MOTIF AND PATHWAY IDENTIFICATION WITH DESIGNING NOVEL LIGANDS FOR FILIPPI SYNDROME

Swathi Manjunath[1], Milan Khandelwal*[1,2], Shylesh Murthy IA[2] and Preenon Bagchi[3]
[1]Padmashree Institute of Management and Sciences, Bengaluru, India. what
Corresponding author: Email:milankhandelwal35@gmail.com
[2]Vasishth Academy of Advanced Studies and Research, Bengaluru, India.
[3]Institute of Biosciences and Technology, MGM University, Aurangabad, India.

Abstract : Filippi Syndrome is a rare genetic developmental disorder caused by defective mutations in the Cytoskeleton Associated Protein 2 Like (CKAP2L) gene. This disorder manifests with a range of symptoms, including abnormalities in digits and craniofacial features, intellectual disability, and growth retardation, commonly observed in infants. In this study, we focused on the CKAP2L gene associated with Filippi syndrome and performed motif analysis, peak identification, and pathway identification. We employed ChIP sequencing, a technique used to analyze protein-DNA interactions, to obtain high-quality immunoprecipitation ChIP Peaks. The presence of these peaks indicates a high concentration of DNA-binding protein genes.Furthermore, we utilized molecular docking, a computational technique for drug design, to predict the interactions between macromolecules and micro-molecules. Molecular docking is a widely used in-silico structure-based approach in the field of drug development.Overall, our study provides insights into the genetic basis of Filippi Syndrome through analysis of the CKAP2L gene and highlights the potential of molecular docking in computational drug design.

Keywords: ChiP-seq, Molecular Docking, Ligand, computational drug design, phytocompounds.

Introduction :- Filippi syndrome, or Filippi syndrome type A, is an incredibly rare hereditary developmental condition distinguished by a wide range of symptoms and physical abnormalities. Dr Giancarlo Filippi, an Italian physician, first discussed it in 1994. Defective mutations in the CKAP2L (Cytoskeleton Associated Protein 2) gene are the cause of the condition.Filippi Syndrome is a very uncommon condition, and it is unclear how common it is in the general population. There isn't much data on its incidence are available. The CKAP2L gene mutation is the main cause of Filippi Syndrome. These mutations cause problems with the gene's protein product, which is important in the organisation and control of the cytoskeleton. Filippi Syndrome patients may have high foreheads, widely spaced eyes (hypertelorism), a short nose, a broad nasal bridge, and a thin upper lip. Malformations of the digits, such as missing or additional fingers or toes (polydactyly or oligodactyly), fused digits (syndactyly), or strangely shaped digits, are common.

The majority of Filippi syndrome patients have intellectual difficulties. Growth delays are common, resulting in small stature and a slower rate of physical development. Additional characteristics can include hearing loss, heart defects, bone deformities, and genitourinary malformations in some people. The presence of these extra characteristics might, however, differ significantly across those who are afflicted.

R. Somashekhar et al. (eds.), *Proceedings of the International Conference on Advances in Nano-Neuro-Bio-Quantum (ICAN 2023)*,
Advances in Intelligent Systems Research 69,
https://doi.org/10.2991/978-94-6463-294-1_20

Due to the rarity of Filippi Syndrome and limited understanding of its underlying mechanisms, there is no specific cure or targeted treatment available. With the help of ChIP-seq enables the identification of specific genomic regions where the CKAP2L gene or its associated proteins bind. By mapping the binding sites, researchers can gain insights into the regulation of CKAP2L expression and the potential role of other proteins involved in Filippi Syndrome. ChIP-seq data can be further analyzed to identify DNA sequence motifs that are enriched in the binding sites. Motif analysis helps in understanding the potential transcription factor binding sites and regulatory elements involved in the regulation of CKAP2L.

ChIP-seq data can be integrated with other omics data to identify signaling pathways or biological processes that may be dysregulated in Filippi Syndrome. Than Computational-based drug development involves to identify potential drug targets, design molecules, and predict their properties.

METHODOLOGY :

We begin by retrieving particular sequences SRR12340404.1 and SRR12340404.2 of the fillipin syndrom from the SRA database, and then execute quality control checks with a tool called FASTQC. The sequences are then aligned with the mus musculus genome using Bowtie2, which generates mapped output files.

A tool named RmDUP is used to delete duplicate sequences. The following stage involves peak calling with MACS2, a commonly used method for identifying prominent sites in the genome. This aids in identifying critical areas of genomic activity.

After that, a filter and sort tool is used to perform a quality check on the chip sequencing sample, which is required for further analysis. The chip sequencing data is visualised using the UCSC genome browser.

The UCSC genome browser is used to visualise and detect any questionable peaks in the chip sequencing data. This browser combines the chip sequencing data with other pertinent information such as gene annotations and evolutionary data from different tissues.

To compare and filter distinct sets of genomic characteristics, bed tools and intersect interval tools are utilised. To discover enriched motifs in specific locations, the Galaxy cistron tool is used. Additionally, gene analysis is performed for prospective drug discovery utilising the Swiss-Prot database. Structure analysis is done to help with drug development. Finally, receptor and drug molecules are docked together, and the interactions between them are analysed using docking scores.

Result and discussion:-

FASTQC quality reports was given quality control to H. Sapiens chip-sequences with SRA accession number sequences SRR12340404.1 and SRR12340404.2 of the fillipin syndrom.

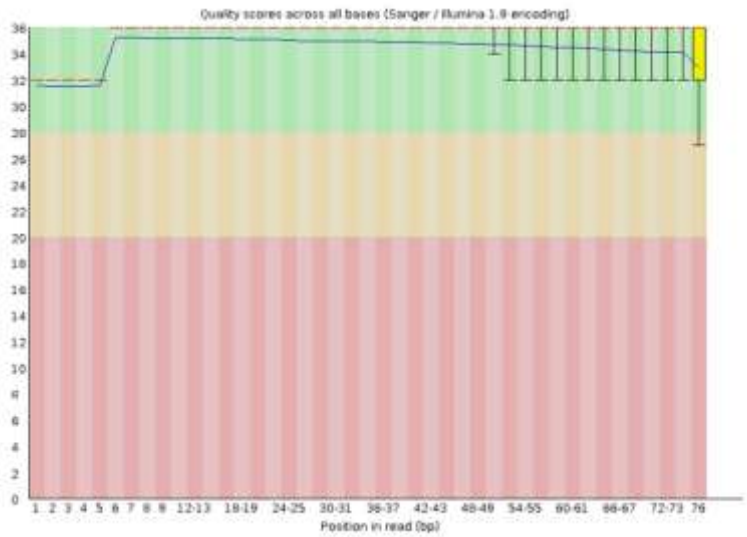


Figure :1 FASTQC Quality reports of SRR12340404

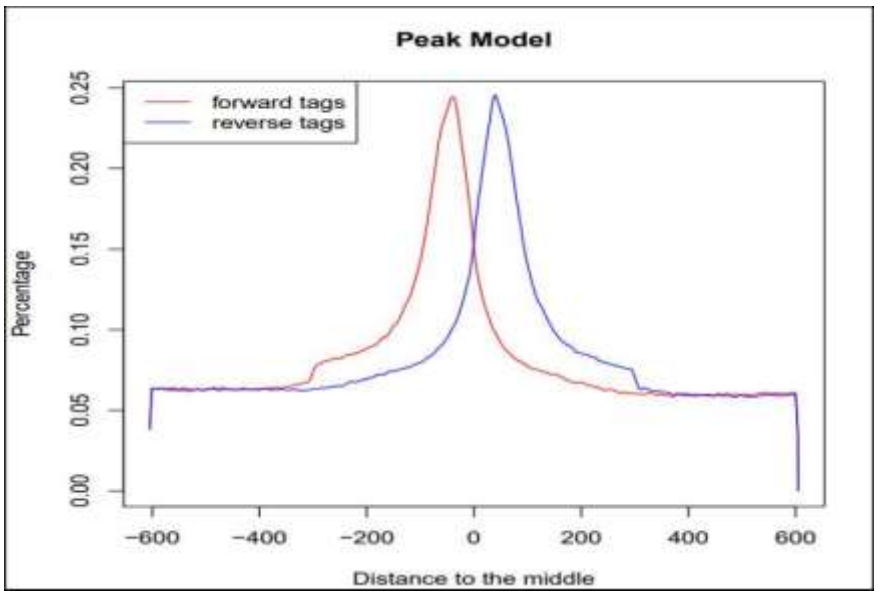


Figure :2 MACS calculated to identify the centre of your peaks. Peaks: bed and peaks: interval provides information about the peaks MACS found in the data. identify the significantly enrich loci in the genome.

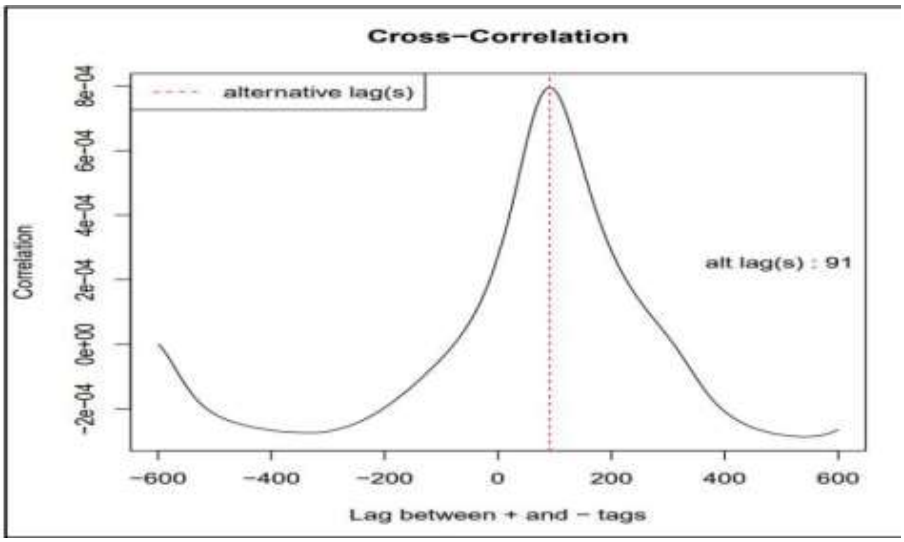
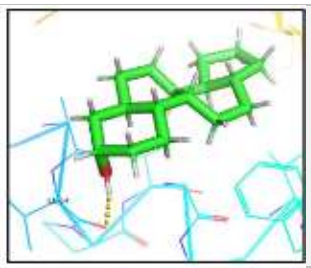
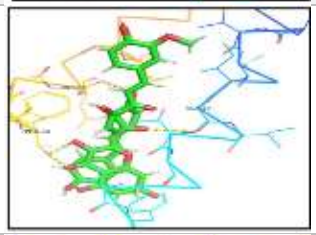
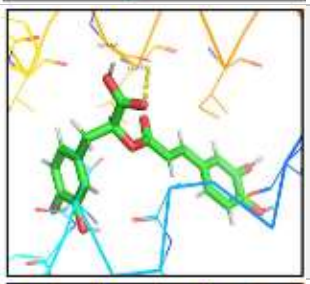
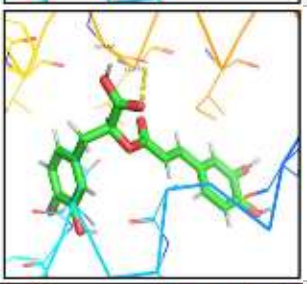
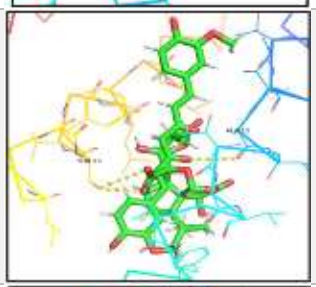
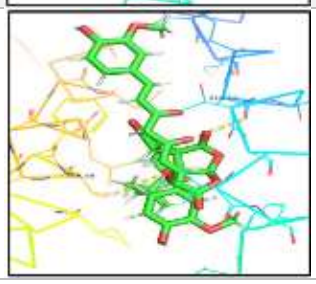


Figure :3 After aligning de-duplicated reads to the genome, peak caller can be used to find regions of the genome with an enrichment of reads, or peaks.

Sr.no	Motif	Function
1.	NP_904335.1	NADH dehydrogenase subunit 3 (mitochondrion) [Mus musculus]
2.	NP_904332.1	ATP synthase F0 subunit 8 (mitochondrion) [Mus musculus].
3.	NP_904333.1	ATP synthase F0 subunit 6 (mitochondrion) [Mus musculus]
4.	NP_904330.1	cytochrome c oxidase subunit I (mitochondrion)
5.	NP_904331.1	cytochrome c oxidase subunit II (mitochondrion) [Mus musculus].
6.	NP_904328.1	NADH dehydrogenase subunit 1 (mitochondrion) [Mus musculus].
7.	XM_030248894.1	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant X2, mRNA.
8.	XM_011239398.3	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant X1, mRNA.
9.	NM_001077514.4	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant 1, mRNA
10.	XM_030248893.1	Mus musculus solute carrier family 1 (glial high affinity glutamate

		transporter), member 2 (Slc1a2), transcript variant X3, mRNA.
11.	NM_001077515.2	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant 2, mRNA
12.	XM_030248894.1	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant X2, mRNA.
13.	XM_011239398.3	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant X1, mRNA.
14.	NM_001077514.4	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant 1, mRNA
15.	XM_030248893.1	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant X3, mRNA.
16.	NM_001077515.2	Mus musculus solute carrier family 1 (glial high affinity glutamate transporter), member 2 (Slc1a2), transcript variant 2, mRNA.
17.	NM_011221.3	Mus musculus purine rich element binding protein B (Purb), mRNA.
18.	NM_027626.1	Mus musculus pleckstrin and Sec7 domain containing 3 (Psd3), transcript variant 3, mRNA.
19.	NM_030263.5	Mus musculus pleckstrin and Sec7 domain containing 3 (Psd3), transcript variant 1, mRNA.
20.	NM_177698.4	Mus musculus pleckstrin and Sec7 domain containing 3 (Psd3), transcript variant 2, mRNA.

Table;1 gene motif and their function.

Serial no.	receptore	Plant name	Phytocompound	Interactions	Docking Score	Interaction image
1	cytochrome oxidase, partial [Cyprinus carpio]	Turmeric	Sterol	ALA-13 ILE-14	-3124	
2	cytochrome oxidase, partial [Cyprinus carpio]	Turmeric	Curcumin	ALA-13 GLY-52 THR-49	-4524	
3	cytochrome oxidase, partial [Cyprinus carpio]	Lemon balm	Roasmarinic acid	THR-49 GLY-52 LEU-53	3976	
4	cytochrome oxidase, partial [Cyprinus carpio]	Ginko biloba	Ginkgolide A	THR-49 ALA-13	-3644	
5	cytochrome oxidase, partial [Cyprinus carpio]	Ginko biloba	Bilobalide	THR-49 ALA-13	-3188	
6	cytochrome oxidase, partial [Cyprinus carpio]	Ginko biloba	Quercetin	GLY-52 THR-49	-3660	

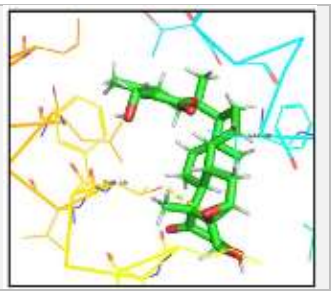
7	cytochrome oxidase, partial [Cyprinus carpio]	ashwagandha	Withanolide	ILE-12 THR-16 GLY-51	-4218	

Table : 2 drug development (receptor and phytocompound docking)

CONCLUSION:

After mapping out of two SRR accession, I used the tool Collect Alignment Summary Metrics tool take the summary of our mapping done above. Both tables contain the alignment summary SRR12340404. Next, I used MAC2 call peak tool to identify areas in the genome that are enriched with the aligned reads. Model-based Analysis of Chip-Seq (MACS) is a commonly used tool for identifying transcription factor binding sites. Then I identified the motifs present in our Mus musculus genome. As per MACS2 peak analysis, the genes identified are XM_01124842, NM_001164681, NM_198672, NM_007818, NM_201642, Curcumin, Roasmarinic acid, Withanolide A. The gene XM_01124842's corresponding receptor shows best interactions with the phytocompounds Curcumin, Roasmarinic acid Withanolide A. These phytocompounds can be selected as best ligands for the gene receptor.

References:

1. Massie CE, Mills IG. Mapping protein-DNA interactions using ChIP-sequencing. Methods Mol Biol. 2012; 809:157-73. doi: 10.1007/978-1-61779-376-9_11. PMID: 22113275.

2. Mimura I, Kanki Y, Kodama T, Nangaku M. Revolution of nephrology research by deep sequencing: ChIP-seq and RNA-seq. Kidney Int. 2014 Jan;85(1):31-8. doi: 10.1038/ki.2013.321. Epub 2013 Aug 28. PMID: 23986147.

3. Mundade R, Ozer HG, Wei H, Prabhu L, Lu T. Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. Cell Cycle. 2014;13(18):2847-52. doi: 10.4161/15384101.2014.949201. PMID: 25486472; PMCID: PMC4614920.

4. Bansal M, Mendiratta G, Anand S, Kushwaha R, Kim R, Kustagi M, Iyer A, Chaganti RS, Califano A, Sumazin P. Direct ChIP-Seq significance analysis improves target prediction. BMC Genomics. 2015;16 Suppl 5(Suppl 5):S4. doi: 10.1186/1471-2164-16-S5-S4. Epub 2015 May 26. PMID: 26040656; PMCID: PMC4460594.

5. Mendoza-Parra MA, Saravaki V, Cholley PE, Blum M, Billoré B, Gronemeyer H. Antibody performance in ChIP-sequencing assays: From quality scores of public data sets

to quantitative certification. F1000Res. 2016 Jan 12;5:54. doi: 10.12688/f1000research.7637.2. PMID: 27335635; PMCID: PMC4893979.

6. Chen X, Bhadauria V, Ma B. ChIP-Seq: A Powerful Tool for Studying Protein-DNA Interactions in Plants. Curr Issues Mol Biol. 2018;27:171-180. doi: 10.21775/cimb.027.171. Epub 2017 Sep 8. PMID: 28885181.

7. Jiang S, Mortazavi A. Integrating ChIP-seq with other functional genomics data. Brief

8. 'Funct Genomics. 2018 Mar 1;17(2):104-115. doi: 10.1093/bfgp/ely002. PMID: 29579165; PMCID: PMC5888983.

9. Suryatenggara J, Yong KJ, Tenen DE, Tenen DG, Bassal MA. ChIP-AP: an integrated analysis pipeline for unbiased ChIP-seq analysis. Brief Bioinform. 2022 Jan 17;23(1):bbab537. doi: 10.1093/bib/bbab537. PMID: 34965583; PMCID: PMC8769893.

10. Avram S, Mernea M, Limban C, Borcan F, Chifiriuc C. Potential Therapeutic Approaches

to Alzheimer's Disease By Bioinformatics, Cheminformatics And Predicted Adme-Tox Tools. Curr Neuropharmacol. 2020;18(8):696-719. doi: 10.2174/1570159X18666191230120053. PMID: 31885353; PMCID: PMC7536829.

11. Scarpino A, Ferenczy GG, Keserű GM. Covalent Docking in Drug Discovery: Scope and Limitations. Curr Pharm Des. 2020;26(44):5684-5699. doi: 10.2174/1381612824999201105164942. PMID: 33155894.

12. Choi J, Lee J. V-Dock: Fast Generation of Novel Drug-like Molecules Using MachineLearning-Based Docking Score and Molecular Optimization. Int J Mol Sci. 2021 Oct 27;22(21):11635. doi: 10.3390/ijms222111635. PMID: 34769065; PMCID: PMC8584000.

13. Sehgal SA, Mannan S, Ali S. Pharmacoinformatic and molecular docking studies reveal potential novel antidepressants against neurodegenerative disorders by targeting HSPB8. Drug Des Devel Ther. 2016 May 6;10:1605-18. doi: 10.2147/DDDT.S101929. PMID: 27226709; PMCID: PMC4866741. 50

14. da Silva Cunha TT, Silva RR, Rodrigues DA, de Sena Murteira Pinheiro P, Kronenberger T, Sant'Anna CMR, Noël F, Fraga CAM. Design, Synthesis and Pharmacological Evaluation of Novel Conformationally Restricted N-arylpiperazine Derivatives Characterized as D2/D3 Receptor Ligands, Candidates for the Treatment of Neurodegenerative Diseases. Biomolecules. 2022 Aug 12;12(8):1112. doi: 10.3390/biom12081112. PMID: 36009006; PMCID: PMC9405847.

15. Kaur A, Mann S, Kaur A, Priyadarshi N, Goyal B, Singhal NK, Goyal D. Multi-targetdirected triazole derivatives as promising agents for the treatment of Alzheimer's disease. Bioorg Chem. 2019 Jun;87:572-584. doi: 10.1016/j.bioorg.2019.03.058. Epub 2019 Mar 23. PMID: 30928879.

16. Chen SG, Gu RX, Dai H, Wei DQ. Virtual screening for alpha7 nicotinic acetylcholine receptor for treatment of Alzheimer's disease. J Mol Graph Model. 2013 Feb;39:98-107. doi: 10.1016/j.jmgm.2012.11.008. Epub 2012 Nov 29. PMID: 23261879.

17. : Pinzi L, Rastelli G. Molecular Docking: Shifting Paradigms in Drug Discovery. Int J Mol Sci. 2019 Sep 4;20(18):4331. doi: 10.3390/ijms20184331. PMID: 31487867; PMCID: PMC6769923.

18. Stanzione F, Giangreco I, Cole JC. Use of molecular docking computational tools in drug discovery. Prog Med Chem. 2021; 60:273-343. doi: 10.1016/bs.pmch.2021.01.004. Epub 2021 May 27. PMID: 34147204.

19. Naqvi AAT, Mohammad T, Hasan GM, Hassan MI. Advancements in Docking and Molecular Dynamics Simulations Towards Ligand-receptor Interactions and Structurefunction Relationships. Curr Top Med Chem. 2018;18(20):1755-1768. doi: 10.2174/1568026618666181025114157. PMID: 30360721.

20. Hussain MS, Battaglia A, Szczepanski S, Kaygusuz E, Toliat MR, Sakakibara S, Altmüller J, Thiele H, Nürnberg G, Moosa S, Yigit G, Beleggia F, Tinschert S, Clayton-Smith J, Vasudevan P, Urquhart JE, Donnai D, Fryer A, Percin F, Brancati F, Dobbie A, Smigiel R, Gillessen-Kaesbach G, Wollnik B, Noegel AA, Newman WG, Nürnberg P. Mutations in CKAP2L, the human homolog of the mouse Radmis gene, cause Filippi syndrome. Am J Hum Genet. 2014 Nov 6;95(5):622-32. doi: 10.1016/j.ajhg.2014.10.008. Epub 2014 Nov 6. PMID: 25439729; PMCID: PMC4225581.

21. de Vries TI, Monroe GR, van Belzen MJ, van der Lans CA, Savelberg SM, Newman WG, van Haaften G, Nievelstein RA, van Haelst MM. Mosaic CREBBP mutation causes overlapping clinical features of Rubinstein-Taybi and Filippi syndromes. Eur J Hum Genet. 2016 Aug;24(9):1363-6. doi: 10.1038/ejhg.2016.14. Epub 2016 Mar 9. PMID: 26956253; PMCID: PMC4989203.

22. Patrick RJ, Weimer J, Davis-Keppen L, Landsverk ML. Novel variants identified in CKAP2L in two siblings with Filippi syndrome. Cold Spring Harb Mol Case Stud. 2022 Mar 24;8(2):a006130. doi: 10.1101/mcs.a006130. PMID: 34921061; PMCID: PMC8958909.