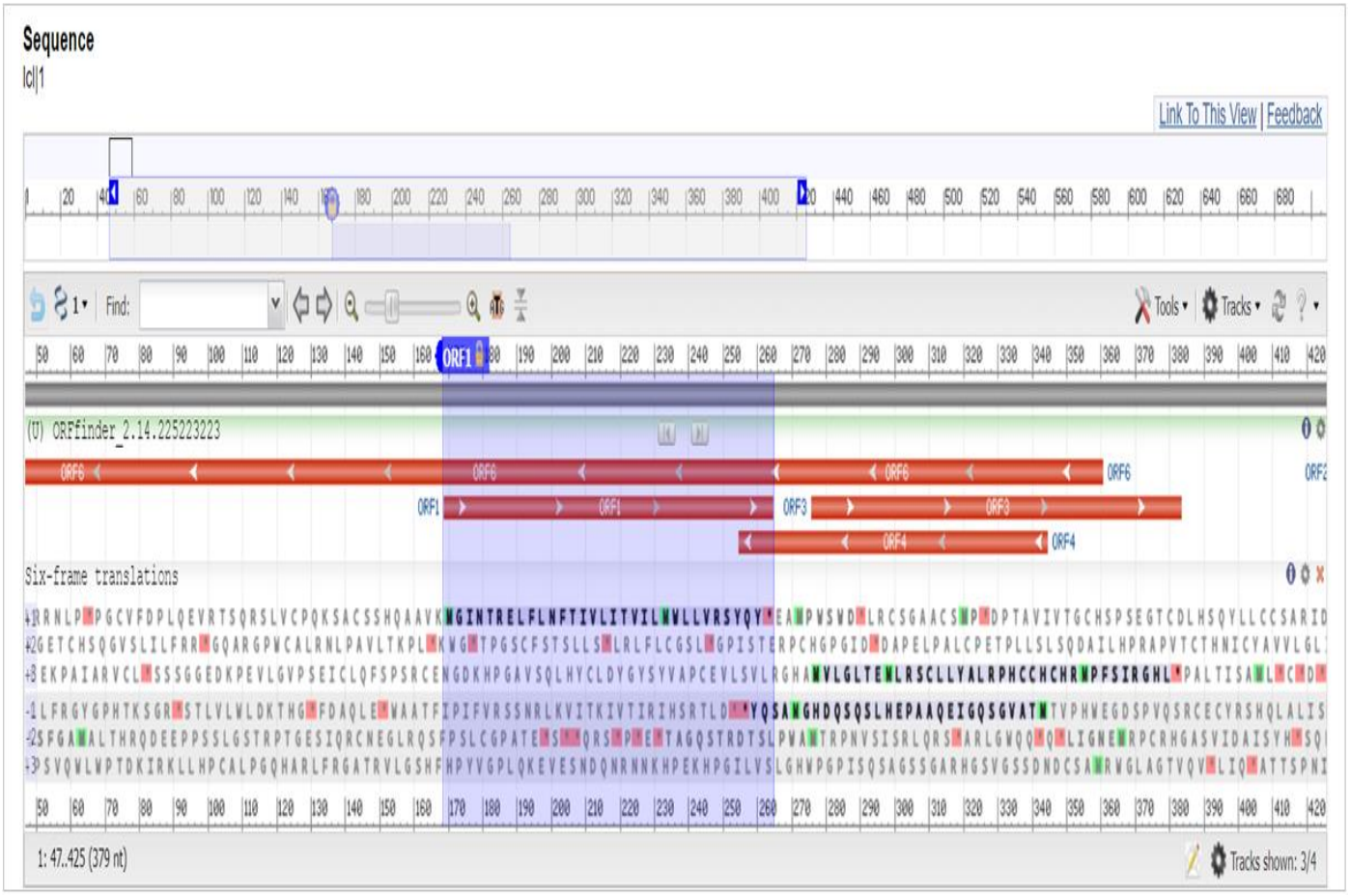
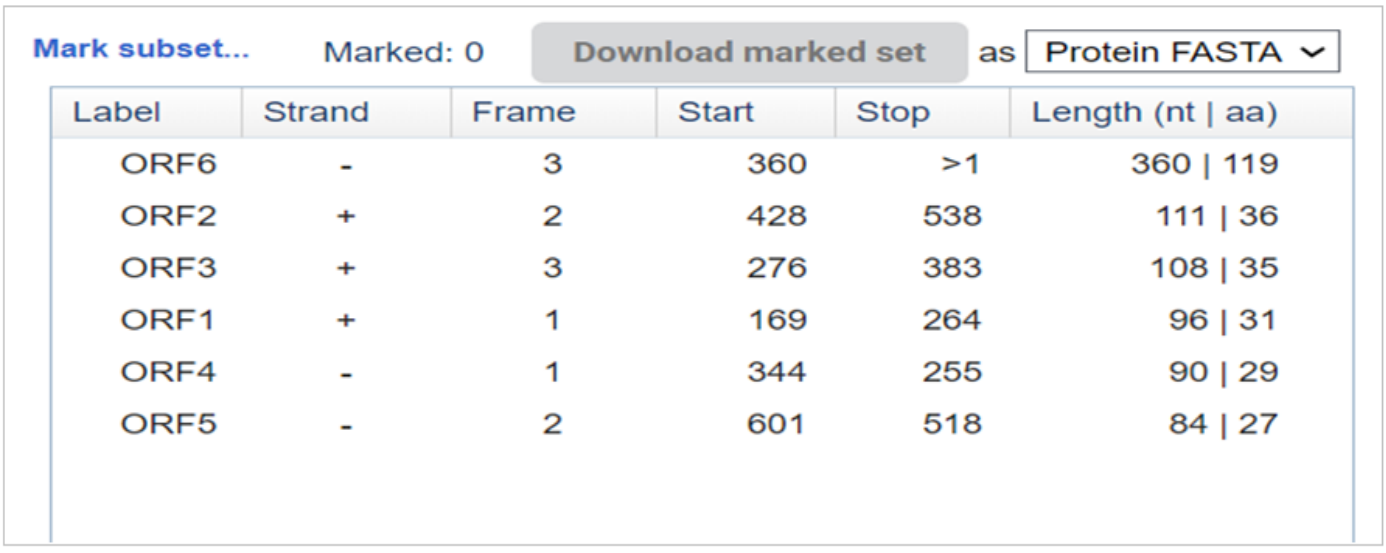


Fig 4 Six-Frame Translation of DNA Sequence



Label	Strand	Frame	Start	Stop	Length (nt | aa)
ORF6	-	3	360	>1	360 | 119
ORF2	+	2	428	538	111 | 36
ORF3	+	3	276	383	108 | 35
ORF1	+	1	169	264	96 | 31
ORF4	-	1	344	255	90 | 29
ORF5	-	2	601	518	84 | 27

Fig 5 Display of Reading-Frame Strand and Length

The result of CDD provides various display options and information for the conserved domains that align to the query sequence. In figure 6, the graphical summary displays the standard result which shows best scoring domain model from each source database. Small triangle in figure indicates specific amino acids involved in conserved features like binding and catalytic sites. Specific hits determine by the e-value of RPS-BLAST hits to be equal or lower than domain-specific threshold e-value. It describes the high confidence association between query sequence and conserved domain i.e. the query sequence is related to the same protein family. Superfamily is a set of conserved domain models of different families which have the same protein sequences and provide structural, functional and evolutionary information for proteins.

ISSN No:-2456-2165

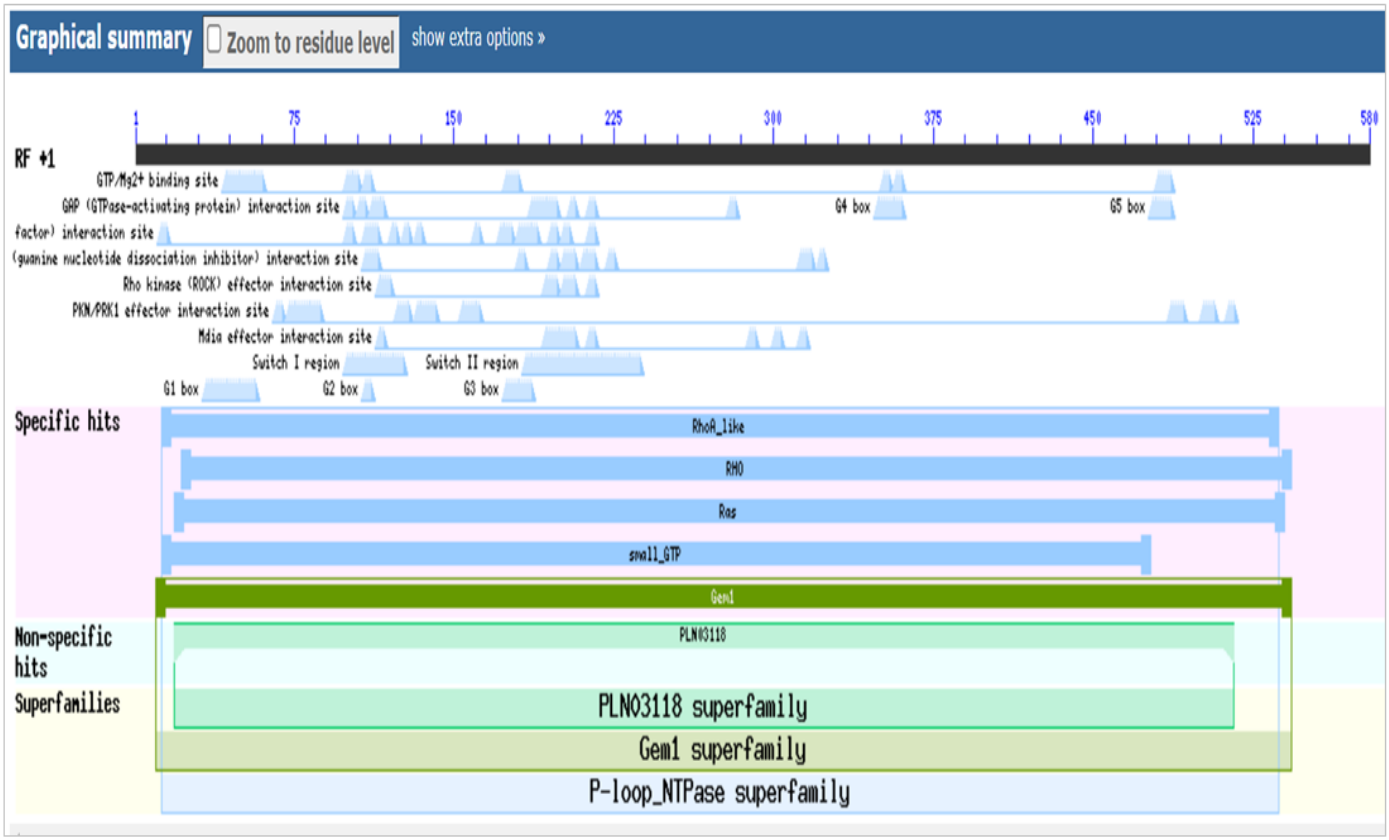


Fig 6 Graphical Representation of CDD-Based Annotation on Query Sequence

Bio.Seq and Bio.SeqUtils modules used here to determine the GC content of nucleotide sequences. GC content is a percentage of nitrogenous base (Guanine or Cytosine) in DNA or RNA molecule. It can be predicted by dividing the number of GC nucleotides by the total number of nucleotides. It also helps in estimating the length of the sequences.

In Biopython module, there are many in-build methods which helps in performing basic and advanced operations on sequences. Some of the advanced operation like complement, reverse complement, transcription, and translation were used to analyze the data of nucleotide sequences of patient with brain tumor.

Complement() function helps to find complement of given DNA sequences while reverse complement() function is used to get original nucleotide sequence using complemented sequence. Transcribe() function is used to convert DNA sequence into RNA sequence. In Biopython, DNA strand is directly converted into mRNA strand by replacing letter T with U. Translate() function is to translate RNA sequence into a protein sequence.

➢ *DNA Seq1*

- >>>from Bio.Seq import Seq
- >>>from Bio.SeqUtils import GC

>>>dna_seq1 = Seq("AGACAGCCTGGGAGGGAGA
AGGAGTTGGAGCTCAAGTTGGAGACAGCGAGGAGA
AACCTGCCATAGCCAGGGTGTGTCTTTGATCCTCTT
CAGGAGGTGAGGACAAGCCAGAGGTCCTTGGTGTG
CCCTCAGAAATCTGCCTGCAGTTCTCACCAAGCCGC
TGTGAAAATGGGGATAAACACCCGGGAGCTGTTTCT
CAACTTCACTATTGTCTTGATTACGGTTATTCTTATG
TGGCTCCTTGTGAGGTCCTATCAGTACTGAGAGGCC
ATGCCATGGTCCTGGGATTGACTGAGATGCTCCGGA
GCTGCCTGCTCTATGCCCTGAGACCCCACTGCTGTC
ATTGTCACAGGATGCCATTCTCCATCCGAGGGCACC
TGTGACCTGCACTCACAATATCTGCTATGCTGTAGT
GCTAGGATTGATTATGTGTTCTCCAAAGATGCTGCT
CCCAAGGGCTGCCAAGTGTTTGCCAGGGAACGGTA
GATTTATTCCCCAACTCTTAACTGAAAATGTGTTAG
ACAAGCCACAAAGTTAAAATTAAACTGGATTCATG
ATGATGTAGGATTGTTACAAGCCCCTGATCTGTCTC
ACCACACATCCCTTCAACCCACACGGTCTGCAACCA
AACTCTAATTCAACCTGCCAGAAGGAATGTTAGAGG
AAGTCTTTGTCAGCCCTTATAGCTATCATGTGAATA
AAGTTAAGTCAACTTC")

- >>>GC(dna_seq1)
- 48.46368715083799
- >>>len(dna_seq1)
- 716
- >>>from Bio.Seq import Seq
- >>>dna_seq1
- >>>dna_seq1.complement()
- Seq('TCTGTCGGACCCTCCCTCTTCCTCAACCTCG AGTTCAACCTCTGTCGCTCCTCT...AAG')
- >>>dna_seq1.reverse_complement()
- Seq('GAAGTTGACTTAACTTTATTCACATGATAGC TATAAGGGCTGACAAAGACTTCC...TCT')
- >>>dna_seq1.transcribe()

ISSN No:- 2456-2165.

- Seq('AGACAGCCUGGGAGGGAGAAGGAGUUGGA GCUCAAGUUGGAGACAGCGAGGAGA...UUC')
- >>>dna_seq1.translate()
- Seq('RQPGREKELELKLETARRNLP*PGCVFDPLQE VRTSQRSLVCPQKSACSSHQAA...KST')

➢ *DNA Seq2*

- >>>from Bio.Seq import Seq
- >>>from Bio.SeqUtils import GC

>>>dna_seq2 = Seq("ATGGCTGCCATCCGGAAGA AACTGGTGATTGTTGGTGATGGAGCCTGTGGAAAGA CATGCTTGCTCATAGTCTTCAGCAAGGACCAGTTCC CAGAGGTGTATGTGCCCACAGTGTTTGAGAACTATG TGGCAGATATCGAGGTGGATGGAAAGCAGGTAGAG TTGGCTTTGTGGGACACAGCTGGGCAGGAAGATTAT GATCGCCTGAGGCCCCTCTCCTACCCAGATACCGAT GTTATACTGATGTGTTTTTCCATCGACAGCCCTGATA GTTTAGAAAACATCCCAGAAAAGTGGACCCCAGAA GTCAAGCATTTCTGTCCCAACGTGCCCATCATCCTG GTTGGGAATAAGAAGGATCTTCGGAATGATGAGCA CACAAGGCGGGAGCTAGCCAAGATGAAGCAGGAGC CGGTGAAACCTGAAGAAGGCAGAGATATGGCAAAC AGGATTGGCGCTTTTGGGTACATGGAGTGTTCAGCA AAGACCAAAGATGGAGTGAGAGAGGTTTTTGAAAT GGCTACGAGAGCTGCTCTGCAAGCTAGACGTGGGA AGAAAAAATCTGGTTGCCTTGTCTTG")

- >>>GC(dna_seq2)
- 49.22279792746114
- >>>len(dna_seq2)
- 579
- >>>from Bio.Seq import Seq
- >>>dna_seq2
- >>>dna_seq2.complement()
- Seq('TACCGACGGTAGGCCTTCTTTGACCACTAAC AACCACTACCTCGGACACCTTTC...AAC')
- >>>dna_seq2.reverse_complement()
- Seq('CAAGACAAGGCAACCAGATTTTTTCTTCCCA CGTCTAGCTTGCAGAGCAGCTCT...CAT')
- >>>dna_seq2.transcribe()
- Seq('AUGGCUGCCAUCCGGAAGAAACUGGUGAUU GUUGGUGAUGGAGCCUGUGGAAAG...UUG')
- >>>dna_seq2.translate()
- Seq('MAAIRKKLVIVGDGACGKTCLLIVFSKDQFPE VYVPTVFENYVADIEVDGKQVE...LVL')

➢ *DNA Seq3*

- >>>from Bio.Seq import Seq
- >>>from Bio.SeqUtils import GC

>>>dna_seq3 = Seq("AGCGTGCGCAAAGTGATCC AGAACCCGGGCCCCTGGCACCCAGAGGCCGCGAGC GCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAG CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAGAGA CTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAG GATGGTTCTCCCCAAGCCCATCGTAGAGGCCCCACA GGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCA CAGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAG AGGTTGCGTCCCAGAGCCTGGAGCCGCCGTGGCCGC CAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTC CGGACGAGGATGACTCAGCTGCCCCATCCACGTTGT CCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCTGCT GCTCCGCTGACCTTAAAGACATCCTGAGCGAGGCCA GCACCATG")

- >>>GC(dna_seq3)
- 66.25766871165644
- >>>len(dna_seq3)
- 489

- >>>from Bio.Seq import Seq
- >>>dna_seq3
- >>>dna_seq3.complement()
- Seq('TCGCACGCGTTTCACTAGGTCTTGGGCC CGGGGACCGTGGGTCTCCGGCGCTCG...TAC')
- >>>dna_seq3.reverse_complement()
- Seq('CATGGTGCTGGCCTCGCTCAGGATGTCT TTAAGGTCAGCGGAGCAGCAGCTTAA...GCT')
- >>>dna_seq3.transcribe()
- Seq('AGCGUGCGCAAAGUGAUCCAGAACCCG GGCCCCUGGCACCCAGAGGCCGCGAGC...AU G')
- >>>dna_seq3.translate()
- Seq('SVRKVIQNPGPWHPEAASAAPPGASLLLL QQQQQQQQQQQQQQQQQQQQQQET...STM')

➢ *DNA Seq4*

- >>>from Bio.Seq import Seq
- >>>from Bio.SeqUtils import GC

>>>dna_seq4 = Seq("TGGATTGGTCCATTTTACAG AGTGCTGATTGNTCCGTTTTTACAGAGTGCTANTTG GTGTGTTTACAAAGCTTTAGCTAGACAGAAAATTTC TCCAAGTCCCCACTGGACACAGGAAGTCCAGCTGGC TTCACCTCTGAAAACTTTTTAGATTAAAAAAATAGA ACAAACTAGTTTTAGTAGACACTTTTAAAATGATAA AGCAACTTGCGTTAATTTAATTCCTATCATTATGAC ATAAATATCTAAGCAATGAAAGATAATATCTTTTAT TATAAAGCTGCATAATGTGAAATCTTGCTGATGGTG TCACATCACTGGACATTACTGACACCTTTTGTTAAA AAACTAACGTTCTACTGATCAGACCAATCCAAATCA CTAGTGAATTCGCG")

- >>>GC(dna_seq4)
- 34.263959390862944
- >>>len(dna_seq4)
- 394
- >>>from Bio.Seq import Seq
- >>>dna_seq4
- >>>dna_seq4.complement()
- Seq('ACCTAACCAGGTAAAATGTCTCACGACTAA CNAGGCAAAAATGTCTCACGATNA...CGC')
- >>>dna_seq4.reverse_complement()

ISSN No:- 2456-2165.

- Seq('CGCGAATTCACTAGTGATTTGGATTGGTCTG ATCAGTAGAACGTTAGTTTTTTA...CCA')
- >>>dna_seq4.transcribe()
- Seq('UGGAUUGGUCCAUUUUACAGAGUGCUGAU UGNUCCGUUUUUACAGAGUGCUANU...GCG')
- >>>dna_seq4.translate()
- Seq('WIGPFYRVLIXPFLQSAXWCVYKALARQKISP SPHWTQEVQLASPLKTF*IKKI...*IR')

IV. CONCLUSION

Biopython, an open-source programming application used in the development of bioinformatics software and in bioinformatics scripting. Computational biology helps to analyze the mechanism process of various diseases or clinical conditions and accelerates the research process. Bioinformatics introduces the application of a brain tumor detection algorithm using machine learning techniques. For this purpose, a number of tools and software were used to understand the genomic and molecular-level information of disease, its specific traits and evolutionary relationships, to study the gene prediction, to study protein annotation, and to perform an exploratory data analysis. In this study explaining the importance of Bioinformatics and Biopython using the data of patients with brain tumor disease.

REFERENCE

[1] Rossum, Guido Van (20 January 2009). "The History of Python: A Brief Timeline of Python". The History of Python. Archived from the original on 5 June 2020. Retrieved 5 March 2021.

[2] Chapman, Brad; Chang, Jeff (August 2000). "Biopython: Python tools for computational biology". ACM SIGBIO Newsletter. 20 (2): 15–19. doi:10.1145/360262.360268. S2CID 9417766.

[3] T. E. Oliphant, "Python for Scientific Computing," in Computing in Science & Engineering, vol. 9, no. 3, pp. 10-20, May-June 2007, doi:10.1109/MCSE .2007.58.

[4] Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JG, Korf I, Lapp H, Lehväslaiho H, Matsalla C, Mungall CJ, Osborne BI, Pocock MR, Schattner P, Senger M, Stein LD, Stupka E, Wilkinson MD, Birney E. The Bioperl toolkit: Perl modules for the life sciences. Genome Res. 2002 Oct;12(10):1611-8. doi: 10.1101/gr.361602. PMID: 12368254; PMCID: PMC187536.

[5] Holland RC, Down TA, Pocock M, Prlić A, Huen D, James K, Foisy S, Dräger A, Yates A, Heuer M, Schreiber MJ. BioJava: an open-source framework for bioinformatics. Bioinformatics. 2008 Sep 15;24(18): 2096-7. doi: 10.1093/bioinformatics/btn397. Epub 2008 Aug 8. PMID: 18689808; PMCID: PMC2530884.

[6] Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJ. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009 Jun 1;25(11):1422-3. doi: 10.1093/bioinformatics/btp163. Epub 2009 Mar 20. PMID: 19304878; PMCID: PMC2682512.

[7] Talevich E, Invergo BM, Cock PJ, Chapman BA. Bio.Phylo: a unified toolkit for processing, analyzing and visualizing phylogenetic trees in Biopython. BMC Bioinformatics. 2012 Aug 21;13:209. doi: 10.1186/1471-2105-13-209. PMID: 22909249; PMCID: PMC3468381.

[8] Dorsey JF, Salinas RD, Dang M, et al. Cancer of the central nervous system. In: Niederhuber JE, Armitage JO, Doroshow JH, Kastan MB, Tepper JE, eds. Abeloff's Clinical Oncology. 6th ed. Philadelphia, PA: Elsevier; 2020:chap 63.

[9] Brain tumor. Cancer.Net. https://www.cancer.net/ cancer -types/braintumor/view-all. Accessed Nov. 1, 2022.

[10] Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. 2021 Aug 2;23(8):1231-1251. doi: 10.1093/neuonc/ noab106. PMID: 34185076; PMCID: PMC8328013.

[11] Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016 Jun;131(6):803-20. doi: 10.1007/s00401-016-1545-1. Epub 2016 May 9. PMID: 27157931.

[12] Louis DN, et al. Classification and pathologic diagnosis of gliomas, glioneuronal tumors and neuronal tumors. https://www.uptodate.com/ contents/ search. Accessed June 10, 2022.

[13] Young RM, Jamshidi A, Davis G, Sherman JH (June 2015). "Current trends in the surgical management and treatment of adult glioblastoma". Annals of Translational Medicine. 3 (9): 121. doi:10.3978 /j.issn.2305-5839.2015.05.10. PMC 4481356. PMID 26207249.

[14] 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, Vol.9, Issue 5, page no.k473-k479, May-2022.

[15] Mukherjee S, Tucker-Burden C, Kaissi E, Newsam A, Duggireddy H, Chau M, Zhang C, Diwedi B, Rupji M, Seby S, Kowalski J, Kong J, Read R, Brat DJ. CDK5 Inhibition Resolves PKA/cAMP-Independent Activation of CREB1 Signaling in Glioma Stem Cells. Cell Rep. 2018 May 8;23(6):1651-1664. doi: 10.1016/j.celrep.2018.04.016. PMID: 29742423; PMCID: PMC5987254.

[16] World Cancer Report 2014. World Health Organization. 2014. Chapter 5.16. ISBN 978-9283204299. Archived from the original on September 19, 2016. https://www.who.int/cancer /publications/ WRC_2014/en/. [Google Scholar]
[17] Gallego O (August 2015). "Nonsurgical treatment of recurrent glioblastoma". Current Oncology. 22 (4): e273–81. doi:10.3747/co.22.2436. PMC 4530825. PMID 26300678.
[18] Vilchez RA, Kozinetz CA, Arrington AS, Madden CR, Butel JS (June 2003). "Simian virus 40 in human cancers". The American Journal of Medicine. 114 (8): 675–84. doi:10.1016/S0002-9343(03)00087-1. PMID 12798456.
[19] Crawford JR, Santi MR, Thorarinsdottir HK, Cornelison R, Rushing EJ, Zhang H, et al. (September 2009). "Detection of human herpesvirus-6 variants in pediatric brain tumors: association of viral antigen in low grade gliomas". Journal of Clinical Virology. 46 (1): 37–42. doi:10.1016/j.jcv.2009.05.011. PMC 2749001. PMID 19505845.
[20] Chi J, Gu B, Zhang C, Peng G, Zhou F, Chen Y, et al. (November 2012). "Human herpesvirus 6 latent infection in patients with glioma". The Journal of Infectious Diseases. 206 (9): 1394–98. doi:10.1093/infdis/jis513. PMID 22962688.
[21] McFaline-Figueroa JR, Wen PY (February 2017). "The Viral Connection to Glioblastoma". Current Infectious Disease Reports. 19 (2): 5. doi:10.1007/s11908-017-0563-z. PMID 28233187. S2CID 30446699.
[22] Reifenberger G, Liu L, Ichimura K, Schmidt EE, Collins VP. Amplification and overexpression of the MDM2 gene in a subset of human malignant gliomas without p53 mutations. Cancer Res. 1993 Jun 15;53(12):2736-9. PMID: 8504413.
[23] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou CM, Hayes DN; Cancer Genome Atlas Research Network. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010 Jan 19;17(1):98-110. doi: 10.1016/j.ccr.2009.12.020. PMID: 20129251; PMCID: PMC2818769.
[24] Uma Kumari “Insilico analysis and computer aided drug designing approach for mutant cancer gene” (IJBTR) DEC 2021 (Impact factor 6.6, ICV 61.5, NASS RATING 3.8).
[25] Shubhi Bindal, Uma Kumari et al “Homology Modeling and Drug Designing Approach for Prospective Malignant Brain Cancer” IJIRT 2022 Volume 9, Issue 6, Pages 528-534.