Establishing Phylogenetic Profile and Identification of Function and Ligands for *Mycobacterium Ulcerans* Microbiome

G. Arpitha[1,2(✉)] and Preenon Bagchi[1,2,3]

[1] Padmashree Institute of Management and Science, Bangalore, India
arpithag107@gmail.com
[2] Vasishth Academy of Advanced Studies and Research (Sarvasumana Association), Bangalore, India
[3] MGM Institute of Biosciences and Technology, Aurangabad, India

Abstract. Buruli ulcer is caused primarily by *Mycobacterium ulcerans.* Due to its high association with the Buruli ulcer, *Mycobacterium ulcerans* microbiota were used in the study. *This microbiota taxon and functional profile have been identified thanks to advances in* Metatranscriptomic Sequencing. Then there are several genes associated with the disease that are identified and used for the further docking study. There are many Ayurvedic medicinal plants, such as *Jatropha curcas L., Aloe vera (L.) Burm. f.,* and *Capsicum annum L., that* are used for medicine.

Keywords: *Mycobacterium ulcerans* · Buruli ulcer · Microbiome · Metatranscriptomics · phytocompounds · Molinspiration software · Patchdock tool · Docking · ADME analysis

1 Introduction

The environmental pathogen *Mycobacterium ulcerans* primarily causes buruli ulcers, which are infections of the subcutaneous fat. The Buruli ulcer clinical lesion typically begins as a lump that later develops into an ulcer with recognizable undermined margins. Mycolactone, a toxin produced by *Mycobacterium ulcerans*, has cytotoxic and immunosuppressive effects that cause necrosis and ulceration. Some individuals experience joint lesions as well as bone (osteomyelitis) lesions. Buruli ulcer is the third most prevalent Mycobacterial infection in humans, after leprosy and tuberculosis. The nodular, plaque, and oedematous forms of Buruli ulcer have more challenging clinical diagnoses, despite the ulcerative form's simple clinical diagnosis [1]. *Mycobacterium ulcerans* has been studied in some depth with regard to its dissemination, but it is still unclear exactly how this bacteria spreads, which makes it difficult to regulate. *Mycobacterium ulcerans* is regarded as an environmental pathogen because it is connected to lentic environments and rarely spreads from person to person. However, a number of mammals and invertebrates also carry the virus, and these animals may act as important reservoirs and mechanical vectors, respectively [2]. The diagnosis of Buruli ulcer

R. Somashekhar et al. (Eds.): ICBDS 2022, AHSR 58, pp. 65–78, 2023.
https://doi.org/10.2991/978-94-6463-164-7_6

is typically relied on clinical factors and infrequently verified. Ziehl-Neelsen staining is thought to be 40–80% sensitive, histopathology and PCR is over 90% sensitive whereas culture method is 20–60% sensitive. The quickest and most accurate diagnosis tool is PCR, but implementing it in the field is technically challenging. The Buruli ulcer disease burden includes high costs, major surgery, protracted hospital stays, and emergence of disabling consequences. Early detection and rapid treatment of Buruli ulcer is the best way to reduce morbidity and costs of the disease. Early detection and treatment will not decrease transmission, as human-to-human transmission is very rare. Inexpensive prevention strategies such as wearing protective clothing and immediate cleansing of traumatic skin injuries might also help [1].

Metatranscriptomic analysis using Galaxy

The non-Sanger-based sequencing technology known as next-generation sequencing (NGS) is an improved variant that enables us. Galaxy is a web-based, open source platform for cutting-edge computational biomedical research [3]. Metagenomics' transcriptomics dataset provides valuable information regarding the entire gene utterance profiling of a complex microbial ecosystem component. While metatranscriptomics provides the variety of the active genes within such a community, their expression profile, and how these levels alter as a result of changes in environmental conditions, metagenomic studies primarily focus on the genomic content and recognition of microbes nearby within a community. Metatranscriptomics is a viable method for searching the eukaryotic gene pool for genes with biotechnological significance because it relies on the use of mRNA isolated from environmental samples [4].

Computer aided drug design

Drug design refers to the entire process of developing a new compound or medicinal molecule. Here, the molecular structure is constructed, displayed, simulated, and examined using a structure-based drug development technique. with the help of the SWISS-MODEL tool [5] for modelling the proteins responsible for Buruli ulcer disease. Then, ADME drug analysis is performed to determine the toxicity of the drug, followed by molecular docking to determine the interaction between the ligand and receptor [6, 7].

2 Materials and Method

Mycobacterium ulcerans fastq sequences SRR16611551.1.1 and SRR16611551.1.2 obtained from the SRA database FASTQC [8] confirmed the sequence's quality score.MultiQC [9] is used to combine the reports of FastaQC. Sequences were trimmed using cutadapt.

Table 1. MetaPhlAn: Taxonomic profiling

kingdom	phylum	phylum_id	class	class_id	order	order_id	Family	family_id	genus	genus_id	species
Bacteria								100.0			
Bacteria	Actinobacteria	201174							100.0		
Bacteria	Actinobacteria	201174	Actinobacteria	1760						100.0	
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007					100.0
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007	Mycobacteriaceae	1762			
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007	Mycobacteriaceae	1762	Mycobacterium	1763	
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007	Mycobacteriaceae	1762	Mycobacterium	1763	*Mycobacterium ulcerans*
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007	Mycobacteriaceae	1762	Mycobacterium	1763	Mycobacterium pseudoshottsii
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007	Mycobacteriaceae	1762	Mycobacterium	1763	Mycobacterium liflandii
Bacteria	Actinobacteria	201174	Actinobacteria	1760	Corynebacteriales	85007	Mycobacteriaceae	1762	Mycobacterium	1763	Mycobacterium marinum

The SortMeRNA tool [10–14] is then used for local sequence alignment and sequence mapping and clustering.Next, using the FASTQ INTERLACE tool [15], we paired end-to-end FASTQ reads from two different segregated files.

MetaPhlAn [16] profiles microbial communities (Bacteria, Archaea, and Eukaryotes) in our research microbiota.

Krona tool [17, 18] was used to show the microbial communities in the diagrammatic image. Further, the HUMAnN [19] pipeline was used to efficiently identify the abundance of microbial pathways in our microbiota.

Next, our disease genes were identified, and their 3D structure was obtained using SWISS-MODEL [5].

The properties and actions of phytocompounds are identified, and their structures are downloaded.

Using molinspiration software [20], used for the identification of the chemical properties of phytocompounds.

Further, docking was performed using patchdock [22, 23].

3 Results and Discussion

Metagenome, having Accession number SRR16611551, for *Mycobacterium ulcerans* was downloaded from SRA Database.

We proceed with trimming the sequence because of the poor sequence quality, as indicated by the results of FASTQC and MultiQC's per-base sequence quality tests.

CUTADAPT tool [24] is used for trimming. It detects and removes from our data any adaptor sequences, primers, poly-A tails, and other sorts of undesirable sequence. When it locates the adapter, it looks through all reads and removes it. In addition, the sequence quality of the output from the cutadapt is examined using FASTQC and MultiQC, and it is discovered to be within the acceptable range.

SortMeRNA tool removes any reads identified as rRNA from our sequence. Taxonomic profiling [25] was obtain from the using MetaPhlAn tool (Table 1). The output is visualized using Krona (Fig. 1).

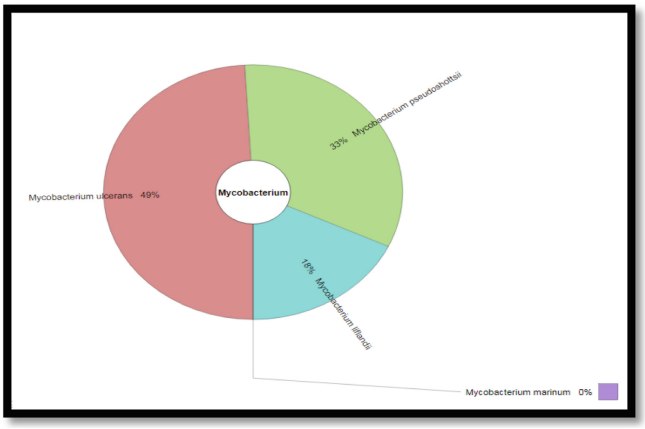


Fig. 1. Visualization of Taxonomic profile in Krona

We next go on to functional knowledge about our microbiome after taxonomy generation. Functional information of this microbiome community [24] was obtain from the using HUMAnN pipeline (Table 2).

Table 2. Functional information of microbiome

# Gene Family	humann_Abundance-RELAB
UNMAPPED	0.126682
UniRef90_A0PR89	0.00948495
UniRef90_A0PR89|g__Mycobacterium.s__Mycobacterium_ulcerans	0.0035743
UniRef90_A0PR89|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00301456
UniRef90_A0PR89|g__Mycobacterium.s__Mycobacterium_liflandii	0.00289609
UniRef90_L7V8L9	0.00819922
UniRef90_L7V8L9|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00752194
UniRef90_L7V8L9|g__Mycobacterium.s__Mycobacterium_liflandii	0.000677284
UniRef90_X8FKY7	0.00599434
UniRef90_X8FKY7|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00599434
UniRef90_X8FGU9	0.00567548
UniRef90_X8FGU9|g__Mycobacterium.s__Mycobacterium_marinum	0.00420406
UniRef90_X8FGU9|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00147142
UniRef90_A0A100IGI4	0.00546528
UniRef90_A0A100IGI4|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00546528
UniRef90_A0A2Z5YHQ7	0.00525508
UniRef90_A0A2Z5YHQ7|g__Mycobacterium.s__Mycobacterium_marinum	0.00525508
UniRef90_X8F9Q5	0.00516928
UniRef90_X8F9Q5|g__Mycobacterium.s__Mycobacterium_ulcerans	0.0037854
UniRef90_X8F9Q5|g__Mycobacterium.s__Mycobacterium_liflandii	0.00138388
UniRef90_L7V896	0.00458588
UniRef90_L7V896|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00351825
UniRef90_L7V896|g__Mycobacterium.s__Mycobacterium_liflandii	0.000590374
UniRef90_L7V896|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000477261
UniRef90_L7VC62	0.00420406
UniRef90_L7VC62|g__Mycobacterium.s__Mycobacterium_liflandii	0.00315305
UniRef90_L7VC62|g__Mycobacterium.s__Mycobacterium_marinum	0.00105102
UniRef90_X8FJI4	0.00409356
UniRef90_X8FJI4|g__Mycobacterium.s__Mycobacterium_liflandii	0.00409356
UniRef90_A0A124BUD0	0.00336325
UniRef90_A0A124BUD0|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00336325
UniRef90_L7V714	0.00336325
UniRef90_L7V714|g__Mycobacterium.s__Mycobacterium_marinum	0.00189183
UniRef90_L7V714|g__Mycobacterium.s__Mycobacterium_liflandii	0.00147142
UniRef90_L7V5V5	0.00319209
UniRef90_L7V5V5|g__Mycobacterium.s__Mycobacterium_liflandii	0.00225791
UniRef90_L7V5V5|g__Mycobacterium.s__Mycobacterium_ulcerans	0.000542469
UniRef90_L7V5V5|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000391709
UniRef90_X8FLQ5	0.00294856
UniRef90_X8FLQ5|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00294856
UniRef90_A0PKM2	0.0028964
UniRef90_A0PKM2|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00180324
UniRef90_A0PKM2|g__Mycobacterium.s__Mycobacterium_liflandii	0.00109315
UniRef90_A0A100ICA9	0.00273264

(*continued*)

Table 2. (*continued*)

UniRef90_A0A100ICA9|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00273264
UniRef90_X8F225	0.00269174
UniRef90_X8F225|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00187628
UniRef90_X8F225|g__Mycobacterium.s__Mycobacterium_liflandii	0.000815461
UniRef90_L7V4W1	0.00264305
UniRef90_L7V4W1|g__Mycobacterium.s__Mycobacterium_liflandii	0.00264305
UniRef90_L7V7K7	0.00262292
UniRef90_L7V7K7|g__Mycobacterium.s__Mycobacterium_liflandii	0.00262292
UniRef90_A0A100I9D0	0.00255747
UniRef90_A0A100I9D0|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00255747
UniRef90_A0A1B4Y824	0.00253337
UniRef90_A0A1B4Y824|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00253337
UniRef90_L7V8B6	0.00241733
UniRef90_L7V8B6|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00241733
UniRef90_L7V4Z3	0.00235808
UniRef90_L7V4Z3|g__Mycobacterium.s__Mycobacterium_liflandii	0.00111466
UniRef90_L7V4Z3|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00102848
UniRef90_L7V4Z3|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000214942
UniRef90_X8FDU5	0.00234877
UniRef90_X8FDU5|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00234877
UniRef90_A0A2Z5TBT5	0.00231223
UniRef90_A0A2Z5TBT5|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00231223
UniRef90_L7V7K4	0.00217513
UniRef90_L7V7K4|g__Mycobacterium.s__Mycobacterium_liflandii	0.00217513
UniRef90_A0A117DZK8	0.0021721
UniRef90_A0A117DZK8|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.0021721
UniRef90_A0PPK7	0.00180036
UniRef90_A0PPK7|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00180036
UniRef90_L7VDK6	0.00176909
UniRef90_L7VDK6|g__Mycobacterium.s__Mycobacterium_liflandii	0.00176909
UniRef90_A0PMP6	0.00172366
UniRef90_A0PMP6|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00172366
UniRef90_L7VBD4	0.00171513
UniRef90_L7VBD4|g__Mycobacterium.s__Mycobacterium_liflandii	0.00171513
UniRef90_A0A117DW64	0.00168162
UniRef90_A0A117DW64|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00101598
UniRef90_A0A117DW64|g__Mycobacterium.s__Mycobacterium_marinum	0.000665643
UniRef90_A0A117DX51	0.00164659
UniRef90_A0A117DX51|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00164659
UniRef90_L7UZQ8	0.00161156
UniRef90_L7UZQ8|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00161156
UniRef90_L7V660	0.00161139
UniRef90_L7V660|g__Mycobacterium.s__Mycobacterium_liflandii	0.00161139
UniRef90_A0A100I0D9	0.00154149
UniRef90_A0A100I0D9|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00154149

(*continued*)

Table 2. (*continued*)

UniRef90_L7V5A8	0.0015364
UniRef90_L7V5A8|g__Mycobacterium.s__Mycobacterium_liflandii	0.000675323
UniRef90_L7V5A8|g__Mycobacterium.s__Mycobacterium_ulcerans	0.000497108
UniRef90_L7V5A8|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000363971
UniRef90_A0A2Z5TZT9	0.00151553
UniRef90_A0A2Z5TZT9|g__Mycobacterium.s__Mycobacterium_liflandii	0.000798772
UniRef90_A0A2Z5TZT9|g__Mycobacterium.s__Mycobacterium_marinum	0.000441426
UniRef90_A0A2Z5TZT9|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000273264
UniRef90_A0A2Z5TZT9|unclassified	2.07097e-06
UniRef90_A0A2Z5TDE7	0.00147142
UniRef90_A0A2Z5TDE7|g__Mycobacterium.s__Mycobacterium_ulcerans	0.000820793
UniRef90_A0A2Z5TDE7|g__Mycobacterium.s__Mycobacterium_marinum	0.000650628
UniRef90_A0A124BX73	0.00138352
UniRef90_A0A124BX73|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00126122
UniRef90_A0A124BX73|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.0001223
UniRef90_A0A117DWW6	0.00136632
UniRef90_A0A117DWW6|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00136632
UniRef90_A0PQY3	0.00135769
UniRef90_A0PQY3|g__Mycobacterium.s__Mycobacterium_marinum	0.000385372
UniRef90_A0PQY3|g__Mycobacterium.s__Mycobacterium_liflandii	0.000385372
UniRef90_A0PQY3|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000350338
UniRef90_A0PQY3|g__Mycobacterium.s__Mycobacterium_ulcerans	0.000236606
UniRef90_A0PN30	0.00133462
UniRef90_A0PN30|g__Mycobacterium.s__Mycobacterium_ulcerans	0.000457775
UniRef90_A0PN30|g__Mycobacterium.s__Mycobacterium_marinum	0.000400387
UniRef90_A0PN30|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000266257
UniRef90_A0PN30|g__Mycobacterium.s__Mycobacterium_liflandii	0.000210203
UniRef90_A0PLJ9	0.00133129
UniRef90_A0PLJ9|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00133129
UniRef90_L7VA12	0.00132428
UniRef90_L7VA12|g__Mycobacterium.s__Mycobacterium_marinum	0.000777751
UniRef90_L7VA12|g__Mycobacterium.s__Mycobacterium_liflandii	0.000546528
UniRef90_A0A2Z5TPV5	0.00131625
UniRef90_A0A2Z5TPV5|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00105102
UniRef90_A0A2Z5TPV5|g__Mycobacterium.s__Mycobacterium_ulcerans	0.000180083
UniRef90_A0A2Z5TPV5|g__Mycobacterium.s__Mycobacterium_marinum	6.05803e-05
UniRef90_A0A2Z5TPV5|g__Mycobacterium.s__Mycobacterium_liflandii	2.45712e-05
UniRef90_L7V7S9	0.00129697
UniRef90_L7V7S9|g__Mycobacterium.s__Mycobacterium_liflandii	0.00129697
UniRef90_A0A2Z5T8B2	0.00129625
UniRef90_A0A2Z5T8B2|g__Mycobacterium.s__Mycobacterium_liflandii	0.00129625
UniRef90_A0A100I7C5	0.00122618
UniRef90_A0A100I7C5|g__Mycobacterium.s__Mycobacterium_marinum	0.000840812
UniRef90_A0A100I7C5|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000385372
UniRef90_A0PX23	0.00122118

(*continued*)

Table 2. (*continued*)

UniRef90_A0PX23	0.00122118
UniRef90_A0PX23|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00122118
UniRef90_L7V6V6	0.0011783
UniRef90_L7V6V6|g__Mycobacterium.s__Mycobacterium_liflandii	0.000875866
UniRef90_L7V6V6|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000302432
UniRef90_A0A117DV50	0.0011351
UniRef90_A0A117DV50|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.0011351
UniRef90_B2HLJ0	0.00111107
UniRef90_B2HLJ0|g__Mycobacterium.s__Mycobacterium_ulcerans	0.00111107
UniRef90_A0A124BVG6	0.00109306
UniRef90_A0A124BVG6|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00100897
UniRef90_A0A124BVG6|g__Mycobacterium.s__Mycobacterium_marinum	8.40812e-05
UniRef90_A0A100IA97	0.00105102
UniRef90_A0A100IA97|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00105102
UniRef90_A0A2Z5T8J9	0.00101658
UniRef90_A0A2Z5T8J9|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.000840812
UniRef90_A0A2Z5T8J9|g__Mycobacterium.s__Mycobacterium_marinum	0.000111544
UniRef90_A0A2Z5T8J9|g__Mycobacterium.s__Mycobacterium_ulcerans	6.42287e-05
UniRef90_A0A117DWZ1	0.00100897
UniRef90_A0A117DWZ1|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00100897
UniRef90_A0A2Z5TVN3	0.00100515
UniRef90_A0A2Z5TVN3|g__Mycobacterium.s__Mycobacterium_pseudoshottsii	0.00100515

Table 3. Protein with its NCBI Accession number

Sl. no.	Protein Name	NCBI Accession Number	Homologous Template
1.	SDR family oxidoreductase [*Mycobacterium ulcerans*]	WP_156091109.1	2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase

Structure based drug designing for Mycobacterium ulcerans

Since, Buruli ulcer is a bacterial disease, we further go ahead towards designing novel drug for the disease. From NCBI, we retrieve reads of Whole Genome Sequencing of *Mycobacterium ulcerans* and run BLASTx (Translate nucleotide to Protein). Select the first proteins' accession number in description of BLAST result. Obtain FASTA sequence of the protein and enter it in SWISS-MODEL search box and execute Build model.

Homology Modeling

SWISS-MODEL server is used to model the homology of the above mentioned receptor. The receptor model is shown, along with the associated ramachandran plot data. Template used for modeling is given in Table 3 (Fig. 2) [21].

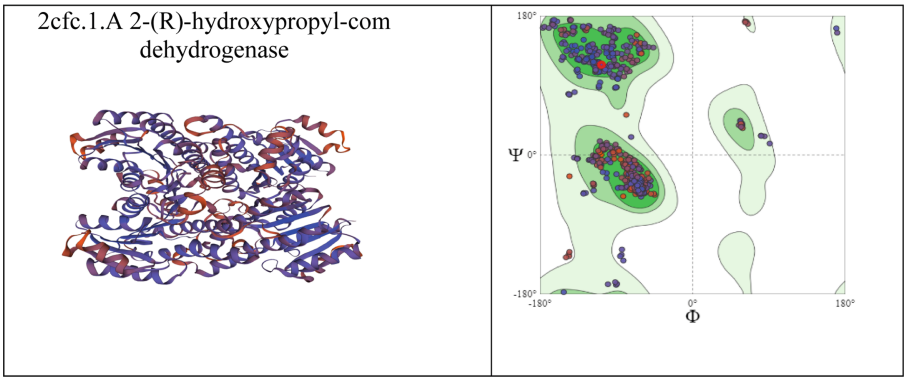


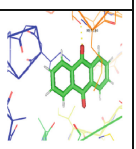
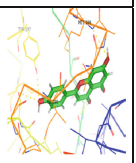
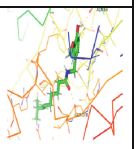
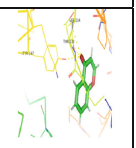
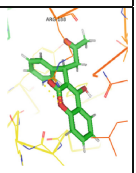
Fig. 2. Swiss-model generated receptor model with its ramachandran plot

Table 4. ADME results of Phytocompounds used to treat Buruli ulcer

Sl no.	Phytocompound	miLogP	TPSA	natoms	Mol wt.	nON	nOHNH	nrotb	Volume	nviolations
1	1-Docosanol	9.17	20.23	23	326.61	1	1	20	390.05	1
2	1-Triacontanol	9.98	20.23	31	438.82	1	1	28	524.47	1
3	Anthraquinone	3.67	34.14	16	208.22	2	0	0	182.58	0
4	Apigenin	2.46	90.89	20	270.24	5	3	1	224.05	0
5	Capsaicinoids	3. 10	58.56	22	305.42	4	2	9	310.37	0
6	Capsorubin	8.46	74. 60	44	600.88	4	2	12	624.19	2
7	Chromone	1.82	30.21	11	146.15	2	0	0	128.59	0
8	Conessine	4.79	6.48	26	356. 60	2	0	1	379. 40	0
9	Coumarins	3.03	67.51	23	308.33	4	1	4	277.19	0
10	Cryptoxanthin	9.64	20.23	41	552.89	1	1	10	600.01	2
11	Holadysamine	4.41	32.26	24	329.53	2	2	2	345.52	0
12	Holarrhenine	3.87	26. 70	27	372. 60	3	1	1	387.45	0
13	Kurchamine	4.07	38.05	24	330.56	2	3	2	355.01	0
14	Lignin	-1.19	188.45	31	463.45	11	2	5	347.93	1
15	Lutein	9.31	40.46	42	568.89	2	2	10	608.08	2
16	Progesterone	3.81	34.14	23	314.47	2	0	1	319.07	0
17	Salicylic acid	1.87	57.53	10	138.12	3	2	1	119.06	0
18	Sapogenins	7. 50	18.47	29	400.65	2	0	0	417.55	1
19	Saponins	0. 10	388.05	79	1131.27	24	13	16	1016.89	3
20	Sitosterol	8.62	20.23	30	414.72	1	1	6	456.52	1
21	Stearic acid	8.07	37. 30	20	284.48	2	1	16	325.03	1
22	Stigmasterol	7.87	20.23	30	412. 70	1	1	5	450.33	1

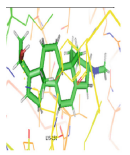
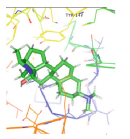
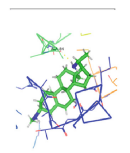
Ayurvedic Medicinal plants such as *Pupalia lappacea, Holarrhena floribunda, Aloe vera (L.) Burm. f., Jatropha curcas L., Capsicum annum L.* are traditionally apply to

Table 5. Docking of 2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase gene receptor

Protein	Plant name	Phytocompounds	Interacting Amino acids	No. of interactions	Docking Score (-Kcal/mol)	Docking
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Aloe vera (L.) Burm. f.*	Anthraquinone	MET-184	1	-4022	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Jatropha curcas L.*	Apigenin	TYR-157, MET-164, ILE-180	3	-4334	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Capsicum annum L.*	Capsaicinoids	ALA-84, GLY-178	2	-5742	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Aloe vera (L.) Burm. f.*	Chromone	TYR-147, THR-133, SER-134	3	-3126	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Holarrhena floribunda*	Conessine	-	-	-	-
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Pupalia lappacea*	Coumarins	ARG-188	1	-4662	

(*continued*)

Table 5. (*continued*)

2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Holarrhena floribunda*	Holadysamine	LYS-151	1	-5096	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Holarrhena floribunda*	Holarrhenine	TYR-147	1	-5572	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Holarrhena floribunda*	Kurchamine	-	-	-	-
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Holarrhena floribunda*	Progesterone	ALA-84	1	-5434	
2cfc.1.A 2-(R)-hydroxypropyl-com dehydrogenase	*Aloe vera (L.) Burm. f.*	Salicylic acid	TYR-147	1	-2826	

treatment of disease and inflammation, ecrosis ablation, wound curing, and exorcism. The potency of their phytocompounds in treating Buruli ulcer is studied here.

According to Lipinski's rule of ADME, we check the drug likeness of the above phytocompounds (Table 4).

It is seen that phytocompounds 1-Docosanol, 1-Triacontanol, Lignin, Sapogenins, Sitosterol, Stearic acid, Stigmasterol, Vitexin have 1 violations each and Zeaxanthin, Lutein, Cryptoxanthin, Capsorubin have 2 violations each and Saponins has 3 violations and Tannins has 4 violations number, so according to Lipinski's rule and it now can not be considered as ligands for the disease.

Molecular Docking

Next docking is performed in Table 5.

As per docking studies it is seen that phytocompounds Apigenin, Chromone, and Capsaicinoids from *Jatropha curcas L., Aloe vera (L.) Burm. f., and Capsicum annum L.* respectively have good interactions with the Buruli ulcer gene receptors.

4 Conclusion

RNA sequence in *Mycobacterium ulcerans* microbiome was analyzed by Metatranscriptomics by using Galaxy Australia tool. This tool identifies the quality of the sequence in bacteria and also helps in generation of functional taxonomy pathway. Taxonomic profiling was done using MetaPhlAn tool. The output is Visualized using Krona. The output as phylogenetic tree was observed. It depicts that Mycobacteriaceae, a family of *Mycobacterium* has evolutionary connection with *Mycobacterium ulcerans*. To analyze the action of proteins on homology modeling docking studies was carried out to find novel ligands against Buruli ulcer. Again, as per docking studies and ADME analysis it is seen that phytocompounds Apigenin, Chromone, and Capsaicinoids from *Jatropha curcas L., Aloe vera (L.) Burm. f., and Capsicum annum L.* respectively have good interactions with the Buruli ulcer gene receptors. They work as a possible ligands for the receptors of Buruli ulcer, and the ADME analysis and docking studies prove that this ligand shows a potent impact against the disease. For this reason, we are investigating the above phytocompounds in vitro and in vivo to determine their potential impact on the Buruli ulcer.

References

1. Sizaire, V., Nackers, F., Comte, E., & Portaels, F. (2006). *Mycobacterium ulcerans* infection: control, diagnosis, and treatment. The Lancet infectious diseases, 6(5), 288-296. https://doi.org/10.1016/S1473-3099(06)70464-9
2. Muleta AJ, Lappan R, Stinear TP, Greening C (2021) Understanding the transmission of *Mycobacterium ulcerans*: A step towards controlling Buruli ulcer. PLOS Neglected Tropical Diseases 15(8): e0009678. https://doi.org/10.1371/journal.pntd.0009678
3. Blankenberg, D., & Hillman-Jackson, J. (2014). Analysis of next-generation sequencing data using Galaxy. In *Stem cell transcriptional networks* (pp. 21–43). Humana Press, New York, NY.
4. Mukherjee, A., & Reddy, M. S. (2020). Metatranscriptomics: an approach for retrieving novel eukaryotic genes from polluted and related environments. 3 Biotech, 10(2), 71. https://doi.org/10.1007/s13205-020-2057-1
5. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, A P de Beer T, Rempfer C, Bordoli L, Lepore R and Schwede T, (2018), SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res.; 46(Web Server issue): W296–W303.
6. Vyas VK, Ukawala RD, Ghate M, and Chintha C (2012), Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives, Indian J Pharm Sci.74(1): 1–17.
7. Morris GM, Lim-Wilby M, (2008), Molecular docking, Methods Mol Biol, 443:365-82.
8. Andrews, S. (n.d.). FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

9. Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics, 32(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354
10. Kopylova, E., Noé, L., &Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics, 28(24), 3211–3217. https://doi.org/10.1093/bioinformatics/bts611
11. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(D1), D590–D596. https://doi.org/10.1093/nar/gks1219
12. Burge, S. W., Daub, J., Eberhardt, R., Tate, J., Barquist, L., Nawrocki, E. P., … Bateman, A. (2012). Rfam 11.0: 10 years of RNA families. Nucleic Acids Research, 41(D1), D226–D232. https://doi.org/10.1093/nar/gks1005
13. Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461
14. Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., &Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. Nature Biotechnology, 30(5), 434–439. https://doi.org/10.1038/nbt.2198
15. Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., & and, A. N. (2010). Manipulation of FASTQ data with Galaxy. Bioinformatics, 26(14), 1783–1785. https://doi.org/10.1093/bioinformatics/btq281
16. Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., … Segata, N. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. ELife, 10. https://doi.org/10.7554/elife.65088
17. Ondov, B. D., Bergman, N. H., &Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. BMC Bioinformatics, 12(1). https://doi.org/10.1186/1471-2105-12-385
18. Cuccuru, G., Orsini, M., Pinna, A., Sbardellati, A., Soranzo, N., Travaglione, A., … Fotia, G. (2014). Orione, a web-based framework for NGS analysis in microbiology. Bioinformatics, 30(13), 1928–1929. https://doi.org/10.1093/bioinformatics/btu135
19. Abubucker, S., Segata, N., Goll, J., Schubert, A. M., Izard, J., Cantarel, B. L., … Huttenhower, C. (2012). Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. PLoS Computational Biology, 8(6), e1002358. https://doi.org/10.1371/journal.pcbi.1002358
20. https://www.molinspiration.com, SlovenskyGrob, Slovakia
21. Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound Docking of Rigid Molecules. In Gusfield et al., Ed. Proceedings of the 2'nd Workshop on Algorithms in Bioinformatics (WABI) Rome, Italy, Lecture Notes in Computer Science 2452, pp. 185-200, Springer Verlag, 2002
22. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl. Acids. Res. 33: W363-367, 2005.
23. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal, 17(1), 10. https://doi.org/10.14806/ej.17.1.200
24. Jagtap P, Mehta S, Sajulga R, Batut B, Leith E, Kumar P, Hiltemann S, 2021 Metatranscriptomics analysis using microbiome RNA-seq data (Galaxy Training Materials).
25. Agarwal, S., & Mehrotra, R. J. J. C. (2016). An overview of molecular docking. JSM chem, 4(2), 1024-1028.