Establishing the Taxa with Phylogenetic Profile and *in-silico* Ayurvedic Remedy of Cervicitis Microbiome

Milan Khandelwal[1,2(✉)], I. A. Shylesh Murthy[2], and Preenon Bagchi[1,2,3]

[1] Padmashree Institute of Management and Science, Bangalore, India
milankhandelwal35@gmail.com
[2] Vasishth Academy of Advanced Studies and Research (Sarvasumana Association), Bangalore, India
[3] MGM Institute of Biosciences and Technology, Aurangabad, India

Abstract. Cervicitis is an inflammation of the cervix, the lower, narrow end of the uterus that opens into the vagina. Basically, it causes inflammation of the uterine cervix. Cervicitis is a clinical syndrome characterised by inflammation of primarily the columnar epithelium of the uterine endocervix. The bacteria associated with cervicitis is *Mycoplasm hominis,* which is a small bacterium that has the ability to cause cervicitis in the female and male genital areas. Mycoplasma hominis has a triple-layered membrane, no cell wall, and a highly pleomorphic shape and size. They lack even cell wall precursors like muramic acid or diaminopimelic acid. Because of their ability to inhibit cell wall synthesis, commonly used antibiotics are generally ineffective. One study suggested that up to 21% of men and 54% of women assessed with a sexually transmitted disease clinically showed signs and symptoms of cervicitis, indicating the importance of improving treatment strategies for women and men. Due to its association with serious complications like infertility, cervicitis should be promptly investigated and treated. The development of next-generation sequencing (NGS) techniques has enabled researchers to study and understand the world of microorganisms from broader and deeper perspectives. 16S rDNA sequencing, also known as metagenomics, is a fundamental sequencing strategy used in the identification and characterization of spices. Metagenomics has paved the way for drug development against the genetic potential of microorganisms that cannot be grown in vitro or in the laboratory but can infect humans. Molecular docking methodologies are of great importance in the planning and design of new drugs. The goal of these methods is to predict the experimental binding mode and affinity of a small molecule within the binding site of the receptor target of interaction as well as to identify a pathway.

Keywords: Cervicitis · Microbiome · Metagenomics · Taxonomic · Gene Receptor · Docking · Pathway Analysis

1 Introduction

Myoplasm hominis is a species of bacteria that can cause cervicitis in the uterine cervix. There is considerable evidence that Myoplasm hominis is associated with severe and

diverse diseases in the human body, but it is specifically found in the genital area of men and women with normal microbiota in the body, but this bacteria demonstrates its original pathogenicity in the immunocompromised state of the body. [1, 2]. *Mycoplasma hominis* is ordinarily found as part of the normal flora in the female genital tract, but several studies have shown that it may be involved in a variety of urogenital infections [2, 3]. *Myoplasm hominis* commonly causes cervicitis, a clinical syndrome described by the presence of purulent and mucopurulent discharge from the endocervical canal, in addition to other clinical signs such as inflammation, cervical bleeding, or post-coital vaginal bleeding. This opportunistic microbe can cause a severe and chronic infection in the cervix. There are several data points that prove that *Myoplasm hominis* is recombined there genetic material with the host genetic system and allows the other bacteria in-vivo to alter their antigenic membrane structure, this event allows for more successful transfer of *Myoplasm hominis* from one person to another and increases their resistance to antibiotics [3, 4]. There is also evidence linking *Myoplasm hominis* predominantly with cervical cancer or some inflammatory disorder. Several genes are associated with cervicitis, such as Kras, hras, tlr4, and tlr9, which are highly activated and associated with this disease. If the *myoplasm of humans* causes a mutation in this gene can suppress the transcriptional activity of oncogene resulting reduce the apoptosis of damage cell. [1, 5] Myoplasm have a triple-layered membrane and lack cell wall and are highly pleomorphic with no fixed shape and size. They lack even cell wall precursor like muramic acid or diaminopimelic acid. *Myoplasm hominis is* also genetically resistant to beta-lactam group antibiotics and macrolides, as well as a mutation in 23S RNA that shows more resistance against ciprofloxacin and ofloxacin. Hence, ordinarily used antibiotics are ineffective for the treatment of cervicitis. That's why cervicitis and *myoplasm in humans* should be promptly investigated and treated. [6, 7]. These bacteria cause many health problems, such as pelvic inflammatory disease, bacterial vaginosis, post-partum fever, and infertility in females [7, 8]. Moreover, *Myoplasma hominis* is proficient in causing diseases of the central nervous system in newborn babies and is connected with prostate cancer, cervicitis, and cervical cancer. Because cervicitis and myoplasm in humans should be promptly investigated and treated. [9] The progression of next-generation sequencing (NGS) techniques has empowered researchers to study and understand the world of microorganisms from a broader and deeper perspective. The advancement of DNA sequencing technology has not only aided in the fine characterization of bacterial genomes, but has also allowed for a greater level of taxonomic identification for complex microbiomes. [1, 6, 7] 16S rDNA sequencing is a basic metagenomic sequencing strategy used in the taxonomic identification and characterization of species and also provides comprehensive information on the entire repertoire of genes, the structure and organisation of the genome, the microbial community structure, and the evolutionary relationships present in the sample. Metagenomic sequencing has cleared the path for drug design against genetically potential microorganisms that are not cultivable in vitro but have the potential to infect humans. [7] Advances in structural and functional metagenomics have paved the way for the discovery of novel genes and metabolic pathways for disease-specific drugs. [6] Natural therapies, such as the use of plant-derived products in the treatment of cervicitis, may reduce the negative side

effects [10]. Plant phytochemicals were considered a potential medicine against cervicitis in ancient days [10]. According to the study, there are numerous phytocompounds in plants that may work against the inhibition of microbe growth in disease [1].

Gene that Associated with Cervicitis

Several genes and chromosomal regions have been found to be associated with cervicitis, such as TLR-4, TLR-9, Kras, and Hras, which are highly activated and associated with this disease [8, 11]. Because of *Myoplasma hominis,* when it interacts with TLR-4 and TLR-9, it leads to an innate and adaptive response in the body that causes the autoimmune disorder, such as activation of LY96 and CD14. This is mediated of innate immune response to bacterial lipopolysaccharide [12], which is also activated by the signalling pathway regulator NMI, which acts as damage-associated molecular patterns (DAMPs) in response to cell injury or pathogen invasion, therefore promoting nuclear factor NF-kappa-B activation. They also induce cytokine secretion and the inflammatory response in the body. Mutations in the Hras and Kras proteins are frequently observed in chronic cervicitis, which eventually leads to cervical cancer. Kras and Hras proteins, which bind GDP/GTP and possess intrinsic GTPase activity, play an important role in the regulation of cell proliferation and promoting oncogenic events by inducing the transcriptional silencing of tumour suppressor genes [13, 14].

2   Materials and Methods

Galaxy, a tutorial by Hiltemann S. and Batut B., 2020, analyses metagenomic data and provides comprehensive knowledge on the entire collection of genes, genome structure and systematic arrangement, and evolutionary relationships found in Myoplasma hominis. [15–17] *Myoplasm Hominies FASTA Sequences* 1.SRR14208169, 2.SRR14208170, 3.SRR14208171. FASTAQz, reverse and forward sequences, were retrieved from the SRA database. We paired the sequences first, then used quality parameters to summarise the sequences based on their name, group, and align report, and removed undesired sequences using the tools Unique.seqs, Count.seqs, Screen.seqs, Align.seqs, Screen.seqs, Filter.seqs, and Pre.cluster.

For further analysis, we used classify.seq's tool for removing errors and creating sequences, and we used the output and information of classify.seq to identify the abundances of the different found taxa. In Mothur's MiSeq SOP**,** the cluster.split command and assign are used to identify the OTUs (operational taxonomy units). [15] The following design-shared tool is used to take a list and rebound files for each group. It becomes a phylogeny after the classify.otu command assigns sequences to the selected taxonomy outline. This taxonomy was visualised using a Krona pie chart.

Next, we use the Metaphalen2 tool, which uses 1 million unique clade-specific marker genes identified from 17,000 reference (bacterial, archeal, viral, and eukaryotic) genomes and extracts the taxonomical information. [16] After the HUMAnN2 tool is used to determine functional information and provide gene family abundance, coverage, and abundance of pathways as output,

For drug discovery, we obtain the 3D structure of associated genes in cervicitis such as Kras, Hras, TLR4, and TLR9 that are highly associated with cervicitis and cause autoimmune symptoms in disease; this 3D is obtained with the help of Swiss Modle. Then our desired and interested phytocompound analyses and structures are downloaded from MOL-INSPIRATION and PUBCHEM, respectively. Patchdock is used to dock additional phytocompounds with our disease receptors [15–17].

3   Result and Discussion

Mycoplasm hominis fasta sequences seqs 1.SRR14208169, 2.SRR14208170, 3.SRR14208171.FASTAQz data summary as per summary.seq is given in Table 1.

Sequence Alignment of our data was done with an alignment of the V4 variable region of the 16s r RNA against the silva reference database (Table 2).

With the help of summary sequence output we summarize and understand the quality of alignment sequence given in the Table 3.

The output of Classify.seqs gave the classification of taxonomic data is given in Table 4.

This taxonomical information was visualize in krona, venn and pinch, phylogenetic tree

All the diagram are given below (Figs. 1, 2 and 3).

The taxonomy of the microbiome follows the taxonomic classification. We use shotgun metagenomic sequencing to understand and comprehensively sample all genes in our microbiome, as well as the abundance of a Microbe's functional information, and to evaluate bacterial diversity in our microbiome.

For further investigation of our target sequence SRR14208169, we used the MetaPhlAN2 tool, which produces a tabular file containing the community structure.

Table 1. Summary report of sequences.

	Start	End	NBases	Ambigs	Polymer	NumSeqs
Minimum:	1	35	35	0	1	1
2.5%-tile:	1	116	116	0	4	54181
25%-tile:	1	261	261	0	6	541804
Median:	1	325	325	0	6	1083608
75%-tile:	1	392	392	0	7	1625411
97.5%-tile:	1	480	480	52	8	2113034
Maximum:	1	502	502	103	250	2167214
Mean:		321.223	321.223	3.59132	6.30624	
# of Seqs:	2167214					

Table 2. Alignment Sequence Output

Query Name	Query Length	Template Name	Template Length	Search Method	Alignment Method
1_SRR14208169	317	AY171918.1	292	kmer	needleman
2_SRR14208169	342	AB038367.1	293	kmer	needleman
4_SRR14208169	344	AB038367.1	293	kmer	needleman
5_SRR14208169	253	AF418964.1	293	kmer	needleman
6_SRR14208169	308	AB038367.1	293	kmer	needleman
7_SRR14208169	301	AB038367.1	293	kmer	needleman
8_SRR14208169	265	AF276077.1	292	kmer	needleman
9_SRR14208169	420	AB015262.1	293	kmer	needleman
10_SRR14208169	219	AY328553.1	293	kmer	needleman
11_SRR14208169	299	EF495229.1	293	kmer	needleman
12_SRR14208169	358	AY907777.1	293	kmer	needleman
13_SRR14208169	286	AB038367.1	293	kmer	needleman
14_SRR14208169	280	AF289152.1	293	kmer	needleman
15_SRR14208169	294	AF353226.1	293	kmer	needleman
16_SRR14208169	286	AB015262.1	293	kmer	needleman
17_SRR14208169	279	AB038367.1	293	kmer	needleman
18_SRR14208169	367	AB038367.1	293	kmer	needleman
19_SRR14208169	150	AF445690.1	295	kmer	needleman
20_SRR14208169	413	U75254.1	293	kmer	needleman
21_SRR14208169	273	AY491599.1	277	kmer	needleman
22_SRR14208169	284	AF419658.1	297	kmer	needleman
23_SRR14208169	387	AF507714.1	293	kmer	needleman
24_SRR14208169	281	AB038367.1	293	kmer	needleman
25_SRR14208169	435	AB057592.1	293	kmer	needleman
26_SRR14208169	358	EF554364.1	293	kmer	needleman

Table 3. Summary Of Alignment Sequence Output

	Start	End	NBases	Ambigs	Polymer	NumSeqs	
Minimum:	0	0	0	0	1	1	
2.5%-tile:	0	0	0	0	1	47224	
25%-tile:	11546	13425	2	0	1	472237	
Median:	13400	13425	4	0	2	944474	
75%-tile:	13422	13425	8	0	2	1416710	
97.5%-tile:	13425	13425	23	0	4	1841723	
Maximum:	13425	13425	295	0	13	1888946	
Mean:	10115	10462.6	6.15034	0	1.75828		
# of unique seqs:		1778658					
total # of seqs:		1888946					

Table 4. Taxonomical Information Of *Mycoplasm Hominis* Out Put Of classify.seq.

Taxlevel	rankID	taxon	Daughter levels	Daughter levels
Taxonomy	total	SRR14208169	SRR14208170	SRR14208171
Root	27693	12370	10863	4460
Bacteria;Bacteria_unclassified; Bacteria_unclassified;Bacteria_unclassified; Bacteria_unclassified;Bacteria_unclassified;	23149	10207	9198	3744
Bacteria;Proteobacteria;Gammaproteobacteria; Pseudomonadales;Pseudomonadaceae;Pseudomonas;	735	323	294	118
Bacteria;tenericutes;mollicutes;mollicutes_Mollicutes_unclassified; Mollicutes_unclassified	2	1	1	0
nknown;unknown_unclassified; unknown_unclassified; unknown_unclassified; unknown_unclassified; unknown_unclassified;	3807	1839	1370	598

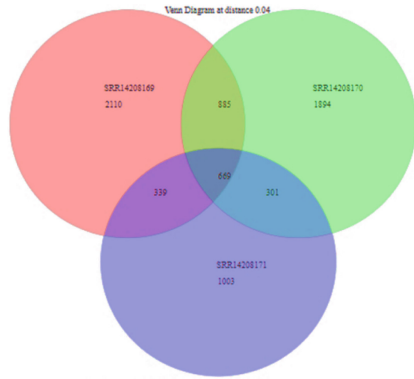**Fig. 1.** Krona Pie Chart Visualization Of The Taxonomy.

Fig. 2. Venn Diagram To Compare The Richness Shared Among SRR14208169, SRR14208170, SRR14208171 Groups (conclude that total richness of all the groups is 7201, and total shared richness between groups is 669,)

This file is fed into the HUMAnN2 tool. Table 5 shows the pathways and abundance file (output of the HUMAnN2 tool) and normalised gene family abundance table of the SRR14208169 sequence (first few lines of the output), and Tables 6 and 7, respectively, show the SRR14208170 or SRR14208171 sequence.

Structure Based Drug Designing of CERVICITIS Disease
Since, cervicitis is infection caused by mycoplasm hominis, we further go ahead towards

Taxonomy Bar Chart	Attributes Column Chart
Bubble Chart	Phylogram
Sankey Diagram	Donut Partition

Fig. 3. Pinch Visualization of Taxonomy Data and Phylogram.

Table 5(a). Pathways and their abundance file of SRR14208169.

# Pathway	humann2
# Pathway	humann2
UNMAPPED	270073.3193132755
UNINTEGRATED	930134.1723980459
UNINTEGRATED|g__Mycoplasma.s__Mycoplasma_hominis	930134.1723980462
PWY0–1296: purine ribonucleosides degradation	3051.1953748126
PWY0–1296: purine ribonucleosides degradation|g__Mycoplasma.s__Mycoplasma_hominis	3051.1953748126

designing novel drug for the disease. The gene receptors are highly associate with the cervicitis are taken from NCBI for our work (Table 8).

Homology Modelling: Homology modeling of the above receptors are done using SWISS-MODEL server. The receptor model and corresponding ramachandran plot results are given in figure. Template used for modeling is given in Table 9.

Table 5(b). Normalized the gene family Abundance stable of SRR14208169seq.

# Gene Family	humann2
# Gene Family	humann2
UNMAPPED	541761.0000000000
UniRef50_D1J7X6	56294.4769993214
UniRef50_D1J7X6|g__Mycoplasma.s__Mycoplasma_hominis	56294.4769993214
UniRef50_unknown	26514.6601152719
UniRef50_unknown|g__Mycoplasma.s__Mycoplasma_hominis	26514.6601152719
UniRef50_D1J8B7	19299.9936262039
UniRef50_D1J8B7|g__Mycoplasma.s__Mycoplasma_hominis	19299.9936262039
UniRef50_Q98Q97: 30S ribosomal protein S16	18925.3319259415
UniRef50_Q98Q97: 30S ribosomal protein S16|g__Mycoplasma.s__Mycoplasma_hominis	18925.3319259415
UniRef50_UPI00037C7492: hypothetical protein	17791.0823671760
UniRef50_UPI00037C7492: hypothetical protein|g__Mycoplasma.s__Mycoplasma_hominis	17791.0823671760
UniRef50_D1J8I8	16881.9213694268
UniRef50_D1J8I8|g__Mycoplasma.s__Mycoplasma_hominis	16881.9213694268
UniRef50_C1A930: 50S ribosomal protein L27	15645.9050516226
UniRef50_C1A930: 50S ribosomal protein L27|g__Mycoplasma.s__Mycoplasma_hominis	15645.9050516226
UniRef50_UPI00037C90F7: hypothetical protein	15331.2751326166
UniRef50_UPI00037C90F7: hypothetical protein|g__Mycoplasma.s__Mycoplasma_hominis	15331.2751326166
UniRef50_D1J7K5	14693.2488536514
UniRef50_D1J7K5|g__Mycoplasma.s__Mycoplasma_hominis	14693.2488536514
UniRef50_Q6KHJ9: UPF0154 protein MMOB4450	10249.1674230581
UniRef50_Q6KHJ9: UPF0154 protein MMOB4450|g__Mycoplasma.s__Mycoplasma_hominis	10249.1674230581
UniRef50_E8UJN8: Transposase	9751.2802906564
UniRef50_E8UJN8: Transposase|g__Mycoplasma.s__Mycoplasma_hominis	9751.2802906564
UniRef50_D1J8H3	9644.3868439346
UniRef50_D1J8H3|g__Mycoplasma.s__Mycoplasma_hominis	9644.3868439346
UniRef50_D1J8Q4	9247.2224370083
UniRef50_D1J8Q4|g__Mycoplasma.s__Mycoplasma_hominis	9247.2224370083
UniRef50_B3PLY2: Virulence-associated protein D	9208.9861205918
UniRef50_B3PLY2: Virulence-associated protein D|g__Mycoplasma.s__Mycoplasma_hominis	9208.9861205918
UniRef50_Q6KHC4: 30S ribosomal protein S20	8996.1058356335
UniRef50_Q6KHC4: 30S ribosomal protein S20|g__Mycoplasma.s__Mycoplasma_hominis	8996.1058356335

Table 6(a). Pathways and their abundance file of SRR14208170.

# Pathway	humann2
# Pathway	humann2
UNMAPPED	244014.4342506090
UNINTEGRATED	841377.1399033436
UNINTEGRATED|g__Mycoplasma.s__Mycoplasma_hominis	841377.1399033435
PWY0–1296: purine ribonucleosides degradation	2792.1886322864
PWY0–1296: purine ribonucleosides degradation|g__Mycoplasma.s__Mycoplasma_hominis	2792.1886322864

Table 7(a). Pathways and their abundance file of SRR14208171.

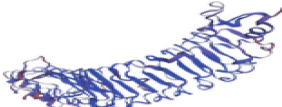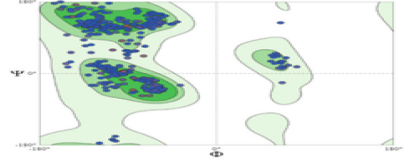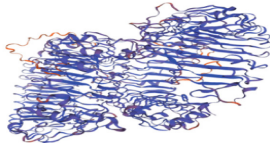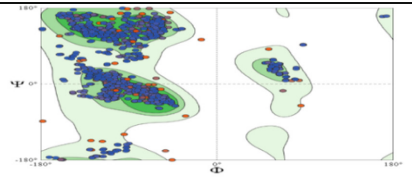| # Pathway	humann2
# Pathway	humann2
UNMAPPED	102445.9820137181
UNINTEGRATED	345776.1893098740
UNINTEGRATED|g__Mycoplasma.s__Mycoplasma_hominis	345776.1893098741
PWY0–1296: purine ribonucleosides degradation	1066.4041115522
PWY0–1296: purine ribonucleosides degradation|g__Mycoplasma.s__Mycoplasma_hominis	1066.4041115522
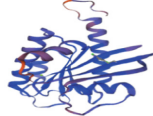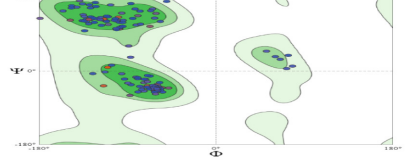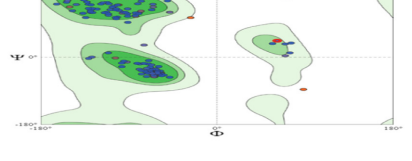
PROTEIN STRUCTURE	RAMACHANDRAN PLOT
TLR-4 (4g8a.1.A)	
TLR-9(3wpc.1.A)	
Hras (4dst.1.A)	
Kras (4q21.1.A)	

Table 7(b). Normalized The Gene Family Abundance Stable Of SRR14208170.

# Gene Family	humann2
# Gene Family	humann2
UNMAPPED	508307.0000000000
UniRef50_D1J7X6	61181.3514969257
UniRef50_D1J7X6|g__Mycoplasma.s__Mycoplasma_hominis	61181.3514969257
UniRef50_D1J8H3	29648.0440313384
UniRef50_D1J8H3|g__Mycoplasma.s__Mycoplasma_hominis	29648.0440313384
UniRef50_unknown	25595.7096517333
UniRef50_unknown|g__Mycoplasma.s__Mycoplasma_hominis	25595.7096517333
UniRef50_Q98Q97: 30S ribosomal protein S16	16441.6225636761
UniRef50_Q98Q97: 30S ribosomal protein S16|g__Mycoplasma.s__Mycoplasma_hominis	16441.6225636761
UniRef50_D1J8I8	15472.1573442926
UniRef50_D1J8I8|g__Mycoplasma.s__Mycoplasma_hominis	15472.1573442926
UniRef50_D1J7K5	14041.2536101078
UniRef50_D1J7K5|g__Mycoplasma.s__Mycoplasma_hominis	14041.2536101078
UniRef50_C1A930: 50S ribosomal protein L27	13877.5094551445
UniRef50_C1A930: 50S ribosomal protein L27|g__Mycoplasma.s__Mycoplasma_hominis	13877.5094551445
UniRef50_Q6KHC4: 30S ribosomal protein S20	11332.3539012506
UniRef50_Q6KHC4: 30S ribosomal protein S20|g__Mycoplasma.s__Mycoplasma_hominis	11332.3539012506
UniRef50_D1J8Q4	9574.2404385798
UniRef50_D1J8Q4|g__Mycoplasma.s__Mycoplasma_hominis	9574.2404385798
UniRef50_Q4A5D0: 30S ribosomal protein S17	9434.5287947750
UniRef50_Q4A5D0: 30S ribosomal protein S17|g__Mycoplasma.s__Mycoplasma_hominis	9434.5287947750
UniRef50_E8UJN8: Transposase	9226.3226129723
UniRef50_E8UJN8: Transposase|g__Mycoplasma.s__Mycoplasma_hominis	9226.3226129723
UniRef50_R5G3S2: 30S ribosomal protein S15	7787.1693245193
UniRef50_R5G3S2: 30S ribosomal protein S15|g__Mycoplasma.s__Mycoplasma_hominis	7787.1693245193
UniRef50_B3PMP3: 30S ribosomal protein S19	7648.4977505322
UniRef50_B3PMP3: 30S ribosomal protein S19|g__Mycoplasma.s__Mycoplasma_hominis	7648.4977505322
UniRef50_B3PLY2: Virulence-associated protein D	7000.3700755278
UniRef50_B3PLY2: Virulence-associated protein D|g__Mycoplasma.s__Mycoplasma_hominis	7000.3700755278

(*continued*)

Table 7(b). (*continued*)

# Gene Family	humann2
UniRef50_D1J7W5	6954.4087156260
UniRef50_D1J7W5|g__Mycoplasma.s__Mycoplasma_hominis	6954.4087156260
UniRef50_D1J7F7	6692.5234335007
UniRef50_D1J7F7|g__Mycoplasma.s__Mycoplasma_hominis	6692.5234335007
UniRef50_D1J8K3: Thioredoxin	6591.7836865483
UniRef50_D1J8K3: Thioredoxin|g__Mycoplasma.s__Mycoplasma_hominis	6591.7836865483
UniRef50_UPI00037C7492: hypothetical protein	6466.4017731263
UniRef50_UPI00037C7492: hypothetical protein|g__Mycoplasma.s__Mycoplasma_hominis	6466.4017731263
UniRef50_B9ZXE6: Type I restriction enzyme, truncation	6390.4126452531
UniRef50_B9ZXE6: Type I restriction enzyme, truncation|g__Mycoplasma.s__Mycoplasma_hominis	6390.4126452531
UniRef50_D1J7H8: Holo-[acyl-carrier-protein] synthase	6346.4488060746
UniRef50_D1J7H8: Holo-[acyl-carrier-protein] synthase|g__Mycoplasma.s__Mycoplasma_hominis	6346.4488060746
UniRef50_Q88WN5: 50S ribosomal protein L21	6334.4295421563
UniRef50_Q88WN5: 50S ribosomal protein L21|g__Mycoplasma.s__Mycoplasma_hominis	6334.4295421563
UniRef50_D1J8B7	6298.7380851998
UniRef50_D1J8B7|g__Mycoplasma.s__Mycoplasma_hominis	6298.7380851998
UniRef50_D1J8W3	6114.6552996736
UniRef50_D1J8W3|g__Mycoplasma.s__Mycoplasma_hominis	6114.6552996736
UniRef50_P47707: Probable cytosol aminopeptidase	6095.6648768615
UniRef50_P47707: Probable cytosol aminopeptidase|g__Mycoplasma.s__Mycoplasma_hominis	6095.6648768615
UniRef50_D1J7Q1: 30S ribosomal protein S18	5996.7695617374
UniRef50_D1J7Q1: 30S ribosomal protein S18|g__Mycoplasma.s__Mycoplasma_hominis	5996.7695617374
UniRef50_D1J7S0	5973.8323574308
UniRef50_D1J7S0|g__Mycoplasma.s__Mycoplasma_hominis	5973.8323574308
UniRef50_B3PLV3: ATP synthase subunit c	5856.8882146441
UniRef50_B3PLV3: ATP synthase subunit c|g__Mycoplasma.s__Mycoplasma_hominis	5856.8882146441
UniRef50_Q6KH14: Ribonuclease P protein component	5831.3021872142
UniRef50_Q6KH14: Ribonuclease P protein component|g__Mycoplasma.s__Mycoplasma_hominis	5831.3021872142
UniRef50_D1J891	

Table 7(c). Normalized The Gene Family Abundance Table Of SRR14208171.

# Gene Family	humann2
# Gene Family	humann2
UNMAPPED	204494.0000000000
UniRef50_D1J7X6	15877.7752324638
UniRef50_D1J7X6|g__Mycoplasma.s__Mycoplasma_hominis	15877.7752324638
UniRef50_D1J8H3	11884.4578378154
UniRef50_D1J8H3|g__Mycoplasma.s__Mycoplasma_hominis	11884.4578378154
UniRef50_unknown	9791.0796478526
UniRef50_unknown|g__Mycoplasma.s__Mycoplasma_hominis	9791.0796478526
UniRef50_D1J7K5	7495.9368879197
UniRef50_D1J7K5|g__Mycoplasma.s__Mycoplasma_hominis	7495.9368879197
UniRef50_D1J8B7	6864.7525487024
UniRef50_D1J8B7|g__Mycoplasma.s__Mycoplasma_hominis	6864.7525487024
UniRef50_D1J8I8	5823.4663310604
UniRef50_D1J8I8|g__Mycoplasma.s__Mycoplasma_hominis	5823.4663310604
UniRef50_C1A930: 50S ribosomal protein L27	5822.2408401448
UniRef50_C1A930: 50S ribosomal protein L27|g__Mycoplasma.s__Mycoplasma_hominis	5822.2408401448
UniRef50_UPI00037C90F7: hypothetical protein	5368.9614111350
UniRef50_UPI00037C90F7: hypothetical protein|g__Mycoplasma.s__Mycoplasma_hominis	5368.9614111350
UniRef50_Q98Q97: 30S ribosomal protein S16	4922.1208551373
UniRef50_Q98Q97: 30S ribosomal protein S16|g__Mycoplasma.s__Mycoplasma_hominis	4922.1208551373
UniRef50_E8UJN8: Transposase	4862.5721897251
UniRef50_E8UJN8: Transposase|g__Mycoplasma.s__Mycoplasma_hominis	4862.5721897251
UniRef50_UPI00037C7492: hypothetical protein	4493.5251465781
UniRef50_UPI00037C7492: hypothetical protein|g__Mycoplasma.s__Mycoplasma_hominis	4493.5251465781
UniRef50_Q6KHC4: 30S ribosomal protein S20	3783.0686310953

Natural therapies, such as the use of plant-derived products in cervicitis treatment, may reduce the adverse effect. Plants such as Ocimum sanctum and Ginkgobiloba and their phytocompounds were considered potential medicines by inhibiting microbial growth and showing anti-inflammatory effects against the disease. According to Lipinski's rule of five [ADME adsorption, distribution, metabolism, and excretion], we check the drug likeness of the above phytocompound.

Table 8. Genes with their NCBI Accession number and Abbreviations of genes

Sr. No	GENES	NCBI ACCESSION NO.	TAMPLETE
01	TLR4	NP612567	4g8a.Pdb
02	TLR9	ACQ41824	ACQ41824
03	Kras	NP_001356715XP_00671932	4dst.1.A
04	Hras	CAG38816	4q2a.Pdb

TLR4	Toll like receptor 4
TLR9	Toll like receptor 9
Kras	Kirsten rat sarcoma viral oncogene
Hras	Harvey Rat sarcoma viral oncogene,
ADME	Adsorption,distribution and metabolism excretion

Table 9(a). ADME Study Of Ocimum Sanctum.

SL.no	Compound	Mi logp	TPSA	NATOMS	Mw	nON	noHNH	N violation	Nrotb	volume
01	Rosemerinic acid	1.63	144.52	26	360.32	8	5	0	7	303.54
02	Apigenin	2.46	90.89	20	270.24	5	3	0	1	224.05
03	Caffiec acid	0.94	77.75	13	180.16	4	3	0	2	154.50
04	Eugenol	2.10	29.46	12	164.20	2	1	0	3	162.14
05	Chrysoriol	2.28	100.13	22	300.27	6	3	0	2	249.59

Table 9(b). ADME Study Of Lowsonia Inermis

Sl.no.	Compound	Mi logp	TPSA	NATOMS	Mw	nON	noHNH	n violation	nrotb	Volume
01	luteolin	1.97	111.12	21	286.24	6	4	0	1	232.07
02	2-Butoxysuccinic acid	0.48	83.83	13	190.19	5	2	0	7	176.22
03	tricin	2.30	109.36	24	330.29	7	3	0	3	275.14
04	Kampferol	2.17	111.12	21	286.24	6	4	0	1	232.07
05	Quercetin	1.68	131.35	22	302.24	7	5	0	1	240.08
06	Isocutellarin	2.51	100.13	22	300.27	6	3	0	2	249.59

Table 9(c). ADME Study Of Ginkgobiloba.

SL.no	Compound	Mi logp	TPSA	NATOMS	Mw	nON	noHNH	N violation	Nrotb	volume
01	Ginkgolide A	-1.46	128.60	29	408.40	9	2	0	1	339.84
02	Ginkgolide B	-2.38	148.83	30	424.40	10	3	0	1	347.88
03	Isorhamnetin	1.99	120.36	23	316.26	7	4	0	2	257.61
04	Protoatehunic acid	0.88	77.75	11	154.12	4	3	0	1	127.08

Table 10(a). TLR-4 Docking With Phytocompound Of *Ocimum Santum.*

Sr.no	Protein (receptor)	Ligands	Docking score kcal/mol	Interacting amino acid	No. of interaction
1	TLR-4	Rosemerinic acid	-4378	LYS-162, THR-119, ARG-64, ASN-139, SER-117,ARG-34	7
2	TLR-4	Apienin	-3412	ASN-139, THR-119, TYR-92	3
3	TLR-4	Caffeic acid	-2712	ASP-336, GLN-310, PHE-333	3
4	TLR-4	Eugenol	-3366	THR-119	1
5	TLR-4	Chrysoriol	-3674	ASN-139, SER-117	2

Table 10(b). Tlr-4 Docking With Phytocompound Of *Lowsonia inermis.*

Sr.no	Protein (receptor)	Ligands	Docking scores kcal/mol	Interacting amino acid	No. of interaction
1	TLR-4	Luteoline	-3482	LEU-12	1
2	TLR-4	2-butoxy succinate	-3088	ASR-94, THR-119, ASN-139	3
3	TLR-4	Tricin	-3984	LYS-162, SER-117, ASN-139, ARG-34, ARG-64,	5
4	TLR-4	Kampeferol	-3614	GLN-310, LEU-311, TYR-351, SER-334	4
5	TLR-4	Quercetin	-3414	SER-334, ASN-139,	2
6	TLR-4	Isocutellarin	-3788	LYS-62, ASN-139, ARG-64, TYR-96	4

Table 10(c). Tlr-4 Docking With Phytocompound Of *Ginkgobiloba.*

Sr.no	Protein (receptor)	Ligands	Docking score kcal/mol	Interacting amino acid	No. of interaction
1	TLR-4	Ginkgolide A	-4084	TH-119, ASN-139,LY-162,TYR-92, ARG-64	5
2	TLR-4	Ginkgolide B	-3932	ASN-139, LYS-162, ARG-64, TYR-92, ASP-94	5
3	TLR-4	Isorhamnetin	-3712	LYS-141, SER-117, ARG-64	3
4	TLR-4	Protoatechunic acid	-2478	LYS-162, ARG-164, TYR-92, ASR-94, THR-119	5

Table 11(a). Tlr-9 Docking With Phytocompound Of *Ocimum Sanctum.*

Sr.no	Protein (receptor)	Ligands	Docking scores kcal/mol	Interacting amino acid	No. of interaction
1	TLR-9	Rosemerinic acid	-4888	PRO-263, ASN-262, ASN-229, GLU-266,MET-265	5
2	TLR-9	Apigenin	-4264	ASN-468,	1
3	TLR-9	Caffeic acid	--3374	APG-480,GLN-556,SER-508,ALA-510, ARG-425	5
4	TLR-9	Eugenol	-3568	GLN-558	1
5	TLR-9	Chrysoriol	-4652	GLU-463, ASN-468, THR-471	3

Molecular Docking

Further docking is performed with the receptors in Table 9 with the above phytocompounds. Docking scores, interacting amino acids along with number of interactions are noted in Table 10.

As per docking studies it is seen that phytocompounds rosemerinic acid, caffeic acid, kampferol, qurecetine, Isorhamnetin docks with good interactions with the gene receptors involved in cervicitis i.e., Tlr-4,Tlr-9,Hars,kras (Tables 11, 12, 13).

Table 11(b). TLR-9 Docking With Phytocompound Of *Lowsonia Inermis.*

Sr.no	Protein (receptor)	Ligands	Docking scores kcal/mol	Interacting amino acid	No. of interaction
1	TLR-4	Luteoline	-3482	LEU-12	1
2	TLR-4	2-butoxy succinate	-3088	ASR-94, THR-119, ASN-139	3
3	TLR-4	Tricin	-3984	LYS-162, SER-117, ASN-139, ARG-34, ARG-64,	5
4	TLR-4	Kampeferol	-3614	GLN-310, LEU-311, TYR-351, SER-334	4
5	TLR-4	Quercetin	-3414	SER-334, ASN-139,	2
6	TLR-4	Isocutellarin	-3788	LYS-62, ASN-139, ARG-64, TYR-96	4

Table 11(c). TLR-9 Docking With Phytocompound Of *Ginkgobiloba.*

Sr.no	Protein (receptor)	Ligands	Docking scores kcal/mol	Interacting amino acid	No.of interaction
1	TLR-9	Ginkgolide A	-4892	ASN-209, TYR-179,	2
2	TLR-9	Ginkgolide B	-4672	TYR-553, GLN-556, HIS-529, HIS-504,	4
3	TLR-9	Isorhamnetin	-4558	GLU-547, ASN-468, THR-471,	3
4	TLR-9	Protoatehunic	-2906	PEP-744, LEU-726	2

Table 12(a). Kras Docking With Phytocompound Of *Ocimumm Santum.*

Sr.no	Protein (receptor)	Ligands	Docking scores kcal/mol	Interacting amino acid	No. Of interaction
1	Kras	Rosmerinic acid	-4830	VAL-29, GLU-31, ALA-18, TYR-32	4
2	Kars	Apigenin	-4062	ASP-33, THR-35, SER-17, GLY-13, LYS-16	5
3	Kars	Caffeic acid	-2960	ASP-33, SER-17, GLY-13	3
4	Kras	Eugenol	-3564	THR-35, SER-17	2
5	Kras	Chrysoriol	-4442	SER-17, THR-35,TYR-32, LYS-117,GLY-15	5

Table 12(b). 2 Kras Docking With Phytocompound Of *Lowsonia Inermis.*

SR.no.	Protein(receptor)	Ligands	Docking score kcal/mol	Interacting amino acid	No. of interaction
1	Kras	Luteoline	-4188	GLY-13,VAL-14,LYS-16,THR-35,SER-17, ASP-33	6
2	Kras	2-butoxysuccinate	-3304	LYS-117,GLY-13	2
3	Kras	Kampferol	-3980	VAL-29,GLU-31,ASP-30,LYS-117	4
4	Kras	Tricin	-4460	GLY-13,LYS-117,THR-35	3
5	Kras	Quercetin	-3872	ASP-30, ALA-18, ASP-119	3
6	Kars	Isocutellarin	-4520	GLU-31,ASP-33,THR-35,SER-17, LYS-16,GLY-13,GLY-15	7

Table 12(c). Kars Docking With Phytocompound Of *Ginkgobiloba.*

Sr.no.	Protein (receptor)	Ligands	Docking score kcal/mol	Interacting amino acid	No. Interaction
1	Kras	Ginkgolide A	-4406	ASP-33, GLU-31,ASP-30,VAL-29, LYS-29, GLY-13,	5
2	Kras	Ginkgolide B	-4330	ASP-30,LYS-117,	2
3	Kras	Isorhamnetin	-4552	ASP-33,THR-35,LYS-16,GLY-13,SER-17	5
4	Kras	Protoatrhunic acid	-2526	THR-35,GLY-60,GLY-15,LYS-16	4

Table 13(a). H Ras Docked With The Phytochemical Of *Ocimum Sanctum.*

Hras	Protein(receptor)	Ligand	Docking scores kcal/mol	Interacting amino acid	No. of interaction
1	H ras	Rosmerinic acid	-4966	GLU-62, ALA-59, THR-58	3
2	Hras	apigenin	-3960	GLN-61, GLY-60, TYR- 64, THR -58, GLY – 10	5
3	Hras	Caffeic acid	-3012	TYR – 64,G LU- 62. GLY -10, THR-58, GLY-60, GLN-61	6
4	Hras	Eugenol	-3424	ARG-68, GLN-99	2
5	Hras	chrysoriol	-4258	GLN-99, GLN-61	2

Table 13(b). H Ras Docking With The Phytochemical Of *Lowsonia Inermis.*

Sr.no	Protein (receptore)	Lingands	Docking scores kcal/mol	Interacting amino acid	Interaction
1	Hras	Luteoline	-4046	GLN-61,GLY-10, THR-58,TYR-64	4
2	Hras	2-butoxysuccinate	-3484	GLY-60,GLU-62	2
3	Hras	Kampferol	-4194	GLN-99, TYR-64, GLU-62	3
4	Hras	Tricin	-4494	TYR-64, GLU-62, GLN-61,THR-58, GLN-99	5
5	Hras	Quercetin	-4238	GLU-62,TYP-64,GLU-61,THR-58,GLN-99	5
6	Hras	Isocutellarin	-4198	GLN-99,GLU-62, TYP-64, THR-58	4

Table 13(c). Hras Docking With Phytochemical Of *Ginkgobiloba.*

Sr.no	Protein (receptor)	Ligands	Docking score kcal/mol	Interacting amino acid	No. of interaction
1	Hras	Ginkgolide A	-4458	ARG-68, GLU-63, ALA-59, THR-58, TYR- 96, GLU-62	5
2	Hras	Ginkgolide B	-4474	GLU-62,THR-58, TYR-96	3
3	Hras	Isorhamnetin	-4520	THR-58, GLU-62	2
4	Hras	Protoatechunic acid	-2478	LYS – 162, ARG- 64, TYR-92, ASR-94, THR-19	5

4   Conclusion

From the metagenomics study, the phylogeny and taxa of the cervicitis metagenome were identified. In this study, the compounds rosemerinic acid and caffeic acid from the plant

Ocimum sanctum docked best with all selected cervicitis receptors. The compounds kampferol and qurecetine from the plant *Lowsonia inermis* dock best with all selected receptors of cervicitis in this work, and the compound isorhamnetin from the plant Ginkgobiloba docks best with all selected receptors of cervicitis in this work. These phytocompounds also satisfy the Lipinskian rule of five for drugs based on the ADME properties; hence, the compounds can be successfully used as ligands for cervicitis receptors. Further in-vitro receptor-ligand binding studies can be done to establish the efficiency of the above ligands as drugs in treating cervicitis.

References

1. Waites, K. B., Katz, B., & Schelonka, R. L. (2005). Mycoplasmas and ureaplasmas as neonatal pathogens. *Clinical microbiology reviews*, *18*(4), 757–789. https://doi.org/10.1128/CMR.18.4.757-789.2005

2. Goret, J., Béven, L., Faustin, B., Contin-Bordes, C., Le Roy, C., Claverol, S., Renaudin, H., Bébéar, C., & Pereyre, S. (2017). Interaction of Mycoplasma hominis PG21 with Human Dendritic Cells: Interleukin-23-Inducing Mycoplasmal Lipoproteins and Inflammasome Activation of the Cell. *Journal of bacteriology*, *199*(15), e00213-17. https://doi.org/10.1128/JB.00213-17

3. Ladefoged SA. Molecular dissection of Mycoplasma hominis. APMIS Suppl. 2000;97:1-45. PMID: 10721331.

4. Razin S. Mycoplasmas. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 37. Available from: https://www.ncbi.nlm.nih.gov/books/NBK7637/

5. Christofolini, D. M., Leuzzi, L., Mafra, F. A., Rodart, I., Kayaki, E. A., Bianco, B., & Barbosa, C. P. (2012). Prevalence of cases of *Mycoplasma hominis*, *Mycoplasma genitalium*, *Ureaplasma urealyticum* and *Chlamydia trachomatis* in women with no gynecologic complaints. *Reproductive medicine and biology*, *11*(4), 201–205. https://doi.org/10.1007/s12522-012-0132-y

6. Chernova, O. A., Medvedeva, E. S., Mouzykantov, A. A., Baranova, N. B., & Chernov, V. M. (2016). Mycoplasmas and Their Antibiotic Resistance: The Problems and Prospects in Controlling Infections. *Acta naturae*, *8*(2), 24–34.

7. Lee, J. Y., & Yang, J. S. (2020). Prevalence and Antimicrobial Susceptibility of Mycoplasma hominis and *Ureaplasma* Species in Nonpregnant Female Patients in South Korea Indicate an Increasing Trend of Pristinamycin-Resistant Isolates. *Antimicrobial agents and chemotherapy*, *64*(10), e01065-20. https://doi.org/10.1128/AAC.01065-20

8. Kairys N, Garg M. Bacterial Vaginosis. [Updated 2021 Jul 18]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459216/

9. Benedetti, F., Curreli, S., & Zella, D. (2020). Mycoplasmas-Host Interaction: Mechanisms of Inflammation and Association with Cellular Transformation. *Microorganisms*, *8*(9), 1351. https://doi.org/10.3390/microorganisms8091351

10. Nabimeybodi, R., Zareshahi, R., Tansaz, M., Vahid Dastjerdi, M., & Hajimehdipoor, H. (2019). Scientific Evaluation of Medicinal Plants Used for the Treatment of Cervicitis (Qorohe- Rahem) in Iranian Traditional Medicine. *Iranian journal of pharmaceutical research : IJPR*, *18*(4), 1884–1901. https://doi.org/10.22037/ijpr.2019.1100852

11. Chauhan A, Pandey N, Desai A, Raithatha N, Patel P, Choxi Y, Kapadia R, Khandelwal R, Jain N. Association of TLR4 and TLR9 gene polymorphisms and haplotypes with cervicitis

susceptibility. PLoS One. 2019 Jul 31;14(7):e0220330. doi: https://doi.org/10.1371/journal.pone.0220330. PMID: 31365550; PMCID: PMC6668796.

12. Hussain, R. Z., Cravens, P. C., Doelger, R., Dentel, B., Herndon, E., Loof, N., Tsai, P., Okuda, D. T., Racke, M. K., & Stüve, O. (2018). TLR3 agonism re-establishes CNS immune competence during α4-integrin deficiency. *Annals of clinical and translational neurology*, 5(12), 1543–1561. https://doi.org/10.1002/acn3.664

13. Sheikh, Adnan & Vimalachandran, Dale & Thompson, Christopher & Jenkins, Rosalind & Nedjadi, Taoufik & Shekouh, Ali & Campbell, Fiona & Dodson, Andrew & Prime, Wendy & Crnogorac-Jurcevic, Tatjana & Lemoine, Nicholas & Costello, Eithne. (2007). The expression of S100A8 in pancreatic cancer-associated monocytes is associated with the Smad4 status of pancreatic cancer cells. Proteomics. 7. 1929-40. https://doi.org/10.1002/pmic.200700072.

14. Zaravinos, Apostolos. (2017). Oncogenic RAS: From Its Activation to Its Direct Targeting. Critical reviews in oncogenesis. 22. https://doi.org/10.1615/CritRevOncog.2017024695.

15. Saskia Hiltemann, Bérénice Batut, 2020 Analyses of metagenomics data - The global picture (Galaxy Training Materials). https://training.galaxyproject.org/training-material/topics/metagenomics/tutorials/general-tutorial/tutorial.html Online; accessed Tue Jan 25 2022

16. Batut et al., 2018 Community-Driven Data Analysis Training for Biology Cell Systems https://doi.org/10.1016/j.cels.2018.05.012

17. Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., … Weber, C. F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology*, 75(23), 7537–7541. https://doi.org/10.1128/aem.01541-09

18. 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.

19. https://www.molinspiration.com, Slovensky Grob, Slovakia

20. 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.

21. Hegde PL & Harini A, 2014, A text book of Dravyaguna Vijnana, Chaukhambha publications