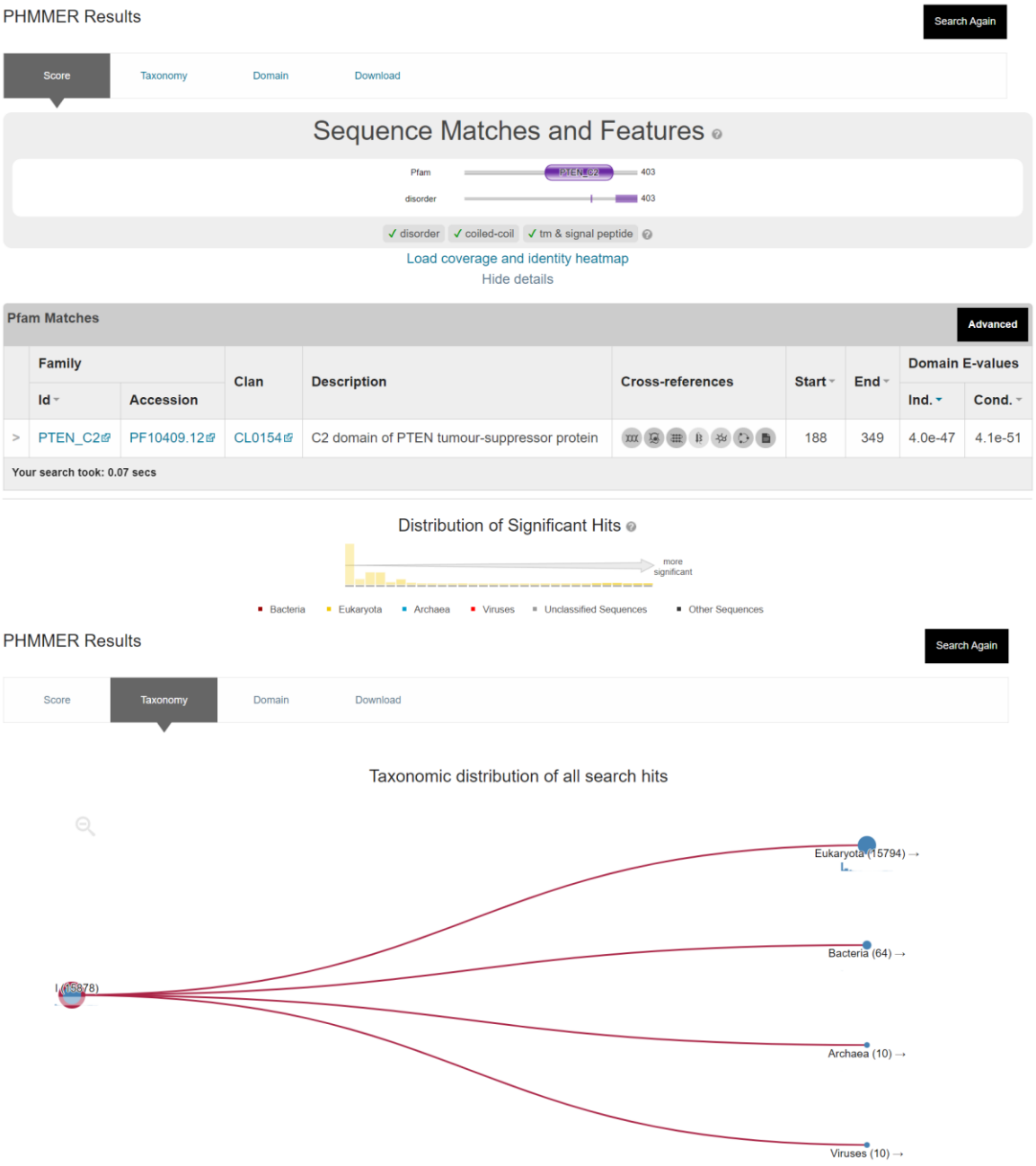
Day 2:

Gene Annotation

Protein Name: phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN isoform PTEN [Homo sapiens]
Protein ID: P60484 (UniProt)
Accession no.: NP_000305
Motif: PDZ domain binding motif (401-403)

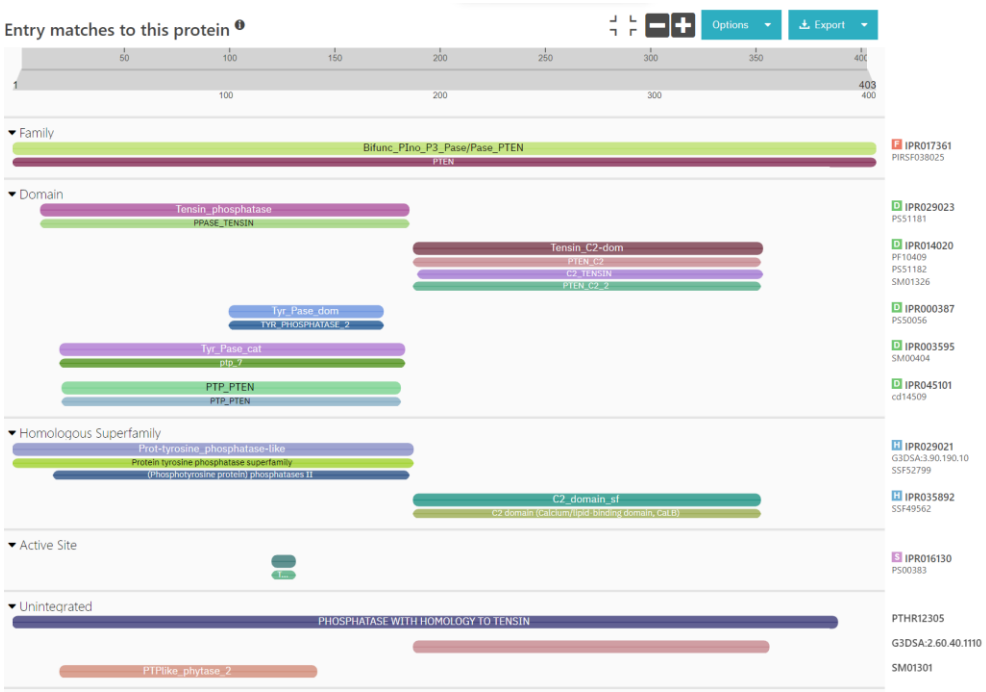
HMMER:



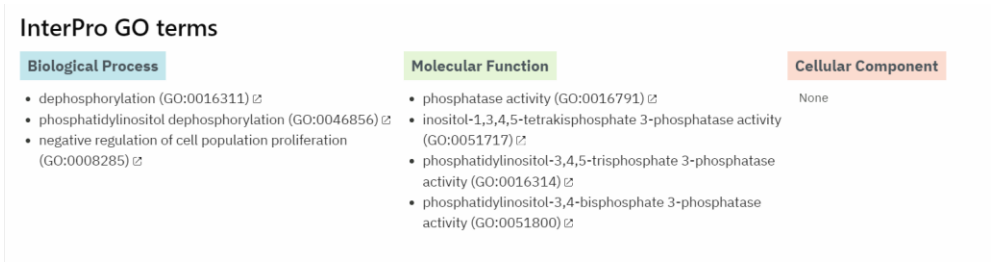
InterProScan:

Family – Bifunctional phosphatidylinositol trisphosphate phosphatase/dual specificity phosphatase PTEN

Domains – Tensin-type phosphatase domain; Tensin phosphatase, C2 domain; Tyrosine-specific protein phosphatases domain; Protein-tyrosine phosphatase, catalytic domain; PTEN, phosphatase domain

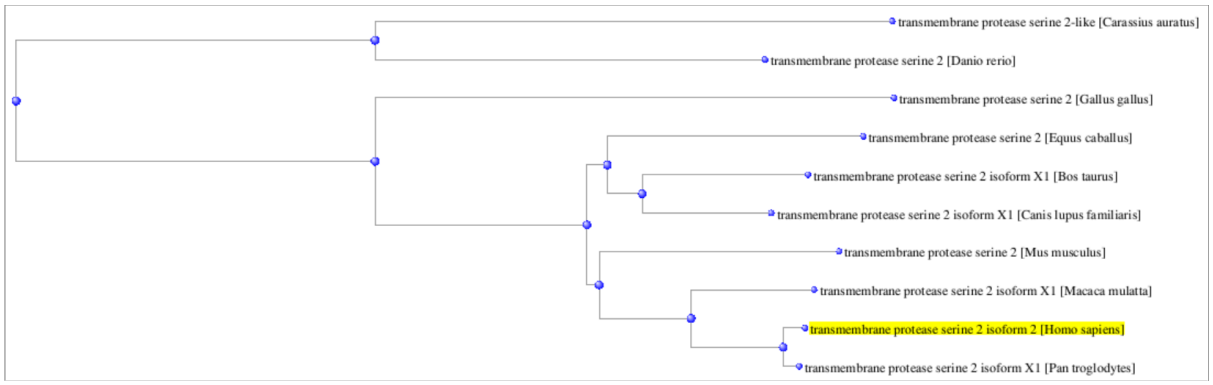


Functions:



Day 3: Phylogenetics

Constructing a Phylogenetic tree for component of Corona virus.

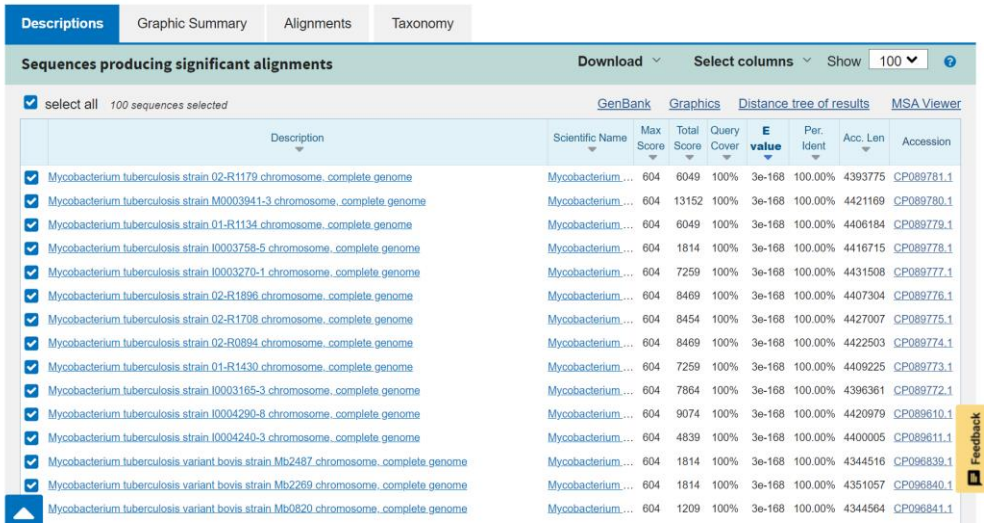


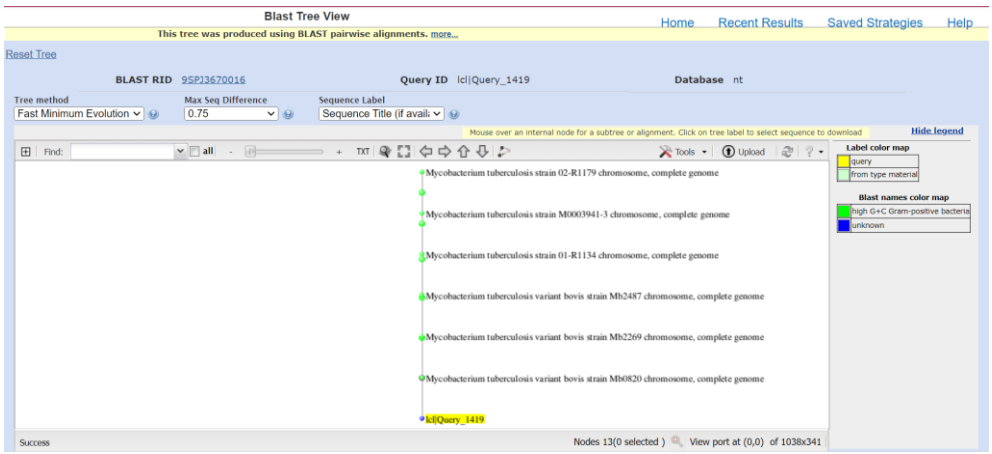
Day 4:

Genome name: Mycobacterium tuberculosis
From RAST results:

1) Nucleotide sequence - atgtcaggtggttcatcgaggaggtacccgccggagctgcgtgagcgggcggtgcggatggtcgcagagatccgcg gtcagcacgattcggagtgggcagcgatcagtgaggtcgcccgtctacttggtgttggctgcgcggagacggtgcgta agtgggtgcgccaggcgcaggtcgatgccggcgcacggcccgggaccacgaccgaagaatccgctgagctgaa gcgcttgcggcgggacaacgccgaattgcgaagggcgaacgcgattttaaagaccgcgtcggctttcttcgcggccg agctcgaccggccagcacgctaa
2) Location on the genome - AL123456.3_889072_889398
3) Start nucleotide – 889072
4) End nucleotide – 889398
5) Function - Insertion element IS6110 (Mycobacterium tuberculosis) transposase

Perform a BLAST on the nucleotide sequence and paste a screenshot of the obtained BLAST results:

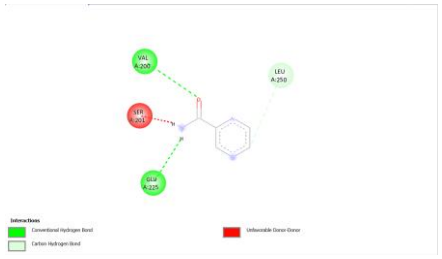
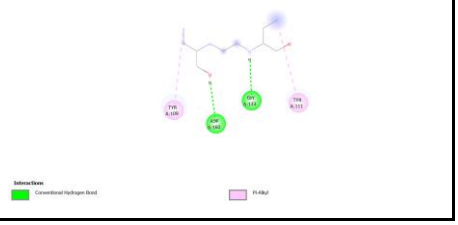




Day 5 & 6:

Molecular Docking

Protein Name: TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide
Protein ID – 2Z80

Ligand Name	Ligand ID	Follows Lipinski Rule?	Energy value	Dock Image
Isoniazid	3767	Yes	96.25	
Pyrazinamide	1046	Yes	61.32	
Ethambutol	14052	Yes	159.47	

Day 7:

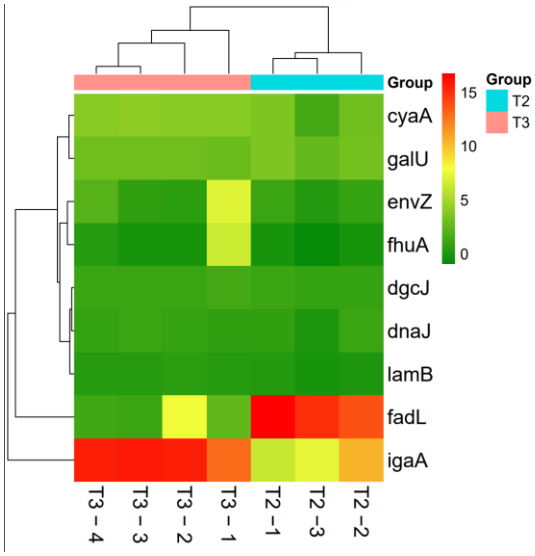
Objective: To plot a heat map and understand the differential expression based on numbered data.
Problem statement: High-throughput mapping of the phage resistance landscape in E. coli.
Reference: doi – https://doi.org/10.1371/journal.pbio.3000877

Input details
1. Gene(s) Name: cyaA, dgcJ, dnaJ, envZ, fadL, fhuA, galU, igaA, lamB (E.coli K-12)
2. Excel data sheet for E. coli K-12:
https://docs.google.com/spreadsheets/d/1pdK6ManQDTXCW_bGQk5HnKoL7KoN-R9J/edit?usp=drive_link&ouid=113843252016005719442&rtpof=true&sd=true

Input data Table:

group	T2	T2	T2	T3	T3	T3	T3
phage	**T2 - 1**	**T2 - 2**	**T2 - 3**	**T3 - 1**	**T3 - 2**	**T3 - 3**	**T3 - 4**
cyaA	3.58	3.12	1.46	3.89	3.96	4.14	3.89
dgcJ	1.13	0.91	0.85	1.41	1.16	1.12	1.02
dnaJ	0.78	1.17	0.04	0.83	1.01	1.05	0.91
envZ	1.03	0.91	0.15	7.13	0.59	0.80	2.19
fadL	16.75	13.88	15.11	2.60	7.90	1.10	1.33
fhuA	-0.19	-0.08	-0.93	6.45	-0.05	-0.18	0.39
galU	3.65	3.28	2.61	2.87	3.05	3.10	2.99
igaA	6.30	10.46	7.26	12.96	15.52	15.80	15.56
lamB	0.23	0.04	-0.14	0.44	0.53	0.36	0.43

Output heatmap:



Discussion points:

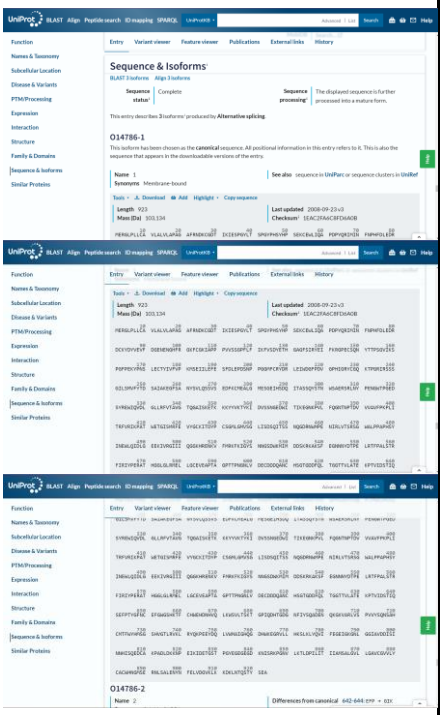
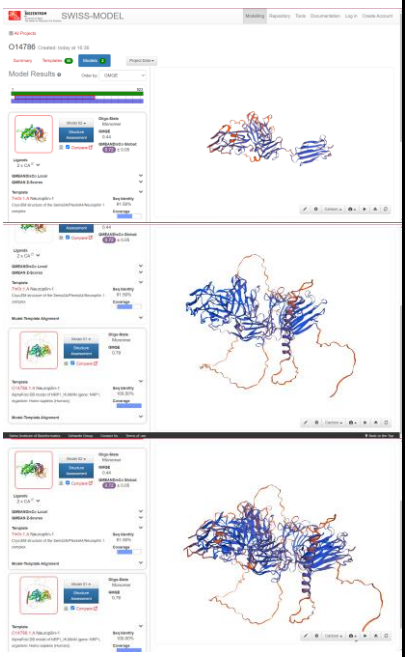
1. Large values have red colour
2. Small values have green colour
3. Darkness of the colour indicates the extremity of the values

Five interpretation points understood:

1. envZ and fhuA values are put together due to their similar trend
2. Phages T3-3 and T3-4 almost have the similar values and hence put together
3. igaA gene seems to have similar resistance to both phages T3-3 and T3-4
4. igaA gene is found to have more resistance to phage T3-2 than fadL gene
5. fhuA shows darker shade of green due to being the least resistant to T2-3 and fadL shows darker shade of red due to being most resistant to T2-1

Day 8 & 9: Homology Modelling:

Problem statement: To visualize the 3D structure of Neuropilin-1 by homology modelling
Protein: Neuropilin-1
Gene: NRP1
PDB: O14786 (First Isoform)
Hypothesis: Homology modelling of Neuropilin-1 will provide insights into its interaction with the SARS-CoV-2 spike protein and the CendR motif RRAR, enabling a better understanding of the mechanism by which Neuropilin-1 enhances SARS-CoV-2 infection.
Purpose: The purpose of this study is to develop a 3D structure of Neuropilin-1 using homology modelling to elucidate the molecular details of its interaction with the SARS-CoV-2 spike protein. By gaining a structural understanding of this interaction, we aim to investigate the role of Neuropilin-1 as a host factor for SARS-CoV-2 infection and the specific binding mechanism involving the CendR motif RRAR.
Outcome: The outcome of this study will be a reliable 3D model of Neuropilin-1, which can be used to identify key residues involved in the recognition and binding of the CendR motif RRAR on the SARS-CoV-2 spike protein S1. This structural information can potentially aid in the development of therapeutics that target the interaction between Neuropilin-1 and SARS-CoV-2, aiming to disrupt or inhibit this interaction and reduce the infectivity of the virus. Additionally, understanding the structural basis of the Neuropilin-1/SARS-CoV-2 interaction may also provide insights into the general mechanisms of viral infection and host factors involved in viral entry, potentially leading to the development of broader antiviral strategies.

Target	Sequence	Result
		

Day 10: GitHub account link – github.com/Arun0364