doi.org/10.36721/PJPS.2021.34.4.REG.1359-1367.1

Comparative proteome-wide study for *in-silico* identification and characterization of indispensable hypothetical proteins of food borne-pathogen *Campylobacter jejuni* (CJJ) by subtractive genomics approach

Reaz Uddin[1]*, Alina Arif[1], Noor-ul-Ain Zahra[1] and Muhammad Sufian[2]
[1]Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan
[2]Department of Biotechnology, Mirpur University of Science and Technology (MUST)

Abstract: *Campylobacter jejuni* (CJJ) is a source of bacterial foodborne diarrhea globally. Mostly found prevalent in children in the developing countries that may lead to mortality. The upsurge in antimicrobial resistance is causing hindrance in the treatment, as highlighted by CDC and WHO. The study hypothesized the application of subtractive genomics approach coupled with metabolic pathway to reveal unidentified essential proteins that could serve as potential drug target (s). The approach was employed to model the druggable proteome of *C. jejuni* resistant strain 81-176. We obtained 728/1744 non-homologous essential proteins by performing sequence similarity search against host proteome and DEG server, respectively. The KAAS annotated metabolic pathway information; PSORTb predicted their sub cellular localization and SVMPro functional annotated 104 hypothetical proteins while the Drug Bank for the druggability analysis. We found 04/104 protein druggable viz. synaptic vesicular amine transporter, Uracil-DNA glycosylase, Laccase domain protein YfiH, and Phosphoenolpyruvate protein phosphor transferase. The study has revealed a formerly uncharacterized pool of *C. jejuni* proteins that can play a significant role in controlling CJJ infection and presented previously uncharacterized four proteins as potential drug targets. These potential drug targets can further be explored employing structure-based and other biochemical methods by the scientific community.

Keywords: Uncharacterized proteins, subtractive genomics, antimicrobial resistance, drug targets.

INTRODUCTION

The first known *Campylobacter* (Gram-negative bacilli) infection occurred in farm animals in the beginning of the 20th century. Before the start of 1990, it became evident that *Campylobacter* spp. are majorly causing bacterial diarrhea globally (Allos 2001). Infection caused by *Campylobacter* spp. annually affects more than 1.3 million people. The symptoms include diarrhea (often bloody lasting several weeks), nausea, vomiting, fever, and abdominal cramps. The Campylobacter occasionally causes bacteremia in immune-compromised patients leading to life-threatening condition (CDC 2019). Campylobacter infections are mostly initiated from gastrointestinal tract from where they spread to gallbladder (cholecystitis), pancreas (pancreatitis), peritoneum (peritonitis) and may cause gastrointestinal bleeding. Being zoonotic, the *Campylobacter* spp. along with other bacteria are very difficult to be eliminated from the poultry flocks especially chickens (Allos 2001).

Generally the antibiotics azithromycin and ciprofloxacin are prescribed by the physicians to the patients of Campylobacter infection (CDC 2019). Infections caused by antibiotic resistant bacteria imply serious concerns to the public especially by increasing burden to healthcare expenses. Frequent prescriptions and over-usage of antibiotics in humans and animal farming has led to increased prevalence of antibiotic-resistant bacteria including *C. jejuni* (Iovine 2013). That's why the Fluoroquinolone was later disapproved by FDA for its usage in poultry due to emergence of resistance in Campylobacter infected persons. Similar increase in resistance was observed in *C. jejuni* isolates (CDC 2019). This dramatic increase in the antibiotic resistance in *C. jejuni* demands a dire need to discover alternative better treatment both to the human and animals.

A widespread genomic variation is found in *C. jejuni* due to intragenomic conflict and horizontal gene transfers. The genome of *C. jejuni* harbors homopolymeric tracts as hypervariable sequences (Parkhill *et al.,* 2000). Majority of these homopolymeric tracts are located within the genes synthesizing or modifying outer membrane carbohydrates such as the capsule and lipooligosaccharide (LOS). Molecular mechanisms such as phase variation, frameshifts, point mutations, indels, and gene duplications within aforementioned genes causes modification in these outer membrane carbohydrates (Young *et al.,* 2007).

Hypothetical (or uncharacterized) proteins are an important aspect for drug discovery research against bacterial pathogens. These are the genes for which the

**Corresponding author:* e-mail: mriazuddin@iccs.edu

sequence information is accessible but their structure and/or functional characterization is elusive (Shawan, Mahmud *et al.,* 2016). Earlier studies have separately reported forty-nine and seven hypothetical proteins of *C. jejuni* pointing to their significance in structural and functional role (Stahl and Stintzi 2011, Watson *et al.,* 2014). Comparative genomics of four phylogenetically divergent strains of *C. jejuni* encountered many hypothetical proteins especially in hypervariable Plasticity Region (PR) 2, 3 and 7 which also harbors genes of outer membrane and periplasmic space (Clark *et al.,* 2018). More research on these genes of Campylobacter is needed to identify their adaptive mechanism to different environments (Pearson *et al.,* 2003).

The current study is focused on short-listing hypothetical proteins of *C. jejuni* that are essential for the growth of the pathogen and non-homologous to the proteins of human host using the methodology of previous work by the same research group (Uddin and Jamil 2018, Siddiqui *et al.,* 2018, Siddiqui *et al.,* 2019). Furthermore, subcellular location, metabolic pathway, functional families, and druggability of these proteins were identified and later predicted as potential drug targets against *C. jejuni*.

MATERIALS AND METHODS

A computational subtractive genomics approach has been applied to predict potential drug target against *Campylobacter jejuni* (CJJ). The overall methodology that is applied in this study is presented in fig.1.

NCBI Protein BLAST

NCBI BLAST standalone version 2.8.1 was obtained from ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/ LATEST/ to perform the BLASTp of the pathogenic CJJ strain (81-167) protein sequences as query against the database of host proteome *(H. sapiens)* (Mahram and Herbordt 2015). The BLASTp searched the protein sequences of *Homo sapiens* against our query protein sequences of pathogenic bacterium to retrieve the non-host/non-homologous proteins. The BLASTp was also performed to retrieve the homologous protein sequences from the databases of essential genes and DrugBank.

Retrieval of Complete Proteomes

We extracted the complete proteome of pathogenic bacteria *Campylobacter jejuni* (*C. jejuni* 81–176) from UniProtKB database (Apweiler, Bairoch *et al.,* 2004). The complete proteome *of H. sapiens* was retrieved from NCBI ftp server last updated on 3-31-2018. We extracted the complete proteomes of pathogenic bacteria *Campylobacter jejuni* (*C. jejuni* 81-176) from UniProtKB database (Apweiler, Bairoch *et al.,* 2004) and *H. sapiens* from NCBI ftp server last updated on 3-31-2018.

Extraction of Non Paralogous Sequences

We identified paralogous or redundant sequences by using CD-HIT standalone (http://weizhongli-lab.org/cd-hit/), where sequence identity cut off was taken as 0.8 (80%) (Fu *et al.,* 2012). The redundant sequences were extracted out from the complete proteome that resulted in sequences that are only non-paralogous.

Identification of Protein Sequences Non-homologous to the Human Proteome

To retrieve the non-homologous protein sequences, the BLASTp was applied against human proteome with e-value 10^{-3}. The sequences showing significant similarity with human host were eliminated, resulting in only non-homologous sequences that are retrieved for further analysis.

Retrieval of Non-Homologous Essential Proteins in Campylobacter Jejuni (C. Jejuni 81-176)

We subjected the non-homologous protein sequences to BLASTp at an e-value of 10^{-5} against Database of Essential Genes (DEG), obtained from http://www.essentialgene.org/ (Gao *et al.,* 2015). From the BLASTp results, we extracted the protein sequences that are non-homologous and showed significant similarities indicating as essential proteins. The DEG is comprise of all the indispensable genes that are required for a micro-organism to survive (Gao *et al.,* 2015). Therefore, the set of protein homologs retrieved are indispensable/essential for our pathogen.

KAAS Metabolic Pathway Analysis

KAAS stands for KEGG Automatic Annotation Server http://www.genome.jp/kegg/kaas/ based on bi-directional best hit information and sequence similarity (Moriya *et al.,* 2007). We submitted non-homologous essential protein sequences to KAAS that performed similarity searches of query protein sequences against all the protein sequences taking part in the metabolic pathways of the selected pathogenic bacterial strain (*C. jejuni* 81-176). KAAS output file comprised of KO list, enzymes names and enzymes commission numbers (EC) present in each metabolic pathways.

Subcellular Localization of Non-Homologous Essential Proteins

In order to predict the subcellular localization of non-homologous essential proteins, we used a computational tool PSORTb version 3.0.2 (https://www.psort.org/ psortb/) (Yu *et al.,* 2010). The retrieved non-homologous essential protein sequences were subjected to PSORTb. The outcome predicted the sub cellular localization that may include extra cellular, cell-wall, cytoplasmic membrane, cytoplasmic and unknown. The sub-cellular localization is useful for predicting the function of query protein sequences and their role in biological processes.

Functional Family prediction of Uncharacterized, Essential Non-host Proteins

It is important to know the protein functions for the prediction of potential drug targets against pathogens but there exist some uncharacterized proteins in the proteome of pathogenic bacteria with unknown function. Hence, to predict the functional families of those uncharacterized proteins, we used SVMProt web server to classify a protein into functional families from its primary sequences. The protein families consist of all main classes of RNA-binding, DNA-binding, enzymes, channels, receptors and transporters.

Potential druggability of shortlisted uncharacterized proteins

The shortlisted non-homologous essential but un-characterized proteins were subjected to BLASTp against the database of Drug Bank with an e-value 10^{-3}. The BLASTp searched the significant similar sequences against the reported target sequence in Drug Bank. The resultant sequences gave the potential protein targets with respect to Drug IDs approved by the FDA (Wishart *et al.,* 2017).

RESULTS

Retrieval of Primary Proteome Dataset

The UniProtKB database was mined to retrieve the complete protein dataset of *Campylobacter jejuni* (CJJ) (Apweiler *et al.,* 2004). The complete proteome of *Campylobacter jejuni* (strain 81-176) was comprised of 1,748 proteins and it was downloaded in FASTA format. The clustering analysis using CD-hit tool was performed to remove the redundancy of the primary proteome dataset. It was found that the total number of proteins was reduced to 1,744. The threshold cut-off value was set to be 0.8 (80%) which explains that sequences having 80 percent similarity are clustered in one group and sequences having more than 80 percent similarity are eliminated from the proteome dataset (Li and Godzik, 2006).

Comparative non-host protein analysis

A comparative analysis of complete proteome of pathogen with the complete proteome set of Host (*Homo sapiens)* was performed by using standalone BLASTp. The non-redundant dataset generated was subjected to sequence similarity search using BLASTp with an e-value $1x10^{-3}$ against host proteome. The resultant proteome was screened to retrieve the pathogen's sequences sharing no similarity with the host. The non-homology analysis resulted in 1,259 proteins that are not similar to the host proteins.

Essential Gene Analysis

The minimum set of genes required for pathogen survival is considered as essential genes. The identification of essential genes in the proteome of CJJ was performed by essentiality analysis with the help of Database of Essential Genes (DEG) (Gao *et al*., 2015). The DEG covers all experimentally validated available entries of bacterial, eukaryotes and archaea protein coding genes and non-coding RNAs genomic elements (Gao *et al*., 2015). The non-homologous dataset of proteins was subsequently subjected to BLASTp against DEG at an e-value $1x10^{-5}$. The protein sequences with significant similarities were retrieved as predicted and essential genes. These 728 protein sequences are predicted as indispensable set of genes comprising of both characterized proteins and uncharacterized proteins.

Subcellular localization prediction analysis

Several bioinformatics tools are available to determine the subcellular localization of proteins. In this study, an online tool PSORTb has been applied that contains known subcellular location of 11,692 proteins from bacteria and archaea (Hossain *et al*., 2017).

The pool of non-host indispensable proteins sequences were subjected to PSORTb which predicted that 50% proteins were cytoplasmic, 26% proteins were membrane bound and distribution of remaining fraction of proteins are shown in the diagram (fig. 2). However, there were 19% proteins which PSORTb could not able to characterize and their location was predicted as unknown.

KAAS Metabolic Pathway Analysis

The non-homolog and essential protein dataset obtained from subtractive genomics to identify potential drug target against CJJ was subjected to the online KAAS server to perform the molecular pathways analysis. Total numbers of 728 sequences of protein obtained from DEG essentiality analysis were submitted to KAAS online server. The KAAS is a rapid method for annotation and reconstruction of the KEGG metabolic pathways based on sequence similarities (Moriya *et al.,* 2007). The KAAS resulted in 533 hits and the distributions of proteins in different vital metabolic pathways of pathogen are shown in fig. 3. The metabolic pathways of pathogen involving majority of the essential proteins are, carbohydrate metabolism (5%), energy metabolism (5%), nucleotide metabolism (1%), lipid metabolism (1%), amino acid metabolism (6%), glycan biosynthesis and metabolism (3%), metabolism of terpenoids and polyketides (1%), metabolism of cofactors and vitamins (5%), replication and repair (4%), translation (2%), replication and repair (4%), membrane transport (4%), signal transduction (3%), cell growth and death (1%), cellular motility (3%), signaling and cellular processes (17%), genetic information processing (14%), protein families: metabolism (7%), metabolism of other amino acids (1%), biosynthesis of other secondary metabolites (1%), and folding sorting and degradation (2%) etc.

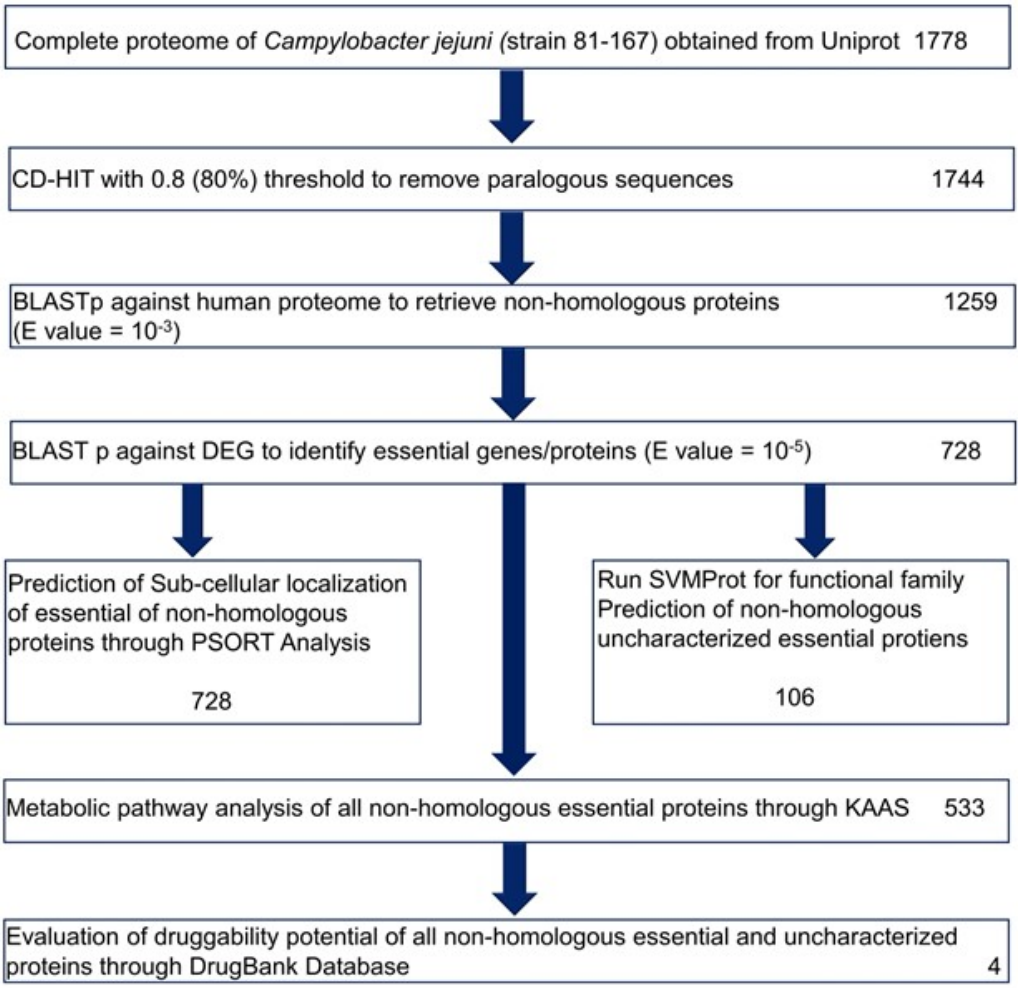

Fig. 2: Complete methodology steps followed in the current study.

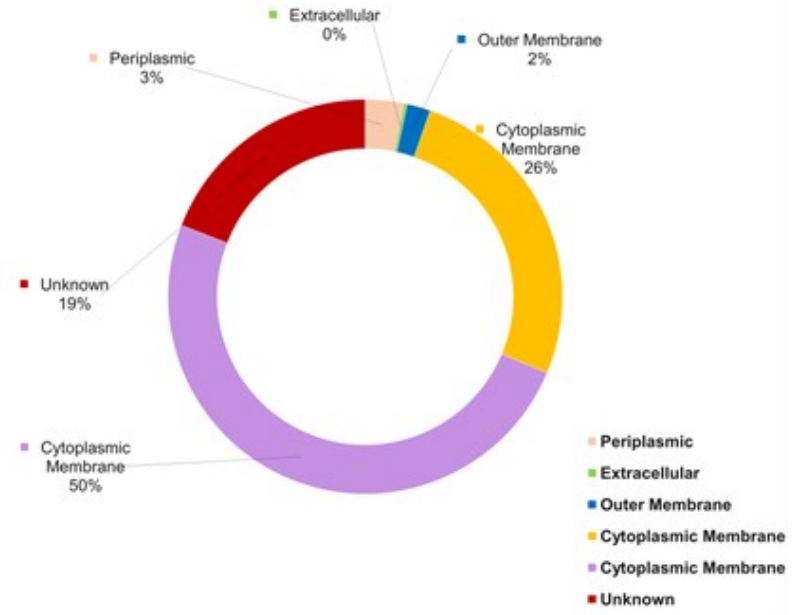

Fig. 2: Sub-cellular localization of non-homologous essential proteins of *C. jejuni* reflected in percentage.

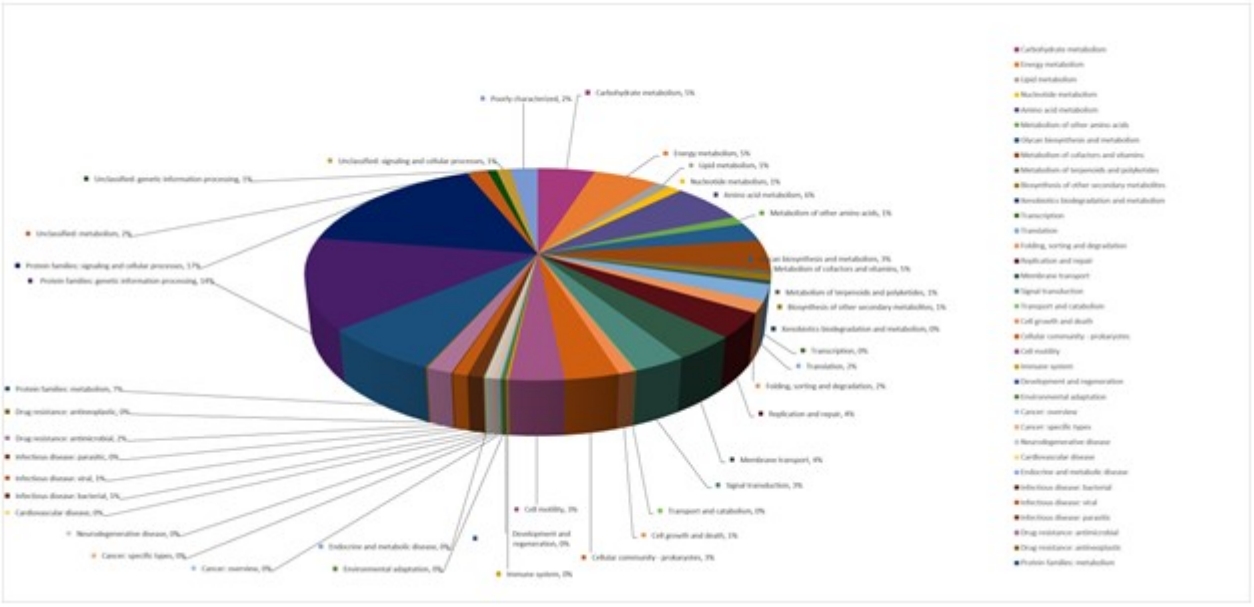

Fig. 4: Distribution of non-homologous essential protein of *C. jejuni* in different metabolic pathways obtained from KAAS

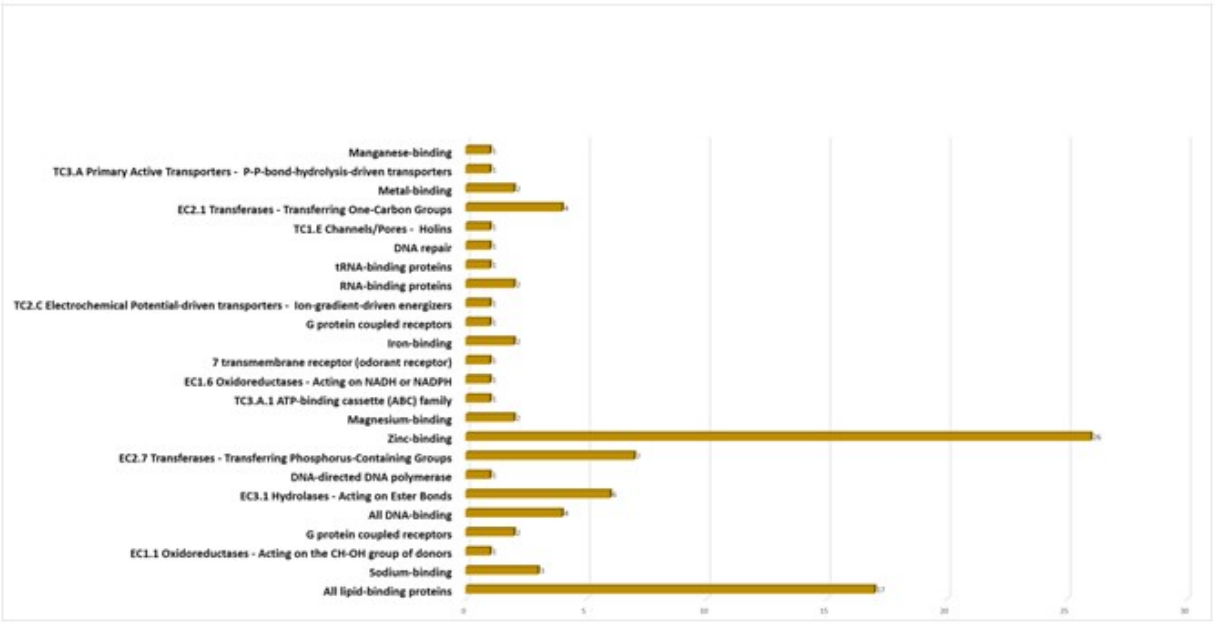

Fig. 4: Functional family prediction of uncharacterized non-homologous essential proteins of *C. jejuni* by using SVMProt.

Table 1: DrugBank targets of the essential, non-homologous, uncharacterized proteins of C. jejuni.

S. no	UniProt IDs	Name of the protein in *Campylobacter jejuni*	DrugBank target name and DrugBank ID
1	A0A0H3PHD4	Uncharacterized protein, A0A0H3PHD4_CAMJJ	Synaptic vesicular amine transporter (DB00182)
2	A0A0H3P983	Uncharacterized protein, A0A0H3P983_CAMJJ	Uracil-DNA glycosylase (DB03419)
3	A0A0H3PAI2	Uncharacterized protein, A0A0H3PAI2_CAMJJ	P33644 Laccase domain protein YfiH (DB04272)
4	A0A0H3PE88	Uncharacterized protein, A0A0H3PE88_CAMJJ	Q6FEW8 Phosphoenolpyruvate-protein phosphotransferase (DB08357)

Functional Family Characterization Analysis

After identification of subcellular localization and metabolic pathway analysis of the essential, non-homolog proteome dataset, we only retrieved 104 uncharacterized proteins or hypothetical proteins to perform their functional family annotation. The protein sequences were uploaded to an online server SVMProt that is frequently used for functional family characterization of hypothetical proteins (Li *et al*., 2016). The results of SVMProt are shown in fig. 4 that represents the most common families and these are zinc-binding, lipid-binding, transferases and hydrolases.

Druggability Analysis for Prioritization of Potential Drug Targets

The uncharacterized proteins obtained from the steps of subtractive genomic analysis were further assessed to identify the potential drug targets by performing the druggability analysis (Birhanu, Lee, Park, Song, & Park, 2018). The database of DrugBank was used to identify the druggability of the essential uncharacterized proteins or hypothetical proteins. The DrugBank database includes all information related to FDA approved drugs and their targets (Wishart *et al*., 2017). The binding potential of shortlisted proteins with drug like molecules was evaluated by subjecting the protein sequences to similarity search against the DrugBank database with default parameters. A total of 04 hypothetical proteins in CJJ showed significant hits against the FDA approved drugs given in table 1. These 04 potential drug targets proteins were found to be similar with synaptic vesicular amine transporter DB00182, Uracil-DNA glycosylase DB03419, Laccase domain protein YfiH DB04272, and phosphoenolpyruvate protein phosphotransferase DB08357. The 02 proteins uracil DNA glycosylase and laccase domain YfiH protein are also involved in metabolic pathways of CJJ. The KAAS results showed that the uracil DNA glycosylase is involved in repair, replication and genetic information processing pathways whereas laccase domain is involved in an unclassified metabolism sharing the enzymatic activity of polyphenol oxidase.

DISCUSSION

The characterization or annotation of uncharacterized proteome and identification of potential drug targets of *Campylobacter jejuni* (strain 81-176) was the primary aim of the current study. The subtractive genomics approach was employed to the whole proteome of the pathogen through utilizing various computational tools and software. The uprising threat of multi drug resistant strains of the *Campylobacter jejuni* (CJJ) demands the identification of potential drug targets (Zhao *et al*., 2016). In current study the identification of 04 predicted potential drug targets is performed by applying the combination of systems biology approaches with bioinformatics tools. The study has enabled us to understand the complex biological, molecular and cellular functions performed by the predicted targets. It also helped to unfold the complex mechanism of resistance and pathogenesis of CJJ. The increasing levels of resistance (multi-drug and extended drug) are posing serious threats to the global population. Therefore, identification of potential therapeutic drug targets by using *in silico* subtractive genomics approach has now been routinely applied. The prioritized targets are from the uncharacterized/hypothetical proteome of CJJ. Yet, these 04 proteins are essential and non-homolog to the host proteins characterizing them as potential drug targets to overcome the current antibacterial resistance.

The subtractive genomic approach applied to identify the potential drug target against CJJ, identified host non-homologous and subsequently the essential proteins from the non-redundant complete proteome of CJJ. The metabolic pathway analysis coupled with subtractive genomic analysis was employed to shortlist the uncharacterized proteins and perform their characterization using annotation tools. In general, one of the properties of an ideal drug target is that it must not share sequence similarity to the host proteins to prevent the adverse effects and cross-reactivity (Uddin & Azam, 2019). The identification of non-homolog proteins is performed to circumvent the binding of drugs to the binding sites of host proteins that could result in unwanted effects (Pourhajibagher & Bahador, 2016). Targeting essential genes could be crucial as they are indispensable for the organism's survival. These genes play vital role in the fulfillment of the minimum requirements of survival and functionality. Thus, such genes can serve as potential drug targets. The current study has identified the host non-homologous essential proteins that were subjected to subcellular localization identification step followed by the metabolic pathway analysis. The proteins perform their optimum functions in the particular compartment of the cell. The identification of the subcellular localization of the proteins aid in predicting their functions, and interacting networks (Hu and Janitz 2009). It is important to identify the cellular positions of the possible drug targets because the drugs need to bind with the target proteins to produce the particular therapeutic responses. Prior knowledge about the drug target's location is useful in the drug designing and development. The particular location of the proteins inside the cell holds an important factor in the selection and prioritization of suitable potential drug targets (Yu *et al*., 2010). The pathway analysis resulted in 533 essential and non-homologous proteins distributed among metabolic processes of the pathogen. The highest proportion of essential proteins accounted for the signaling, cellular processes, protein metabolism, and genetic information processing, carbohydrate and energy metabolism inferring the indispensability of essential genes for the pathogen.

The proteins which are not fully characterized are referred to as hypothetical proteins (Uddin and Jamil, 2018). The genome of pathogenic bacteria also consists of number of hypothetical proteins. Therefore, one strategy in order to predict the functions of such protein sequences is to perform their functional family characterization based on the sequence similarity. The purpose of the current study was to characterize the uncharacterized proteome of CJJ and identify the druggable proteins among them. The 04 hypothetical proteins were shortlisted as potential drug targets proteins after performing the sequence similarity analysis against the DrugBank Database. The A0A0H3PHD4 (uncharacterized protein) primarily annotated as major facilitator superfamily (MFS) domain containing protein and it is involved in the membrane transport of small solutes in response to chemiosmotic ion gradient. Such membrane transporters are also involved in the extrusion of cytotoxic drugs from multidrug resistant cells imparting failure of the drug based therapy against the diseases or infections caused by pathogenic bacteria (Harwood and Parales 1994). The bioinformatics strategy applied in current study has identified MFS domain-containing protein as one of the potential therapeutic target against CJJ. As described earlier, these transporters are contributing in gaining the resistance against the drug molecules rendering this protein crucial to develop antimicrobial agents or could help in reversing the antimicrobial resistance of CJJ. The second protein A0A0H3P983 (Uncharacterized protein) is annotated as uracil-DNA glycosylase (UDG) domain-containing protein. It is involved in DNA repair mechanism specifically performs the base excision repair (BER) mechanism. The role of proteins involved in DNA repair in CJJ is largely unknown; however one study identified the presence of functional uracil-DNA glycosylase (UGD) in BER mechanism (Gaasbeek *et al.,* 2009). The UDG initiates and catalyzes the first step by recognizing specific DNA mutations and removes the damaged bases by cleaving the N-glycosylic bonds. Targeting DNA repair proteins could be promising targets for drug development against CJJ as the DNA repair mechanism is vital for its survival (Coronado and Tayler, 2011; (Khanam & Ramachandran, 2014). The third potential drug target A0A0H3PAI2 (uncharacterized protein) belongs to the multicopper oxidase (MCO) Yfih/RL5 or metal binding family. In other *C. jejuni* (11168) strain similar sequence protein has been characterized as polyphenol oxidase. It is often annotated as laccase domain in bacterial genomes (Silva *et al.,* 2012). The protein is involved in copper hemostasis. It has oxidoreductase activity and performs metal acquisition such as iron, zinc, copper and magnesium. These metals are essential for the bacterial growth. Metal acquisition has been marked as one of the virulence factor and this area has been extensively studied recently. Disturbing the metal hemostasis could be lethal for bacterial survival (Hall and Kelly 2008). Therefore, targeting such enzymes can provide help in developing the robust therapeutic intervention against the infections caused by CJJ. A0A0H3PE88 (uncharacterized protein) is the fourth shortlisted protein this has kinase activity and harbors pyruvate phosphate dikinase, PEP/pyruvate binding domain and PEP utilizer domain. It is characterized as L-glutamine kinase in NCTC11168 strain of *C. jejuni,* where it catalyzes the *O*-methyl phosphoramidate (MeOPN) modification on specific sugar residues and produces a unique capsular polysaccharide (CPS) (Taylor *et al.,* 2017). The phosphorous nitrogen bonds are uncommon in nature making CPS modification unusual. This reaction is ATP dependent phosphorylation, in which amide nitrogen of L-glutamine is phosphorylated to form L-glutamine phosphate. In our study, KEGG orthology (KO) has also grouped this uncharacterized protein as glutamine kinase (table1 S1). The CPS envelops the bacteria, provides protection against environmental condition and plays a critical part for invasion and colonization of host organisms (Chamberlain and Raushel, 2018). Hence, targeting enzymes involved in biochemical synthesis of CPS could be crucial in the drug discovery and development against the CJJ. The predicted target proteins identified are found to be playing crucial roles in pathogen's life cycle. The proteins have potential to be considered as drug targets for designing antimicrobial agents with novel mechanism of action. Therefore, the target proteins can be further explored by applying structural bioinformatics approaches in follow up studies

CONCLUSION

This study provides new avenues into the metabolic pathways governed by *C. jejuni* strain 81–176, and possibly other related intestinal pathogens. The current study identified the uncharacterized essential proteins that could be target for new antimicrobial agents. The further characterization of four identified protein targets predicted their probable role in the resistance (membrane transporter), metal hemostasis (laccase domain protein YfiH), and pathogen specific unusual biosynthesis of CPS (pyruvate phosphate dikinase/glutamine kinase). However, Uracil-DNA glycosylase is involved in multiple metabolic pathways related to growth of pathogen such as DNA repair mechanism and DNA processing. Designing or discovering new compounds having activity against these proteins should be proceeded in a knowledgeable method to cure Campylobacter infections. It is concluded that the antimicrobial agents designed against these shortlisted proteins will only be active against the pathogen and with negligible off-target activity in the host. Our efforts mainly focused on determining the importance of *in silico* subtractive genomics approach in characterizing druggable bacterial genomes for deciphering new potential target candidates for the antimicrobials.

ACKNOWLEDGEMENT

The authors would like to acknowledge the funds provided by the Pakistan Science Foundation (Project# Med 431)

REFERENCES

Allos BM (2001). *Campylobacter jejuni* Infections: Update on emerging issues and trends. *Clin. Infect Dis.,* **32**(8): 1201-1206.

Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R and Magrane M (2004). UniProt: The universal protein knowledgebase. *Nucleic Acids Res.,* **32**(suppl. 1): D115-D119.

Birhanu BT, Lee SJ, Park NH, Song JB and Park SC (2018). *In silico* analysis of putative drug and vaccine targets of the metabolic pathways of Actinobacillus pleuropneumoniae using a subtractive/comparative genomics approach. *J. Vet. Sci.,***19**(2): 188-199.

CDC (2019, August 6, 2019). Campylobacter (Campylobacteriosis)https://www.cdc.gov/campylobacter/index.html.

Clark CG, Chen CY, Berry C, Walker M, McCorrister SJ, Chong PM and Westmacott GR (2018). Comparison of genomes and proteomes of four whole genome-sequenced *Campylobacter jejuni* from different phylogenetic backgrounds. *PLoS One,* **13**(1): e0190836.

Coronado LM, De La Guardia CI, González YS, Restrepo CM and Tayler NM (2011). DNA repair mechanisms as drug targets in prokaryotes. *Curr. Trends Biotechnol. Pharm.,* **5**(3): 1206-1232.

Fu L, Niu B, Zhu Z, Wu S and Li W (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics,* **28**(23): 3150-3152.

Gaasbeek EJ, Van der Wal FJ, van Putten JP, de Boer P, Van der Graaf-van Bloois L, de Boer AG, Vermaning BJ and Wagenaar JA (2009). Functional characterization of excision repair and RecA-dependent recombinational DNA repair in *Campylobacter jejuni. J. Bacteriol.,* **191**(12): 3785-3793.

Gao F, Luo H, Zhang CT and Zhang R (2015). Gene essentiality analysis based on DEG 10, an updated database of essential genes. Gene Essentiality, Springer, pp.219-233.

Hall SJ, Hitchcock A, Butler CS and Kelly DJ (2008). A multicopper oxidase (Cj1516) and a CopA homologue (Cj1161) are major components of the copper homeostasis system of *Campylobacter jejuni*. *J. Bacteriol.,* **190**(24): 8075-8085.

Harwood CS, Nichols NN, Kim MK, Ditty JL and Parales RE (1994). Identification of the pcaRKF gene cluster from *Pseudomonas putida*: Involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. *J. Bacteriol.,* **176**(21): 6479-6488.

Hossain T, Kamruzzaman M, Choudhury TZ, Mahmood HN, Nabi A and Hosen M (2017). Application of the subtractive genomics and molecular docking analysis for the identification of novel putative drug targets against *Salmonella enterica* subsp. enterica serovar Poona. *BioMed Res. Int.,* 1-9.

Hu Y, Lehrach H and Janitz M (2009). Comparative analysis of an experimental subcellular protein localization assay and in silico prediction methods. *J. Mol. Histol.,* **40**(5-6): 343-352.

Iovine NM (2013). Resistance mechanisms in *Campylobacter jejuni*. *Virulence,* **4**(3): 230-240.

Khanam T and Ramachandran R (2014). Exploiting bacterial DNA repair systems as drug targets: A review of the current scenario with focus on mycobacteria. *J. Indian InstSci.,* **94**(1): 149-168.

Li W and Godzik A (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics,* **22**(13): 1658-1659.

Li YH, Xu JY, Tao L, Li XF, Li S, Zeng X, Chen SY, Zhang P, Qin C and Zhang C (2016). SVM-Prot 2016: A web-server for machine learning prediction of protein functional families from sequence irrespective of similarity. *PLoS One,* **11**(8): e0155290.

Mahram A and Herbordt MC (2015). NCBI BLASTP on high-performance reconfigurable computing systems. *ACM Trans. Reconfigurable Technol. Syst. (TRETS),* **7**(4): 33.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC and Kanehisa M (2007). KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.,* **35**(suppl. 2): W182-W185.

Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S and Barrell BG (2000). The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature,* **403**(6770): 665-668.

Pearson BM, Pin C, Wright J, I'Anson K, Humphrey T and Wells JM (2003). Comparative genome analysis of *Campylobacter jejuni* using whole genome DNA microarrays. *FEBS Lett.,* **554**(1-2): 224-230.

Pourhajibagher M and Bahador A (2016). Designing and *in silico* analysis of PorB protein from *Chlamydia* trachomatis for developing a vaccine candidate. *Drug Res.,* **66**(09): 479-483.

Shawan M, Mahmud H, Gope PS, Salauddin NM, Rahman MH, Ahmed MA, Nafiz TN, Imran KM, Rahman MN and Rahman SB (2016). *Campylobacter jejuni* ATCC 700819: An *in silico* approach to identify and categorize probable drug targets by subtractive genome analysis. *J. Pure Appl. Microbiol.,* **10**(1): 229-241.

Silva CS, Durao P, Fillat A, Lindley PF, Martins LO and Bento I (2012). Crystal structure of the multicopper oxidase from the pathogenic bacterium *Campylobacter jejuni* CGUG11284: Characterization of a metallo-oxidase. *Metallomics,* **4**(1): 37-47.

Stahl M and Stintzi A (2011). Identification of essential genes in *C. jejuni* genome highlights hyper-variable plasticity regions. *Funct Integr Genomics,* **11**(2): 241-257.

Taylor ZW, Brown HA, Narindoshvili T, Wenzel CQ, Szymanski CM, Holden HM and Raushel FM (2017). Discovery of a glutamine kinase required for the biosynthesis of the O-methyl phosphoramidate modifications found in the capsular polysaccharides of *Campylobacter jejuni. J. Am. Chem. Soc.,* **139**(28): 9463-9466.

Taylor ZW, Chamberlain AR and Raushel FM (2018). Substrate specificity and chemical mechanism for the reaction catalyzed by glutamine kinase. *Biochemistry,* **57**(37): 5447-5455.

Uddin R and Azam SS (2019). Identification of glucosyl-3-phosphoglycerate phosphatase as a novel drug target against resistant strain of Mycobacterium tuberculosis (XDR1219) by using comparative metabolic pathway approach. *Comput. Biol. Chem.,* **79**: 91-102.

Uddin R and Jamil F (2018). Prioritization of potential drug targets against *P. aeruginosa* by core proteomic analysis using computational subtractive genomics and Protein-Protein interaction network. *Comput. Biol. Chem.,* **74**: 115-122.

Uddin R, Siddiqui QN, Azam SS, Saima B and Wadood A (2018). Identification and characterization of potential druggable targets among hypothetical proteins of extensively drug resistant *Mycobacterium tuberculosis* (XDR KZN 605) through subtractive genomics approach. *Eur. J. Pharm. Sci.,* **114**(1): 13-23.

Uddin R, Siddiqui QN, Sufian M, Azam SS and Wadood A (2019). Proteome-wide subtractive approach to prioritize a hypothetical protein of XDR-*Mycobacterium tuberculosis* as potential drug target. *Genes Genomics*, **41**(11):1281-1292.

Watson E, Sherry A, Inglis NF, Lainson A, Jyothi D, Yaga R, Manson E, Imrie L, Everest P and Smith DG (2014). Proteomic and genomic analysis reveals novel *Campylobacter jejuni* outer membrane proteins and potential heterogeneity. *EuPA. Open Proteom.,* **4**: 184-194.

Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C and Sayeeda Z (2017). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.,* **46**(D1): D1074-D1082.

Young KT, Davis LM and Dirita VJ (2007). *Campylobacter jejuni*: Molecular biology and pathogenesis. *Nat. Rev. Microbiol.,* **5**(9): 665-679.

Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M and Foster LJ (2010). PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics,* **26**(13): 1608-1615.

Zhao S, Tyson GH, Chen Y, Li C, Mukherjee S, Young S, Lam C, Folster J, Whichard JM and McDermott PF (2016). Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in Campylobacter spp. *Appl. Environ. Microbiol.,* **82**(2): 459-466.