

## NOVEL THERAPEUTIC PROTEIN TARGETS IDENTIFICATION IN ARCOBACTER CRYAEROPHILUS IN POST-GENOMIC ERA

<sup>1</sup>Sonidia Adil, <sup>2</sup>Dr. Syed Shah Hassan, <sup>3</sup>Saliha Salam, <sup>4</sup>Syeda Sakina,

<sup>5</sup>Muhammad Muzamil Arshad, <sup>6</sup>Hamad Ullah

<sup>1</sup>Islamia College, University Peshawar, Pakistan

<sup>2</sup>JRC Genome Research, PCMD-ICCBS, University of Karachi 75270

<sup>3</sup>Lecturer Biochemistry, Sarhad Institute of Allied Health Sciences, SUIT Peshawar, Pakistan

<sup>4</sup>Lecturer Human Nutrition and Dietetics, Sarhad Institute of Allied Health Sciences, SUIT Peshawar, Pakistan

<sup>5</sup>Lecturer Biochemistry, Sarhad Institute of Allied Health Sciences, SUIT Peshawar, Pakistan

<sup>6</sup>Lecturer Biochemistry, Sarhad Institute of Allied Health Sciences, SUIT Peshawar, Pakistan

<sup>1</sup>[adilsanan99@gmail.com](mailto:adilsanan99@gmail.com) <sup>2</sup>[hassanchemist83@gmail.com](mailto:hassanchemist83@gmail.com) <sup>3</sup>[Saliha.siahs@suit.edu.pk](mailto:Saliha.siahs@suit.edu.pk)

<sup>4</sup>[Sakina.siahs@suit.edu.pk](mailto:Sakina.siahs@suit.edu.pk) <sup>5</sup>[muzamil.siahs@suit.edu.pk](mailto:muzamil.siahs@suit.edu.pk) <sup>6</sup>[hamadullah.siahs@suit.edu.pk](mailto:hamadullah.siahs@suit.edu.pk)

DOI: <https://doi.org/>

### Keywords:

Arcobacter Cryaerophilus, Post-Genomic Era and Reverse Vaccinology

### Article History

Received on 21 Feb, 2026

Accepted on 09 March, 2026

Published on 11 March, 2026

Copyright @Author

Corresponding Author: \*

Saliha Salam

### Abstract

*Arcobacter cryaerophilus*, a gram-negative, curved or helical bacillus primarily known as a bovine and porcine pathogen. *Arcobacter cryaerophilus* is found in the gastrointestinal tract of healthy animals such as poultry, pigs, sheep and sheep, but is associated with enteritis in animals and may play a pathogenic role in reproductive disorders in cattle, pigs and sheep. *Arcobacter* transmission to humans is mediated by contaminated foods and water, and poultry products act as an important reservoir. Here, we applied a reverse vaccinology and subtractive genomic approach for in silico prediction of potential vaccines and drug targets for seven strains from the *A. cryaerophilus* species. Next, the core genome (942) was predicted, followed by non-host homology analysis, resulting in 662 target proteins. By screening for essential genes, 35 genes can be obtained. After screening for essential genes, we also checked their intracellular localization, functionality, structural modeling, and protein-protein interactions. 6 proteins (Tryptophan synthase alpha chain, Homoserine kinase, Anthranilate phosphoribosyl transferase, Tryptophan synthase beta chain, Imidazole glycerol phosphate synthase subunit HisF and Phenylalanine-tRNA ligase beta subunit beta chain) are considered highly conserved non-host essential homologues with 3D modeled structures and were further subjected to druggability analysis, followed by toxicity factors prediction. The proposed approach facilitates the study of putative targets for selected *Arcobacter* species for the development of a broad spectrum of new drugs and vaccines based on experimental validation.

## Introduction

In 1991, the term *Arcobacter* was proposed. The first isolation of an *Arcobacter* species was from bovine fetuses in the late 1970s. *Arcobacter* is a Gram-negative, spiral-shaped bacterium belonging to the Epsilon proteobacteria class. It has an extraordinarily broad range of environments, and some of the species are human and animal diseases [1]. *A. butzleri*, *A. cryaerophilus*, *A. skirrowii*, *A. nitrofigilis*, and *A. sulfidicus* are the five species that make up this genus. *A. butzleri*, *A. cryaerophilus*, *A. skirrowii*, and *A. thereius* have all been found in human infectious processes, particularly diarrhea [2]. *Arcobacter cryaerophilus* is a new enteropathogen and zoonotic threat that may be spread by food and water. It's a Gram-negative curving rod that's known to cause food and waterborne illness. Because of its DNA base makeup and overall appearance, this bacterium was formerly described as aero tolerant *Campylobacter*-like, distinguishing itself from other *Campylobacter spp.* by its ability to thrive at temperatures as low as 15 C. Vandamme and De Ley suggested bacteria and the *Arcobacter* genus more than two decades ago. *Arcobacter* species have been found all over the world. They've been found in a variety of animal-based items, including chicken, beef, pig, seafood, and milk [3]. Isolation from contaminated drinking and feces water has also been recorded [4]. *Arcobacter* is derived from the Latin *arcus*, which means "bow," and the Greek *bacter*, which means "rod," and means "bow-shaped rod" or "curved rod." This refers to the curved form that most *Arcobacter* cells have. Gram-negative *Arcobacter spp.* are 0.2–0.9 mm broad and 1–3 mm long, slightly curved, S-shaped, or helical rods [5]. A single polar, unsheathed flagellum provides a distinctive darting or corkscrew-like movement. Depending on the species, *Arcobacters* develop white or greyish smooth-rounded colonies of varied sizes. They can multiply at 15 and 37 degrees Celsius, but not all of them can multiply at 41.5 degrees Celsius. *Arcobacters* can grow under aerobic conditions, however micro aerobic settings with 3–10% O<sub>2</sub> and 5–10% CO<sub>2</sub> produce the best results. When hydrogen is present, *A. skirrowii* grows better [6]. *Arcobacter's* most distinguishing characteristics are their capacity to grow in air at 30 degrees Celsius and their ability to thrive at temperatures as low as 15 degrees Celsius [7]. According to the World Health Organization, different *Arcobacter* species can cause different diseases in animals and humans, and these bacteria are classified intestinal pathogens. It has been demonstrated that this category of chemicals may cause sickness in humans and animals. *A. butzleri*, *A. cryaerophilus*, *A. skirrowii*, and *A. thereius* have all been found in human infectious processes, particularly diarrhea [8]. Cross-contamination during food handling, intake of contaminated animal food, polluted drinking water, or direct contact with feces are all possible ways to become infected. *Arcobacter butzleri*, *Arcobacter cryaerophilus*, and *Arcobacter skirrowii* are three species of *Arcobacter* that have been linked to a variety of disorders in people and animals [9].

*Arcobacter butzleri* is the most often reported species in humans as a cause of gastroenteritis and bacteremia, with *Arcobacter cryaerophilus* and *Arcobacter skirrowii* being less prevalent. The most common symptom reported is diarrhea [10]. They've been discovered as a cause of diarrhea in malnourished children in impoverished nations, in particular. Although both *Arcobacter spp.* and *Campylobacter jejuni* produce diarrhoea with comparable symptoms, *A. butzleri* infections are more often associated with persistent, watery diarrhoea, and *C. jejuni* infections are more usually associated with bloody diarrhea [11]. Infections with *Arcobacter butzleri* are also linked to stomach discomfort, both with and without diarrhoea, nausea and vomiting, and fever. Some people, however, may be asymptomatic. Chronic diarrhoea has been linked to the bacteria *Arcobacter skirrowii*. Enteritis, mastitis, and reproductive problems, including abortions, are all linked to *Arcobacter butzleri*, *A. cryaerophilus*, and *A. skirrowii* in cattle, sheep, and pigs. In non-human primates, such as macaques, *Arcobacter butzleri* is linked to diarrhea [12]. Although *Arcobacters* are linked to a variety of illnesses in people and animals, our understanding of their pathogenicity and toxin generation mechanisms is still restricted. Some strains produce cytotoxins and cytolethal distending factors, while others hemagglutinate human and animal

erythrocytes and have the ability to attach to cell lines and invade them [13]. *Arcobacter* species, notably *A. butzleri* and *A. cryaerophilus*, are emerging pathogens in humans that cause gastroenteritis. Infections with *Arcobacter* can cause reproductive problems, mastitis, and diarrhoea in animals, and the bacterium can also be isolated from healthy carriers [14]. In humans, severe instances have been documented, including protracted watery diarrhoea with stomach pains, bacteremia, endocarditis, and peritonitis as a result of *Arcobacter* infection [15]. Human infections with *A. butzleri* and *A. cryaerophilus* are the most common species isolated from human specimens, while human infections with *A. skirrowii* and *A. thereius* are uncommon. *A. skirrowii* is a pathogen that is well-known but seldom found. Slow growth on culture media and overrun by other bacteria might be the major causes [16]. For the detection and identification of this infection, most conventional microbiology lab procedures must be modified. Because of the bacteria's fastidious growth, antimicrobial susceptibility testing of *A. skirrowii* is best done using the gradient strip method [17]. Antibiotic susceptibility predictions based on WGS data should be approached with care. Resistance to erythromycin, tetracycline, and streptomycin was found in *A. butzleri*. Only streptomycin was resistant to *A. skirrowii* [18]. Macrolides (such as erythromycin) are the preferred antibiotics for *Campylobacter* infections, but they aren't always the first option for *Arcobacter* infections, for which tetracycline was only recommended in severe instances [19]. The unrestricted use of antibiotics in animal husbandry is a major role in the development of antimicrobial resistance. Antimicrobial resistance data monitoring and reporting, as well as WGS data analysis of *Campylobacter* and *Arcobacter* from domestic animals, are critical for tracking *antimicrobial* resistance evolution and optimizing diagnoses. The *Arcobacter* genus has 27 species with significant genetic diversity and rising antibiotic resistance [20]. *Arcobacter* is a genus of Gram-negative bacteria that belongs to the *Campylobacteraceae* family and the *Epsilonproteobacteria* class. They are non-spore-forming, spiral-shaped, motile, and fastidious [21]. The genus was first established in 1991 as a group of aerotolerant bacteria. Until recently, the genus has been made up of 27 species with significant genetic diversity and rising antibiotic resistance [22].

## 2. Material and methods

### 2.1 Genome Selection of *Arcobacter cryaerophilus*

We decided on the genomes of *A. cryaerophilus* due to the fact maximum of the genomes have already been sequenced and have handiest been sequenced recently, demonstrating the significance of this *Arcobacter* species to be had in GOLD (Genomes Online Database). This database is open supply for complete get entry to records on genome and metagenome sequencing initiatives and associated metadata across the world [36].

### 2.2 Data Retrieval of Pathogens

The seven stains of *A. cryaerophilus* had been covered on this study. The whole genome, genes and protein sequences of those traces had been retrieved from NCBI (the National Center for Biotechnology Information) the use of ftp file (<ftp://ftp.ncbi.nih.gov/genomes/Bacteria>) [36].

### 2.3 Prediction of Core Genome

To collect the center genome of *A. cryaerophilus*, an excessive throughput carrier known as PATRIC (the Pathosystem Resource Integration Center ([www.patricbrc.org/](http://www.patricbrc.org/)), become used to expect the center genome with the aid of using deciding on randomly one strain (Ac\_16CS0369-1-AR-4, .faa file) as our reference and the closing six lines have been in comparison with this reference strain. Genes that were common in all the strains were selected for further analysis [37].

### 2.4 Identification of Non host homologous

The record changed into subjected to NCBI-BLASTp ( $E\text{-value}=0.0001$ ,  $bit\ score = 100$  and identification 25%) ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) in opposition to the human genome for filtering pathogen non-host homologous, after the prediction of core genome. The contrast of proteins

with human host protein finds the non-hit proteins lists that denote non-human homologous proteins of the pathogens. This will assist to layout the pathogen specific therapeutics drugs [38].

### **2.5 Analyses of Essential Genes**

A subtractive genomics approach was followed to identify conserved targets essential for bacteria. All the conserved core proteins of *Arcobacter cryaerophilus* were submitted to the Database of Essential Genes (DEG) for homology analysis. DEG contains experimentally validated data on bacteria, archaea and eukaryotes that include essential genomic elements currently reported, including genes encoding proteins that are essential for supporting cell life. The threshold values used for BLASTp were: E value = 0.0001, bit score = 100 and identity = 25% [39].

### **2.6 Comparative Subcellular Localization**

Proteins designated as non-redundant, non-homologous humans in the previous phase were analyzed for subcellular localization to execute the proteins that make up the exoproteome and secretome of the pathogen. The exoproteome and secretome are considered an excellent source of vaccine candidates because of their frequent contact with biotic and abiotic factors in the extracellular environment. The subcellular localization of the proteins was based on a comparative approach with two subcellular localization tools of the line PsortB and CELLO2GO carried out [40].

### **2.7 Prediction of Protein-Protein Interaction**

A protein-protein interaction network of the selected proteins was constructed using STRING, which includes more than 1100 completely sequenced organisms. The Non host homologous critical genes are subjected to String to test their interplay with different protein. The proteins which can be displaying the more than five interplay are the critical proteins. In String the protein series is submitted and offers the desired result. <https://string-db.org/cgi/input> [41].

### **2.8 Detection of protein biological pathways**

Metabolic pathways that are present only in pathogen and not in the host can be used in a targeted manner for effective drug design. KEGG (Kyoto Encyclopedia of Genes and Genomes) is a pathway database, which is a metabolic pathway for non-homologous proteins. We manually compared the metabolic pathways of host and the pathogen to identify pathways that exist only in the pathogen and not in the human host. Added a list of proteins that play a role in unique signaling pathways. Proteins were also isolated according to their role only in pathogen-specific, unique and common pathways present in both the pathogen and the host [42]

### **2.9 Detection of biological functions**

A freely accessible database of protein sequences and functional information was used to understand the biological and molecular function of UniProt proteins (<https://www.uniprot.org>). Contains a wealth of information on the biological function of proteins derived from research literature [42]

### **2.10 Three-Dimensional (3D) Protein Structure Prediction**

We submitted the fasta protein sequences to the Swiss Model database ([swissmodel.expasy.org/](http://swissmodel.expasy.org/)) one by one and as a result the protein systems got early by uploading the PDB (protein database) documents to from the Swiss version, after which they were consulted. These 3-dimensional (3D) structures through the use of PyMOL [43].

### **2.11 Protein structure validation**

The protein PDB file was submitted to an online database tool called SAVES v6.0 (<http://saves.mbi.ucla.edu>) to confirm protein model validation and the valid model was submitted to a more in-depth analysis.

### **2.12 Visualization of protein model**

To visualize the confirmed protein models, PYMOL, a three-dimensional structure visualization application that led to the 3D structure of each protein one by one, was used.

### **2.14 Molecular weight Determination**

The molecular weight (MW) of each of the potential targets was determined using online tools, followed by compliance with the available literature. Virulent proteins have been classified according to molecular weight. ProtParm is a tool which calculates various physicochemical properties of a protein. This tool is available at (<http://web.expasy.org/protparam>). In this study, protein sequences were submitted to ProtParm to analyze molecular weight [44].

### 2.13 Drug ability analysis

DoGSiteScorer (<https://proteins.plus/>) was used to find the catalytic pocket of essential proteins with the specific drug score. DoGSiteScorer is an automated pocket detection tool and calculates the capacity of protein cavities. The 3D model or the protein ID PDB can be provided as input for the analysis [45].

### 2.15 Molecular docking and screening

To detect drug interactions with the key targets we need, the targets were applied to the MOE computerized database, which resulted in molecular screening and comparison of target targets with organ libraries to match target targets. This step led to the creation of a set of drug targets. It will help us find drugs for the protein we are studying.

## Results

In the present study, we identify potential therapeutic targets in *A. cryaerophilus* through comparative and subtractive genome analyzes. We used a systematic hierarchical approach that included the use of various computer tools, database searches, and prioritization analysis of drug targets.

### 3.1 Data Retrieval

NCBI (National Center for Biotechnology Information) is a database resource that provides access to biomedical and genomic information used to review the genome construction and annotation report of *A. cryaerophilus* of 7 complete strains of *A. cryaerophilus* one at a time using an ftp file and all of them were analyzed in the present study.

**Table-2: General information about the 11 strains of *A. cryaerophilus* used in this work**

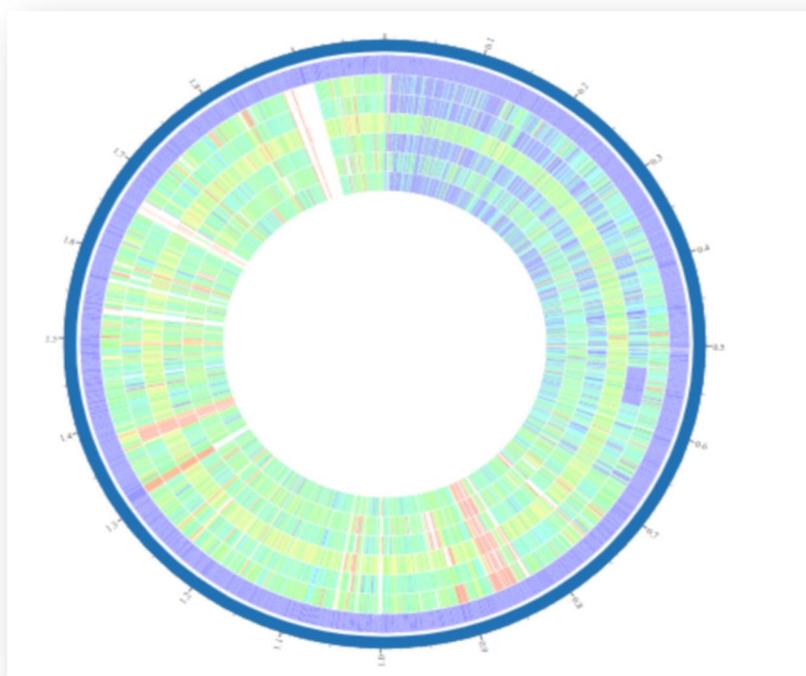
S. No.	Organism Name	Strain	Bio Sample	Assembly	Size (Mb)	%GC	Replicons	Proteins
1	<i>Arcobacter cryaerophilus</i>	16CS036 9-1-AR-4	16CS1285-4	GCA_0143019 85.1	2.0 2	27.5 0	• chromosome: NZ_CP060264.1 / CP060264.1	2,030
2	<i>Arcobacter cryaerophilus</i>	G13RTA	SAMN08391 433	GCA_0110454 15.1	2.1 2	27.4 0	• chromosome: NZ_CP026655.1 / CP026655.1	2,114
3	<i>Arcobacter cryaerophilus</i>	ATCC 43158	SAMN03737 948	GCA_0036601 05.1	2.1 1	27.3 6	• chromosome: NZ_CP032823.1 / CP032823.1	2,076
4	<i>Arcobacter cryaerophilus</i>	D2610	SAMN03737 949	GCA_0036600 85.1	2.0 6	27.5 0	• chromosome: NZ_CP032825.1 / CP032825.1	1,998

5	<i>Arcobacter cryaerophilus</i>	16CS083 0-1	SAMN14846 593	GCA_0143528 95.1	2.0 5	27.5 0	• chromoso me: NZ_CP060692.1 / CP060692.1	2,029
6	<i>Arcobacter cryaerophilus</i>	16CS129 2-4	SAMN14846 675	GCA_0143529 35.1	2.0 2	27.6 0	• chromoso me: • NZ_CP060694.1 / CP060694.1	2,023
7	<i>Arcobacter cryaerophilus</i>	16CS128 5-4	PRJNA63272 0	GCA_0143529 15.1	2.1 4	27.7 0	chromosome: NZ_CP060693.1 / CP060693.1	2,143

### 3.2 Prediction of Core Proteome

A high throughput service called PATRIC was used to predict the core genome file by selecting the proteome comparison option in the service path of the 7 *A. cryaerophilus* strains, one strain (Ac\_16CS0369-1-AR-4) was used as the reference strain and six other strains. They were compared to our reference and then the data was transmitted. The coding DNA sequences that all strains have in common are part of the nuclear genome and after a while we got the results by checking the workspace in PATRIC and thus downloading the Genome Compare.txt file, which corresponds to 2,026 gene sequence.

Bidirectional best hit	100	99.9	99.8	99.5	99	98	95	90	80	70	60	50	40	30	20	10
Unidirectional best hit	100	99.9	99.8	99.5	99	98	95	90	80	70	60	50	40	30	20	10



*Figure 3. Circular genome representation of A. cryaerophilus generated through PATRIC server ([www.patricbrc.org](http://www.patricbrc.org))*

### List of tracks, from outside to inside

- 1 *A. cryaerophilus*\_16CS0369-1-AR-4
- 2 *A. cryaerophilus*\_G13RTA
- 3 *A. cryaerophilus*\_D2610
- 4 *A. cryaerophilus*\_ATCC\_43158
- 5 *A. cryaerophilus*\_16CS1292-4
- 6 *A. cryaerophilus*\_16CS1285-4
- 7 *A. cryaerophilus*\_16CS0830-1

### 3.3 Omission of Redundant Data

In this step, we opened the master genome file in Excel and omitted the incomplete date (i.e., incomplete percent identity, genome function, and sequence coverage). This filter reduced our file to 942 targets.

### 3.4 Identification of Non- Host Homologous and Intra-species Conserved Proteins

The core genome comparison file containing 942 genomes was then subjected to NCBI BLASTp (*E-value* = 0.0001, *Bit score* 100 and 25% identity) against the human genome to filter for non-host pathogen homologues. Among these gene sequences, taking into account the human genome as the host genome, they found that 662 were non-homologous host proteins.

Localization	Amount
Extracellular	1
Outer membrane	1
Periplasmic	1
Inner membrane	4
Cytoplasmic	31

This step is important in order to avoid undesirable cross-reactivity of the active substance due to its binding to the active sites of homologous proteins in the host.

### 3.5 Analysis of Essential Genes

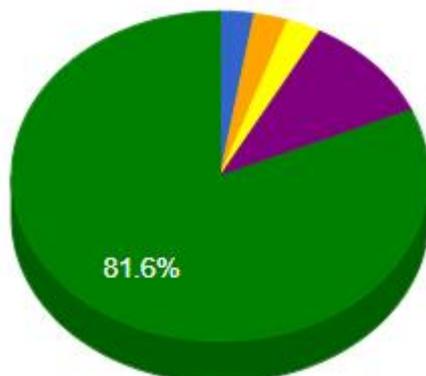
We downloaded the essential genetic files of eukaryotes, prokaryotes and archaea from DEG (<http://tubic.tju.edu.cn/deg/>). Therefore, the first BLASTp search was done on the Archaea file from which we get 89 essential proteins, according to BLASTp the research was done on the eukaryotic file from which our data was reduced to 35 essential proteins, and then thirdly, we ran the BLASTp search on the prokaryotic file from which we get the same 35 essential proteins (done in NCBI BLASTp and [https://fasta.bioch.virginia.edu/fasta\\_www2/fasta\\_list2.shtml](https://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml) using the Perl script with a threshold Evaluated as 10e4, a bit score of 100 and a sequence identity  $\geq$  30%) The results show that 35 proteins were essential for *A. cryaerophilus*.

### 3.6 Comparative SubCellular Localization

The subcellular localization of proteins in a cell is an important feature that can determine their potential functions and identify suitable and effective drug targets. Cytoplasmic proteins are more beneficial as therapeutic drug targets because membrane-localized proteins are difficult to purify. The essential proteins of *Arcobacter cryaerophilus* were further characterized based on other essential properties such as: Accessibility value of a target protein secreted, putative surface exposed (PSE) and membrane proteins based on the presence or absence of signal peptides, retention signals and presence of transmembrane helices. From the 35 protein targets of the genome, we obtained the following results, which are shown in the table 3 below

Periplasmic  
Inner membrane  
Cytoplasmic

Figure 4. Comparative Cellular Localization Prediction.



### 3.7 Protein-Protein Interaction

35 essential proteins non-homologous to the host are obtained. These proteins exhibit a protein-protein interaction with several proteins. Out of 35 proteins, only 22 proteins show multiple interactions with other proteins. These proteins were subjected to STRING. It is also believed that a highly interactive protein is metabolically important and can act as a potent drug.

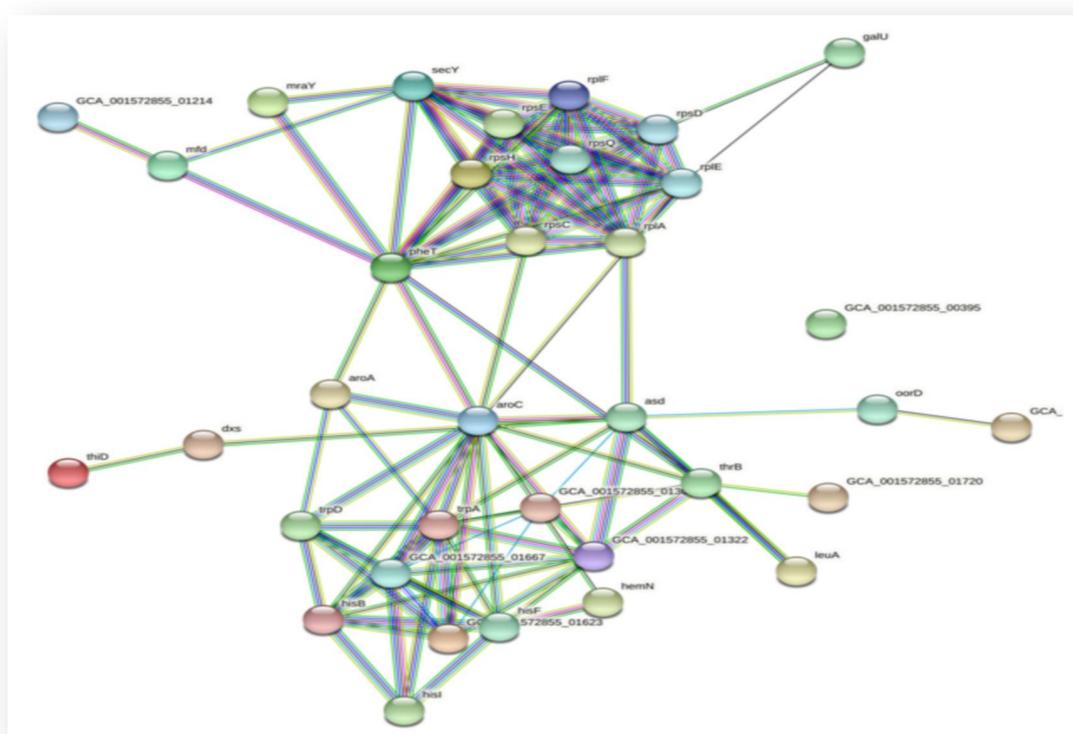


Figure 5. Interaction of essential non host homologous proteins (<https://string-db.org>)

### 3.8 Modelome Prediction

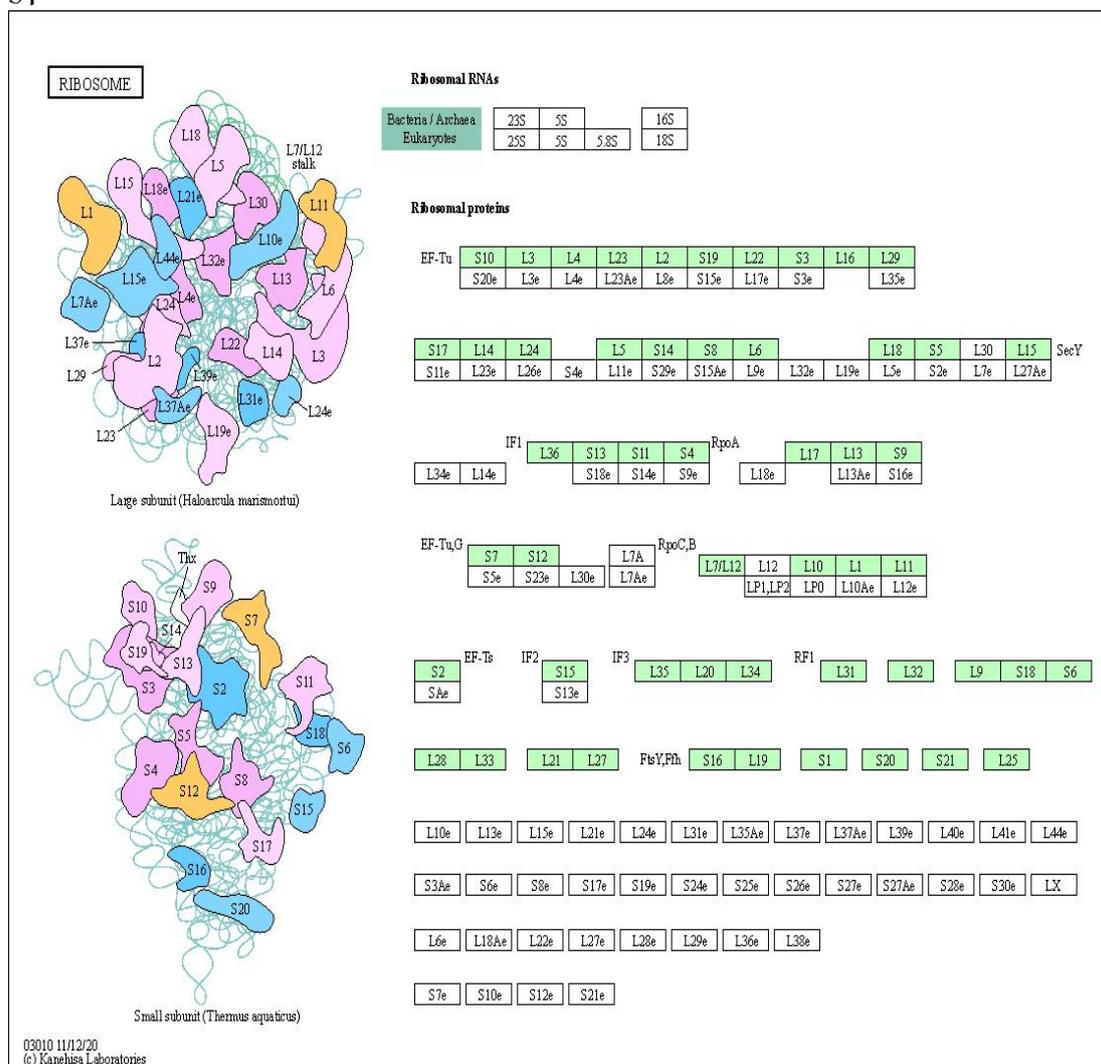
The focus of this study was on finding candidate target vaccines. The identified 22 conservative non-homologous *Arcobacter cryaerophilus* proteins were transferred to the online tool MHOLline for model prediction. The results obtained were carefully examined and among them, the proteins belonging to the very high fasta group were selected in order to obtain the

best accurate results, which is the fundamental need for drugable targets. With the help of this step, a list of 8 proteins was extracted which was then examined for further results.

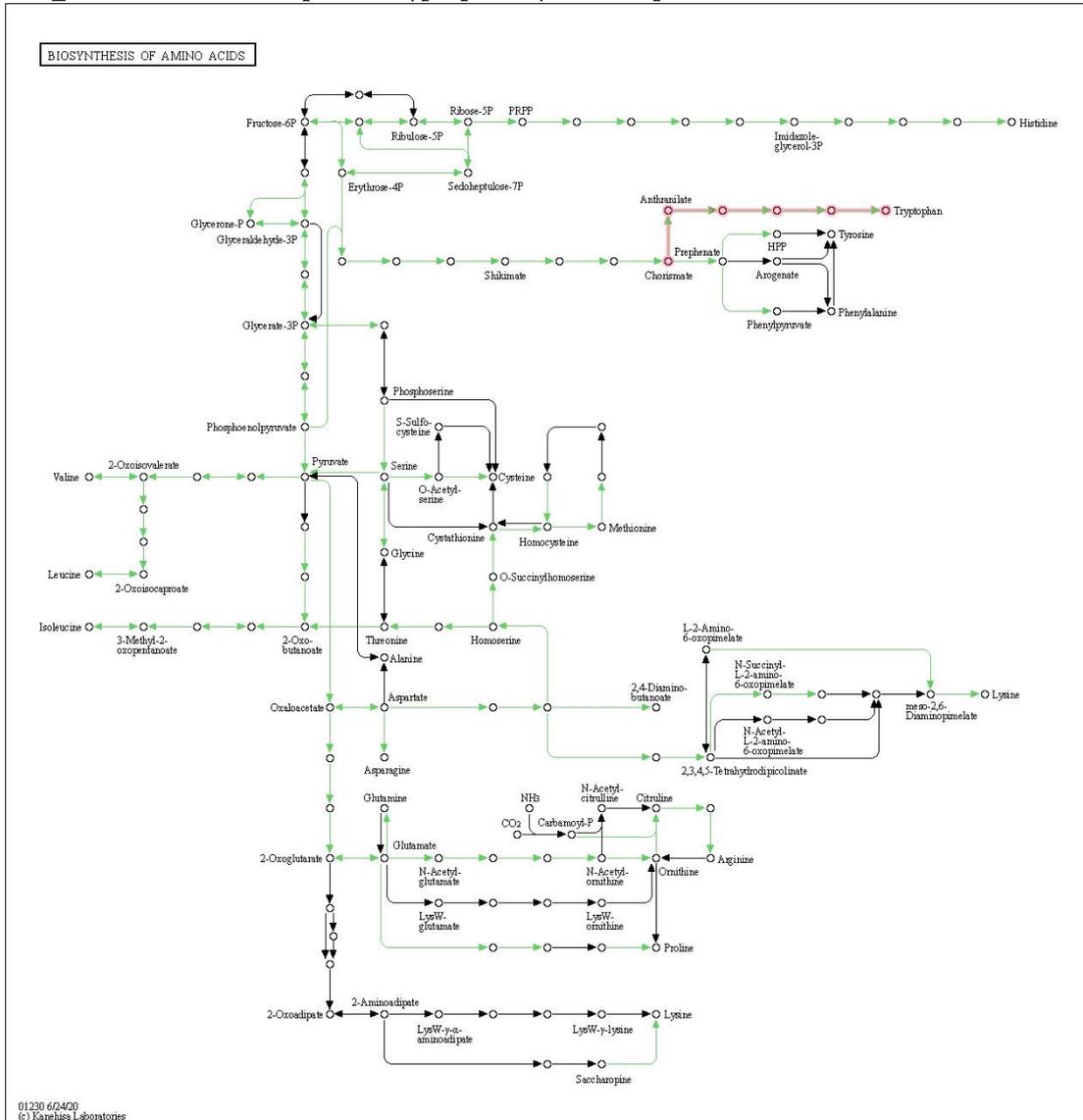
### 3.9 Pathogen Specific Metabolic Pathways

Using computational, comparative and subtractive genomic analyzes of various *A. cryaerophilus* metabolic pathways, a list of possible drug and vaccine targets was identified, the aim being to obtain information about proteins involved in various *A. cryaerophilus* metabolic pathways, but in theirs Host absent in order to minimize possible side effects. The various metabolic pathways involved in the pathogen were taken from the KEGG database. By the help of this step 6 potential targets were obtained and rest of proteins were filtered because in similar pathways.

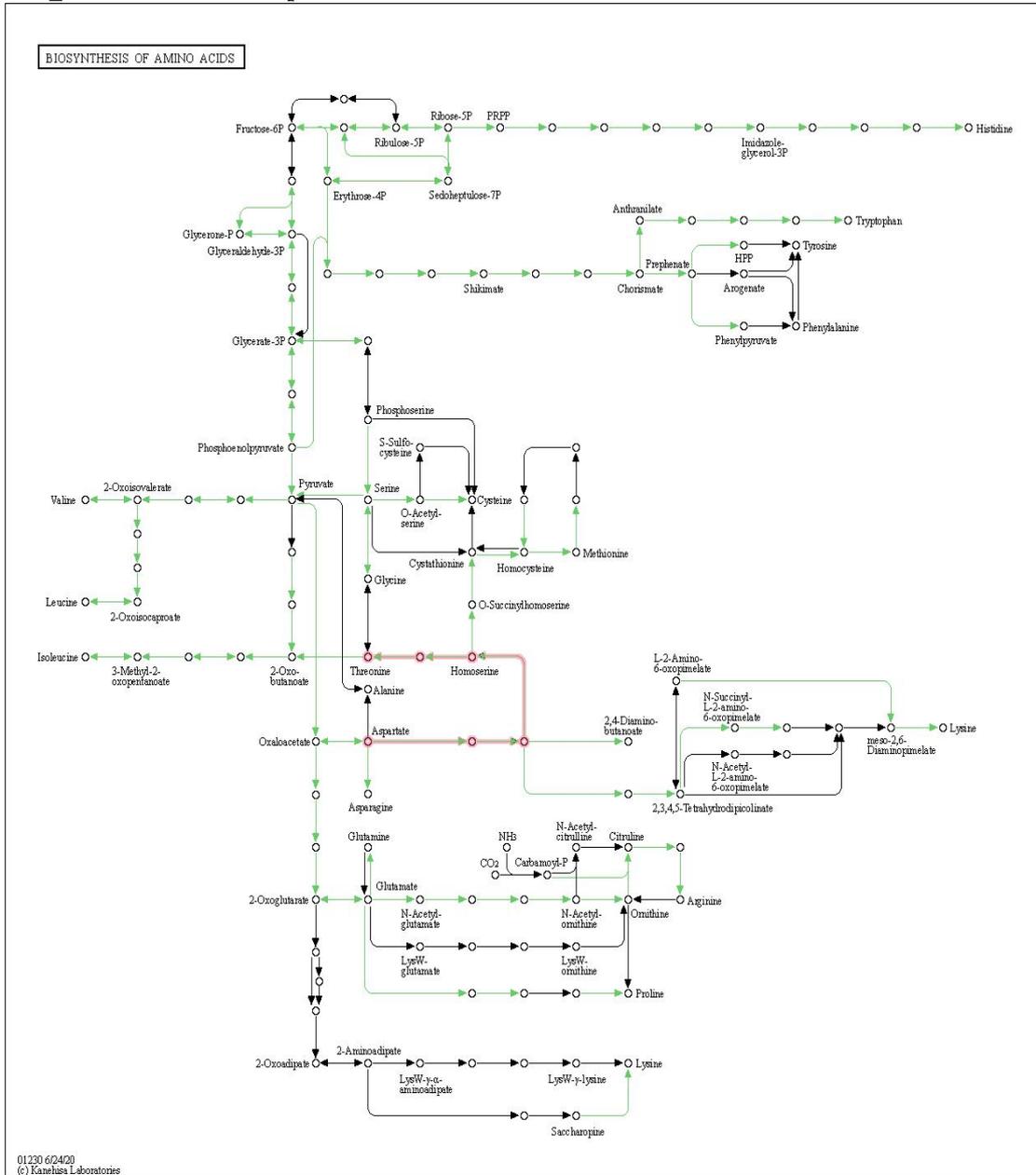
WP\_066152584.1      **MULTISPECIES:**      30S      ribosomal      protein  
S4



WP\_066153395.1 Multispecies Tryptophan synthase alpha chain

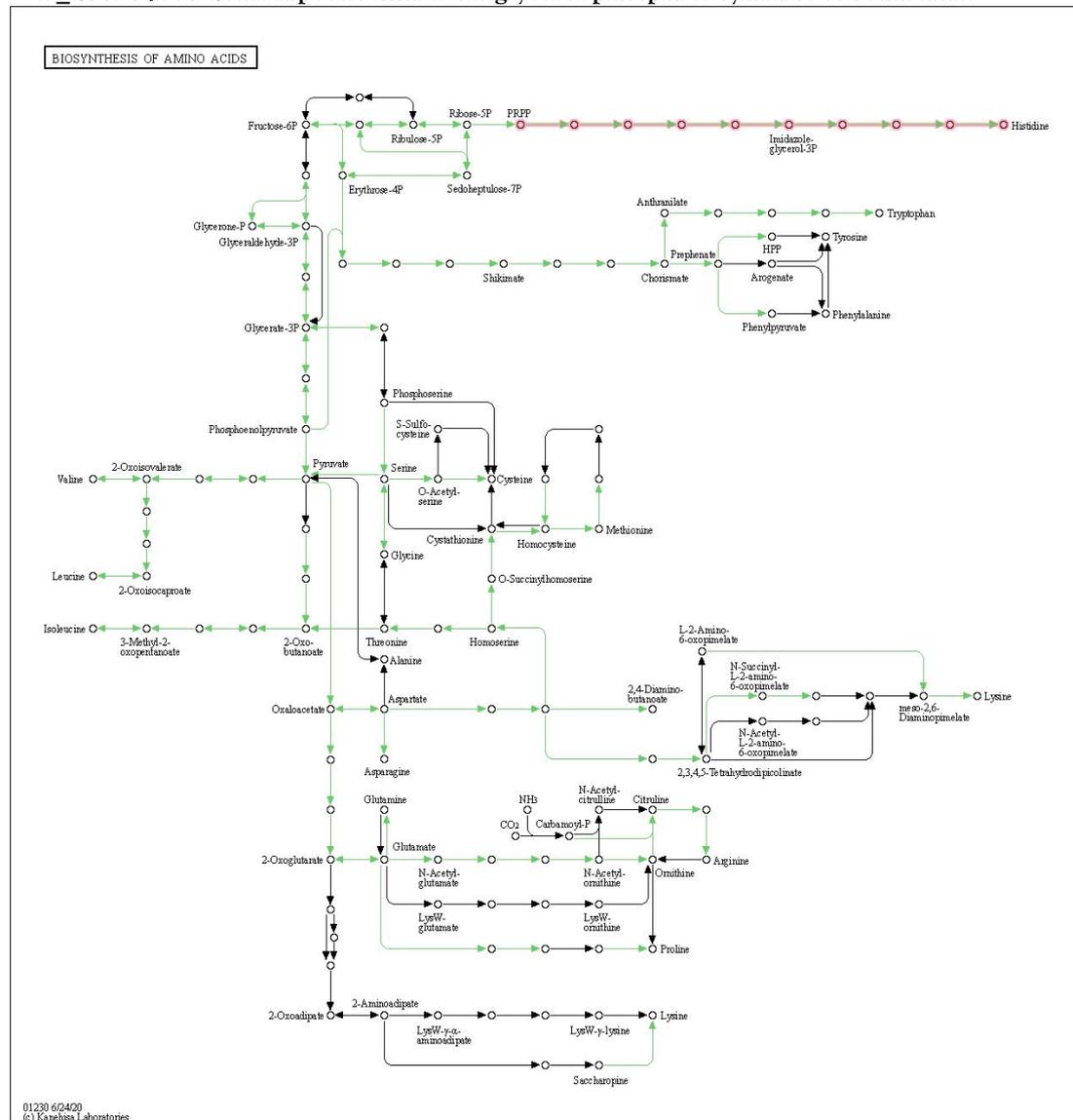


WP\_066156502.1 Multispecies Homoserine kinase

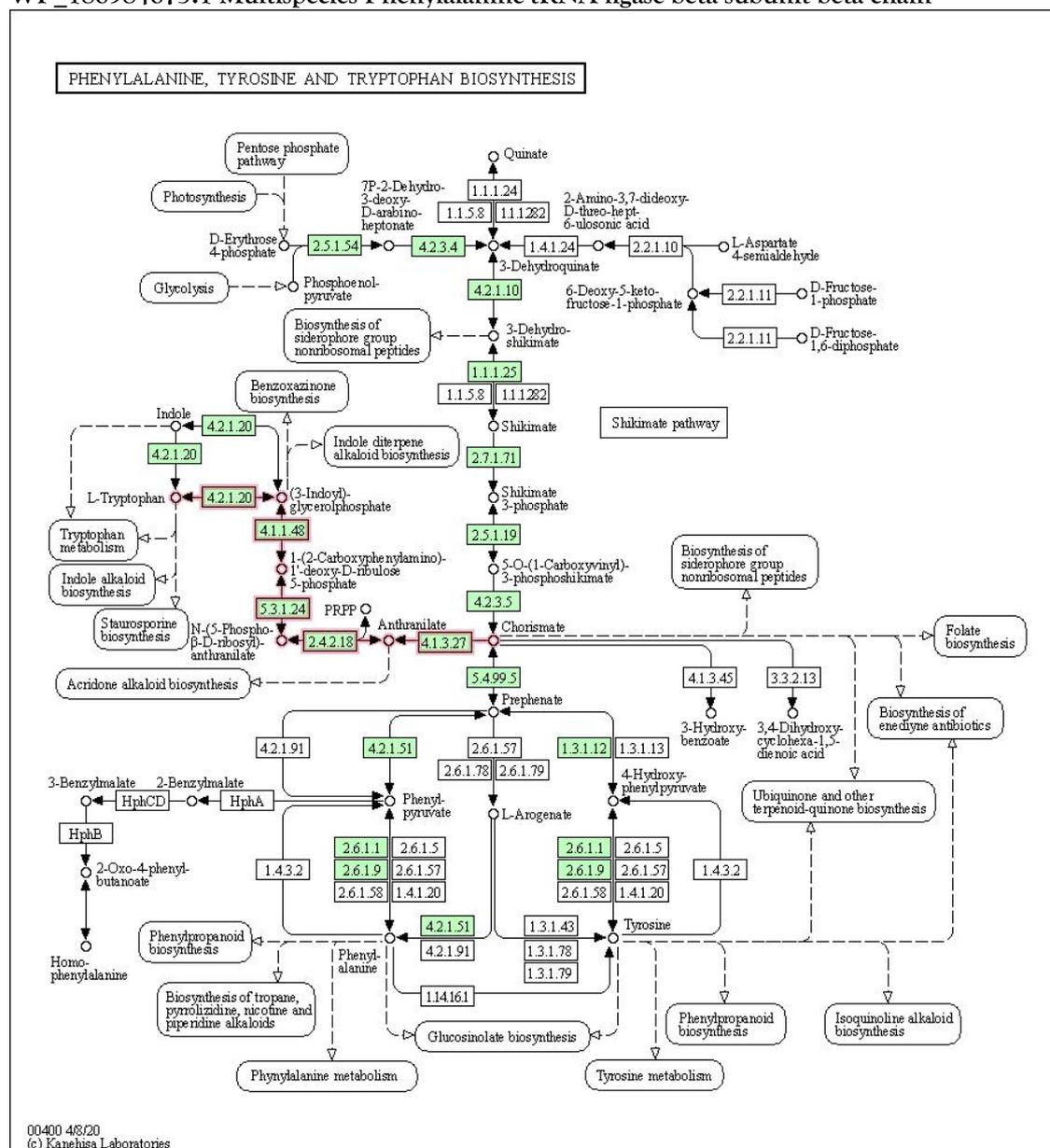




WP\_186984668.1 Multispecies Imidazole glycerol phosphate synthase subunit HisF

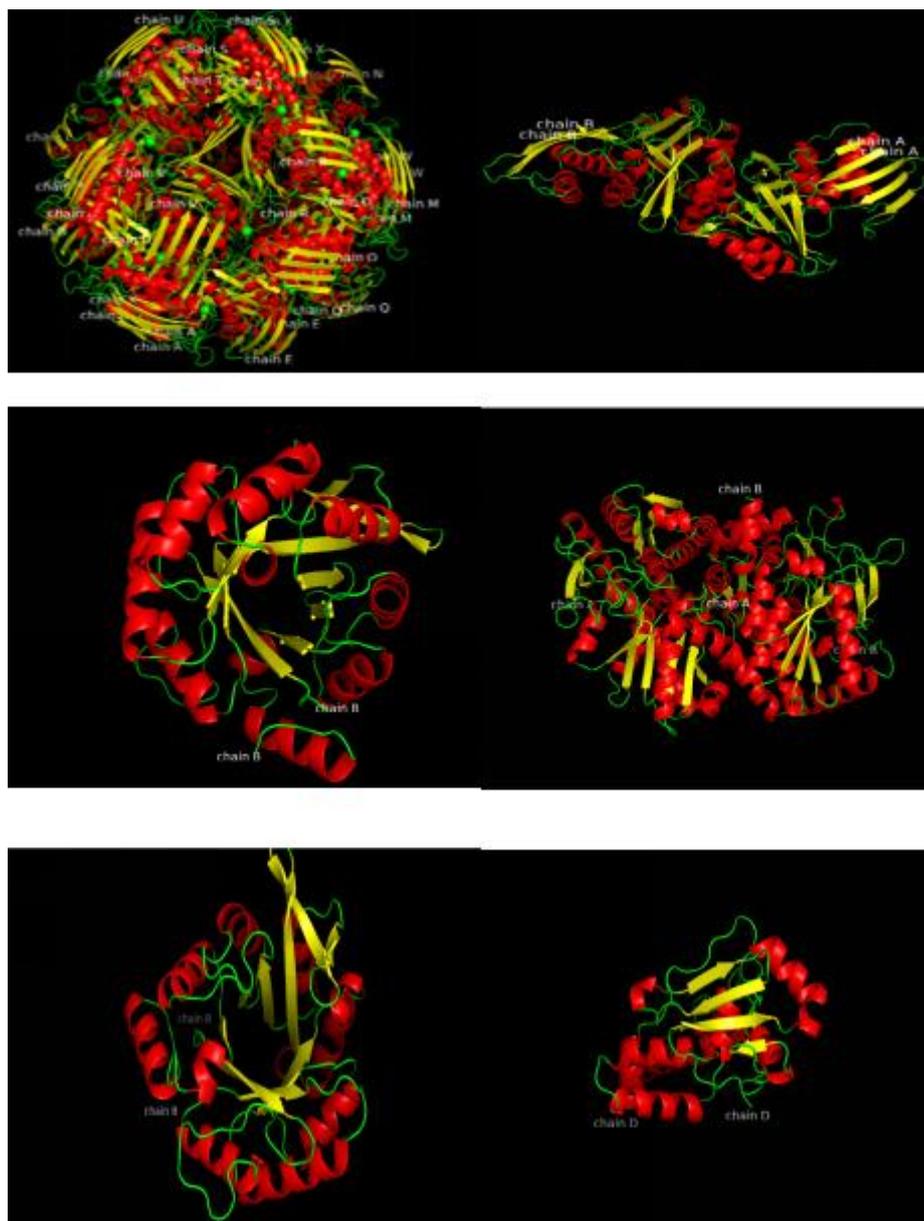


WP\_186984675.1 Multispecies Phenylalanine-tRNA ligase beta subunit beta chain



### 3.10 Prediction of 3D Structures of Target Proteins

The results were then submitted to the Swiss model to detect the three-dimensional (3D) structures of the protein. Here we have uploaded some good quality fasta file and submitted the data. After uploading the data, we selected the structure assessment option and then uploaded the PDB format files of the proteins with Ramachandran score above 92% as well as the Ramachandran plots, which is important for identify the quality of the protein structure. Out of 8, we obtained 6 protein structures that have Ramachandran scores greater than 92% using the server (<http://swissmodel.expasy.org>). These files in PDB format were then opened in PyMOL software then displayed the 3D protein. Next comes the verification phase of validation of some protein models, for this process SAVES v6.0 (<http://saves.mbi.ucla.edu>) was used which provided confirmation of the required protein models. Other valid protein models were visualized using an efficient and reliable tool called PYMOL, which resulted in the creation of the required 3D models of the required drug targets.



**Figure 7. Three dimensional structures (3D) of high-quality sequences**

- (a) Tryptophan synthase alpha chain
- (b) Homoserine kinase
- (c) Anthranilate phosphoribosyl transferase
- (d) Tryptophan synthase beta chain
- (e) Imidazole glycerol phosphate synthase subunit His F
- (f) Phenylalanine-tRNA ligase beta subunit beta chain

### 3.11 Prioritization Parameters for Drug Targets and Vaccine Candidates

During the reverse Vaccinology process, we considered all *A. cryaerophilus* genomic sequences. Next, we analyzed the genes conserved between the different genomes that are essential to the pathogen and non-host counterparts. We then identified six proteins that were non-host homologous, essential and with good quality 3D structures. These could be new therapeutic targets for *A. cryaerophilus*, which were then subjected to further analysis.

### 3.12 Molecular Weight Determination

In this step, we further characterized the target genomes based on their molecular weight. After cellular localization, these proteins were subjected to a molecular weight search calculator, i.e.

[https://www.bioinformatics.org/sms/prot\\_mw.html](https://www.bioinformatics.org/sms/prot_mw.html) to prioritize target proteins based on their molecular weight and the following results were obtained, shown in Table

**Table 4: Characterization of Proteins Based on their Molecular Weight**

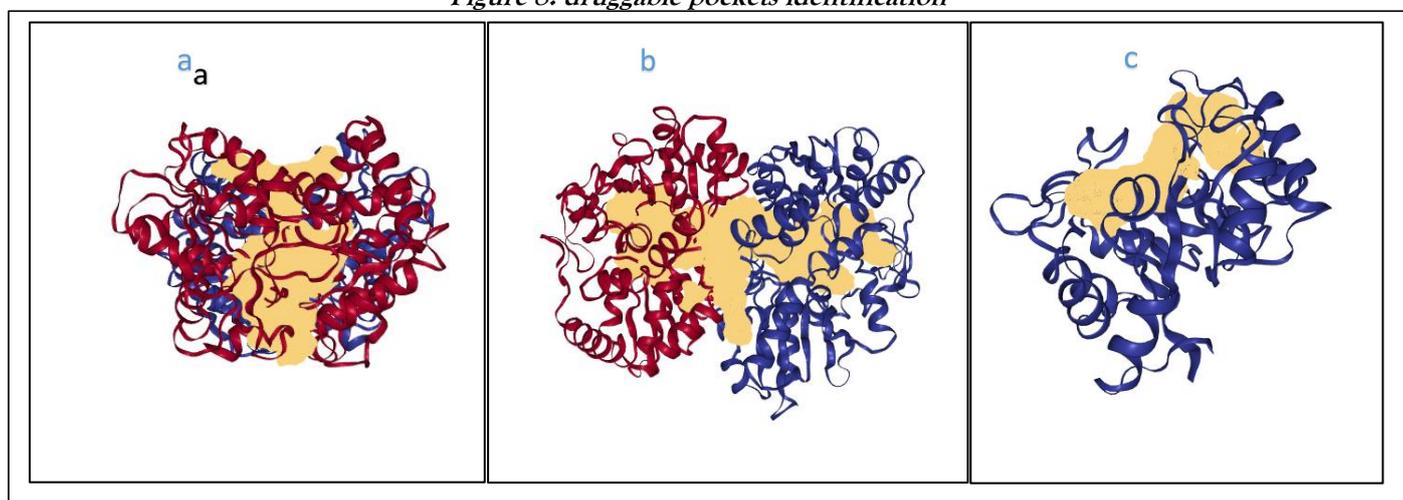
S. No.	Accession Number	Protein name	Molecular Weight
1	WP_066153395.1	Tryptophan synthase alpha chain	27633.76
2	WP_066156502.1	Homoserine kinase	32757.65
3	WP_164469403.1	Anthranilate phosphoribosyltransferase	35588.95
4	WP_186984447.1	Tryptophan synthase beta chain	44931.30
5	WP_186984668.1	Imidazole glycerol phosphate synthase subunit HisF	27188.03
6	WP_186984675.1	Phenylalanine-tRNA ligase beta subunit beta chain	43873.16

### 3.13 Identification of Druggable Pockets and Druggability Analysis

Information obtained from 3D structures and Druggability analyzes are important features for prioritizing and authenticating putative pathogenic targets. As mentioned above, for Druggability analyses, the final list of good quality non-host essential protein targets was submitted to DogSite Scorer in PDB format. There were 5 good quality sequences, but only 3 were highly drugged. Subsequently, the Target Pathogen database was used to analyze Druggability and other biochemical functions.

For these three non-host pathogenic homologs with a score >0.80, the predicted cavity numbers with the respective Druggability scores were as follows:

*Figure 8. druggable pockets identification*



S. No.	Query id	Protein name	Gene name	Volume	Surface area	Drug score	Total druggable cavities	Cavities DS > 0.080	Ramachandran scores
1	WP_186984447.1	Tryptophan synthase beta chain	trpB	5100.34	4766.73	0.8	5	1	96.71%
2	WP_186984668.1	Imidazole glycerol phosphate synthase subunit HisF	hisF	1124.74	1418.25	0.81	23	14	96.79%

3	WP_164 469403.1	Anthranilate phosphoribosyl transferase	trpD	914.86	1065.85	0.82	24	2	94.75%
---	--------------------	--	------	--------	---------	------	----	---	--------

**Table 5. Druggability scores**

**3.14 Virtual Screening and Molecular Docking**

Molecular docking is the concept which deals with the interaction of two or more molecular structures (e.g. drug and enzyme or protein) into each other. simply docking is a molecular modeling technique that is used to predict how a protein (enzyme) interacts with small molecules (ligands). This process involves two basic steps, prediction of the ligand conformation as well as its position and orientation within these sites (usually referred to as pose) and assessment of the binding affinity. For this process most important protein targets of our study were selected and its interaction with top five compounds of our druggable library were selected as ideal drugs as because they had a high drug score and were less entropic values; The results are listed below in table

**Table 6: Ligand Predication**

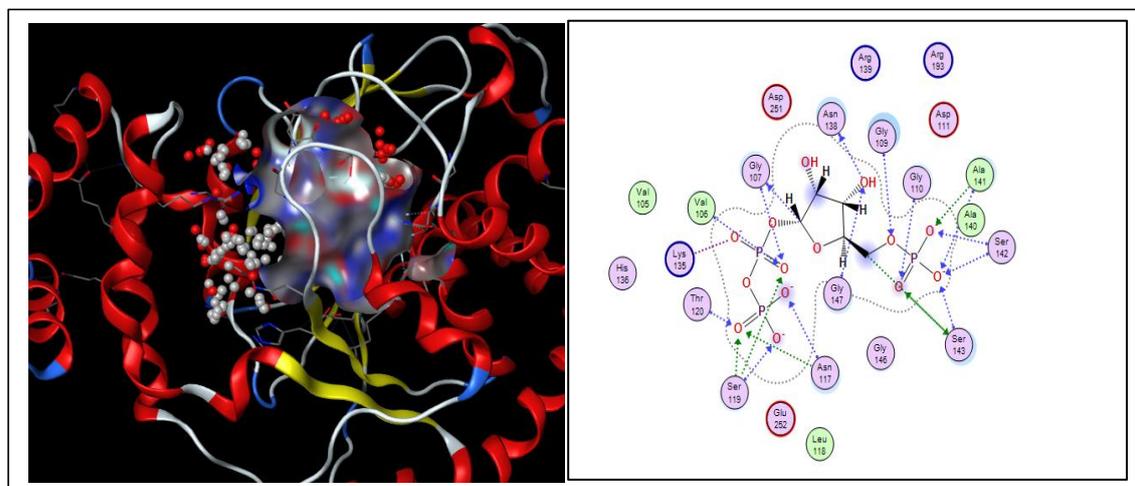
Protein id	Protein name	Template Selected	Ligand	Possible hits	Ligand Interaction
WP_186984447.1	Tryptophan synthase beta chain	1C29	HE1	172	Asp 60, Arg179, Thr183, Gly184, Phe212, Gly213, Gly234, Ser235, Glu350
WP_164469403.1	Anthranilate phosphoribosyl transferase	3TWP	PRP	249	Gly, Ser, Ala, Arg, Asn,

**Table 7: Docking Score And Top 10 Docked Compounds**

**A. Anthranilate phosphoribosyl transferase**

S. No.	Compound molecular formula	Docking sore	Rmsd value
1	<chem>O=C(Nc1nc(C(=O)Nc2nc(c(=O)O)n(C)c2)n(C)c1)C</chem>	-7.57	1.15
2	<chem>O=C(C(C#N)=C([O-])NC)C1CN(C(=O)C1)C1CCCCC1</chem>	-7.38	1.17
3	<chem>P(=O)(OC(C)C)OC(C)C1N(C(=O)O)CC(O)C1</chem>	-7.26	1.73
4	<chem>O=C(Nc1nc(C(=O)Nc2nc(C(=O)O)n(C)c2)n(C)c1)C</chem>	-7.19	1.02
5	<chem>O=C(OCc1cccc1)N1NC(C(=O)O)CCC1</chem>	-7.12	2.20
6	<chem>O=C(Nc1nc(C(=O)Nc2nc(C(=O)O)n(C)c2)n(C)c1)C</chem>	-7.10	1.01
7	<chem>S(C(=N)N)C1=C(C(=O)CC)C([O-])=C(C(=O)OC)C(C)C1</chem>	-7.04	3.54
8	<chem>O=C([O-])CN(C)c1sc2c(C)nn(c3cccc3)c2n1</chem>	-6.96	3.05
9	<chem>O=C([O-])c1[nH]nc(N=NC=2C(=O)N(c3cccc3)NC=2C)n1</chem>	-6.91	2.51
10	<chem>FC1ccc(N2C(=O)C=CC(OC(C(=O)O)CC)N2)cc1</chem>	-6.90	1.67

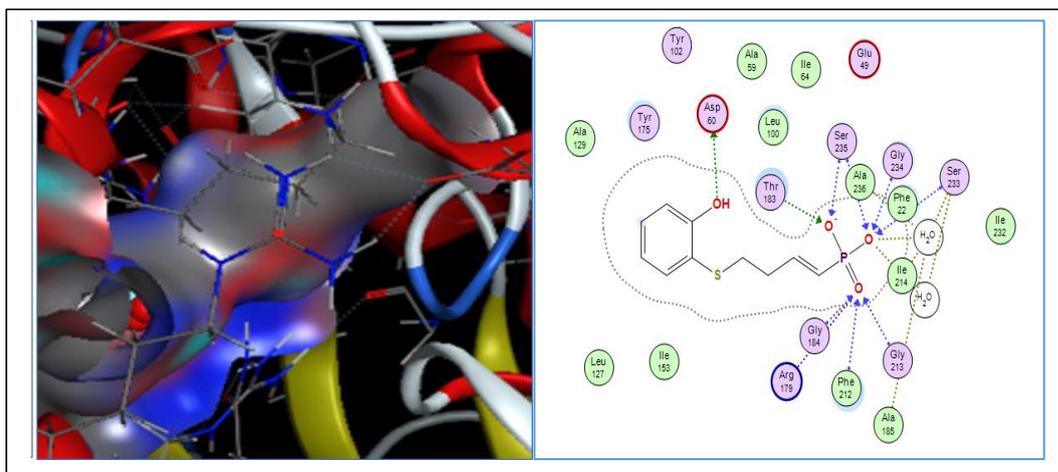
Figure 9. 2D and 3D images of final target proteins and their identified novel inhibitors  
1- Showing 2D and 3D images of top 1 drug for Anthranilate phosphoribosyl transferase



2) Tryptophan synthase beta chain

S. No.	Compound molecular formula	Docking score	RMSD value
1	<chem>O=C([O-])CC1CN(C2=CC(=O)N(CCOC)N=C2)CCC1</chem>	-8.23	1.10
2	<chem>O=C([O-])CC1CN(C2=CC(=O)N(CC3CC3)N=C2)CCC1</chem>	-7.75	1.12
3	<chem>O=C(Nc1cc2c(C(=O)[O-])cc(N3CC[N+H](CC(=O)NCCC)CC3)nc2cc1)C(C)C</chem>	-7.63	1.95
4	<chem>Fc1c(C[N+H]2CCN(c3nc4c(c(C(=O)[O-])c3)cc(NC(=O)C(C)C)cc4)CC2)cccc1</chem>	-7.16	1.44
5	<chem>O=C([O-])c1cc(CNC(=O)c2c(C3CC[N+H](Cc4cccc4)CC3)nc(C)nc2)ccc1</chem>	-7.15	2.19
6	<chem>O=C(Nc1cc2c(C(=O)[O-])cc(N3CC[N+H](Cc4c(OC)cccc4)CC3)nc2cc1)CC(C)C</chem>	-7.10	3.46
7	<chem>O=C(Nc1cc2c(C(=O)[O-])cc(N3CC[N+H](CC(=O)NC(C)C)CC3)nc2cc1)CCC</chem>	-7.10	2.13
8	<chem>O=C(NCCCC)C[N+H]1CCN(c2cc(C(=O)[O-])c(NC(=O)C3CCC3)cc2)CC1</chem>	-7.06	2.31
9	<chem>O=C(C(C)C)N1CCC2(C(=O)N(C)c3c2cc(C(=O)NCc2cc(C(=O)[O-])ccc2)cc3)CC1</chem>	-7.06	0.59
10	<chem>O=C(Nc1cc2c(C(=O)[O-])cc(N3CC[N+H](C4CCCC4)CC3)nc2cc1)C(C)(C)C</chem>	-7.04	1.67

## 2- Showing 2D and 3D images of top 1 drug for Tryptophan synthase beta chain



### USION

Bacterial resistance to antibiotics is a major concern today. The bacterial machinery has the ability to copy any environment and condition through its efficient cellular mechanism, simple genomics, and different variations in biochemical processes. The increase in the number of antibiotics as well as their high use against infections has caused bacteria to modify their cellular machinery and their biochemical processes in such a way that they are resistant to antibiotics, which is of great concern today. Researchers around the world are busy and working hard to gain knowledge not only about the causes of bacterial resistance, but also to find new therapeutic targets in the bacterial machinery that would lead to the cure of serious bacterial infections. In this sense, IT tools are of greater importance today. This tool has not only made the job easy but also quick which is needed nowadays. Similarly, in this work, various online tools were used for data collection and to acquire important information about *Arcobacter cryaerophilus*, including the Genome Online Database (GOLD) for data retrieval and database. PATRIC data to obtain a single proteome file. The NCBI (Blast P) was used for comparative genomics which provided important information on the target's homology with humans. The DEG online database was then used, which provided an approximation of the minimum essential proteins that can be used as therapeutic targets. Since all proteins are interconnected with each other by single or multiple paths, an online database was used to find out its STRING interrelationship. The cellular localization of proteins provides essential information about the function and role of this protein, which is why CELLO2GO was used for this purpose. In addition, MHOLLINE was used, which is an important software regarding the essentiality of drug targets for bacterial cells. Among all these PROTPARM, SAVESVS6.0 and KEGG were also used for the assembly of essential data on our targets. For structural annotation of drug able targets SWISSMODEL was used which accounted for 3D structures. This is because a good quality protein 3D structure is essential for giving information about function and interaction with possible ligands. The DoGSiteScorer performed identification of drug able cavities in any target protein, the larger the druggability score ( $> 0.80$ ), were considered ideal for drug target. In this study an approach is made to not only highlight possible targets for drugs in *Arcobacter cryaerophilus* but also an effort is made to know its potency to infect human body. Also, these filtered targets will account for future in vitro and in vivo experimentation for drug as well as vaccine therapies, which is need of the day.

### 5. REFERENCES

1. Giacometti, F.; Lucchi, A.; Di Francesco, A.; Delogu, M.; Grilli, E.; Guarniero, I.; Stancampiano, L.; Manfreda, G.; Merialdi, G.; Serraino, A. *Arcobacter butzleri*, *Arcobacter cryaerophilus*, and *Arcobacter skirrowii* circulation in a dairy farm and sources of milk contamination. *Appl. Environ. Microbiol.* 2015, 81, 5055–5063.

2. Lehner, A.; Tasara, T.; Stephan, R. Relevant aspects of *Arcobacter* spp. as potential foodborne pathogen. *Int. J. Food Microbiol.* 2005, 102, 127–135. [PubMed]
3. Vandamme, P.; De Ley, J. Proposal for a New Family, Campylobacteraceae. *Int. J. Syst. Bacteriol.* 1991, 41, 451–455.
4. Fernandez, H., Villanueva, M. P., Mansilla, I., Gonzalez, M., & Latif, F. (2015). *Arcobacter butzleri* and *A. cryaerophilus* in human, animals and food sources, in southern Chile. *Brazilian Journal of Microbiology*, 46(1), 145–147. doi:10.1590/s1517-838246120140095
5. Ramees, T.P.; Dhama, K.; Karthik, K.; Rathore, R.S.; Kumar, A.; Saminathan, M.; Tiwari, R.; Malik, Y.S.; Singh, R.K. *Arcobacter*: An emerging food-borne zoonotic pathogen, its public health concerns and advances in diagnosis and control-A comprehensive review. *Vet. Q.* 2017, 37, 136–161.
6. Banting, G.; Figueras Salvat, M.J. *Arcobacter*. In *Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project), Part 3: Specific Excreted Pathogens: Environmental and Epidemiology Aspects–Section 2: Bacteria*; Rose, J.B., Jiménez-Cisneros, B., Pruden, A., Ashbolt, N., Miller, J., Eds.; Michigan State University Press: East Lansing, MI, USA, 2017.
7. ICMSF Microorganisms in Food. Microbiological testing in food safety management. International Commission on Microbiological Specifications for Foods. In *Microorganisms in Foods 7*; Kluwer Academic Publishing: New York, NY, USA; Springer International Publishing: Cham, Switzerland, 2002; pp. 1–30.
8. Collado, L.; Figueras, M.J. Taxonomy, epidemiology, and clinical relevance of the genus *Arcobacter*. *Clin. Microbiol. Rev.* 2011, 24, 174–192.
9. Kim, H.M.; Hwang, C.Y.; Cho, B.C. *Arcobacter marinus* sp. nov. *Int. J. Syst. Evol. Microbiol.* 2010, 60, 531–536.
10. Levican, A.; Collado, L.; Figueras, M.J. *Arcobacter cloacae* sp. nov. and *Arcobacter suis* sp. nov., two new species isolated from food and sewage. *Syst. Appl. Microbiol.* 2013, 36, 22–27.
11. Park, S.; Jung, Y.T.; Kim, S.; Yoon, J.H. *Arcobacter acticola* sp. nov., isolated from seawater on the East Sea in South Korea. *J. Microbiol.* 2016, 54, 655–659. [PubMed]
12. Jalava, K.; Rintala, H.; Ollgren, J.; Maunula, L.; Gomez-Alvarez, V.; Revez, J.; Palander, M.; Antikainen, J.; Kauppinen, A.; Räsänen, P.; et al. Novel Microbiological and Spatial Statistical Methods to Improve Strength of Epidemiological Evidence in a Community-Wide Waterborne Outbreak. *PLoS ONE* 2014, 9, e104713. [PubMed]
13. Talay, F.; Molva, C.; Atabay, H.I. Isolation and identification of *Arcobacter* species from environmental and drinking water samples. *Folia Microbiol.* 2016, 61, 479–484. [PubMed]
14. Merga, J.Y.; Royden, A.; Pandey, A.K.; Williams, N.J. *Arcobacter* spp. isolated from untreated domestic effluent. *Lett. Appl. Microbiol.* 2014, 59, 122–126. [PubMed]
15. Lee, C.; Agidi, S.; Marion, J.W.; Lee, J. *Arcobacter* in Lake Erie beach waters: An emerging gastrointestinal pathogen linked with human-associated fecal contamination. *Appl. Environ. Microbiol.* 2012, 78, 5511–5519.
16. Fong, T.T.; Mansfield, L.S.; Wilson, D.L.; Schwab, D.J.; Molloy, S.L.; Rose, J.B. Massive microbiological groundwater contamination associated with a waterborne outbreak in Lake Erie, South Bass Island, Ohio. *Environ. Health Perspect.* 2007, 115, 856–864. [PubMed]
17. Phillips, C.A. *Arcobacters* as emerging human foodborne pathogens. *Food Control* 2001, 12, 1–6.
18. Wesley, I.V.; Baetz, A.L.; Larson, D.J. Infection of cesarean-derived colostrum-deprived 1-day-old piglets with *Arcobacter butzleri*, *Arcobacter cryaerophilus*, and *Arcobacter skirrowii*. *Infect. Immun.* 1996, 64, 2295–2299. [PubMed]

19. Van Driessche, E.; Houf, K.; Vangroenweghe, F.; Nollet, N.; De Zutter, L.; Vandamme, P.; Van Hoof, J. Occurrence and strain diversity of *Arcobacter* species isolated from healthy Belgian pigs. *Res. Microbiol.* 2004, 155, 662–666.
20. Corry, J.E.L.; Atabay, H.I. Poultry as a source of *Campylobacter* and related organisms. *J. Appl. Microbiol.* 2001, 90, 96S–114S.
21. Houf, K.; On, S.L.W.; Coenye, T.; Mast, J.; Van Hoof, J.; Vandamme, P. *Arcobacter cibarius* sp. nov., isolated from broiler carcasses. *Int. J. Syst. Evol. Microbiol.* 2005, 55, 713–717.
22. Eifert, J.; Castle, R.; Pierson, F.; Larsen, C.; Hackney, C. Comparison of sampling techniques for detection of *Arcobacter butzleri* from chickens. *Poult. Sci.* 2003, 82, 1898–1902.
23. Kabeya, H.; Maruyama, S.; Morita, Y.; Kubo, M.; Yamamoto, K.; Arai, S.; Izumi, T.; Kobayashi, Y.; Katsube, Y.; Mikami, T. Distribution of *Arcobacter* species among livestock in Japan. *Vet. Microbiol.* 2003, 93, 153–158.
24. Rivas, L.; Fegan, N.; Vanderlinde, P. Isolation and characterisation of *Arcobacter butzleri* from meat. *Int. J. Food Microbiol.* 2004, 91, 31–41.
25. Serraino, A.; Giacometti, F.; Daminelli, P.; Losio, M.N.; Finazzi, G.; Marchetti, G.; Zambrini, A.V.; Rosmini, R. Survival of *Arcobacter butzleri* during production and storage of artisan water Buffalo Mozzarella Cheese. *Foodborne Pathog. Dis.* 2013, 10, 820–824.
26. Mottola, A.; Bonerba, E.; Bozzo, G.; Marchetti, P.; Celano, G.V.; Colao, V.; Terio, V.; Tantillo, G.; Figueras, M.J.; Di Pinto, A. Occurrence of emerging food-borne pathogenic *Arcobacter* spp. isolated from pre-cut (ready-to-eat) vegetables. *Int. J. Food Microbiol.* 2016, 236, 33–37. *Foods* 2021, 10, 1673 16 of 16
27. Leoni, F.; Chierichetti, S.; Santarelli, S.; Talevi, G.; Masini, L.; Bartolini, C.; Rocchegiani, E.; Naceur Haouet, M.; Ottaviani, D. Occurrence of *Arcobacter* spp. and correlation with the bacterial indicator of faecal contamination *Escherichia coli* in bivalve molluscs from the Central Adriatic, Italy. *Int. J. Food Microbiol.* 2017, 245, 6–12.
28. Mottola, A.; Bonerba, E.; Figueras, M.J.; Pérez-Cataluña, A.; Marchetti, P.; Serraino, A.; Bozzo, G.; Terio, V.; Tantillo, G.; Di Pinto, A. Occurrence of potentially pathogenic *Arcobacters* in shellfish. *Food Microbiol.* 2016, 57, 23–27. [PubMed]
29. Andersen, M.M.E.; Wesley, I.V.; Nestor, E.; Trampel, D.W. Prevalence of *Arcobacter* species in market-weight commercial turkeys. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 2007, 92, 309–317. [PubMed]
30. Rasmussen, L.H.; Kjeldgaard, J.; Christensen, J.P.; Ingmer, H. Multilocus sequence typing and biocide tolerance of *Arcobacter butzleri* from Danish broiler carcasses. *BMC Res. Notes* 2013, 6, 322.
31. Giacometti, F.; Lucchi, A.; Manfreda, G.; Florio, D.; Zanoni, R.G.; Serraino, A. Occurrence and genetic diversity of *Arcobacter butzleri* in an artisanal dairy plant in Italy. *Appl. Environ. Microbiol.* 2013, 79, 6665–6669.
32. Serraino, A.; Florio, D.; Giacometti, F.; Piva, S.; Mion, D.; Zanoni, R.G. Presence of *Campylobacter* and *Arcobacter* species in in-line milk filters of farms authorized to produce and sell raw milk and of a water buffalo dairy farm in Italy. *J. Dairy Sci.* 2013, 96, 2801–2807.
33. Serraino, A.; Giacometti, F. Short communication: Occurrence of *Arcobacter* species in industrial dairy plants. *J. Dairy Sci.* 2014, 97, 2061–2065.
34. Hausdorf, L.; Neumann, M.; Bergmann, I.; Sobiella, K.; Mundt, K.; Fröhling, A.; Schlüter, O.; Klocke, M. Occurrence and genetic diversity of *Arcobacter* spp. in a spinach-processing plant and evaluation of two *Arcobacter*-specific quantitative PCR assays. *Syst. Appl. Microbiol.* 2013, 36, 235–243.
35. Wang H, Wang H, Yu X, Zhou H, Li B, Chen G, et al. Impact of antimicrobial

- stewardship managed by clinical pharmacists on antibiotic use and drug resistance in a Chinese hospital, 2010-2016: A retrospective observational study. *BMJ Open*. 2019 Aug;9(8):e026072.
36. S.B. Jamal, S.S. Hassan, S. Tiwari, M. V. Viana, L. de J. Benevides, A. Ullah, A.G. Turjanski, D. Barh, P. Ghosh, D.A. Costa, A. Silva, R. Röttger, J. Baumbach, V.A.C. Azevedo, An integrative in-silico approach for therapeutic target identification in the human pathogen *Corynebacterium diphtheriae*, *PLoS One*. 12 (2017) e0186401. <https://doi.org/10.1371/journal.pone.0186401>.
  37. Identify human non-homologous protein by blastp?, (n.d.). [https://www.researchgate.net/post/Identify\\_human\\_non-homologous\\_protein\\_by\\_blastp](https://www.researchgate.net/post/Identify_human_non-homologous_protein_by_blastp) (accessed October 1, 2020).
  38. R. Zhang, H.Y. Ou, C.T. Zhang, DEG: A database of essential genes, *Nucleic Acids Res*. 32 (2004). <https://doi.org/10.1093/nar/gkh024>.
  39. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against *Acinetobacter baumannii*.
  40. Hu Y, Lehrach H, Janitz M. Comparative analysis of an experimental subcellular protein localization assay and in silico prediction methods. *J Mol Histol*. 2009 Oct;40(5-6):343-52.
  41. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. 2009 Jun;30(SUPPL. 1).
  42. Shah, A.H.; Saleha, A.A.; Zunita, Z.; Murugaiyah, M.; Aliyu, A.B.; Jafri, N. Prevalence, Distribution and Antibiotic Resistance of Emergent *Arcobacter* spp. from Clinically Healthy Cattle and Goats. *Transbound. Emerg. Di*
  43. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. 2009 Jun;30(SUPPL. 1).
  44. Stierand K, Rarey M. Consistent two-dimensional visualization of protein-ligand complex series. *J Cheminform*. 2011 Jun;3(6):21.
  45. Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*. 2000;28(6).