



Functional Based Characterization of Hypothetical Proteins from *Candida Auris* to Discover Novel Drug Targets

P.R., KireseeSaghana^{1*}, K. Veerakkumar², R. Priya³, R.Senthil⁴, S. Shanmugavani⁵, J. Kabilan⁵, P. Monika⁶.

^{1*,3,4,5}Department of Bioinformatics, School of Life Science, Vels Institute of Science, Technology & Advanced Studies (VISTAS) Pallavaram, Chennai, 600011

²Department of Bioinformatics, School of Life Science, Bharathidasan University, Trichy - 620024,

^{5,6} Department of Pharmacy practice, School of Pharmaceuticals Science, Vels Institute of Science, Technology & Advanced Studies (VISTAS) Pallavaram, Chennai, 600011.

*Corresponding Author: P.R.KireseeSaghana

(Received: 02 September 2023

Revised: 14 October

Accepted: 07 November)

KEYWORDS

Candida auris, Insilco, candidiasis, hypothetical Proteins (HPs)

ABSTRACT:

Candida auris is a Gram-negative bacterium that causes candidiasis in humans. *C.auris* infections are the most frequent among all communicable diseases. The *Candida auris* (CJI97) B11221_V1 strain of *C. auris* contains 5715 genes and three pseudogenes. Out of these genes, 5521 genes code for proteins while 37 for rRNA, 157 tRNA. The proteome of *C. auris* made of 4689 from 5521 proteins contains 161 Hypothetical proteins (HPs) that are subjected to functional characterization. This study aims to annotate the functions and the structures of these hypothetical protein sequences. After extensive analysis of all the HPs, we have successfully assigned functions to 95 HPs with high precision. The selected proteins were further analyzed on the basis of their physicochemical properties, functional and structural categorization using an integrated bioinformatics approach by means of amino acid sequence and structure based analysis. This study found that these hypothetical protein sequences belong to various classes of proteins which involves in enzymes activity, hydrolase activity, Amino-acid biosynthesis, binding proteins, virulence and other proteins. Hence, the present study may serve as a potential lead for developing inhibitors against drug-resistant *Candida auris*.

INTRODUCTION

Fungal infections are increasingly recognized as a worldwide threat to human health. About 1.7 billion people worldwide suffer from a fungal infection, most of which are superficial infections of the skin and mucosa. In fact, the WHO already warned in its 2007 World Health Report that infectious diseases were spreading and emerging faster than at any time in history [2]. *Candida* species are the predominant cause of nosocomial fungal infections and are the fourth leading cause of all hospital-acquired infections [2]. Annually, there are approximately 400,000 bloodstream infections caused by *Candida* species globally, with mortality rates exceeding 40% [1]. The most frequently encountered *Candida* species is *Candida albicans*; however, the incidence of non-albicans species, such as *Candida tropicalis*, *Candida parapsilosis*, and *Candida glabrata*, has increased over recent decades due to the long-term use and limited

options of antifungal drugs [4; 6]. *Candida auris* (*C.auris*) infection is an emerging global threat since its first identification in Japan in 2009 [7]. Like no other fungal pathogen, *C. auris* has quickly unfolded to become an acutely worrisome infective agent, reaching pandemic proportions and cementing its condition as a superbug just within a decade after its first isolation in 2009 [3; 2]. *C.auris* is the first and so far only fungal pathogen categorized as an urgent health threat by the Centers for Disease Control and Prevention (CDC) [5]. It lead the healthcare and scientific communities to consider *C.auris* as one of the most serious emerging pathogen that critical care physicians should be aware of the description of *C. auris* fungemia outbreaks worldwide illustrates the dramatically increased need to develop new strategies to prevent and control infections due to this multiresistant pathogen. The main problem is that the *C.auris* genome sequence contains many uncharacterized and hypothetical proteins, and it is



unclear whether these proteins are involved in species-specific characteristics that promote its aggressiveness as a pathogen [26].

The proteome of *C.auris*, which consists of 5521 proteins, contains 161 hypothetical proteins (HPs) that are subjected to functional characterization. HPs are the protein repositories of an organism in the database that have functional annotation either by homology-based inference or by experimental data [27; 28].

MATERIAL AND METHOD

Sequence retrieval

The genome of the *Cand_auris_B11221_V1* was retrieved from NCBI's Genome database (<http://www.ncbi.nlm.nih.gov/genome/>). Out of the pool of 5521 proteins, 161 were listed as HPs. The NCBI IDs of all the HPs were retrieved from the NCBI Database (<http://www.ncbi.nlm.nih.gov/>) using their respective protein product ID's (e.g. PIS50921.1).

Physicochemical characterization:

We used Expassy's ProtParam (<https://web.expasy.org/protparam/>) [11] for the physicochemical characterization of the HPs. Estimation of physicochemical parameters gives insight into the biochemical nature of the proteins that serve as a basis for function prediction. Following this strategy, we predicted relative molecular weight, theoretical PI, extinction coefficient, instability index, aliphatic index and grand average of the hydropathicity (GRAVY) of the HPs.

Sub-cellular localization:

The function of a protein is closely related to its sub-cellular localization. Therefore, it is helpful to know the sub-cellular location of a protein in the functional characterization of HPs. Moreover, it also helps in inferring whether an HP is a probable drug target or vaccine target [12; 13; 14]. The cytoplasmic proteins are often drug targets, while surface membrane proteins are putative vaccine targets [15]. We used PSL Pred and CELLO for the prediction of sub-cellular localization of the HPs. PSL pred is used for predicting sub-cellular localization of only Gram-negative bacteria. Therefore, making it more reliable for the purpose. We used SignalP 4.1 [16] for the prediction of the signal peptide and SecretomeP [17] to identify protein's involvement in the non-classical secretory pathway. We used HMMTOP [18] for predicting the number of transmembrane helices in the HPs. This approach helps in the identification of membrane proteins among HPs.

Sequence comparisons for identification of homology:

For sequence similarity search to infer homology between HPs and the proteins of known functions, we

used CDD. It is a well-known tool widely used for finding sequence similarity by doing local alignments. For the identification of signature motifs in the HPs, we used SMART [19] and InterProscan [20]. Proteins showing sequence identities more than 30% and e-value -1 to +1 are referred to as close homologues of the HPs.

Function prediction by domain identification:

Functional domains of the HPs were predicted using various available protein family and domain databases such as Pfam (<http://pfam.xfam.org/>), NCBI CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and Motif, (<https://www.genome.jp/tools/motif/>). Pfam is a consortium of functional families of proteins with well-defined domain organization. Motif is an online resource, which can have known protein domains in the query sequences. It uses similarity search technique on a reference domain sequence database. CDD is the database of conserved domains from known proteins that include manually curated domain models based on the tertiary structures of the proteins.

Virulence factors analysis:

Virulence factors (VFs) are described as potent targets for drug discovery, which help in the invasion and dissemination of the bacteria in the host body [21]. We have used virulentpred [22] to predict VFs amongst the HPs. Virulentpred (<http://bioinfo.icgeb.res.in/virulent/>) is a support vector machine-based method used for bacterial VFs prediction from protein sequences.

RESULTS AND DISCUSSION

Over the past two decades there is generation of large amount of genome sequencing data that is conceptually translated in to protein sequences. This large amount of protein sequence data archived in databases like UniProt both in reviewed or unreviewed form. With the beginning of genomics or proteomics era there is simultaneous increase in development of bioinformatics prediction algorithms or tools. These bioinformatics prediction tools are being utilized successfully to annotate or characterize the large amount of gene or protein sequence data [23] here in this work we have retrieved 5521 HPs of *Candida auris* was done using, SMART, Inter Proscan and motif and CDD servers. We have successfully assigned functions 161 from 4689 HPs with high precision. The study suggests that the functionally assigned HPs may play a vital role in the growth and pathogenesis of this organism. Functional classes of the newly annotated HPs are further described in the following segments.

Physicochemical properties

Physicochemical properties like molecular weight, isoelectric point, aliphatic index and GRAVY score were calculated by ProtParam (Table 1). The aliphatic



index representing the volume occupied by the aliphatic residues (Ala, Val, Leu, and Ile) which consequently indicate the increased thermal stability for globular proteins thereby a greater aliphatic index value represents higher thermally stable protein. Greater aliphatic indices and greater hydrophobicity indices show a positive inclination towards the thermo stability of the proteins. The GRAVY value representing the protein water interaction, the proteins having low gravity score interact better with water and higher GRAVY score indicate higher hydrophobicity

[24]. On the basis of physiochemical parameter we have filtered twelve thermo stable proteins with accession number PIS49097.1, PIS57821.1, PIS54301.1, PIS52727.1, PIS56039.1, PIS53384.1, PIS56305.1, PIS49182.1, PIS56312.1, PIS56493.1, PIS53654.1 and PIS48116.1, the data are listed in Table 1. All twelve HPs have molecular weight less than 30kD, theoretical pI range from 4.2-10.2, aliphatic index value >90 and low GRAVY scores which is comparable with other thermostable proteins [25].

Table 1: The data are listed as following characteristic molecular weight, theoretical PI, extinction coefficient, instability index, aliphatic index and grand average of the hydropathicity (GRAVY) of the HPs.

S.no	ID	Molecular Weight	Theoretical pI	Extinction Coefficient	Instability Index	Aliphatic Index	GRAVY
1.	PIS53738.1	96197.1	5.01	155410	41.72	81.32	-0.52
2.	PIS53809.1	86785.73	8.89	82195	37.66	67.8	-0.933
3.	PIS53852.1	100620.54	5.99	43625	52.59	86.26	-0.46
4.	PIS53863.1	77603.02	5.86	110865	30.12	71.75	-0.361
5.	PIS48955.1	60446.71	8.52	125165	40.56	106.76	0.298
6.	PIS49007.1	121861.95	4.6	57020	41.59	100.77	-0.03
7.	PIS49043.1	101635.46	5.85	40965	30	92.91	-0.127
8.	PIS54201.1	71632.47	9.13	40020	50.13	77.43	-0.578
9.	PIS49083.1	37459.9	5.93	33475	26.5	93.76	-0.117
10.	PIS49097.1	27274.9	6	38055	35.03	86.03	-0.556
11.	PIS49137.1	147041	5.31	121725	46.01	90.58	-0.281
12.	PIS54211.1	33681.12	9.52	58330	49.44	60.82	-1.041
13.	PIS57814.1	45261.91	9.44	52260	46.82	63.52	-0.403
14.	PIS57821.1	23625	7.88	37930	30.62	79.75	-0.534
15.	PIS52624.1	145507.77	6.72	103555	44.34	103.78	-0.12
16.	PIS54263.1	35195.71	4.34	63370	40.01	110.06	0.488
17.	PIS57836.1	56509.62	5.62	93865	32.44	77.67	-0.508
18.	PIS57864.1	63480.21	9.17	141680	39.42	120.31	0.588
19.	PIS52684.1	64381.85	7.95	108790	32.83	97.71	0.363
20.	PIS52688.1	108940.39	5.18	125055	38.84	89.54	-0.11
21.	PIS54301.1	20486.75	8.74	8940	42.08	55.65	-1.596
22.	PIS54327.1	68642.81	6.42	39810	47.36	79.62	-0.358
23.	PIS52727.1	29529.56	9.46	26275	59.7	64.94	-0.696
24.	PIS52750.1	78368.35	8.48	92290	26.68	85.21	-0.244
25.	PIS52798.1	72090.55	9.09	121295	38.7	104.49	0.371
26.	PIS52815.1	95642	5.73	84340	54.97	66.87	-0.644
27.	PIS54423.1	48004.51	6.21	60070	34.48	76.98	-0.497
28.	PIS52834.1	177095.52	8.24	192785	41.17	82.01	-0.447
29.	PIS49264.1	110593.26	5.32	91970	38.29	88.16	-0.252
30.	PIS49271.1	63560.76	7.54	164250	40.68	97.01	0.385
31.	PIS49301.1	142734.68	9.09	190415	42	79.19	-0.573
32.	PIS55854.1	61958.08	9.12	98460	38.67	76.97	-0.364
33.	PIS53001.1	60461.87	4.49	34840	54.5	67.51	-1.043
34.	PIS49417.1	81688.46	7.21	125220	37.76	95.41	0.231
35.	PIS56994.1	70098.43	6.5	92390	37.02	74.9	-0.396
36.	PIS53032.1	50560.06	6.12	71280	34.4	85.46	-0.248
37.	PIS53062.1	59137.82	8.93	75540	37.78	85.13	-0.245
38.	PIS49420.1	162701.7	6.41	125835	42.34	80.89	-0.522
39.	PIS49430.1	35603.58	9.75	29575	36.64	99.3	0.069
40.	PIS55967.1	32906.33	9.78	74370	31.92	78.08	-0.011
41.	PIS53116.1	37263.8	6.09	27640	41	95.21	-0.259
42.	PIS53131.1	34742.93	6.72	13450	39.8	52.28	-0.383



43.	PIS56039.1	20196.82	4.82	3230	52.65	76.78	-0.248
44.	PIS53217.1	135904.59	5.83	142305	40.54	96.71	0.03
45.	PIS57222.1	51078.88	5.76	24785	40.62	87.62	-0.266
46.	PIS56051.1	52039.58	6.26	38445	41.33	66.34	-0.663
47.	PIS53260.1	154855.44	6.07	211370	36.71	96.56	-0.232
48.	PIS57258.1	126211.78	7.02	104765	52.88	67.97	-0.66
49.	PIS53362.1	105424.09	5.61	87515	42.97	94	-0.422
50.	PIS53384.1	29407.61	10.02	18910	31.59	66.53	-1.31
51.	PIS53451.1	112970.83	8.05	92765	49.95	87.32	-0.432
52.	PIS50045.1	137502.94	5.77	141430	41.55	78.18	-0.665
53.	PIS57324.1	36654.44	5.67	59025	56.59	49.53	-1.023
54.	PIS57354.1	45711.11	9.97	38070	42.74	78.49	-0.18
55.	PIS50838.1	15883.1	9.6	9970	34.38	135.97	1.076
56.	PIS56305.1	29495.39	6.31	18450	61.72	67.47	-1.169
57.	PIS53601.1	47050.35	9.33	51130	37.76	86.24	-0.251
58.	PIS50196.1	81937.65	7.28	62355	45.43	73.53	-0.739
59.	PIS50227.1	56794.26	4.92	89840	30.67	75.49	-0.421
60.	PIS50229.1	79095.14	5.59	127700	30.94	72.99	-0.365
61.	PIS50926.1	37868.56	10.36	61420	45.79	96.94	0.089
62.	PIS53610.1	58539.84	5.91	33475	37.07	96.79	-0.121
63.	PIS50668.1	117391.99	4.86	106410	42.48	94.09	-0.156
64.	PIS50698.1	59055.59	9.06	63955	43.98	84.89	-0.293
65.	PIS50964.1	62465.9	5.67	77935	44.77	88.2	-0.384
66.	PIS51002.1	49806.23	8.85	83100	25.57	81.24	-0.426
67.	PIS53667.1	102374.78	5.53	139495	39.02	89.01	-0.384
68.	PIS55112.1	75297.74	4.54	151415	38.62	65.21	-0.596
69.	PIS56400.1	74993.01	5.88	101480	42.11	72.25	-0.403
70.	PIS56435.1	112302.4	6.17	129145	34.61	98.58	-0.018
71.	PIS55323.1	47751.4	7.61	56060	27.29	91.11	-0.295
72.	PIS55332.1	65592.38	6.57	47245	44.49	84.42	-0.398
73.	PIS55422.1	48713.69	5.67	44390	36.95	86.81	-0.305
74.	PIS50791.1	79795.47	5.19	82655	42.79	91.82	-0.346
75.	PIS53765.1	36396.7	6.31	44015	25.99	85.34	-0.372
76.	PIS53779.1	69175.56	6.05	38405	52.38	69.18	-0.582
77.	PIS52508.1	53919.56	6.29	63955	27.33	89.3	-0.055
78.	PIS57411.1	63178.63	8.65	103165	35.35	108.05	0.578
79.	PIS57450.1	178319.78	6.02	172105	49.62	91.75	-0.365
80.	PIS54110.1	62179.28	9.12	113470	38.2	99.96	0.39
81.	PIS57483.1	88249.35	6.45	105365	34.63	86.81	-0.417
82.	PIS57589.1	95649.32	5.22	58720	39.39	100.16	-0.3
83.	PIS57621.1	81418.77	5.64	87040	43.54	89.93	-0.319
84.	PIS57650.1	93399.8	5.6	126310	46.38	75.63	-0.629
85.	PIS54172.1	150297.64	4.93	182940	43.1	76.67	-0.559
86.	PIS54188.1	46594.94	9.49	51465	41.94	88.46	-0.291
87.	PIS57690.1	134684.29	9.64	196780	44.88	81.54	-0.57
88.	PIS49107.1	113581.61	6.77	85065	65.4	52.07	-0.983
89.	PIS49126.1	63734.35	9.14	103750	43.52	102.12	0.194
90.	PIS49138.1	93471.92	5.37	141680	43.95	97.58	0.06
91.	PIS57724.1	251082.65	5.19	200235	56.84	52.65	-1.056
92.	PIS57740.1	51233.56	6.71	62715	45.4	83.83	-0.296
93.	PIS52565.1	73837.18	5.25	35535	61.15	56.47	-1.193
94.	PIS49182.1	20634.26	4.21	500	70.94	64.61	0.062
95.	PIS57790.1	86905.4	8.09	88950	55.99	67.67	-0.854
96.	PIS52683.1	62929.76	6.05	111060	33.51	100.63	0.42
97.	PIS52698.1	54289.73	5.62	54695	48.02	93.4	-0.327
98.	PIS54294.1	67623.22	5.95	81415	57.39	73.29	-0.657
99.	PIS52810.1	52776.96	5.4	27055	49.78	84.83	-0.67
100.	PIS49242.1	89229.83	6.24	76820	60.59	72.68	-0.653
101.	PIS52854.1	35788.51	6.55	23295	26.67	103.49	0.158



102.	PIS52870.1	73396.24	8.72	102680	34.27	82.7	-0.316
103.	PIS57984.1	203846.5	8.73	334085	41.51	93.26	-0.048
104.	PIS57989.1	79391.25	6.24	90690	38.66	85.33	-0.4
105.	PIS58001.1	15977.85	4.29	8480	53.3	83.96	-0.397
106.	PIS49377.1	68321.57	6.34	91135	43.69	80.02	-0.473
107.	PIS49403.1	47687.01	8.84	69915	50.43	88.96	-0.302
108.	PIS57088.1	32090.63	9.46	31650	50.89	76.88	-0.561
109.	PIS57097.1	86709.24	5.75	71210	48.63	77.06	-0.491
110.	PIS49907.1	85219.94	8.55	148460	43.66	83.16	-0.432
111.	PIS57131.1	73341.96	5.76	82320	50.36	101.3	-0.264
112.	PIS57138.1	63551.07	4.66	30955	50.08	87.85	-0.668
113.	PIS57147.1	45365.5	9.47	41620	51.04	80.9	-0.518
114.	PIS53117.1	59824.12	4.8	86595	33.28	81.68	-0.083
115.	PIS53157.1	77441.97	8.6	92750	27.69	91.78	-0.321
116.	PIS53160.1	147614.07	6.45	105170	50.51	92.19	-0.234
117.	PIS53168.1	80044.3	6.28	105615	38.23	70.92	-0.503
118.	PIS49938.1	99564.72	7.14	146945	47.99	83.14	-0.367
119.	PIS57190.1	63632.63	5.48	72895	30.8	91.25	-0.19
120.	PIS56031.1	51387.56	8.79	18910	62.23	50.43	-1.022
121.	PIS49980.1	84489.79	9.01	53080	49.19	63.43	-0.666
122.	PIS53287.1	131010.67	5.86	114765	49.99	84.61	-0.48
123.	PIS57261.1	73078.9	6.88	51590	46.9	85.49	-0.41
124.	PIS53367.1	52571.71	5.62	71530	33.8	79.01	-0.514
125.	PIS53388.1	73203.38	6.98	67645	57.73	85.21	-0.325
126.	PIS53549.1	74198.31	5.54	36120	39.35	84.73	-0.367
127.	PIS56300.1	52828.58	9.5	59500	54.19	68.04	-0.925
128.	PIS56312.1	24682.41	5.11	27515	37.02	82.05	-0.54
129.	PIS50930.1	39261.6	5.43	35200	41.31	87.71	-0.267
130.	PIS53622.1	84966.82	5.31	83615	44.66	81.64	-0.389
131.	PIS53635.1	37078.73	6.17	21430	53.3	80.68	-0.606
132.	PIS50669.1	249042.22	5.7	243300	41.28	93.92	-0.227
133.	PIS50701.1	94353.4	5.92	44350	43.71	71.87	-0.76
134.	PIS50725.1	66417.14	6.09	94950	42.67	99	0.214
135.	PIS50729.1	120189.28	5.85	109615	35.9	87.38	-0.341
136.	PIS56470.1	59693.86	8.43	65125	40.71	82.74	-0.194
137.	PIS56493.1	26759.53	4.25	25440	74.63	70.46	-0.561
138.	PIS56525.1	63551.58	8.21	85425	28.33	110.19	0.571
139.	PIS55322.1	60775.82	5.41	53540	25.83	85.05	-0.201
140.	PIS55421.1	53010.59	8.74	47870	38.83	97.95	0.003
141.	PIS50105.1	46017.2	5.16	47135	35.16	90.73	-0.402
142.	PIS56199.1	36528.97	5.64	60640	42.8	68.53	-0.586
143.	PIS56219.1	59491.3	8.6	36330	64.46	50.97	-0.892
144.	PIS50123.1	74980.45	5.74	68925	44.28	81.5	-0.505
145.	PIS53555.1	72423.6	4.66	59415	42.68	91.72	-0.29
146.	PIS53606.1	63511.76	9.1	109570	36.74	107.16	0.409
147.	PIS50870.1	64838.33	6.56	81860	49.46	86.76	-0.201
148.	PIS50921.1	145153.67	5.89	127270	38.21	90.92	-0.383
149.	PIS53644.1	57571.75	6.11	81485	39.63	90.6	-0.18
150.	PIS50282.1	58602.1	9.07	72560	46.24	81.59	-0.27
151.	PIS50654.1	38751.35	7.74	64875	35.14	86.01	-0.488
152.	PIS51014.1	64427.73	7.48	99530	32.54	98.54	0.363
153.	PIS53654.1	22646.62	5.98	20525	49.35	79.35	-0.403
154.	PIS53693.1	52308.47	4.87	21930	72.64	49.32	-0.876
155.	PIS53714.1	160197.73	5.93	155715	34.14	96.37	-0.038
156.	PIS56387.1	175171.78	8.47	200720	40.2	104.76	0.049
157.	PIS55113.1	75087.65	4.67	148435	43.31	65.13	-0.535
158.	PIS55130.1	64380.74	6.13	111560	30.18	108.05	0.516
159.	PIS55208.1	72438.35	9.6	49655	44.79	71.73	-0.665
160.	PIS56491.1	132882.21	6.2	60530	39.13	88.23	-0.439



161.	PIS48116.1	28016.29	8.09	29255	33.6	85.33	-0.342
------	------------	----------	------	-------	------	-------	--------

SUB-CELLULAR LOCALIZATION

Finding a proteins subcellular localization is important, especially when identifying a target. Additionally, location prediction can provide information regarding the function of a query protein as well as its type, such as cytoplasmic or membrane protein. It is also crucial to identify the presence of trans-membrane helices and signal peptides because a protein's role in extracellular contacts can be further supported by a positive prediction of these two. Table 2, shows the analysis done to 161 selected proteins of this study revealed that 49 proteins to be located at cytoplasmic, 37 proteins at

Outer Membrane, and 7 proteins at Periplasmic, 7 proteins at Extracellular. 21 proteins at Inner Membrane, 9 proteins to be located at Periplasmic/Inner Membrane/Cytoplasmic, 17 proteins to be located at Extracellular/Outer Membrane/Periplasmic/Cytoplasmic, 11 proteins to be located at Outer Membrane/Periplasmic/ Inner Membrane/ Cytoplasmic and 3 proteins to be located at Inner Membrane/Cytoplasmic. The sub-cellular localization predictions of all the 167 HPs are listed in Table 2.

Table 2: The subcellular localization predictions of all the 161 HPs listed

S.no	ID	PSL Pred	CELLO	Secretome P	HMMTOP
1.	PIS53738.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
2.	PIS53809.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
3.	PIS53852.1	Extracellular Protein	Outer Membrane	NO	Nil
4.	PIS53863.1	Extracellular Protein	Extracellular	NO	Nil
5.	PIS48955.1	Inner-membrane Protein	Inner Membrane	NO	13TMH
6.	PIS49007.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
7.	PIS49043.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
8.	PIS54201.1	Extracellular Protein	Outer Membrane	NO	1TMH
9.	PIS49083.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
10.	PIS49097.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
11.	PIS49137.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
12.	PIS54211.1	Periplasmic Protein	Cytoplasmic	NO	Nil
13.	PIS57814.1	Outer Membrane Protein	Outer Membrane	NO	3TMH
14.	PIS57821.1	Cytoplasmic Protein	Periplasmic /Cytoplasmic	NO	Nil
15.	PIS52624.1	Cytoplasmic Protein	Outer Membrane	NO	2TMH
16.	PIS54263.1	Inner-membrane Protein	Inner Membrane	NO	7TMH
17.	PIS57836.1	Periplasmic Protein	Outer Membrane	NO	Nil
18.	PIS57864.1	Inner-membrane Protein	Inner Membrane	NO	11TMH
19.	PIS52684.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
20.	PIS52688.1	Cytoplasmic Protein	Inner Membrane/Cytoplasmic	NO	6TMH
21.	PIS54301.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
22.	PIS54327.1	Extracellular Protein	Outer Membrane	NO	1TMH
23.	PIS52727.1	Extracellular Protein	Extracellular/Outer Membrane/Periplasmic	NO	Nil
24.	PIS52750.1	Cytoplasmic Protein	Periplasmic/Cytoplasmic	NO	Nil
25.	PIS52798.1	Inner-membrane Protein	Inner Membrane	NO	9TMH
26.	PIS52815.1	Extracellular Protein	Extracellular	NO	1TMH
27.	PIS54423.1	Periplasmic Protein	Periplasmic	NO	1TMH
28.	PIS52834.1	Inner-membrane Protein	Outer Membrane	NO	Nil
29.	PIS49264.1	Periplasmic Protein	Cytoplasmic	NO	Nil
30.	PIS49271.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
31.	PIS49301.1	Periplasmic Protein	Outer Membrane	NO	Nil
32.	PIS55854.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
33.	PIS53001.1	Cytoplasmic Protein	Extracellular/Outer Membrane/Cytoplasmic	NO	Nil
34.	PIS49417.1	Inner-membrane Protein	Inner Membrane	NO	11TMH
35.	PIS56994.1	Periplasmic Protein	Periplasmic	NO	Nil
36.	PIS53032.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	Nil
37.	PIS53062.1	Cytoplasmic Protein	Inner Membrane/Cytoplasmic	NO	Nil
38.	PIS49420.1	Periplasmic Protein	Outer Membrane/Cytoplasmic	NO	Nil
39.	PIS49430.1	Cytoplasmic Protein	Periplasmic/Inner Membrane	NO	4TMH



40.	PIS55967.1	Inner-membrane Protein	Periplasmic/Inner Membrane	NO	6TMH
41.	PIS53116.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
42.	PIS53131.1	Extracellular Protein	Extracellular/Outer Membrane	NO	2TMH
43.	PIS56039.1	Cytoplasmic Protein	Periplasmic/Inner Membrane/ Cytoplasmic	NO	Nil
44.	PIS53217.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
45.	PIS57222.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
46.	PIS56051.1	Periplasmic Protein	Extracellular/Outer Membrane/Cytoplasmic	NO	Nil
47.	PIS53260.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	Nil
48.	PIS57258.1	Inner-membrane Protein	Extracellular/Outer Membrane/Periplasmic	NO	Nil
49.	PIS53362.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
50.	PIS53384.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
51.	PIS53451.1	Extracellular Protein	Outer Membrane	NO	Nil
52.	PIS50045.1	Extracellular Protein	Outer Membrane	NO	2TMH
53.	PIS57324.1	Cytoplasmic Protein	Extracellular/Outer Membrane/Cytoplasmic	NO	Nil
54.	PIS57354.1	Cytoplasmic Protein	Inner Membrane/Cytoplasmic	NO	6TMH
55.	PIS50838.1	Inner-membrane Protein	Inner Membrane	NO	5TMH
56.	PIS56305.1	Extracellular Protein	Extracellular/Periplasmic/ Cytoplasmic	NO	Nil
57.	PIS53601.1	Extracellular Protein	Outer Membrane	NO	Nil
58.	PIS50196.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
59.	PIS50227.1	Periplasmic Protein	Extracellular/Outer Membrane	NO	Nil
60.	PIS50229.1	Periplasmic Protein	Periplasmic	NO	Nil
61.	PIS50926.1	Inner-membrane Protein	Periplasmic/Inner Membrane	NO	Nil
62.	PIS53610.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
63.	PIS50668.1	Cytoplasmic Protein	Cytoplasmic	NO	2TMH
64.	PIS50698.1	Cytoplasmic Protein	Inner Membrane/Cytoplasmic	NO	Nil
65.	PIS50964.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
66.	PIS51002.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
67.	PIS53667.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
68.	PIS55112.1	Extracellular Protein	Extracellular	NO	1TMH
69.	PIS56400.1	Periplasmic Protein	Periplasmic	NO	Nil
70.	PIS56435.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	2TMH
71.	PIS55323.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	Nil
72.	PIS55332.1	Inner-membrane Protein	Cytoplasmic	NO	Nil
73.	PIS55422.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
74.	PIS50791.1	Cytoplasmic Protein	Outer Membrane	NO	2TMH
75.	PIS53765.1	Inner-membrane Protein	Cytoplasmic	NO	Nil
76.	PIS53779.1	Extracellular Protein	Extracellular/Outer Membrane	NO	Nil
77.	PIS52508.1	Cytoplasmic Protein	Periplasmic/Cytoplasmic	NO	2TMH
78.	PIS57411.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
79.	PIS57450.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	Nil
80.	PIS54110.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
81.	PIS57483.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
82.	PIS57589.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
83.	PIS57621.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
84.	PIS57650.1	Periplasmic Protein	Outer Membrane	NO	Nil
85.	PIS54172.1	Periplasmic Protein	Cytoplasmic	NO	Nil
86.	PIS54188.1	Periplasmic Protein	Periplasmic	NO	Nil
87.	PIS57690.1	Cytoplasmic Protein	Outer Membrane	NO	Nil
88.	PIS49107.1	Extracellular Protein	Extracellular/Outer Membrane /Periplasmic	NO	1TMH
89.	PIS49126.1	Inner-membrane Protein	Inner Membrane	NO	10TMH
90.	PIS49138.1	Inner-membrane Protein	Inner Membrane	NO	9TMH
91.	PIS57724.1	Extracellular Protein	Extracellular/Periplasmic	NO	5TMH
92.	PIS57740.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil



93.	PIS52565.1	Cytoplasmic Protein	Extracellular/Outer Membrane/Cytoplasmic	NO	1TMH
94.	PIS49182.1	Periplasmic Protein	Extracellular	NO	1TMH
95.	PIS57790.1	Inner-membrane Protein	Extracellular/Outer Membrane/Periplasmic	NO	Nil
96.	PIS52683.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
97.	PIS52698.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
98.	PIS54294.1	Extracellular Protein	Outer Membrane	NO	Nil
99.	PIS52810.1	Extracellular Protein	Outer Membrane	NO	Nil
100.	PIS49242.1	Cytoplasmic Protein	Outer Membrane	NO	1TMH
101.	PIS52854.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
102.	PIS52870.1	Inner-membrane Protein	Periplasmic	NO	Nil
103.	PIS57984.1	Inner-membrane Protein	Inner Membrane	NO	16TMH
104.	PIS57989.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
105.	PIS58001.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
106.	PIS49377.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
107.	PIS49403.1	Cytoplasmic Protein	Periplasmic/Inner Membrane/Cytoplasmic	NO	2TMH
108.	PIS57088.1	Extracellular Protein	Outer Membrane	NO	Nil
109.	PIS57097.1	Extracellular Protein	Outer Membrane	NO	Nil
110.	PIS49907.1	Cytoplasmic Protein	Outer Membrane	NO	2TMH
111.	PIS57131.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
112.	PIS57138.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
113.	PIS57147.1	Inner-membrane Protein	Cytoplasmic	NO	1TMH
114.	PIS53117.1	Cytoplasmic Protein	Outer Membrane	NO	1TMH
115.	PIS53157.1	Inner-membrane Protein	Cytoplasmic	NO	Nil
116.	PIS53160.1	Extracellular Protein	Outer Membrane	NO	Nil
117.	PIS53168.1	Periplasmic Protein	Periplasmic	NO	Nil
118.	PIS49938.1	Extracellular Protein	Outer Membrane	NO	10TMH
119.	PIS57190.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
120.	PIS56031.1	Extracellular Protein	Extracellular	NO	Nil
121.	PIS49980.1	Periplasmic Protein	Outer Membrane	NO	2TMH
122.	PIS53287.1	Inner-membrane Protein	Outer Membrane	NO	1TMH
123.	PIS57261.1	Inner-membrane Protein	Outer Membrane	NO	Nil
124.	PIS53367.1	Periplasmic Protein	Periplasmic/Cytoplasmic	NO	Nil
125.	PIS53388.1	Periplasmic Protein	Outer Membrane/Periplasmic	NO	Nil
126.	PIS53549.1	Extracellular Protein	Outer Membrane	NO	Nil
127.	PIS56300.1	Periplasmic Protein	Outer Membrane	NO	Nil
128.	PIS56312.1	Extracellular Protein	Inner Membrane	NO	Nil
129.	PIS50930.1	Cytoplasmic Protein	Outer Membrane	NO	1TMH
130.	PIS53622.1	Extracellular Protein	Outer Membrane	NO	1TMH
131.	PIS53635.1	Extracellular Protein	Outer Membrane	NO	1TMH
132.	PIS50669.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	Nil
133.	PIS50701.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
134.	PIS50725.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
135.	PIS50729.1	Cytoplasmic Protein	Outer Membrane	NO	Nil
136.	PIS56470.1	Cytoplasmic Protein	Cytoplasmic	NO	2TMH
137.	PIS56493.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
138.	PIS56525.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
139.	PIS55322.1	Cytoplasmic Protein	Outer Membrane	NO	Nil
140.	PIS55421.1	Cytoplasmic Protein	Cytoplasmic	NO	2TMH
141.	PIS50105.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
142.	PIS56199.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
143.	PIS56219.1	Extracellular Protein	Extracellular	NO	1TMH
144.	PIS50123.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	1TMH
145.	PIS53555.1	Cytoplasmic Protein	Cytoplasmic	NO	1TMH
146.	PIS53606.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
147.	PIS50870.1	Cytoplasmic Protein	Periplasmic/Cytoplasmic	NO	2TMH
148.	PIS50921.1	Extracellular Protein	Outer Membrane	NO	Nil



149.	PIS53644.1	Cytoplasmic Protein	Outer Membrane/Cytoplasmic	NO	4TMH
150.	PIS50282.1	Cytoplasmic Protein	Outer Membrane/ Inner Membrane	NO	1TMH
151.	PIS50654.1	Inner-membrane Protein	Cytoplasmic	NO	Nil
152.	PIS51014.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
153.	PIS53654.1	Extracellular Protein	Outer Membrane	NO	Nil
154.	PIS53693.1	Extracellular Protein	Extracellular	NO	Nil
155.	PIS53714.1	Inner-membrane Protein	Inner Membrane	NO	10TMH
156.	PIS56387.1	Inner-membrane Protein	Outer Membrane/Inner Membrane	NO	14THM
157.	PIS55113.1	Outer Membrane Protein	Extracellular/ Outer Membrane	NO	1TMH
158.	PIS55130.1	Inner-membrane Protein	Inner Membrane	NO	12TMH
159.	PIS55208.1	Extracellular Protein	Outer Membrane	NO	Nil
160.	PIS56491.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil
161.	PIS48116.1	Cytoplasmic Protein	Cytoplasmic	NO	Nil

Functional annotation of the hypothetical protein

The initial protein domain was achieved from the CDD of NCBI. The region of the domain, superfamily, and family classifications have been determined by the servers CDD, SMART and Interpro. The domain, superfamily, and family were selected based on the lowest e-value of the following domain. The higher e-value has been filtered out from the selection

procedure. The e-value 9.73E-68 of *C. auris* Protein from smart, 9.65E-25 of *C. auris* superfamily macrodomain from CDD, Casein Kinase 2, subunit of *C. auris* family from interproscan indicate extremely good protein alignment respectively. The results of SMART, CDD and Inter Proscan are summarized in Table3.

Table 3: The results of CDD, SMART and Inter Proscan are summarized.

S.no	SMART		CDD		INTERPROSCAN
	Function	e-value	Function	e-value	
1.	o		Glycosyl hydrolase family 63 C-terminal domain	0.00E+00	Glycoside_hydrolase_63
2.	DNA Topoisomerase I	0e+00	DNA Topoisomerase I (eukaryota)	0.00E+00	DNA topoisomerase I
3.	minichromosome maintenance proteins	0.00E+00	minichromosome maintenance proteins	0.00E+00	Mini-chromosome maintenance protein
4.	0		Glycosyl hydrolase family 81	0.00E+00	Endo-1,3(4)-beta-glucanase
5.	0		Peptidase family M49;	0.00E+00	TLC-dom
6.	0		40S ribosomal S2	1.04e-86	DASH complex subunit Dad4
7.	0		Transient receptor potential (TRP) ion	0.00E+00	0
8.	Protein kinase C conserved region 2 (CalB)	1.19E-24	Ca ²⁺ -dependent lipid-binding protein, contains C2 domain [General function prediction only]	0.00E+00	16S/18S rRNAaminocarboxypropyltransferase Tsr3
9.	0		TCP-1 (CTT or eukaryotic type II)	0.00E+00	Ribosomal protein S35, mitochondrial
10.	0		Dolichyl-phosphate-mannose-protein mannosyltransferase	1.03E-56	GPI mannosyl transferase
11.	Importin beta N-terminal domain;	1.00E-12	Importin-beta N-terminal domain	1.00E-12	Importin beta family
12.	0		Formate--tetrahydrofolate ligase	0.00E+00	FFormate-tetrahydrofolate ligase, FTHFS
13.	0		Major Facilitator Superfamily	0.00E+00	Proton-dependent oligopeptide transporter family



14.	macrodomain superfamily	1.04E-11	macrodomain superfamily	9.65E-25	0
15.	Kinesin motor, catalytic domain. ATPase	1.03E-103	Myosin and Kinesin motor domain	2.38E-124	Kinesin-like protein
16.	0		Major Facilitator Superfamily;	0.00E+00	Proton-dependent oligopeptide transporter family (
17.	0		P-loop containing Nucleoside Triphosphate	0.00E+00	RHD3/Sey1
18.	0		Fructose-1,6-bisphosphatase	5.96E-177	Fructose-1,6-bisphosphatase class 1
19.	Phosphoglycerate mutase family;	1.37E-33	Histidine phosphatase domain found in a functionally diverse set of proteins	1.78E-44	Histidine phosphatase superfamily, clade-1
20.	0		IKI3 family	0.00E+00	Elongator complex protein 1
21.	DnaJ molecular chaperone homology domain;	1.03E-06	chaperone protein DnaJ;	1.68E-09	0
22.	Src homology 3 domains	7.83E-13	Peroxin 13, N-terminal region;	1.03E-59	Peroxin 13
23.	0		Dihydrofolatereductase (DHFR)	6.78E-47	Dihydrofolatereductase
24.	0		Rapamycin-insensitive companion of mTOR, middle domain	1.04E-20	Pianissimo family
25.	Repeated motif present between transmembrane helices in cystinosin	2.10E-06	PQ loop repeat;	1.02E-13	PQ-loop_rpt
26.	C1 Peptidase family	4.63E-03	Peptidase C1B subfamily	0.00E+00	Peptidase C1B, bleomycin hydrolase
27.	0		ALG6, ALG8 glycosyltransferase family	0.00E+00	Glycosyltransferase, ALG6/ALG8
28.	0		Amino acid permease	0.00E+00	Amino acid permease, fungi (
29.	0		V-type ATPase 116kDa subunit family	0e+00	V-type ATPase, V0 complex, 116kDa subunit family (
30.	0		cwf21 domain found in fungal complexed with CEF1 protein 21	1.00E-14	cwf21 domain
31.	0		Nitrogen permease regulator 2;	0e+00	Nitrogen permease regulator 2
32.	0		Solute carrier families 5 and 6-like	1.04E-71	transmembrane domain
33.	zinc finger binding to DNA consensus sequence	1.04E-13	zinc finger binding to DNA consensus sequence	1.04e-13	Znf_GATA
34.	0		Acyl-CoA dehydrogenase;	0.00E+00	Acyl-CoA oxidase
35.	0		Endomembrane protein 70;	0.00E+00	EMP70
36.	GYF domain:	2.68E-10	GYF domain;	1.04e-15	GYF



37.	0		Exonuclease- Endonuclease- Phosphatase domain	2.09E-85	Endonuclease/exonuclease/phosphatase
38.	DNA polymerase type-B family	2.56E-71	DNA polymerase type-B zeta subfamily catalytic domain.	0.00E+00	DNA-directed DNA polymerase, family B
39.	0		PDZ-like domain;	1.61E-38	PDZ-like_dom
40.	0		Amino acid permease [	0.00E+00	Amino acid permeases
41.	DNA polymerase A domain;	1.95E-51	Family A polymerase primarily fills DNA gaps that arise during DNA repair,	0.00E+00	DNA-directed DNA-polymerase, family A, mitochondria
42.	0		cytochrome P450 (CYP) superfamily;	0.00E+00	Cytochrome P450
43.	0		Apoptosis antagonizing transcription factor;	1.04E-23	Protein AATF/Bfr2 (
44.	Two tandem repeats of the cystathionine beta-synthase (CBS pair) domains superfamily;	1.12E-04	CLC voltage-gated chloride channel.	0.00E+00	Chloride channel, voltage gated (
45.	0		The Ntn hydrolases (N- terminal nucleophile)	1.02E-91	Proteasome, subunit alpha/beta
46.	0		Choline/Carnitine o- acyltransferase;	0.00E+00	Acyltransferase Cho Actase/COT/CPT (
47.	0		Zinc peptidases M18, M20, M28, and M42;	0.00E+00	Peptidase M18
48.	0		cytochrome P450 (CYP) superfamily;	0.00E+00	Cytochrome P450 (
49.	DNA polymerase type-B family;	4.84E-83	DNA polymerase (pol2);	0.00E+00	DNA-directed DNA polymerase, family B
50.	0		Tricarboxylate carrier;	0.00E+00	Tricarboxylate/iron carrier (
51.	RNA recognition motif (RRM) superfamily;	5.12E-04	RNA12 protein	0.00E+00	Mitochondrial escape protein 2 (
52.	0		Mitochondrial carrier protein	1.05E-07	Mitochondrial carrier protein
53.	Alpha mannosidase middle domain;	1.59E-31	Alpha-mannosidase	0.00E+00	glycosyl hydrolases families 57
54.	P-loop containing Nucleoside Triphosphate Hydrolases;	8.67E-11	DNA polymerase III, gamma/tau subunits	1.69E-34	P-loop NTPase fold
55.	WSC domain;	5.78E-10	WSC domain	1.05E-10	WSC domain
56.	0		Adenylate forming domain, Class I superfamily;	0.00E+00	AMP-binding
57.	Mpr1p, Pad1p N- terminal (MPN) domains;	9.42E-05	Mpr1p, Pad1p N- terminal (MPN) domains;	1.05E-62	Eukaryotic translation initiation factor 3 subunit F
58.	Src Homology 3 domain superfamily;	2.97E-13	The Bin/Amphiphysin/Rvs (BAR) domain	1.17E-19	SH3_domain



59.	0		2Fe-2S iron-sulfur cluster binding domain.	2.75E-57	2Fe-2S_ferredoxin-type
60.	0		Haloacid Dehalogenase-like Hydrolases;	0.00E+00	P-type ATPase
61.	0		Glycosyltransferase family A (GT-A)	1.04E-88	Chitin synthase (
62.	0		Zinc peptidases M18, M20, M28, and M42;	0.00E+00	Peptidase M18
63.	DnaJ domain or J-domain.	1.02E-21	DnaJ domain or J-domain.	1.02E-21	DnaJ homolog subfamily A member 1/2-like
64.	0		RAVE protein 1 C terminal;	0.00E+00	WD40-repeat-containing domain superfamily
65.	Protein kinase C-related kinase	1.05E-10	Protein Kinases, catalytic domain;	0.00E+00	Prot_kinase_dom
66.	0		Eukaryotic porin family that forms channels in the mitochondrial outer membrane;	0.00E+00	Mitochondrial distribution and morphology protein 10
67.	0		Vacuolar protein sorting-associated protein 35;	0.00E+00	Vacuolar protein sorting-associated protein 35
68.	Inositol 1,4,5-trisphosphate/ryanodine receptor;	1.42E-06	Dolichyl-phosphate-mannose--protein O-mannosyltransferase	0.00E+00	Glycosyltransferase 39-like (
69.	Inositol 1,4,5-trisphosphate/ryanodine receptor;	2.35E-03	Dolichyl-phosphate-mannose--protein O-mannosyltransferase	0.00E+00	Glycosyltransferase 39-like (
70.	0		Utp11 protein;	1.01E-75	Small-subunit processome, Utp11
71.	0		Nup93/Nic96;	0.00E+00	Nucleoporin interacting component Nup93/Nic96
72.	Pleckstrin homology-like domain;	1.17E-08	Oxysterol-binding protein;	0.00E+00	Oxysterol-binding protein (
73.	Mpr1p, Pad1p N-terminal (MPN) domains;	7.13E-29	26S proteasome regulatory subunit;	9.07E-135	26S Proteasome non-ATPase regulatory subunit 7/8
74.	0		Mitochondrial peculiar membrane protein 1;	1.02E-49	Mitochondrial peculiar membrane protein 1 (
75.	0		Mitochondrial carrier protein; Mitochondrial carrier protein.	1.02E-23	Mitochondrial carrier protein
76.	0		phospholipid-translocating P-type ATPase, flippase;	0.00E+00	P-type ATPase
77.	0		NAD(P)+-dependent aldehyde dehydrogenase superfamily;	0.00E+00	Aldehyde_DH_dom
78.	0		Got1/Sft2-like family;	1.00E-26	Vesicle transport protein, G-1/GTPase 11
79.	basic Helix Loop Helix (bHLH) domain superfamily;	1.05E-09	basic Helix Loop Helix (bHLH) domain superfamily;	1.44E-34	Helix-loop-helix DNA-binding domain
80.	WD40 domain,	1.03E-03	WD40 domain	1.03E-03	WD40
81.	0		Utp25, U3 small nucleolar RNA-associated SSU processome protein 25;	0.00E+00	U3 small nucleolar RNA-associated protein 25



82.	Glycosyl hydrolases family 32;	1.3174E-129	Glycosyl hydrolases family 32;	1.3174E-129	Glycoside hydrolase, family 32
83.	0		Glycosyl hydrolase family 81;	0.00E+00	Endo-1,3(4)-beta-glucanase
84.	0		60Kd inner membrane protein;	1.04E-12	Membrane insertase YidC/ALB3/OXA1/COX18
85.	0		Amino Acid Kinases (AAK) superfamily, catalytic domain;	0.00E+00	Asp_kinase
86.	0		Aspartate aminotransferase (AAT) superfamily (fold type I)	3.43E-81	Cys/Met metabolism, pyridoxal phosphate-dependent enzyme
87.	0		Type 1 glutamine amidotransferase (GATase1)-like domain;	0.00E+00	Biotin-prot_ligase_N
88.	Importin-beta N-terminal domain;	1.08E-13	Importin, protein involved in nuclear import	0.00E+00	Importin-beta_N
89.	0		cytochrome P450 (CYP) superfamily;	0.00E+00	Cytochrome P450
90.	0		Choline/ethanolamine kinase;	1.05E-56	Choline_kin_N
91.	Sec14p-like lipid-binding domain	6.55E-32	CRAL/TRIO, N-terminal domain;	1.04E-05	CRAL/TRIO, N-terminal domain
92.	0		Gluzincin Peptidase family (thermolysin-like proteinases,	0.00E+00	Peptidase M3A/M3B
93.	0		glycosyltransferase family 1 and related proteins with GTB topology;	0.00E+00	Glycosyltransferase, family 35
94.	RNA polymerase I subunit A N-terminus;	4.387E-108	Largest subunit of RNA polymerase (RNAP), N-terminal domain;	0.00E+00	DNA-directed RNA polymerase, subunit beta-prime
95.	0		Beta-glucan synthesis-associated protein (SKN1);	0.00E+00	Beta-glucan synthesis-associated protein Skn1/Kre6/Sbg1
96.	0		Copper amine oxidase, enzyme domain;	0.00E+00	Copper amine oxidase (
97.	WD40 domain,	3.12E-05	WD40 domain	1.04E-06	WD40_repeat
98.	0		Aspartate aminotransferase (AAT) superfamily (fold type I)	0.00E+00	Aspartate/other aminotransferase
99.	Protein Kinases, catalytic domain;	6.63E-87	Protein Kinases, catalytic domain;	1.01E-141	Protein_kinase_ATP_BS
100.	0		Zinc peptidases M18, M20, M28, and M42	0.00E+00	Peptidase M20 (
101.	0		CDC45-like protein;	0.00E+00	CDC45 family
102.	calponin homology (CH) domain superfamily	1.07E-06	calponin homology (CH) domain superfamily	1.03E-36	Ras GAP_dom



103.	Protein Kinases, catalytic domain;	2.0965E-103	Protein Kinases, catalytic domain;	0.00E+00	Protein_kinase_ATP_BS
104.	0		Cyclin box fold superfamily;	1.01E-35	Cyclin PHO80-like (I
105.	0		NAD(P)+-dependent aldehyde dehydrogenase superfamily;	0.00E+00	Aldehyde dehydrogenase NAD(P)-dependent
106.	WD40 repeats	4.71E-05	Guanylate-binding protein, C-terminal domain	1.02E-03	WD repeat-containing protein WDR18/Ipi3/RID3
107.	0		Amino acid permease	0e+00	Amoino_acid_permease_CS
108.	Myosin. Large ATPases	0.00E+00	Myosin heavy chain	0.00E+00	Myosin_head_motor_dom
109.	0		Amino acid permease [Amino acid transport and metabolism];	1.02E-176	Amoino_acid_permease_CS
110.	0		Peptidase M3 mitochondrial intermediate peptidase (MIP);	0e+00	Peptidase M3A/M3B
111.	0		Exocyst complex component Sec10;	0.00E+00	Exocyst complex component Sec10-like
112.	0		ChAPs (Chs5p-Arf1p-binding proteins);	0e+00	Chs5p-Arf1p binding
113.	Histidine kinase-like ATPases;	1.00E-14	Histidine kinase-like ATPase domain;	2.85E-33	PDK/BCKDK protein kinase
114.	0		Trehalase;	0.00E+00	Glycoside hydrolase, family 37
115.	meprin and TRAF homology;	5.42E-07	Ubiquitin carboxyl-terminal hydrolase	0e+00	Endopeptidase, NLPC/P60 domain
116.	0		LETM1-like protein;	1.04E-07	LETM1/MDM38-like
117.	0		Sfi1 spindle body protein;	1.04e-19	Sfi1_dom
118.	Middle domain of eukaryotic initiation factor 4G	5.80E-32	Eukaryotic translation initiation factor 4G1	1.05E-15	Initiation factor 4G (
119.	0		Acetyl-coenzyme A transporter 1	0e+00	Acetyl-coenzyme A transporter 1-like
120.	0		V-type ATPase 116kDa subunit family	0.00E+00	V-type ATPase, V0 complex, 116kDa subunit family
121.	0		Sec23-binding domain of Sec16	1.02E-84	Sec16_CCD
122.	0		Indoleamine 2,3-dioxygenase	0.00E+00	Indoleamine 2,3-dioxygenase (
123.	0		Chromatin assembly factor 1 subunit A;	1.02e-16	Chromatin assembly factor 1 subunit A (I
124.	CFEM domain;	1.80E-05	CFEM domain	1.04E-08	Extracellular_membr_CFEM_dom
125.	Extension to Ser/Thr-type protein kinases;	1.22E-10	Protein Kinases, catalytic domain;	0e+00	Protein_kinase_ATP_BS
126.	0		Amino acid permease [Amino acid transport and metabolism];	0.00E+00	Amino acid permease, fungi
127.	0		IMP-specific 5'-nucleotidase	0.00E+00	IMP-specific 5-nucleotidase
128.	Protein Kinases, catalytic domain;	3.69E-34	Protein Kinases, catalytic domain	1.08E-61	Prot_kinase_dom



129.	0		Universal stress protein family	1.03E-20	Universal stress protein A family
130.	Protein Kinases, catalytic domain;	5.62E-29	Protein Kinases, catalytic domain;	2.51E-31	Prot_kinase_dom
131.	0		glycosyltransferase family 1	2.01E-162	Glyco_trans_4-like_N
132.	0		Adenylate forming domain, Class I superfamily	0.00E+00	Medium/long-chain-fatty-acid--CoA ligase FadD17/FadD6-like
133.	0		1,3-beta-glucan synthase component;	0.00E+00	Glycosyltransferase, family 48
134.	0		Glycogen synthase;	0.00E+00	Glycogen synthase
135.	0		Mitochondrial import receptor subunit Tom22	1.04E-48	Mitochondrial import receptor subunit Tom22
136.	0		Metal-independent alpha-mannosidase (GH125)	0.00E+00	Metal-independent alpha-mannosidase
137.	N-terminal domain of Fpg (formamidopyrimidine-DNA glycosylase, MutM)_Nei (endonuclease VIII) base-excision repair DNA glycosylases;	7.39E-15	Formamidopyrimidine-DNA glycosylase H2TH domain;	1.36E-14	FPG_cat
138.	0		alpha/beta hydrolases;	1.00E-28	Steryl acetyl hydrolase 1-like
139.	0		Nucleoporin subcomplex protein binding to Pom34;	0.00E+00	Nucleoporin Nup188
140.	0		ATP10 protein;	1.05E-98	ATPase assembly factor ATP10
141.	0		Peptidase family S64;	0.00E+00	Peptidase S64, Ssy5
142.	0		Patatins and Phospholipases;	0.00E+00	AcylTrfase/lyso P Lipase
143.	0		Leucine rich repeat;	1.03E-03	Leu-rich_rpt
144.	PWI domain;	2.39E-04	PWI domain	1.02E-04	PWI_dom
145.	Protein Kinases, catalytic domain;	1.00E-79	cyclin-dependent protein kinase;	1.01E-80	Prot_kinase_dom
146.	0		Aspartate aminotransferase (AAT) superfamily (fold type I)	0.00E+00	Cys/Met metabolism, pyridoxal phosphate-dependent enzyme
147.	0		Stretch-activated Ca2+-permeable channel component;	1.03E-173	Stretch-activated cation channel Mid1/Calcium influx-promoting protein Ehs1
148.	0		NAD(P)+-dependent aldehyde dehydrogenase superfamily;	0.00E+00	Succinate semialdehyde dehydrogenase
149.	0		Adenylate forming domain,	0.00E+00	AMP-dep_Synth/Lig
150.	0		Transport protein Trs120 or TRAPPC9, TRAPP II complex subunit;	0.00E+00	TRAPP II complex, Trs120



151.	0		primary-amine oxidase;	0.00E+00	Copper amine oxidase
152.	0		Autophagy protein Apg9;	0.00E+00	Autophagy-related protein 9
153.	STI1 domain;	3.15E-05	STI1 domain;	1.03E-16	Heat shock protein Sti1-like
154.	0		Zinc peptidases M18, M20, M28, and M42;	0.00E+00	Peptidase M20
155.	0		Amino acid permease	0.00E+00	AA-permease/SLC12A_dom
156.	basic Helix Loop Helix (bHLH) domain superfamily	2.39E-09	basic Helix Loop Helix (bHLH) domain superfamily;	3.43E-11	HLH_DNA-bd_sf
157.	SNARE motif;	2.71E-04	SNARE motif;	1.05E-17	GOSR2/Membrin/Bos1
158.	0		Arginase-like and histone-like hydrolases;	0.00E+00	Histone deacetylase family
159.	Protein Kinases, catalytic domain;	3.92E-92	Protein Kinases, catalytic domain;	0.00E+00	Prot_kinase_dom
160.	0		Major Facilitator Superfamily;	0.00E+00	Prot_kinase_dom
161.	Protein Kinases, catalytic domain;	1.01E-75	Stk1 family PASTA domain-containing	2.40E-21	Prot_kinase_dom
162.	GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain;	1.23E-09	GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain;	1.05E-09	Zn(2)-C6 fungal-typeDNA-binding domain
163.	Protein kinase C terminal domain; Protein kinase C terminal domain.	1.05E-15	Protein Kinases, catalytic domain;	0.00E+00	Prot_kinase_dom
164.	0		C1 Peptidase family (MEROPS database nomenclature)	0.00E+00	Peptidase C1B, bleomycin hydrolase
165.	0		fungal transcription factor regulatory middle homology region;	1.03E-05	Fungal transcription factor
166.	0		NAD(P)+-dependent aldehyde dehydrogenase superfamily;	0.00E+00	Succinate semialdehyde dehydrogenase
167.	0		DNA polymerase alpha subunit B N-terminal;	1.01E-41	DNA polymerase alpha, subunit B
168.	0		NAD(P)+-dependent aldehyde dehydrogenase superfamily;	0.00E+00	Aldehyde dehydrogenase NAD(P)-dependent
169.	Protein Kinases, catalytic domain;	3.20E-26	Protein Kinases, catalytic domain	0.00E+00	Casein kinase I
170.	Rab subfamily of small GTPases;	1.10E-98	Ras family	1.05E-85	Small GTPase
171.	0		Maintenance of mitochondrial	0.00E+00	Maintenance of mitochondrial morphology protein 1
172.	0		Choline/Carnitine o-acyltransferase;	0.00E+00	Acyltransferase ChoActase/COT/CPT
173.	SNARE motif;	7.35E-12	t-SNARE complex subunit,	1.02E-11	Syntaxin
174.	0		Proteasome-substrate-size regulator, mid region	0.00E+00	Proteasome activator complex subunit 4



175.	RNA recognition motif (RRM) superfamily	2.95E-07	RNA recognition motif (RRM) domain [Translation, ribosomal structure and biogenesis]	1.01E-07	RRM_dom
176.	0		Major Facilitator Superfamily;	0.00E+00	Proton-dependent oligopeptide transporter family
177.	Bacterial NAD-glutamate dehydrogenase	2.14E-17	Bacterial NAD-glutamate dehydrogenase;	0.00E+00	NAD-dependent glutamate dehydrogenase, eukaryotes
178.	0		cytochrome P450 (CYP) superfamily;	0.00E+00	Cytochrome P450
179.	Ubiquitin-conjugating enzyme E2, catalytic (UBCc) domain.	1.53E-48	Ubiquitin-conjugating enzyme E2, catalytic (UBCc) domain	4.45E-55	UBQ-conjugating_AS
180.	0		Amino acid permease;	0.00E+00	AA-permease/SLC12A_dom
181.	0		The alpha-D-phosphohexomutase superfamily	0.00E+00	Alpha-D-phosphohexomutase superfamily
182.	GAL4-like Zn2Cys6 binuclear cluster DNA-binding domain	5.73E-12	fungal transcription factor regulatory middle homology region;	1.04E-03	Fungal transcription factor
183.	Rossmann-fold NAD(P)(+)-binding protein	1.05E-04	cis-2,3-dihydrobiphenyl-2,3-diol dehydrogenase;	1.19E-10	Short-chain dehydrogenase/reductase SDR
184.	PCI domain;	1.97E-15	26S proteasome regulatory complex component,	1.09E-142	26S proteasome regulatory subunit Rpn7/COP9 signalosome complex subunit 1
185.	0		NAD(P)+-dependent aldehyde dehydrogenase superfamily;	0.00E+00	Aldehyde_DH_dom
186.	0		Coproporphyrinogen III oxidase	0.00E+00	Coproporphyrinogen III oxidase, aerobic
187.	0		TATA Binding Protein (TBP) Associated Factor 4 (TAF4)Transcription Factor IID (TFIID) complex;	1.02E-52	Transcription initiation factor TFIID component TAF4
188.	RNA polymerases M/15 Kd subunit;	4.23E-03	C-terminal zinc ribbon domain of RNA polymerase intrinsic transcript cleavage subunit;	1.05E-28	DNA-directed RNA polymerase subunit/transcription factor S
189.	Protein Kinases, catalytic domain	1.58E-57	Protein Kinases, catalytic domain;	1.30E-170	Prot_kinase_dom
190.	0		RNA polymerase I specific transcription initiation factor RRN3;	0.00E+00	RNA polymerase I specific transcription initiation factor RRN3
191.	0		Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily;	0.00E+00	ATPase_NBD
192.	0		Amino acid permease;	0.00E+00	Amino acid permease, fungi



193.	0		Gpi16 subunit, GPI transamidase component;	0.00E+00	GPI transamidase component PIG-T
194.	0		Ankyrin repeat [Signal transduction mechanisms];	1.00E-07	Glycerophosphodiesterphosphodiesterase Gde1
195.	0		Cyclases involved in the biosynthesis of lantibiotics, and similar proteins;	0.00E+00	Glycoside hydrolase family 47
196.	0		Yeast mitochondrial distribution and morphology (MDM) proteins	0.00E+00	Mitochondrial distribution and morphology protein family 31/32, fungi
197.	0		cytochrome P450 (CYP) superfamily;	0.00E+00	Cytochrome P450
198.	Protein Kinases, catalytic domain;	9.73E-68	Protein Kinases, catalytic domain;	0.00E+00	Casein Kinase 2, subunit alpha
199.	0		Amino acid permease;	0.00E+00	Amino acid permease, fungi
200.	0		Synaptobrevin/VAMP-like protein	1.04E-56	Synaptobrevin
201.	N-terminal helicase domain of the DEAD-box helicase superfamily	1.15E-18	Superfamily II DNA or RNA helicase, SNF2 family	1.52E-92	Helicase_ATP-bd
202.	The superfamily of RING finger (Really Interesting New Gene) domain and U-box domain	7.38E-07	RING-H2 zinc finger domain	1.05E-06	E3 ubiquitin-protein ligase RFWD3
203.	P-loop containing Nucleoside Triphosphate Hydrolases	2.18E-09	ABC transporter transmembrane region;	0.00E+00	Oligomycin resistance ATP-dependent permease Yor1
204.	P-loop containing Nucleoside Triphosphate Hydrolases	2.61E-04	ABC transporter transmembrane region	0.00E+00	ABC_transporter-like_ATP-bd
205.	0		Beta-glucan synthesis-associated protein (SKN1)	0.00E+00	Beta-glucan synthesis-associated protein Skn1/Kre6/Sbg1
206.	0		amino acid permease (yeast);	0.00E+00	Amino acid permease, fungi
207.	Phosphatidylinositol phosphate kinase (PIPK) catalytic domain family	1.3019E-121	Phosphatidylinositol phosphate kinase (PIPK) catalytic domain family	0.00E+00	Phosphatidylinositol-4-phosphate 5-kinase
208.	K homology (KH) RNA-binding domain, type I	1.07E-05	polynucleotide phosphorylase/polyadenylase;	1.04E-03	KH_dom_type_1_sf
209.	metallophosphatase superfamily, metallophosphatase domain	1.2712E-118	Calcineurin-like phosphoesterase	1.01E-26	Calcineurin-like_PHP_ApaH



Function prediction by domain identification

HP PIS53062.1 is identified as Cytochrome p450s comprise a superfamily of heme-thiolate proteins named for the spectral absorbance peak of their carbon-monoxide-bound species at 450 nm. Having been found in every class of organism, including Archaea, the p450 superfamily is believed to have originated from an ancestral gene that existed over 3 billion years ago. Repeated gene duplications have subsequently given rise to one of the largest of multigene families. These enzymes are notable both for the diversity of reactions that they catalyze and the range of chemically dissimilar substrates upon which they act.

Cytochrome p450s support the oxidative, peroxidative and reductive metabolism of such endogenous and xenobiotic substrates as environmental pollutants, agrochemicals, plant allelochemicals, steroids, prostaglandins and fatty acids. In humans, cytochrome p450s are best known for their central role in phase I drug metabolism where they are of critical importance to two of the most significant problems in clinical pharmacology: drug interactions and inter individual variability in drug metabolism. Hp PIS57821.1 is identified as Dihydrofolatereductase is an enzyme that

converts dihydrofolate to tetrahydrofolate and is involved in purines and thymidylate synthesis. It is encoded by the human DHFR gene [30].

AAK_AK:PIS53610.1 Amino Acid Kinase Superfamily (AAK), Aspartokinase (AK); this CD includes the N-terminal catalytic domain of aspartokinase (4-L-aspartate-4-phosphotransferase). AK is the first enzyme in the biosynthetic pathway of the aspartate family of amino acids (lysine, threonine, methionine, and isoleucine) and the bacterial cell wall component, meso-diaminopimelate. It also catalyzes the conversion of aspartate and ATP to aspartylphosphate and ADP. One mechanism for the regulation of this pathway is by the production of several isoenzymes of aspartokinase with different repressors and allosteric inhibitors.

PIS53852.1 is identified as Minichromosome maintenance (MCM) is a family of six proteins (Mcm2–7, molecular masses of 101, 91, 97, 82, 93, and 81 kDa, respectively) with highly conserved amino acid sequences between the six different polypeptides.

They are essential for DNA replication in yeast and are considered to function as DNA helicases. Table 4, shows the functional classification of all the 161 HPs.

Table 4: Predicted Functions of 161 HPs from *C. auris*.

S.no	NCBI ID	FUNCTION
1.	PIS53738.1	Glycosyl hydrolase family 63 C-terminal domain
2.	PIS53809.1	DNA Topoisomerase I (eukaryota)
3.	PIS53852.1	Mini-chromosome maintenance protein
4.	PIS53863.1	Glycosyl hydrolase family 81 C-terminal domain
5.	PIS48955.1	Dolichyl-phosphate-mannose-protein mannosyltransferase
6.	PIS49007.1	Importin-beta N-terminal domain
7.	PIS49043.1	Formate--tetrahydrofolate ligase
8.	PIS54201.1	Kinesin motor domain
9.	PIS49083.1	Fructose-1-6-bisphosphatase, N-terminal domain
10.	PIS49097.1	Histidine phosphatase superfamily
11.	PIS49137.1	IKI3 family
12.	PIS54211.1	DnaJ molecular chaperone homology domain
13.	PIS57814.1	Peroxin 13, N-terminal region
14.	PIS57821.1	Dihydrofolatereductase (DHFR)
15.	PIS52624.1	Rapamycin-insensitive companion of mTOR, middle domain
16.	PIS54263.1	PQ loop repeat
17.	PIS57836.1	Peptidase C1-like family
18.	PIS57864.1	ALG6, ALG8 glycosyltransferase family
19.	PIS52684.1	Amino acid permease
20.	PIS52688.1	V-type ATPase 116kDa subunit family
21.	PIS54301.1	cwf21 domain
22.	PIS54327.1	Nitrogen permease regulator 2
23.	PIS52727.1	GATA zinc finger
24.	PIS52750.1	Acyl-CoA dehydrogenase, N-terminal domain
25.	PIS52798.1	Endomembrane protein 70
26.	PIS52815.1	GYF domain
27.	PIS54423.1	Exonuclease-Endonuclease-Phosphatase domain
28.	PIS52834.1	DNA polymerase family B, exonuclease domain
29.	PIS49264.1	PDZ domain
30.	PIS49271.1	Amino acid permease



31.	PIS49301.1	DNA polymerase family A
32.	PIS55854.1	Cytochrome P450
33.	PIS53001.1	Apoptosis antagonizing transcription factor
34.	PIS49417.1	Two tandem repeats of the cystathionine beta-synthase (CBS pair) domains superfamily
35.	PIS56994.1	Choline/Carnitine o-acyltransferase
36.	PIS53032.1	Zinc peptidases M18
37.	PIS53062.1	cytochrome P450 (CYP) superfamily
38.	PIS49420.1	DNA polymerase family B
39.	PIS49430.1	Tricarboxylate carrier
40.	PIS55967.1	Mitochondrial carrier protein
41.	PIS53116.1	P-loop containing Nucleoside Triphosphate Hydrolases
42.	PIS53131.1	WSC domain
43.	PIS56039.1	2Fe-2S iron-sulfur cluster binding domain
44.	PIS53217.1	haloaciddehalogenase-like hydrolase
45.	PIS57222.1	Zinc peptidases M18
46.	PIS56051.1	DnaJ domain or J-domain
47.	PIS53260.1	RAVE protein 1 C terminal
48.	PIS57258.1	Protein kinase C terminal domain
49.	PIS53362.1	Vacuolar protein sorting-associated protein 35
50.	PIS53384.1	Utp11 protein
51.	PIS53451.1	Nucleoporin interacting component Nup93/Nic96
52.	PIS50045.1	Oxysterol-binding protein
53.	PIS57324.1	Mitochondrial peculiar membrane protein 1
54.	PIS57354.1	Mitochondrial carrier protein
55.	PIS50838.1	Got1/Sft2-like family
56.	PIS56305.1	basic Helix Loop Helix (bHLH) domain superfamily
57.	PIS53601.1	WD40 domain
58.	PIS50196.1	U3 small nucleolar RNA-associated protein 25
59.	PIS50227.1	Glycosyl hydrolases family 32
60.	PIS50229.1	Glycosyl hydrolase family 81 C-terminal domain
61.	PIS50926.1	60Kd inner membrane protein
62.	PIS53610.1	Amino Acid Kinases (AAK) superfamily
63.	PIS50668.1	Importin-beta N-terminal domain
64.	PIS50698.1	Cytochrome P450
65.	PIS50964.1	Choline/ethanolamine kinase
66.	PIS51002.1	CRAL/TRIO, N-terminal domain
67.	PIS53667.1	glycosyltransferase family 1 and related proteins
68.	PIS55112.1	Beta-glucan synthesis-associated protein (SKN1)
69.	PIS56400.1	Copper amine oxidase, enzyme domain
70.	PIS56435.1	WD40 domain
71.	PIS55323.1	Aspartate/other aminotransferase
72.	PIS55332.1	Protein Kinases, catalytic domain
73.	PIS55422.1	Peptidase family M20
74.	PIS50791.1	CDC45-like protein
75.	PIS53765.1	Protein Kinases, catalytic domain
76.	PIS53779.1	Cyclin box fold superfamily
77.	PIS52508.1	NAD(P)+-dependent aldehyde dehydrogenase superfamily
78.	PIS57411.1	Amino acid permease
79.	PIS57450.1	Myosin head (motor domain)
80.	PIS54110.1	Amino acid permease [Amino acid transport and metabolism]
81.	PIS57483.1	Peptidase M3 mitochondrial intermediate peptidase (MIP)
82.	PIS57589.1	Exocyst complex component Sec10
83.	PIS57621.1	ChAPs (Chs5p-Arf1p-binding proteins)
84.	PIS57650.1	Trehalase
85.	PIS54172.1	Ubiquitin carboxyl-terminal hydrolase
86.	PIS54188.1	LETM1-like protein
87.	PIS57690.1	Sfi1 spindle body protein
88.	PIS49107.1	Eukaryotic translation initiation factor 4G1
89.	PIS49126.1	Acetyl-coenzyme A transporter 1



90.	PIS49138.1	V-type ATPase 116kDa subunit family
91.	PIS57724.1	Sec23-binding domain of Sec16
92.	PIS57740.1	Indoleamine 2,3-dioxygenase
93.	PIS52565.1	Chromatin assembly factor 1 subunit A
94.	PIS49182.1	CFEM domain
95.	PIS57790.1	Protein kinase domain
96.	PIS52683.1	Amino acid permease
97.	PIS52698.1	IMP-specific 5'-nucleotidase
98.	PIS54294.1	Protein Kinases, catalytic domain
99.	PIS52810.1	Universal stress protein family
100.	PIS49242.1	Protein Kinases, catalytic domain
101.	PIS52854.1	glycosyltransferase family 1
102.	PIS52870.1	Adenylate forming domain, Class I superfamily
103.	PIS57984.1	1,3-beta-glucan synthase component
104.	PIS57989.1	Protein Kinases, catalytic domain
105.	PIS58001.1	Mitochondrial import receptor subunit Tom22
106.	PIS49377.1	Metal-independent alpha-mannosidase (GH125)
107.	PIS49403.1	alpha/beta hydrolases
108.	PIS57088.1	ATP10 protein
109.	PIS57097.1	Peptidase family S64
110.	PIS49907.1	Patatins and Phospholipases
111.	PIS57131.1	Leucine rich repeat
112.	PIS57138.1	PWI domain
113.	PIS57147.1	Protein Kinases, catalytic domain
114.	PIS53117.1	Stretch-activated Ca2+-permeable channel component
115.	PIS53157.1	Adenylate forming domain, Class I superfamily
116.	PIS53160.1	Transport protein Trs120 or TRAPPC9, TRAPP II complex subunit
117.	PIS53168.1	Copper amine oxidase, enzyme domain
118.	PIS49938.1	Autophagy protein ATG9
119.	PIS57190.1	Zinc peptidases M18, M20, M28, and M42
120.	PIS56031.1	basic Helix Loop Helix (bHLH) domain superfamily
121.	PIS49980.1	Protein Kinases, catalytic domain
122.	PIS53287.1	Protein Kinases, catalytic domain
123.	PIS57261.1	Protein Kinases, catalytic domain
124.	PIS53367.1	C1 Peptidase family
125.	PIS53388.1	fungus transcription factor regulatory middle homology region
126.	PIS53549.1	DNA polymerase alpha subunit B N-terminal
127.	PIS56300.1	Protein kinase domain
128.	PIS56312.1	Ras family
129.	PIS50930.1	Maintenance of mitochondrial morphology protein 1
130.	PIS53622.1	Choline/Carnitine o-acyltransferase;
131.	PIS53635.1	SNARE motif
132.	PIS50669.1	Proteasome-substrate-size regulator, mid region
133.	PIS50701.1	RNA recognition motif (RRM) superfamily
134.	PIS50725.1	Major Facilitator Superfamily
135.	PIS50729.1	Bacterial NAD-glutamate dehydrogenase
136.	PIS56470.1	Cytochrome P450
137.	PIS56493.1	Ubiquitin-conjugating enzyme E2, catalytic (UBCc) domain
138.	PIS56525.1	Amino acid permease
139.	PIS55322.1	alpha-D-phosphohexomutase superfamily
140.	PIS55421.1	fungus transcription factor
141.	PIS50105.1	PCI domain
142.	PIS56199.1	Coproporphyrinogen III oxidase
143.	PIS56219.1	Transcription initiation factor TFIID component TAF4 family
144.	PIS50123.1	Protein Kinases, catalytic domain
145.	PIS53555.1	RNA polymerase I specific transcription initiation factor RRN3
146.	PIS53606.1	Amino acid permease
147.	PIS50870.1	Gpi16 subunit, GPI transamidase component
148.	PIS50921.1	Ankyrin repeats (3 copies)



149.	PIS53644.1	Yeast mitochondrial distribution and morphology (MDM) proteins
150.	PIS50282.1	Cytochrome P450
151.	PIS50654.1	Protein Kinases, catalytic domain
152.	PIS51014.1	Amino acid permease
153.	PIS53654.1	Synaptobrevin
154.	PIS53693.1	RING-H2 zinc finger domain
155.	PIS53714.1	ABC transporter transmembrane region
156.	PIS56387.1	ABC transporter transmembrane region
157.	PIS55113.1	Beta-glucan synthesis-associated protein SKN1/KRE6/Sbg1
158.	PIS55130.1	Amino acid permease
159.	PIS55208.1	Phosphatidylinositol phosphate kinase (PIPK) catalytic domain family
160.	PIS56491.1	K homology (KH) RNA-binding domain, type I
161.	PIS48116.1	Calcineurin-like phosphoesterase

Virulence prediction:

Bacterial pathogens tend to enter the host cells with the help of various active and passive mechanisms. There they produce specific virulence factors that promote propagation throughout the body of the host. Microbial products like adhesions, toxins (endotoxins and exotoxins), porins, etc., constitute the leading machine of bacterial virulence by helping the bacteria in host cell invasion, adhesion, ensuing dissemination

throughout the host body. Strategically, we consider these virulence factors as possible drug targets for the development of effective drugs against the pathogenesis of the bacteria. We used VirulentPred results for predicting virulence factors among the 161 HPs and found 95 HPs that give positive virulence score and can be used as potent drug targets for designing a drug. The results of virulent factor analysis have been summarised in Table 5.

Table 5: Predicted virulence factors from *C. auris*.

S. No	NCBI ID	PREDICTION RESULTS	PREDICTED SCORES
1.	PIS53852.1	Virulent	1.0146
2.	PIS53863.1	Virulent	0.9989
3.	PIS54201.1	Virulent	1.0941
4.	PIS49137.1	Virulent	0.3086
5.	PIS57821.1	Virulent	1.0551
6.	PIS52624.1	Virulent	0.8559
7.	PIS54263.1	Virulent	0.5549
8.	PIS57836.1	Virulent	0.1824
9.	PIS52688.1	Virulent	0.9447
10.	PIS54327.1	Virulent	1.0334
11.	PIS52727.1	Virulent	1.001
12.	PIS52798.1	Virulent	0.4951
13.	PIS52815.1	Virulent	1.0274
14.	PIS52834.1	Virulent	1.0487
15.	PIS49301.1	Virulent	0.9388
16.	PIS53001.1	Virulent	0.2648
17.	PIS49420.1	Virulent	0.3545
18.	PIS49430.1	Virulent	0.3278
19.	PIS56039.1	Virulent	0.8702
20.	PIS53260.1	Virulent	1.0523
21.	PIS57258.1	Virulent	0.9038
22.	PIS53362.1	Virulent	1.0555
23.	PIS53384.1	Virulent	0.8374
24.	PIS53451.1	Virulent	1.0281
25.	PIS50045.1	Virulent	0.2987
26.	PIS57324.1	Virulent	0.9863
27.	PIS57354.1	Virulent	1.0451
28.	PIS56305.1	Virulent	1.1122
29.	PIS53601.1	Virulent	1.0706
30.	PIS50196.1	Virulent	0.2195
31.	PIS50227.1	Virulent	0.9922
32.	PIS50229.1	Virulent	1.1561



33.	PIS50668.1	Virulent	0.2362
34.	PIS50964.1	Virulent	1.0256
35.	PIS55112.1	Virulent	0.9091
36.	PIS56435.1	Virulent	0.9131
37.	PIS50791.1	Virulent	0.976
38.	PIS53779.1	Virulent	1.0445
39.	PIS57450.1	Virulent	1.0294
40.	PIS57483.1	Virulent	1.0542
41.	PIS57589.1	Virulent	0.3777
42.	PIS57621.1	Virulent	0.7792
43.	PIS54188.1	Virulent	0.9979
44.	PIS57690.1	Virulent	0.8019
45.	PIS49107.1	Virulent	0.9685
46.	PIS49138.1	Virulent	0.3232
47.	PIS57724.1	Virulent	1.0129
48.	PIS52565.1	Virulent	0.9279
49.	PIS49182.1	Virulent	1.1076
50.	PIS57790.1	Virulent	0.9417
51.	PIS54294.1	Virulent	0.5866
52.	PIS52810.1	Virulent	1.0524
53.	PIS49242.1	Virulent	1.0748
54.	PIS52854.1	Virulent	1.007
55.	PIS57984.1	Virulent	1.0799
56.	PIS58001.1	Virulent	1.1214
57.	PIS49377.1	Virulent	0.9951
58.	PIS57088.1	Virulent	1.1117
59.	PIS57097.1	Virulent	1.0654
60.	PIS49907.1	Virulent	1.0476
61.	PIS57131.1	Virulent	1.0528
62.	PIS53117.1	Virulent	1.0751
63.	PIS53160.1	Virulent	1.0068
64.	PIS53168.1	Virulent	0.9178
65.	PIS49938.1	Virulent	0.9033
66.	PIS56031.1	Virulent	0.943
67.	PIS49980.1	Virulent	0.9704
68.	PIS53287.1	Virulent	1.0283
69.	PIS57261.1	Virulent	0.8981
70.	PIS53367.1	Virulent	0.3402
71.	PIS53388.1	Virulent	1.0105
72.	PIS53549.1	Virulent	0.9726
73.	PIS56300.1	Virulent	0.7982
74.	PIS56312.1	Virulent	1.0597
75.	PIS50930.1	Virulent	1.026
76.	PIS53622.1	Virulent	1.1134
77.	PIS53635.1	Virulent	1.0789
78.	PIS50669.1	Virulent	1.138
79.	PIS50729.1	Virulent	0.4231
80.	PIS55421.1	Virulent	0.964
81.	PIS50105.1	Virulent	1.0137
82.	PIS56219.1	Virulent	0.971
83.	PIS50123.1	Virulent	1.0037
84.	PIS53555.1	Virulent	1.0037
85.	PIS50870.1	Virulent	0.3879
86.	PIS50921.1	Virulent	1.1489
87.	PIS53644.1	Virulent	0.7921
88.	PIS50654.1	Virulent	0.7761
89.	PIS53654.1	Virulent	1.0092
90.	PIS53693.1	Virulent	1.033
91.	PIS53714.1	Virulent	1.0157



92.	PIS56387.1	Virulent	0.794
93.	PIS55113.1	Virulent	0.9552
94.	PIS55208.1	Virulent	1.122
95.	PIS48116.1	Virulent	1.09

CONCLUSION:

An *insilico* approach finally we predicted the functions of 161 from 4689 HPs with high precision. We could not find enough evidence for characterization of remaining 4528 HPs, therefore only 161 HPs are chosen. We predicted 95 HPs as virulence proteins, which are important for pathogenesis and survival of this organism. These findings may facilitate the drug discovery process to bring forward effective drugs against the pathogenesis of *C. auris*. For this reason, it is important to quickly identify *C. auris* characteristic. Moreover, by knowing the function, it is helpful for identify the drug discovery and developmental use, really it's mandatory for identify and easily cure the patient.

Reference

- Brown GD, Denning DW, Gow NA, Levitz SM, Netea MG, White TC. 2012. Hidden killers: human fungal infections. *SciTransl Med.* 4(165): 165rv13. <https://doi.org/10.1126/scitranslmed.3004404> PMID: 23253612.
- Wisplinghoff H, Bischoff T, Tallent SM, Seifert H, Wenzel RP, Edmond MB. 2004. Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. *Clin Infect Dis.* 39(3):309–317.
- Du H, Bing J, Hu T, Ennis CL, Nobile CJ, and Huang G 2020. *Candida auris*: Epidemiology, biology, antifungal resistance, and virulence. *PLoSPathog* 16(10): e1008921.
- Ahmad S, and Alfouzan W 2021. *Candida auris*: Epidemiology, Diagnosis, Pathogenesis, Antifungal Susceptibility, and Infection Control Measures to Combat the Spread of Infections in Healthcare Facilities. *Microorganisms* 9(4): 807.
- Vila T, Sultan AS, Montelongo-Jauregui D, and Jabra-Rizk MA 2020. *Candida auris*: a fungus with identity crisis. *Pathog Dis* 78(4): ftaa034.
- CernakovaL, Roudbary M, Bras S, Tafaj S, and Rodrigues CF 2021. *Candida auris*: A Quick Review on Identification, Current Treatments, and Challenges. *Int J MolSci* 22(9): 4470.
- Satoh K, Makimura K, Hasumi Y, Nishiyama Y, Uchida K, Yamaguchi H. 2009. *Candida auris* sp. nov., a novel ascomycetous yeast isolated from the external ear canal of an inpatient in a Japanese hospital. *MicrobiolImmunol.* 53:41–4.
- Sarma S, Upadhyay S. 2017. Current perspective on emergence, diagnosis and drug resistance in *Candida auris*. *Infect Drug Resist.* 10:155–65.
- Kohlenberg A, Struelens MJ, Monnet DL, Plachouras D. 2018. *Candida auris*: epidemiological situation, laboratory capacity and preparedness in European Union and European Economic Area countries, 2013 to 2017. *Euro Surveill.* 23:18–00136.
- Kim SH, Iyer KR, Pardeshi L, Munoz JF, Robbins N, Cuomo CA, 2019. Genetic analysis of *Candida auris* implicates Hsp90 in morphogenesis and azole tolerance and Cdr1 in azole resistance. *mBio.* 10: pii: e02529-18.
- E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, 2003. ExpPASy: the proteomics server for in-depth protein knowledge and analysis, *NucleicAcids Res.* 31 (13) 3784–3788.
- C.S. Yu, Y.C. Chen, C.H. Lu, J.K. Hwang, 2006. Prediction of protein subcellular localization, *Proteins* 64 (3) 643–651.
- M.C. Hung, W. Link, 2011. Protein localization in disease and therapy, *J. Cell Sci.* 124 (Pt. 20) 3381–3392.
- B. Shanmugham, A. Pan, 2013. Identification and characterization of potential therapeutic candidates in emerging human pathogen *Mycobacterium abscessus*: a novel hierarchical in silico approach, *PLoS One* 8 (3) e59126.
- U. Vetrivel, G. Subramanian, S. Dorairaj, 2011. A novel in silico approach to identify potential therapeutic targets in human bacterial pathogens, *Hugo J.* 5 (1–4) 25–34.
- Emanuelsson O., S. Brunak, G. von Heijne, H. Nielsen, 200. Locating proteins in the cell using TargetP, SignalP and related tools, *Nat. Protoc.* 2 (4) 953–971.
- Bendtsen J.D., L. Kiemer, A. Fausboll, S. Brunak, 2005. Non-classical protein secretion in bacteria, *BMC Microbiol.* 5-58.
- Tusnady G.E., I. Simon, 2001. The HMMTOP transmembrane topology prediction server, *Bioinformatics* 17 (9) 849–850.
- Schultz J., R.R. Copley, T. Doerks, C.P. Ponting, P. Bork, 2000. SMART: a web-based tool for the study of genetically mobile domains, *Nucleic Acids Res.* 28 (1) 231–234.
- Finn R.D., J. Clements, S.R. Eddy, 2011. HMMER web server: interactive sequence similarity searching, *Nucleic Acids Res.* 39 (Web Server issue) W29–W37.
- Baron C., B. Coombes, 2007. Targeting bacterial secretion systems: benefits of disarmament in the



- microcosm, Infect. Disord. Drug Targets 7 (1) 19–27.
22. Saha S., G.P. Raghava, 2006.VICMpred: an SVM-based method for the prediction of functional proteins of Gram-negative bacteria using amino acid patterns and composition, Genomics Proteomics Bioinf. 4 (1) 42–47.
23. Sivashankari S, Shanmughavel P. 2006. Functional annotation of hypothetical proteins-A review. Bioinformation 1:335-338.
24. McGinnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res 32:W20-W25.
25. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. 2004. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res 32:D115-D119.
26. Chatterjee, S., Alampalli, S. V., Nageshan, R. K., Chettiar, S. T., Joshi, S., and Tatu, U. S. 2015. Draft genome of a commonly misdiagnosed multidrug resistant pathogen *Candida auris*. BMC Genomics. 16:686.
27. M.A. Gazi, M.G. Kibria, M. Mahfuz, M.R. Islam, P. Ghosh, M.N. Afsar, M.A. Khan, T. Ahmed, 2016. Functional, structural and epitopic prediction of hypothetical proteins of *Mycobacterium tuberculosis* H37Rv: an in silico approach for prioritizing the targets Gene, 591 (2), pp. 442-455.
28. M. Shahbaaz, K. Bisetty, F. Ahmad, I. Hassan, 2015. Functional insight into putative conserved proteins of *Rickettsia rickettsii* and their virulence characterization, Curr. Proteomics, 12 (2), pp. 101-116.
29. M. Shahbaaz, M.I. Hassan, F. Ahmad 2013. Functional annotation of conserved hypothetical proteins from *Haemophilus influenzae* Rd KW20. PLoS One, 8 (12).
30. Liu CT, Francis K, Layfield JP, Huang X, Hammes-Schiffer S, Kohen A, Benkovic SJ 2014. "*Escherichia coli* dihydrofolate reductase catalyzed proton and hydride transfers: temporal order and the roles of Asp27 and Tyr100". Proceedings of the National Academy of Sciences of the United States of America. **111** (51): 18231–6.