



Original article

Proteome based mapping and molecular docking revealed DnaA as a potential drug target against *Shigella sonnei*Farah Shahid^a, Youssef Saeed Alghamdi^b, Mutaib Mashraqi^c, Mohsin Khurshid^d, Usman Ali Ashfaq^{a,*}^a Department of Bioinformatics and Biotechnology, Government College University, Faisalabad, Pakistan^b Department of Biology, Turabah University College, Taif University, P.O.BOX 11099, Taif 21944, Saudi Arabia^c Department of Clinical Laboratory Sciences, College of Applied Medical Science, Najran University, Najran, Saudi Arabia^d Department of Microbiology, Government College University, Faisalabad, Pakistan

ARTICLE INFO

Article history:

Received 11 July 2021

Revised 9 August 2021

Accepted 13 September 2021

Available online 1 October 2021

Keywords:

Shigella sonnei

Chromosomal replication initiator protein

DnaA

Molecular docking

Phytochemicals

Bioinformatics

Drug target

ABSTRACT

Shigella sonnei is one of the major causes of diarrhea and remained a critical microbe responsible for higher morbidity and mortality rates resulting from dysentery every year across the world. Antibiotic therapy of *Shigella* diseases plays a critical role in decreasing the prevalence as well as the fatality rate of this infection. However, the management of these diseases remains challenging, owing to the overall increase in resistance against many antimicrobials. The situation necessitates the rapid development of effective and feasible *S. sonnei* treatments. In the present study, the subtractive genomics approach was utilized to find the potential drug targets for *S. sonnei* strain Ss046. Various tools of bioinformatics were implemented to remove the human-specific homologous and pathogen-specific paralogous sequences from the bacterial proteome. Then, metabolic pathway and subcellular location analysis were performed of essential bacterial proteins to describe their role in various cellular processes. Only one essential protein i.e Chromosomal replication initiator protein DnaA was found in the proteome of the pathogen that could be used as a potent target for designing new drugs. 3D structure prediction of DnaA protein was carried out using Phyre 2. Molecular docking of 5000 phytochemicals was performed against DnaA to identify four top-ranked phytochemicals (Riccioniadin A, Dothistromin, Fustin, and Morin) based on scoring functions and interaction with the active site. This study suggests that these phytochemicals could be used as antibacterial drugs to treat *S. sonnei* infections in the future. To confirm their efficacy and evaluate their drug potency, further in vitro analyses are required.

© 2021 Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Shigellosis is widespread internationally and highly endemic in the least developed countries, with almost 0.25 billion cases every year (Khalil et al., 2018). Approximately 13% of deaths caused by diarrhea are due to the *Shigella* infection (Khalil et al., 2018). About 10 lac cases of shigellosis arise every year in highly industrialized countries, where infants in crowded rooms and childcare centers, soldiers in the field, and people traveling to endemic countries are at a high risk of developing this disease (Pires et al., 2015;

Porter et al., 2017). *Shigella sonnei* prevails as a major source of this infection in high-income countries (Shad & Shad, 2021). *S. sonnei*, for example, has played a substantial role in foodborne outbreaks in both the Canada and the United States, emphasizing the importance of food as a source of infection (Kimura et al., 2004). However, due to evolving mechanisms and diverse infection causing modes in *S. sonnei*, the global trend of *S. sonnei* infections has shifted dramatically over time from industrialized to poor countries (Torraca et al., 2020b). The frequency and intensity of outbreaks vary by region due to a variety of factors including geography, climate, host–pathogen interactions, and controlling strategies. *S. sonnei* has three primary lineages (I, II, and III) that have recently been identified. *S. sonnei* evolved quickly in Europe and spread to other continents as a single multidrug-resistant (MDR) lineage (Anandan et al., 2017; Holt et al., 2012).

Shigella sonnei is a Gram-negative, rod-shaped, and facultative intracellular pathogen. *S. sonnei* evolved to specialize in human gut epithelium intracellular infection from *Escherichia coli*. Its gen-

* Correspondence author.

E-mail address: ashfaqua@gcuf.edu.pk (U.A. Ashfaq).

Peer review under responsibility of King Saud University.



Table 1
Unique metabolic pathways of essential nonhomologous proteins.

Protein Name (ID)	Pathway ID	Common Pathway	Unique Pathway
3-methyl-2-oxobutanoate hydroxymethyltransferase (Q3Z5M6)	ssn00770 ssn01100 ssn01110	Pantothenate and CoA biosynthesis Metabolic pathways	Biosynthesis of secondary metabolites
Glutamate-1-semialdehyde 2,1-aminomutase (Q3Z5K3)	ssn00860 ssn01100 ssn01110 ssn01120	Porphyrin and chlorophyll metabolism Metabolic pathways	Biosynthesis of secondary metabolites Microbial metabolism in diverse environments
Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha (Q3Z5H3)	ssn01100 ssn01110 ssn00061 ssn01200 ssn01212 ssn00620 ssn00640 ssn01120 ssn01130	Metabolic pathways Fatty acid biosynthesis Carbon metabolism Fatty acid metabolism Pyruvate metabolism Propanoate metabolism	Biosynthesis of secondary metabolites Microbial metabolism in diverse environments Biosynthesis of antibiotics
1-deoxy-D-xylulose-5-phosphate synthase (Q3Z4Y9)	ssn01100 ssn01120 ssn01130 ssn00730 ssn00900	Metabolic pathways Thiamine metabolism Terpenoid backbone biosynthesis	Microbial metabolism in diverse environments Biosynthesis of antibiotics
Ferrochelatase (Q3Z4S4)	ssn00860 ssn01100 ssn01110	Porphyrin and chlorophyll metabolism Metabolic pathways	Biosynthesis of secondary metabolites
Phosphate acyltransferase (Q3Z327)	ssn00561 ssn01100 ssn01110	Glycerolipid metabolism Metabolic pathways	Biosynthesis of secondary metabolites
Enoyl-[acyl-carrier-protein] reductase [NADH] (Q3Z136)	ssn00061 ssn01212 ssn01100 ssn00780 ssn01110 ssn01130	Fatty acid biosynthesis Fatty acid metabolism Metabolic pathways Biotin metabolism	Biosynthesis of secondary metabolites Biosynthesis of antibiotics
Glutamyl-tRNA reductase (Q3Z0S8)	ssn00860 ssn01100 ssn01110 ssn01120	Porphyrin and chlorophyll metabolism Metabolic pathways	Biosynthesis of secondary metabolites Microbial metabolism in diverse environments
Glutamate-tRNA ligase (Q3YZD6)	ssn00860 ssn00970 ssn01100 ssn01110 ssn01120	Porphyrin and chlorophyll metabolism Aminoacyl-tRNA biosynthesis Metabolic pathways	Biosynthesis of secondary metabolites Microbial metabolism in diverse environments
4-hydroxy-tetrahydrodipicolinate synthase (Q3YZ74)	ssn01230 ssn00261 ssn00300 ssn01100 ssn01110 ssn01120 ssn01130	Biosynthesis of amino acids Lysine biosynthesis Metabolic pathways	Monobactam biosynthesis Biosynthesis of secondary metabolites Microbial metabolism in diverse environments Biosynthesis of antibiotics
Succinyl-diaminopimelate desuccinylase (Q3YZ81)	ssn01230 ssn00300 ssn01100 ssn01120	Biosynthesis of amino acids Lysine biosynthesis Metabolic pathways	Microbial metabolism in diverse environments
Transketolase (Q3YZ88)	ssn01200 ssn01230 ssn01100 ssn01110 ssn01120 ssn01130 ssn00030	Carbon metabolism Biosynthesis of amino acids Metabolic pathways Pentose phosphate pathway	Biosynthesis of secondary metabolites Microbial metabolism in diverse environments Biosynthesis of antibiotics
Chromosomal replication initiator protein DnaA (Q3YWB2) Membrane protein insertase YidC (Q3YWA8)	ssn02020 ssn02024 ssn03070 ssn03060	Protein export	Two-component system Quorum sensing Bacterial secretion system
Glycerol-3-phosphate dehydrogenase [NAD(P) +] (Q3YVX3)	ssn00564 ssn01110	Glycerophospholipid metabolism	Biosynthesis of secondary metabolites
Glutamine synthetase (Q3YVA3)	ssn01230 ssn01100 ssn02020 ssn00220 ssn00250 ssn00630 ssn00910 ssn01120	Biosynthesis of amino acids Metabolic pathways Arginine biosynthesis Alanine, aspartate and glutamate metabolism Glyoxylate and dicarboxylate metabolism Nitrogen metabolism	Two-component system Microbial metabolism in diverse environments

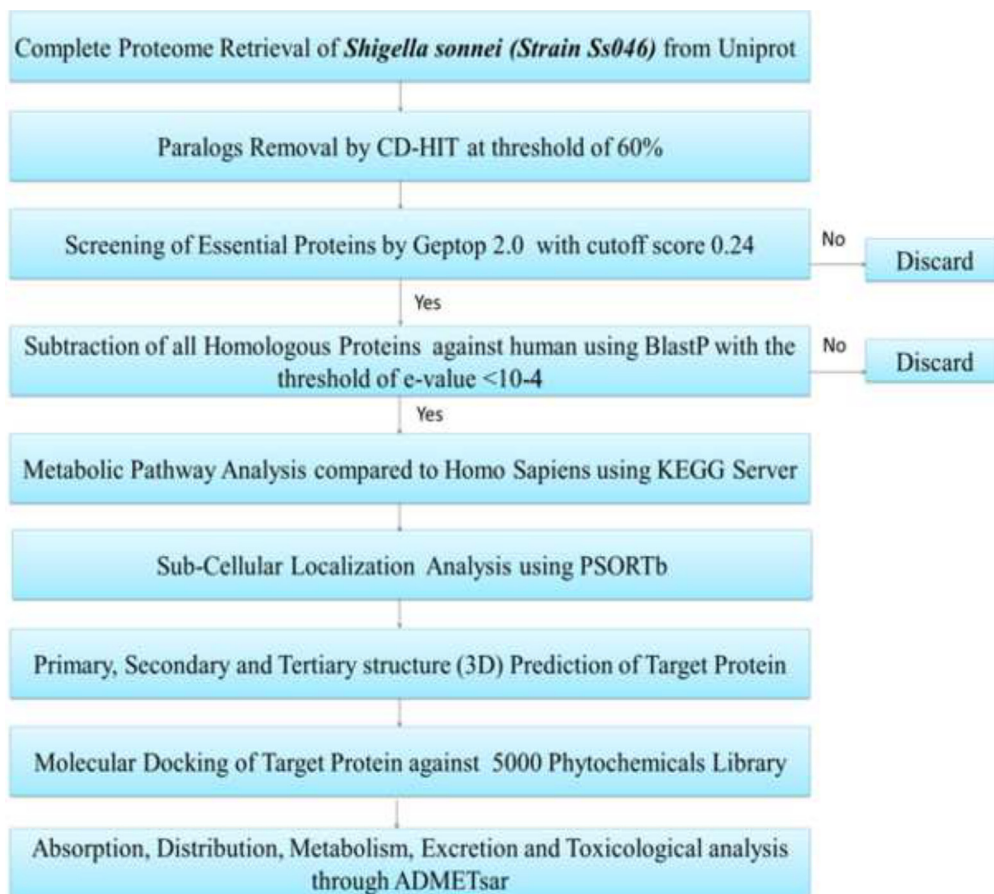


Fig. 1. A schematic representation of the identification of novel drug targets in *Shigella sonnei*.

ome includes a Plasmid invasion (pINV) 216 kbp, and a circular chromosome of 4,99 Mbp, required for virulence (Torraca et al., 2020a). After 1 to 7 days of infection with *S. sonnei*, the person shows symptoms such as cramping rectal pain, acute abdominal pain, acute fever, watery diarrhea, cramping, nausea, blood or mucus or, pus in the stool. Usually, the infection clears up without complications. It can cause serious complications if the diagnosis remains untreated or delayed. The most common of them are severe dehydration that can lead to seizure, reactive arthritis, death, shock, toxic megacolon, and hemolytic uremic syndrome (HUS) (Kotloff et al., 2018).

No vaccines or drugs are available for *S. sonnei*. The best way to prevent shigellosis is washing hands with soap and water thoroughly, repeatedly and cautiously, before handling the food, after and before using the washroom. Strict compliance with standard food and water safety precautions is also important. Water from lakes, untreated swimming pools, or ponds should also not be swallowed. Avoid having sex with diarrhea patients (Prabhurajeshwar, 2018).

S. sonnei is resistant to almost all the antimicrobials, and the management of these infections has become expensive, sometimes challenging, and time-consuming as well, specifically in regions with limited healthcare facilities (Qiu et al., 2013; Taneja & Mewara, 2016). Lately, several mechanisms of antimicrobial resistance have now been identified by investigators, limiting the treatment options for the management of these diseases (Qiu et al., 2013). As a consequence of the significant global burden, medical severity along with the increased reports of evolving antibiotic resis-

tance to first as well as second-line treatments, an effective drug for shigellosis has now become a growing requirement (Gu et al., 2012).

Various approaches like comparative along with subtractive genomics have been utilized to find the targets in several human microbes (Jamal et al., 2017; Mahmud et al., 2019). The main objective is to identify the targeted essential genes that show no homology with the *Homo sapiens* so that the drug targets can be used with minimum off-target effects in humans. Here, we employed an integrated *in-silico* method to the entire proteome of *S. sonnei* strain Ss046 to relate the genomic data with the presumed therapeutic targets based on their 3D structure. The approach can also be exploited for the detection of powerful inhibitors, which could potentially give rise to the development of such compounds that possess the capability of inhibiting bacterial growth. The proteome of *S. sonnei* (Ss046) was retrieved to apply subtractive proteomics. The computational tool Geptop was utilized for the detection of proteins which are considered essential for pathogen survival. Metabolic pathway and subcellular location analysis were carried out to avoid the involvement of bacterial proteins in human metabolism as well as the drugs cross-reactivity with humans, respectively. Chromosomal replication initiator protein DnaA was identified as a drug target. DnaA protein is responsible for the initiation of chromosome replication in bacteria, and its activity is limited to one cell division cycle. DnaA binds to DnaA-boxes, which are high-affinity sites on oriC (the origin of replication), and to other sites on the chromosome with lower affinity. Binding to oriC sites causes the adjacent AT-rich region to unwind, resulting in the opening of the double helix. DnaA

recruits the DnaB helicase and the replisome is assembled, which effectively initiates DNA replication. In bacteria, the initiation of chromosome replication by DnaA protein is a highly regulated event because DNA replication timing is critical for steady-state cell growth (Katayama et al., 2010; Regev et al., 2012). The research was further extended to model the 3D structure of the probable drug target DnaA via Phyre 2 to identify a selective and potent inhibitor using docking studies. The present study would be useful in designing an effective drug target against *S. sonnei* (strain Ss046). Fig. 1 illustrates a comprehensive strategy to identify putative drug targets in *Shigella sonnei* (Strain Ss046).

2. Methodology

2.1. Data collection of proteome

The complete proteome of *Shigella sonnei* (Strain Ss046) was retrieved from Uniprot (<https://www.uniprot.org/>). UniProt (Universal Protein Resource) provides a complete as well as an open-access resource for protein sequences and genome annotations. Protein sequences provided by UniProt are obtained by translating the coding sequences. More than 95% of the protein sequences are retrieved in this way (Renaux, 2018).

2.2. Paralogs removal

The whole proteome of *Shigella sonnei* (Strain Ss046) was run at the CD-HIT suite (http://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi?cmd=cd-hit). The threshold value was set at 60%. CD-HIT is widely utilized for comparing and clustering protein and nucleotide sequences. It removes paralogs or redundant proteins (Huang et al., 2010).

2.3. Screening of essential genes

Essential proteins are needed for organisms for their survival and are considered the basis of life. Essential proteins were retrieved by using an online software Geptop 2.0 server (<http://guolab.whu.edu.cn/geptop/>) (Wei et al., 2013). Geptop serves as a platform for the detection of genes essential in pathogenic species,

comparing query protein orthology and phylogeny with the experimentally established essential gene datasets.

2.4. Identification of non-homologous essential genes

Screened essential genes should not be homologous to *Homo sapiens*. For this purpose, these proteins were submitted to Blast p (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) against host proteome with a threshold of e-value 10^{-4} , and query coverage and identity more than 70 % and 30% respectively (Altschul et al., 1990).

2.5. Analysis of metabolic pathways

Non-homologous essential proteins were further computed for comparative analysis of metabolic pathways. The screened proteins were used to determine the metabolic pathways with which they are associated. This analysis is conducted to identify drug targets based on common and essential pathway enzymes for bacteria (Anishetty et al., 2005). The analysis of metabolic pathways of the non-human homologous essential proteins of *Shigella sonnei* (Strain Ss046) was computed at KEGG (<https://www.genome.jp/kegg/>) (Aoki-Kinoshita & Kanehisa, 2007). Those metabolic pathways were selected that were unique to only *Shigella sonnei* (Strain Ss046) and were not found in humans. Thus, the proteins having unique metabolic pathways were selected for further assessment.

2.6. Subcellular localization analysis

The prediction of subcellular localization provides a fast and relatively cost-effective way to obtain information about a protein particular function. Besides, it has been found that proteins can locate at multiple sites, so localization is a critical aspect of designing any therapeutic agent (Goyal & Citu, 2018). Subcellular localization of the target proteins of *Shigella sonnei* (Strain Ss046) was identified through Psortb (Yu et al., 2010). Psortb (<https://www.psort.org/psortb/>) is an online tool used to find the subcellular location of proteins, whether they are periplasmic proteins, cytoplasmic membrane proteins, or cytoplasmic proteins.

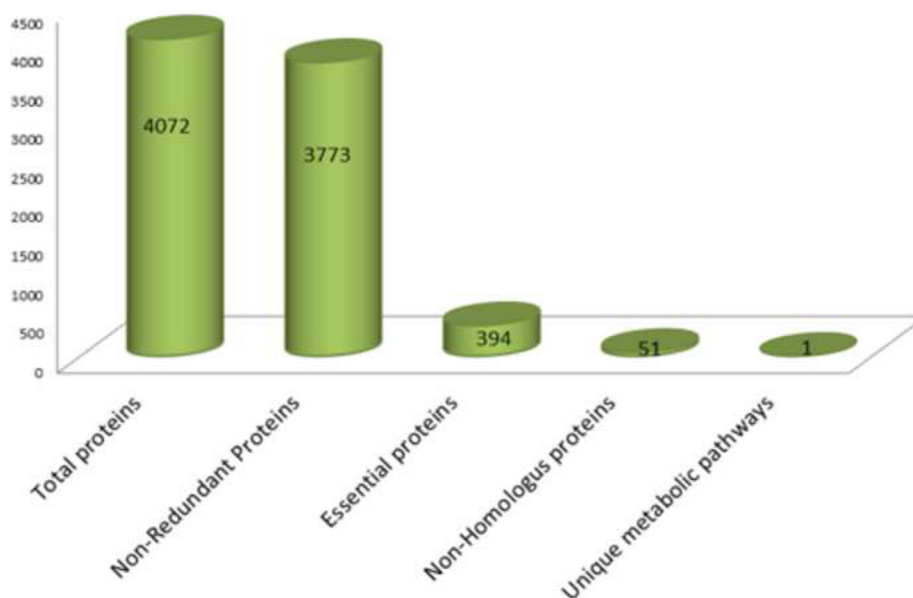


Fig. 2. Summary of subtractive genomics approach employed in this study from core proteome retrieval to essential drug target prediction.

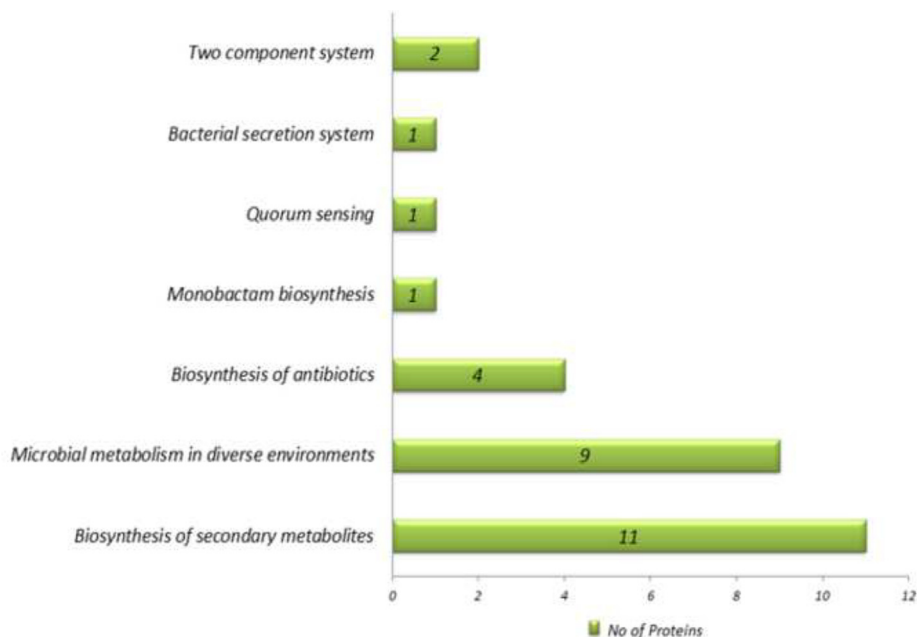


Fig. 3. Frequency Distribution of proteins involved in unique metabolic pathways of *Shigella sonnei* (Strain Ss046).

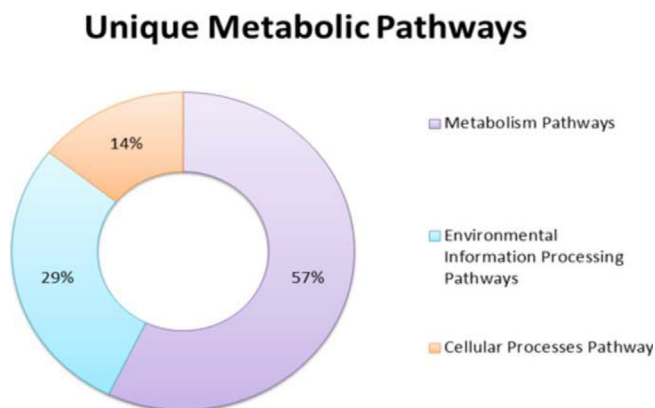


Fig. 4. Classification of seven identified unique metabolic pathways according to biochemical processes.

2.7. Structural analysis of target protein

ProtParam (<https://web.expasy.org/protparam/>) and GOR4 (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html) were used for primary structure analysis and secondary structure analysis, respectively (Garnier et al., 1996; ProtParam, 2017). The signal peptide location and presence can be identified by SignalP-5.0 Server (Armenteros et al., 2019). The Transmembrane Topology of the protein was checked through the TMHMM tool (<http://www.cbs.dtu.dk/services/TMHMM/>) (Krogh et al., 2001).

2.8. Structure prediction and validation

Different online tools such as Phyre 2, Swiss Model as well as Raptor X were employed to determine the 3D protein structure (Källberg et al., 2014; Kelley et al., 2015; Schwede et al., 2003). The predicted structure of the target protein was visualized through Chimera (Pettersen et al., 2004). The 3D model of the target protein was refined using the Galaxy WEB server (Ko et al.,

2012). To confirm the quality and reliability of the 3D structure of the target protein, different online tools were applied. PROSA web to provide a quality score for a specific structure, ERRAT server to analyze non-bonded interactions, OPM-PPM to check the spatial orientation in the cell membrane, and RAMPAGE to construct the Ramachandran plot of 3D structure (Lengths & Angles, 2018; Lomize et al., 2012).

2.9. Molecular docking

The refined structure of the target protein was docked against the phytochemical's library. For docking, 5000 phytochemicals were retrieved from different databases in 2D conformation i.e form PubChem, MPD3, Zinc database, and MAPS database in sdf file format (Ashfaq et al., 2013; Irwin & Shoichet, 2005; Kim et al., 2016; Mumtaz et al., 2017). Ligand library composed of 5000 phytochemicals was optimized for docking. Optimization entails adding partial charges and minimizing the energy of a selected compound using the Protonate-3D and MMFF94X force fields, respectively. Optimized ligand files were saved in a ligand database, which was later used as an input file for docking studies. The site finder tool in MOE software was utilized to locate the active sites on the target protein (Inc., 2016). After the solvent and additional ligands were removed from the receptor, docking was performed. Parameters set for docking were placement: triangle matcher, refinement: forcefield, retain 10 and rescoring: London dg. Ligands were screened out after confirmation with a receptor molecule.

2.10. Physicochemical property profile and toxicity prediction

Molecular descriptors and drug likeliness properties of best docking score phytochemicals were analyzed using the Molinspiration server, which provides a prediction based on the 'rule of five' (Lipinski et al., 1997). The criteria relate to the following molecular properties (AlogP values less than 5, less than 10 hydrogen bond acceptors, less than 5 hydrogen bond donors, and a molecular mass less than 500 Daltons). The pharmacokinetic characteristics e.g. distribution, metabolism, adsorption, toxicity, and excretion of

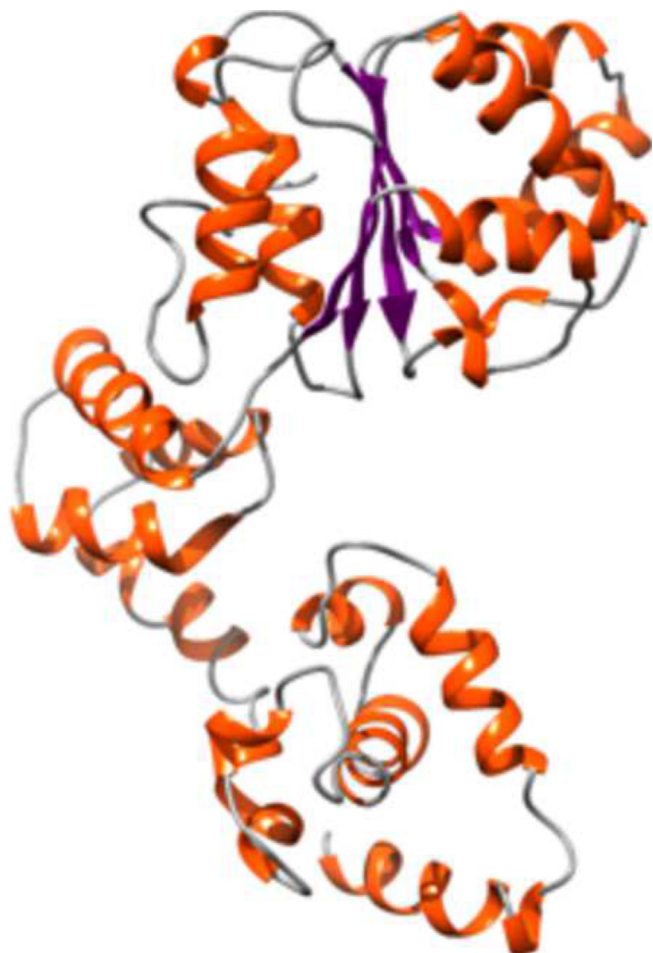


Fig. 5. The three-dimensional structure of the Chromosomal replication initiator protein. Orange color represents alpha helixes, purple color represents beta sheets, and grey color represents loops.

compounds can be determined through admetSAR (Yang et al., 2019). Besides this, various toxicity screening models such as oral acute toxicity, immunotoxicity, organ toxicity, genetic endpoint toxicity were applied to the selected molecules using ProTox-II webserver. The ProTox-II is a server used to predict the toxicity of various chemical compounds with different toxicological endpoints (Banerjee et al., 2018).

3. Results

This study was designed for the identification of novel targets against *Shigella sonnei* (strain Ss046). The approach of subtractive genomics was utilized to identify the therapeutic proteins necessary for bacterial survival but absent within the host (Fig. 2)

3.1. Identification of essential non-homologous genes

The whole proteome retrieved from the UniProt of *Shigella sonnei* (Ss046) contained 4072 proteins. After paralogs removal, using a CD-HIT suite with a 60% threshold, 3773 out of 4072 non-redundant proteins were obtained. Geptop2.0 was employed to predict the essential *Shigella sonnei* (Ss046) proteins. The essential proteins obtained from Geptop2.0 were 394 out of 3773 proteins. To prevent drug interactivity with human proteins, essential proteins were submitted to BlastP with a threshold value less than

10^{-4} , and only 51 proteins were screened as non-homologous to humans.

3.2. Identification of essential non-homologous proteins involved in unique metabolic pathways

The analysis of the metabolic pathway of these 51 proteins showed the involvement of these proteins in 29 pathways. Comparison of the *S. sonnei* pathways with the *H. sapiens* pathways revealed that 7 specific metabolic pathways for *S. sonnei* and 22 resting pathways are common in pathogen and host. 16 essential proteins of *S. sonnei* were associated with these 7 pathways. The distribution of proteins in each unique metabolic pathway is shown in (Fig. 3). The classification of seven unique metabolic pathways according to biochemical processes was represented in (Fig. 4).

These 16 proteins involved in unique metabolic processes were studied individually using the KEGG database. Of these 16 proteins, 15 proteins played a role in several pathways. These 15 proteins were involved in several unique pathways, as well as in some pathways common between *H. sapiens* and *S. sonnei* (Table 1). Therefore, these were not considered for further analysis. Proteins present in unique metabolic pathways can be considered pathogen-specific and serve as potential vaccine and drug targets. Only one protein Chromosomal replication initiator protein DnaA was involved in a unique metabolic pathway i-e two-component regulatory system, so this protein was considered as potential drug target.

Two-component regulatory system consists of histidine kinase along with a response regulator. Bacteria detect and respond to extracellular signals through this system. Such systems enable the cell to adapt by changing cellular physiology to the prevailing conditions, including initiating gene expression programs, catalyzing reactions, and altering protein–protein interaction. Such pathways are also involved in regulating the growth and production of the bacterial cell cycle (Skerker et al., 2005). This pathway is not found within metazoans including homo sapiens, so enzymes associated with this pathway can be identified as a significant target for anti-fungal as well as anti-bacterial drug designing (Barrett & Hoch, 1998).

3.3. Subcellular localization prediction

The subcellular location of the DnaA was identified using Psortb. The subcellular location of protein helps in determining the protein function as well as its potential as a potent target (Goyal & Citu, 2018). DnaA was found to be cytoplasmic protein as predicted by Psortb.

3.4. Structural analysis of target protein

The sequence of DnaA protein was obtained from UniProt (ID: Q3YWB2). Primary structure analysis was predicted by ProtParam. Chromosomal replication initiator protein DnaA protein had a molecular mass of 52550.81 Daltons. The isoelectric point was 8.77. As this protein has an isoelectric point above 7, it is tagged as a negatively charged protein. The Grand average of hydropathy (GRAVY) describes the hydrophilicity of the protein. The GRAVY value calculated was -0.364 . The terminating amino acid at the N-terminus of protein is Methionine (Met). The protein is unstable as the instability index (II) calculated to be is 47.21. The rich quantity of amino acids found in this protein is Alanine (A), Leucine (L), Arginine (R), and Asparagine (N). The secondary structure was predicted by GOR4 and PSIPRED. The secondary structure was reported to have 50.32% of alpha helices, 0 % of beta turns, 9.64% extended strand, and 40.04% of random coils (Fig. 6A). The signal

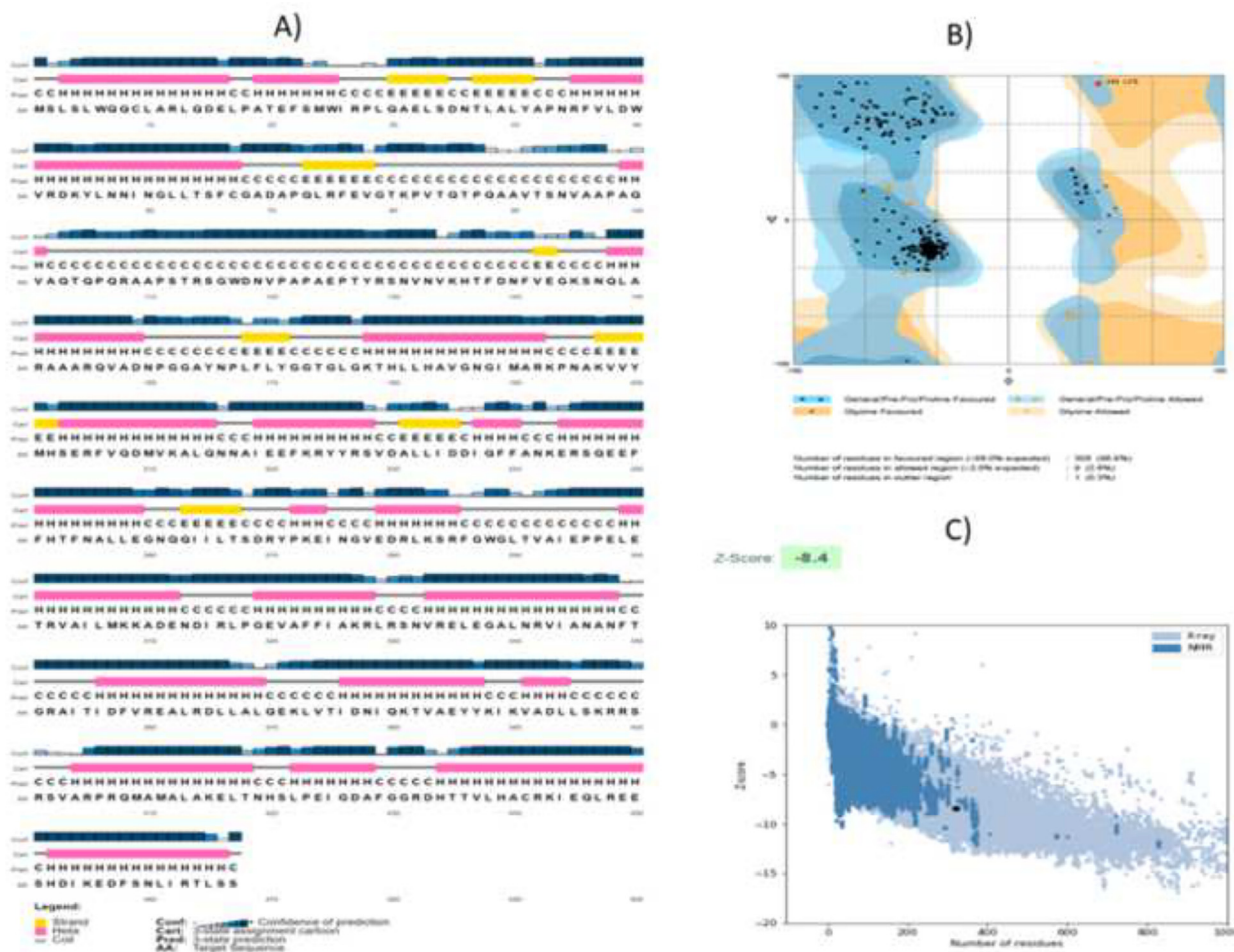


Fig. 6. Structural Analysis of DnaA Protein: (A) the DnaA protein contains α -helix (50.32%, 235), β -strand (9.64%, 45) and random coil (40.04%, 187); (B) the Ramachandran plot of the refined structure shows 98.1%, 1.9% and 0.0% residues in favored, allowed and disallowed region, respectively; (C) the z-score (-8.04) of the DnaA protein.

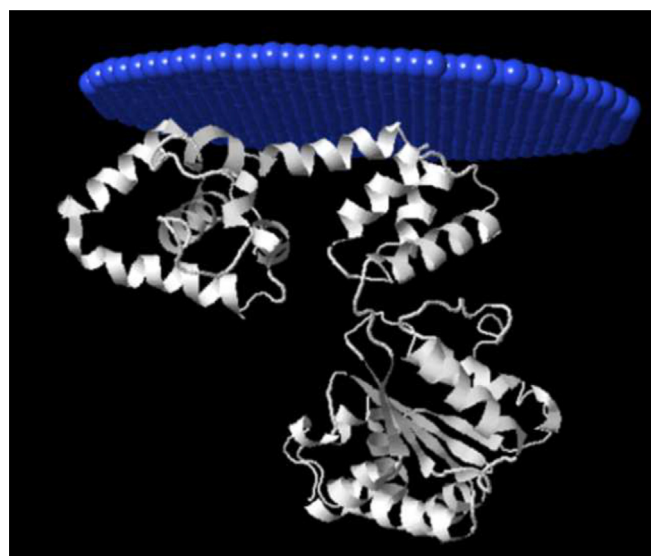


Fig. 7. Topology of DnaA protein predicted by OPM-PPM server.

peptide probability value obtained by SignalP is 0.0127, thus no signal peptide was found in a novel protein target. TMHMM showed that DnaA protein does not have any transmembrane helices. Thus, DnaA is attached to the inside of the cell membrane.

3.5. Structure prediction and validation

As the 3D structure was not found in Protein Data Bank, different tools e.g. Phyre 2, Raptor \times along with the Swiss model were used to find the 3D structure of the protein. Structures obtained from these tools were analyzed by the Ramachandran plot and compared to get the best possible outcome. Phyre 2 model was selected as it was better than other models. The confidence level of the structure was 100% and coverage was 66% (Fig. 5). The obtained structure was refined using the Galaxy server. It generated five models. Model 3 was considered the finest model and selected for further analysis. The improved model has a 95.3% favored region in RAMPAGE and poor rotamers as 0.4%, qrmsd as 0.416, clash score as 18.5, MolProbity as 2.092, and RAMPAGE server analyzes and validates the refined structure by creating Ramachandran plot. In the more favorite region, 98.1% of residues are generated, 1.9% of amino acid residues within the permitted area, and none was present in outer regions according to the vaccine construct Ramachandran plot (Fig. 6B). ProSA-web also gives - 8.04 Z, which is within an acceptable range of values (Fig. 6C). Besides, the refined model showed 3errors with PRO-CHECK. The refined model was 90.3846 in total quality (ERRAT score). These results show that the refined model is of good quality.

OPM-PPM server predicts that chromosomal replication initiator protein DnaA of *Shigella sonnei* (Ss046) was a peripheral protein. Peripheral membrane proteins interact with the lipid bilayer of the cell membrane. They do not cross cell membranes; therefore,

Table 2
Interaction detail of top four bioactive phytochemicals in the proposed site of DnaA protein.

Sr no	PubChem Id	Chemical name	Docking score	Interaction detail	
				rmsd value	Residues
1	441,775	Riccionidin A	−17.6	0.8	Glu 335 Thr 174
2	108,014	Dothistromin	−16.9	1.3	Ser 331 Arg 432 Asp 427
3	5,317,435	Fustin	−16.6	1.5	Ser 331 Arg 432
4	5,281,670	Morin	−16.4	1.3	Glu 335 Ser 331 Arg 432 Asp 427

they do not enter into the hydrophobic spaces of the lipid bilayer. They are temporarily associated with the cell membrane in which they adhere, may be attached to integral membrane proteins, or penetrate the peripheral regions (Fig. 7). The depth/hydrophobic thickness calculated was $0.9 \pm 0.6 \text{ \AA}$, ΔG transfer was -2.6 kcal/mol and tilt angle was $45 \pm 7^\circ$.

3.6. Molecular docking

Molecular docking was carried out through MOE software for the screening of compounds possessing the best interaction of binding residue with the receptor. Compounds were analyzed based on RMSD value, residue binding with ligands, and docking score. Out of 5000 docked molecules, the top four ranked docking poses were chosen. Compounds with low scores i.e. $\text{rmsd} < 3$ and residues with the largest number of interactions were selected. These selected phytochemicals exhibited their minimum binding energy in the range of $-17 \text{ kcal mol}^{-1}$ to $-16 \text{ kcal mol}^{-1}$. The minimum binding energy and scoring function of each docked ligand are shown in Table 2. To get a precise idea of the receptor-ligand interaction with the highest docked complexes, the MOE LigX tool was used to analyze 2D plots of these interactions. Docked complexes were then visualized by the UCSF chimera. Riccionidin A/DnaA pocket complex LigX interaction diagram showed good binding with Glu 335 and Thr 174 via hydrogen bonds with a score of 17.6 kcal/mol (Fig. 8A). LigX interaction diagrams showed that Dothistromin, Fustin, and Morin were found to bind with DnaA protein with a score of $-16.9 \text{ kcal mol}^{-1}$, $-16.6 \text{ kcal mol}^{-1}$, and $-16.4 \text{ kcal mol}^{-1}$ forming hydrogen bonds with the side chains of Ser 331, Arg 432, and Asp 427 (Dothistromin), Ser 331 and Arg 432 (Fustin) and Glu 335, Ser 331, Arg 432 and Asp 427 (Morin), Fig. 8 (B-D) respectively.

However, Dothistromin was ranked top ascribed to its maximum binding score and good binding affinity among the revealed binding residues from the DnaA. Three out of the four phytochemicals, i.e., appear to exhibit high binding affinity to the Ser 331 and Arg 432 amino acids, indicating that these are the most active residues.

3.7. ADMET/Drug scan results

The drug-likeness of four selected compounds was predicted through the Molinspiration server, based on the Lipinski Rules of five. The selected candidates displayed zero violation of the 'rule of five' and exhibited drug-like properties (Table 3). All the candidate compounds were subjected to assessment for pharmacokinetic properties via the admetSAR server to further validate the potential of drug likeliness (Table 4).

3.8. Toxicity assessment results

After the docking test, the four molecules obtained underwent various toxicity modules. Table 5 lists the rat's oral acute toxicity (LD50) as mg/Kg, predicted specific toxicity classes (I–VI) and predicted accuracy in percentage, organ toxicity prediction regarding the findings of liver toxicity. The graphical representation of the expected distribution of dose value for the candidate compounds is shown in Fig. 9A–D. Table 6 enlisted genotoxicity prediction regarding cytotoxicity and mutagenicity. The genotoxicity prediction concerning cytotoxicity and mutagenicity showed that all compounds were obtained cytotoxic inactive with likelihood scores of 0.89, 0.76, 0.98, and 0.98 respectively. In the case of mutagenicity endpoints, all compounds were found to be mutagenicity inactive with a Probability score of 0.55, 0.56, 0.53, and 0.52.

Among four compounds, Riccionidin A and Fustin obtained class IV i–e prescribed as toxic if consumed, with 67.38% and 72.9% prediction accuracy respectively (Fig. 9A & C). While the rest of the two compounds Dothistromin and Morin obtained class V i–e prescribed as can be toxic if consumed, with 68.07% and 69.26% prediction accuracy respectively (Fig. 9B & D). All compounds were identified as hepatotoxic inactive with a likelihood score of 0.77, 0.75, 0.7, and 0.68 respectively for the prediction of organ toxicity regarding liver toxicity.

4. Discussion

Shigella sonnei is a multidrug-resistant (MDR) bacterium (Abbasi et al., 2019). *Shigella sonnei* causes shigellosis which affects millions of people annually and causes several deaths. Mild shigellosis cases are treated without antimicrobial agents and their recovery is quite rapid. Antibiotics control the spread of this disease inside the intestine as well as reduces the duration required for it to enter the entire body. The use of anti-diarrheal medications is not recommended as it may aggravate the disease (Sati et al., 2019). There is currently no effective vaccine for shigellosis, but frequent and thorough hand washing can prevent person-to-person transmission (Sybilski et al.).

To combat this life-threatening situation, there is a dire need to develop drugs against *Shigella sonnei* immediately. In our study, subtractive proteomics was employed to screen the drug targets against MDR *Shigella sonnei*. This approach is utilized to find the targets based on the determination of essential as well as non-homologous proteins within pathogenic organisms. Identifying drug targets is a critical step in the computer-based drug designing procedures (Hosen et al., 2014). Recent advances in the disciplines of bioinformatics as well as computational biology have created a variety of approaches to drug design and *in silico* analysis, reducing the time and expense associated with trial and error of ions devoted to drug development (Barh et al., 2011).

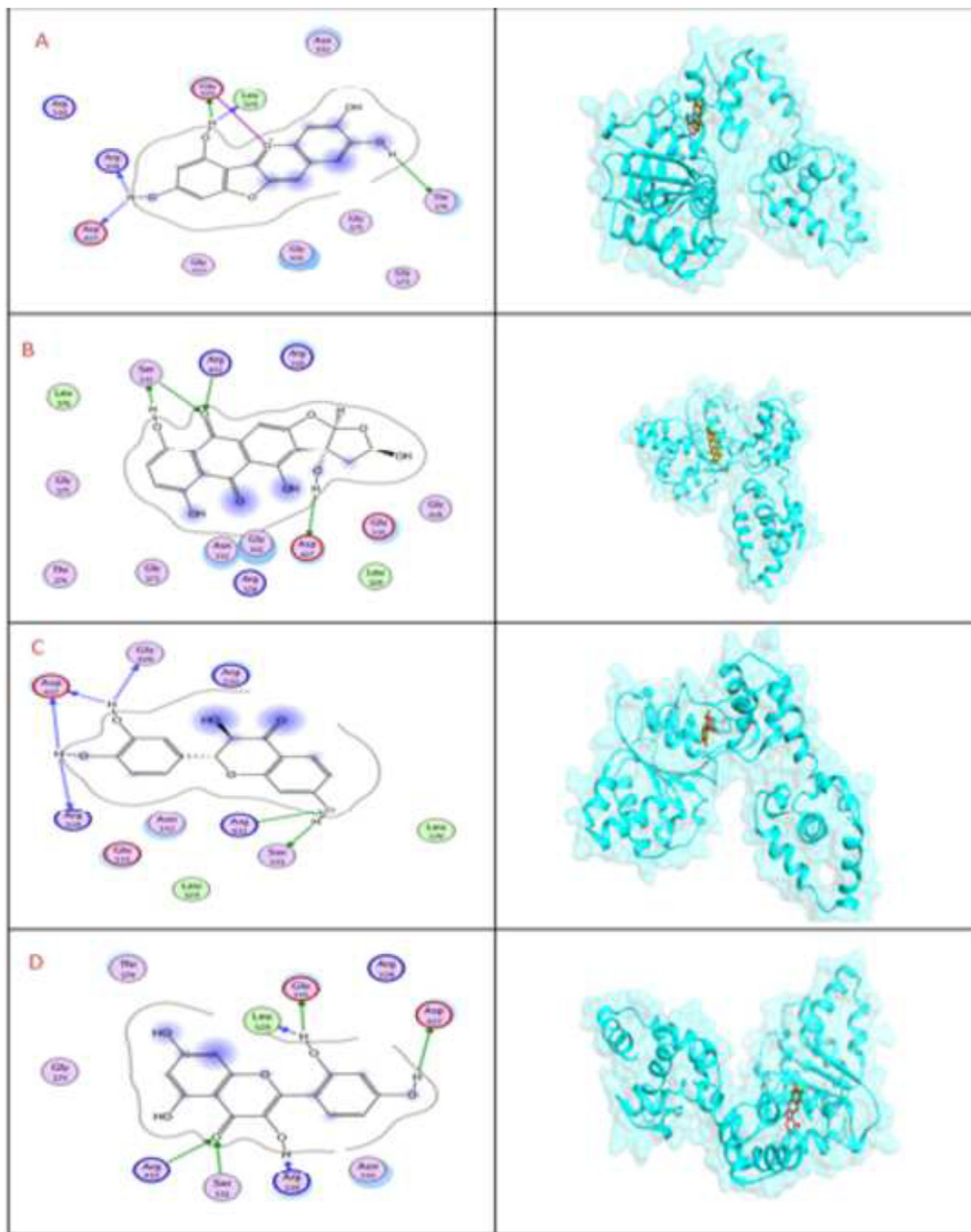


Fig. 8. A: Docked Riccionidin in complex with DnaA; side chains atoms of Glu 335 and Thr 174 making hydrogen bonds, shown in the green line. The docked pose of the compound visualized by Chimera is shown at right. B: Docked Dothistromin in complex with DnaA; side chains atoms of Ser 331, Arg 432, and Asp 427 making hydrogen bonds, shown in the green line. The docked pose of the compound visualized by Chimera is shown at right. C: Docked Fustin in complex with DnaA; side chains atoms of Ser 331 and Arg 432 making hydrogen bonds, shown in the green line. The docked pose of the compound visualized by Chimera is shown at right. D: Docked Morin in complex with DnaA; side chains atoms of Glu 335, Ser 331, Arg 432, and Asp 427 making hydrogen bonds, shown in the green line. The docked pose of the compound visualized by Chimera is shown at right.

Table 3
Results of compounds examined for Lipinski rule.

Compound	Molecular weight (g/mol)	Number of HBA	Number of HBD	MLogP
Lipinski rule of five	<500	<10	<5	<5
Riccionidin A	285.23	6	4	-0.67
Dothistromin	372.29	9	5	1.49
Fustin	288.25	6	4	0.80
Morin	302.24	6	5	1.88

The whole proteome of *Shigella sonnei* (Ss046) that contained 4072 proteins was analyzed through CD-HIT that eliminated all the redundant proteins and provided a group of 3773 non-

redundant proteins. Essential proteins are necessary for the survival of bacteria. If these essential proteins are degraded or mutated, bacteria cannot survive. By targeting these proteins, we

Table 4
ADMET Profiling Enlisting Absorption, Metabolism and Toxicity related drug-like parameters of candidate compounds.

A. ADMET Profiling				
Compounds	Riccionidin A	Dothistromin	Fustin	Morin
A. Absorption				
Blood-Brain Barrier	+	-	-	-
Human Intestinal Absorption	+	+	+	+
P-glycoprotein substrate	-	-	-	-
CYP450 1A2 Inhibitor	Yes	No	Yes	Yes
CYP450 2C9 Inhibitor	No	No	Yes	Yes
CYP450 2D6 Inhibitor	No	No	No	No
CYP450 2C19 Inhibitor	No	No	No	Yes
CYP450 3A4 Inhibitor	No	No	No	Yes
Distribution				
Subcellular localization	Mitochondria	Mitochondria	Mitochondria	Mitochondria
Toxicity				
AMES Toxicity	No	No	No	No

Table 5
Prediction of oral acute toxicity, class and accuracy, organ toxicity and genetic toxicity endpoints of candidate compounds (IA: Inactive).

Sr.	Compounds Name	Oral LD50 Value (mg/kg)	Predicted Toxicity Class	Prediction Accuracy (%)	Hepatotoxicity	Probability	Cytotoxicity	Probability
1	Riccionidin A	2991	IV	67.38%	IA	0.77	IA	0.89
2	Dothistromin	3000	V	68.07%	IA	0.75	IA	0.76
3	Fustin	2000	IV	72.9%	IA	0.7	IA	0.98
4	Morin	3919	V	69.26%	IA	0.68	IA	0.98

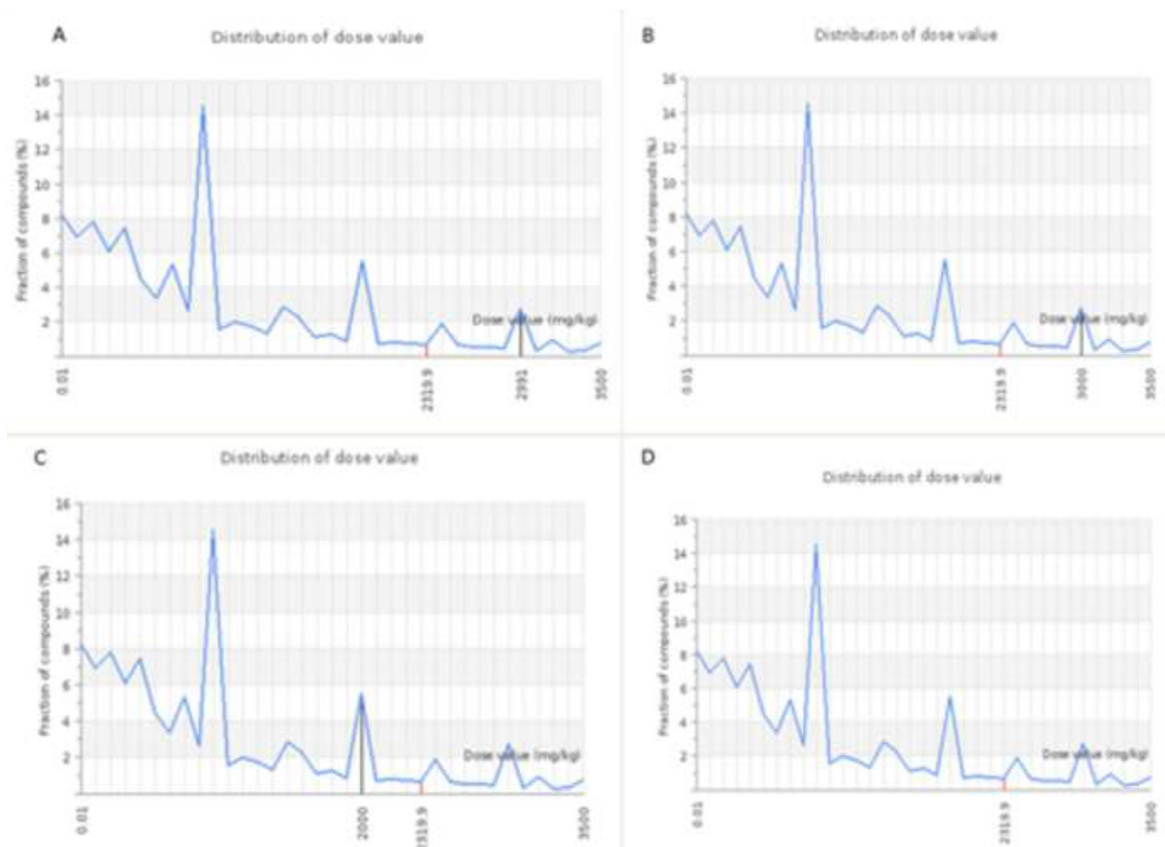


Fig. 9. Graphical representation of predicted dose value distribution for candidate compounds (A = Riccionidin A; B = Dothistromin; C = Fustin; D = Morin;).

can kill bacteria and cure the disease. Essential proteins are preferred targets for vaccine development and antibacterial drugs. Thus 394 essential proteins were screened from non-redundant proteins. Shilpa S., et al identified 807 essential proteins in *Eubacterium nodatum*; Sakharkar et al. found 306 essential genes within *P. aeruginosa* whereas Chan-Eng Chong et al. identified 312 essen-

tial proteins within *Burkholderia pseudomallei* using this method (Chong et al., 2006; Sakharkar et al., 2004; Shiragannavar et al., 2019). These genes could be homologous to humans. Thus, targeting such genes can interfere with human metabolism and prove fatal. The possibility of cross-reactivity, as well as adverse events, can be reduced by the selection of the non-homologous proteins

Table 6
Prediction of genetic toxicity endpoints of candidate compounds (IA: Inactive).

Sr. No	Compounds name	Cytotoxicity	Probability	Mutagenicity	Probability
1	Riccionidin A	IA	0.89	IA	0.55
2	Dothistromin	IA	0.76	IA	0.56
3	Fustin	IA	0.98	IA	0.53
4	Morin	IA	0.98	IA	0.52

that are not found in homo sapiens (Barh et al., 2011). To avoid such undesirable circumstances and toxicity, we screened 51 non-homologous proteins.

Comparative study of human and pathogen metabolic pathways through the KEGG database showed that 7 pathways are specific to pathogen only and 22 pathways are common in both pathogen and host. These 7 pathways include a two-component system, bacterial secretion system, Quorum sensing, Monobactam biosynthesis, metabolism of microbe within diverse environments, and synthesis of the secondary metabolites. The findings of pathways identification which are specific to the pathogen are in line with the results of *L. interrogans*, *A. baumannii*, and *S. saprophyticus* reported by (Amineni et al., 2010; Goyal et al., 2018; Shahid et al., 2020). A total of 16 essential proteins of *S. sonnei* were associated with these 7 pathways. Out of these 16, 15 proteins were involved in some pathways common between *H. sapiens* and *S. sonnei*. Only one protein DnaA was involved in a unique metabolic pathway. DnaA, a ubiquitous protein that is responsible for initiating the replication of chromosomal DNA at specific sites by the unwinding of double-stranded DNA helix. It is very crucial for the replication of DNA in bacteria (Mott & Berger, 2007; Zakrzewska-Czerwińska et al., 2007). It is evident that DnaA is crucial for the replication of the chromosome and survival of *Shigella sonnei* (Ss046), thus it could be a potential drug target against this pathogen. DNA replication proteins are necessary for cell viability and are therefore attractive targets for drug designing.

Different tools were applied to determine the sequence and structural features as well as the function and localization of that protein. DnaA was found to be cytoplasmic protein as predicted by Psortb. The crucial function of this protein in cell viability maintenance renders them more attractive towards drug targets. The protein was 55426.35 kDa in molecular weight. DnaA was basic, according to the theoretical pI value, which can ensure stable physiological pH interaction. The GRAVY has been measured as -0.364 , the negative value refers to its hydrophilic nature, which makes it interact with neighboring water molecules. The tertiary structure of the putative drug target protein was predicted, evaluated, and validated.

Molecular docking was carried out to find the compounds exhibiting the best residue interaction with the target protein. Out of 5000 docked molecules, four top molecules: Riccionidin A, Dothistromin, Fustin, and Morin were selected based on low score i.e. $rmsd < 3$ and many interacting residues. These phytochemicals exhibited minimum binding energy of almost $-17 \text{ kcal mol}^{-1}$ to $-16 \text{ kcal mol}^{-1}$. Based on "Lipinski's Rule of Five" the molecular profile and drug probability of these four compounds were assessed. All of them fulfilled the "Lipinski's Rule of Five" and exhibited no desecration.

The compounds were then tested for penetration of the blood-brain barrier (BBB), HIA (Human Intestinal Absorption), as well as AMES monitoring. Predicting the ADMET properties is a significant indicator of the behavior, toxicity level, and fate of the drug candidate in the human body. It gives a likelihood of the candidate's ability to enter the intestinal absorption, metabolism, blood-brain barrier, subcellular localization, and most importantly, the level of harm it can cause in the body (Lin et al., 2003). The superfamily cytochrome P450 consists of isoforms such as CYP2A6, CYP1A2,

CYP2C9, CYP2D6, CYP2C19, CYP3A4, and CYP2E1 which are involved in drug metabolism as well as hepatic clearance (Vasanthanathan et al., 2009). So, inhibiting the cytochrome P450 isoforms can result in drug-drug interaction that hinders the metabolism of concomitant drugs that cause its accumulation to toxic levels (Lynch & Price, 2007). ADMET profile of compounds indicates that all these compounds have no adverse effects on Absorption. Besides, all compounds showed no toxicity and no mutagenic effects compared to the AMES test.

Various toxicity modules were subjected to the four-hit compounds obtained after the virtual screening. Results of the toxicity evaluation revealed that no compound was found to be cytotoxic, hepatotoxic, and mutagenic. Our study identified four drug-leading inhibitors that could be one therapeutic inhibitor of DnaA by effectively targeting and inhibiting apoptosis.

5. Conclusion

In recent years, diarrhea-causing *S. sonnei* has become multidrug-resistant. Realizing the importance of the development of potent drug targets, this study was designed to carry out the computational analysis of human microbial pathogen *S. sonnei* (strain Ss046) to find out the potential target using several computational software and tools. In the first phase of the study, one protein target was identified based on unique pathways not common to *H. sapiens* and pathogens. In the second phase, structural analysis of the putative target DnaA was performed which was then employed for docking study. The four compounds namely Riccionidin A, Dothistromin, Fustin, and Morin exhibited high binding affinity with the binding pocket of DnaA protein. Thus, this study represents a significant advance in the design of new, potent compounds against *S. sonnei*.

Funding

Taif University Researchers Supporting Project number (TURSP-2020/258), Taif University, Taif, Saudi Arabia

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abbasi, E., Abtahi, H., van Belkum, A., Ghaznavi-Rad, E., 2019. Multidrug-resistant *Shigella* infection in pediatric patients with diarrhea from central Iran. *Infect. Drug Resistance* 12, 1535.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403–410.
- Amineni, U., Pradhan, D., Marisetty, H., 2010. In silico identification of common putative drug targets in *Leptospira interrogans*. *J. Chem. Biol.* 3 (4), 165–173. <https://doi.org/10.1007/s12154-010-0039-1>.
- Anandan, S., Muthuirulandi Sethuvel, D.P., Gajendiren, R., Verghese, V.P., Walia, K., Veeraghavan, B., 2017. Molecular characterization of antimicrobial resistance in clinical *Shigella* isolates during 2014 and 2015: trends in South India. *Germes* 7 (3), 115–122.

- Anishetty, S., Pulimi, M., Pennathur, G., 2005. Potential drug targets in *Mycobacterium tuberculosis* through metabolic pathway analysis. *Comput. Biol. Chem.* 29 (5), 368–378.
- Aoki-Kinoshita, K.F., Kanehisa, M., 2007. Gene annotation and pathway mapping in KEGG. In *Comparative Genomics* (pp. 71–91). Springer.
- Armenteros, J.J.A., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37 (4), 420–423.
- Ashfaq, U.A., Mumtaz, A., Qamar, T.u., Fatima, T., 2013. MAPS Database: Medicinal plant activities, phytochemical and structural database. *Bioinformatics* 9 (19), 993–995.
- Banerjee, P., Eckert, A.O., Schrey, A.K., Preissner, R., 2018. ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* 46 (W1), W257–W263.
- Barh, D., Tiwari, S., Jain, N., Ali, A., Santos, A.R., Misra, A.N., Azevedo, V., Kumar, A., 2011. In silico subtractive genomics for target identification in human bacterial pathogens. *Drug Dev. Res.* 72 (2), 162–177.
- Barrett, J.F., Hoch, J.A., 1998. Two-component signal transduction as a target for microbial anti-infective therapy. *Antimicrob. Agents Chemother.* 42 (7), 1529–1536.
- Chong, C.-E., Lim, B.-S., Nathan, S., Mohamed, R., 2006. In silico analysis of *Burkholderia pseudomallei* genome sequence for potential drug targets. In *Silico Biol.* 6 (4), 341–346.
- Garnier, J., Gibrat, J.-F., Robson, B., 1996. [32] GOR method for predicting protein secondary structure from amino acid sequence. *Methods in enzymology*, Vol. 266. Elsevier, pp. 540–553.
- Goyal, M., Citu, C., Singh, N., 2018. In silico identification of novel drug targets in *Acinetobacter baumannii* by subtractive genomic approach. *Asian J. Pharm. Clin. Res.* 11, 230 <https://doi.org/10.22159/ajpcr.2018.v11i3.22105>.
- Goyal, M., Citu, C., Singh, N., 2018. In silico identification of novel drug targets in *Acinetobacter baumannii* by subtractive genomic approach. *Asian J. Pharm. Clin. Res.* 11 (3), 230. <https://doi.org/10.22159/ajpcr.2018.v11i3.22105>.
- Gu, B., Cao, Y., Pan, S., Zhuang, L., Yu, R., Peng, Z., Qian, H., Wei, Y., Zhao, L., Liu, G., Tong, M., 2012. Comparison of the prevalence and changing resistance to nalidixic acid and ciprofloxacin of *Shigella* between Europe-America and Asia-Africa from 1998 to 2009. *Int. J. Antimicrob. Agents* 40 (1), 9–17. <https://doi.org/10.1016/j.ijantimicag.2012.02.005>.
- Holt, K.E., Baker, S., Weill, F.-X., Holmes, E.C., Kitchen, A., Yu, J., Sangal, V., Brown, D. J., Coia, J.E., Kim, D.W., Choi, S.Y., Kim, S.H., da Silveira, W.D., Pickard, D.J., Farrar, J.J., Parkhill, J., Dougan, G., Thomson, N.R., 2012. *Shigella* whole genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nat. Genet.* 44 (9), 1056–1059.
- Hosen, M.I., Tanmoy, A.M., Mahbuba, D.-A., Salma, U., Nazim, M., Islam, M.T., Akhteruzzaman, S., 2014. Application of a subtractive genomics approach for in silico identification and characterization of novel drug targets in *Mycobacterium tuberculosis* F11. *Interdisc. Sci.: Comput. Life Sci.* 6 (1), 48–56.
- Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26 (5), 680–682.
- Inc., C. C. G. (2016). Molecular operating environment (MOE). In: Chemical Computing Group Inc 1010 Sherbooke St. West, Suite# 910, Montreal
- Irwin, J.J., Shoichet, B.K., 2005. ZINC – a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* 45 (1), 177–182.
- Jamal, S.B., Hassan, S.S., Tiwari, S., Viana, M.V., Benevides, L.d.J., Ullah, A., Turjanski, A.G., Barh, D., Ghosh, P., Costa, D.A., Silva, A., Röttger, R., Baumbach, J., Azevedo, V.A.C., de Brevem, A.G., 2017. An integrative in-silico approach for therapeutic target identification in the human pathogen *Corynebacterium diphtheriae*. *PLoS ONE* 12 (10), e0186401. <https://doi.org/10.1371/journal.pone.0186401>. <https://doi.org/10.1371/journal.pone.0186401.g001>. <https://doi.org/10.1371/journal.pone.0186401.g002>. <https://doi.org/10.1371/journal.pone.0186401.g003>. <https://doi.org/10.1371/journal.pone.0186401.g004>. <https://doi.org/10.1371/journal.pone.0186401.g005>. <https://doi.org/10.1371/journal.pone.0186401.g006>. <https://doi.org/10.1371/journal.pone.0186401.g007>. <https://doi.org/10.1371/journal.pone.0186401.g008>. <https://doi.org/10.1371/journal.pone.0186401.g009>. <https://doi.org/10.1371/journal.pone.0186401.g010>. <https://doi.org/10.1371/journal.pone.0186401.g011>. <https://doi.org/10.1371/journal.pone.0186401.g012>. <https://doi.org/10.1371/journal.pone.0186401.t001>. <https://doi.org/10.1371/journal.pone.0186401.t002>. <https://doi.org/10.1371/journal.pone.0186401.t003>. <https://doi.org/10.1371/journal.pone.0186401.t004>. <https://doi.org/10.1371/journal.pone.0186401.t005>. <https://doi.org/10.1371/journal.pone.0186401.t006>. <https://doi.org/10.1371/journal.pone.0186401.t007>. <https://doi.org/10.1371/journal.pone.0186401.t008>. <https://doi.org/10.1371/journal.pone.0186401.t009>. <https://doi.org/10.1371/journal.pone.0186401.t010>. <https://doi.org/10.1371/journal.pone.0186401.t011>. <https://doi.org/10.1371/journal.pone.0186401.s001>. <https://doi.org/10.1371/journal.pone.0186401.s002>. <https://doi.org/10.1371/journal.pone.0186401.s003>.
- Källberg, M., Margaryan, G., Wang, S., Ma, J., Xu, J., 2014. RaptorX server: a resource for template-based protein structure modeling. In: *Protein Structure Prediction*. Springer, pp. 17–27.
- Katayama, T., Ozaki, S., Keyamura, K., Fujimitsu, K., 2010. Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* 8 (3), 163–170.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10 (6), 845–858.
- Khalil, I.A., Troeger, C., Blacker, B.F., Rao, P.C., Brown, A., Atherly, D.E., Brewer, T.G., Engmann, C.M., Hout, E.R., Kang, G., Kotloff, K.L., Levine, M.M., Luby, S.P., MacLennan, C.A., Pan, W.K., Pavlinac, P.B., Platts-Mills, J.A., Qadri, F., Riddle, M.S., Ryan, E.T., Shoultz, D.A., Steele, A.D., Watson, J.L., Sanders, J.W., Mokdad, A.H., Murray, C.J.L., Hay, S.I., Reiner Jr., R.C., 2018. Morbidity and mortality due to shigella and enterotoxigenic *Escherichia coli* diarrhoea: the Global Burden of Disease Study 1990–2016. *Lancet Infect. Dis.* 18 (11), 1229–1240. [https://doi.org/10.1016/s1473-3099\(18\)30475-4](https://doi.org/10.1016/s1473-3099(18)30475-4).
- Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B.o., Zhang, J., Bryant, S.H., 2016. PubChem substance and compound databases. *Nucleic Acids Res.* 44 (D1), D1202–D1213.
- Kimura, A.C., Johnson, K., Palumbo, M.S., Hopkins, J., Boase, J.C., Reporter, R., Goldoft, M., Stefonek, K.R., Farrar, J.A., Van Gilder, T.J., Vugia, D.J., 2004. Multistate shigellosis outbreak and commercially prepared food, United States. *Emerg. Infect. Dis.* 10 (6), 1147–1149.
- Ko, J., Park, H., Heo, L., Seok, S., Xu, J.J., Polzer, R.J., Winter, S.M., 2012. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res.* 40 (W1), W294–W297.
- Kotloff, K.L., Riddle, M.S., Platts-Mills, J.A., Pavlinac, P., Zaidi, A.K.M., 2018. Shigellosis. *The Lancet* 391 (10122), 801–812.
- Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305 (3), 567–580.
- Lengths, M., Angles, M., 2018. Limitations of structure evaluation tools errata. *Quick Guideline Comput. Drug Design* 16, 75.
- Lin, J., Sahakian, D.C., De Moraes, S., Xu, J.J., Polzer, R.J., Winter, S.M., 2003. The role of absorption, distribution, metabolism, excretion and toxicity in drug discovery. *Curr. Top. Med. Chem.* 3 (10), 1125–1154.
- Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23 (1–3), 3–25.
- Lomize, M.A., Pogozheva, I.D., Joo, H., Mosberg, H.L., Lomize, A.L., 2012. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* 40 (D1), D370–D376.
- Lynch, T., Price, A.L., 2007. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am. Fam. Physician* 76 (3), 391–396.
- Mahmud, A., Khan, M.T., Iqbal, A., 2019. Identification of novel drug targets for humans and potential vaccine targets for cattle by subtractive genomic analysis of *Brucella abortus* strain 2308. *Microbial Pathogenesis* 137, 103731. <https://doi.org/10.1016/j.micpath.2019.103731>.
- Mott, M.L., Berger, J.M., 2007. DNA replication initiation: mechanisms and regulation in bacteria. *Nat. Rev. Microbiol.* 5 (5), 343–354.
- Mumtaz, A., Ashfaq, U.A., ul Qamar, M.T., Anwar, F., Gulzar, F., Ali, M.A., Saari, N., Pervez, M.T., 2017. MPD3: a useful medicinal plants database for drug designing. *Nat. Prod. Res.* 31 (11), 1228–1236.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25 (13), 1605–1612.
- Pires, S.M., Fischer-Walker, C.L., Lanata, C.F., Devleesschauer, B., Hall, A.J., Kirk, M. D., Duarte, A.S.R., Black, R.E., Angulo, F.J., Selvey, L.A., 2015. Aetiology-Specific Estimates of the Global and Regional Incidence and Mortality of Diarrhoeal Diseases Commonly Transmitted through Food. *PLoS ONE* 10 (12), e0142927. <https://doi.org/10.1371/journal.pone.0142927>. <https://doi.org/10.1371/journal.pone.0142927.g001>. <https://doi.org/10.1371/journal.pone.0142927.g002>. <https://doi.org/10.1371/journal.pone.0142927.g003>. <https://doi.org/10.1371/journal.pone.0142927.g004>. <https://doi.org/10.1371/journal.pone.0142927.g005>. <https://doi.org/10.1371/journal.pone.0142927.t001>. <https://doi.org/10.1371/journal.pone.0142927.t002>. <https://doi.org/10.1371/journal.pone.0142927.t003>. <https://doi.org/10.1371/journal.pone.0142927.t004>. <https://doi.org/10.1371/journal.pone.0142927.t005>. <https://doi.org/10.1371/journal.pone.0142927.t006>. <https://doi.org/10.1371/journal.pone.0142927.t007>. <https://doi.org/10.1371/journal.pone.0142927.t008>. <https://doi.org/10.1371/journal.pone.0142927.t009>. <https://doi.org/10.1371/journal.pone.0142927.t010>. <https://doi.org/10.1371/journal.pone.0142927.t011>. <https://doi.org/10.1371/journal.pone.0142927.s001>. <https://doi.org/10.1371/journal.pone.0142927.s002>. <https://doi.org/10.1371/journal.pone.0142927.s003>. <https://doi.org/10.1371/journal.pone.0142927.s004>. <https://doi.org/10.1371/journal.pone.0142927.s005>.
- Porter, C.K., Olson, S., Hall, A., Riddle, M.S., 2017. Travelers' Diarrhea: An Update on the Incidence, Etiology, and Risk in Military Deployments and Similar Travel Populations. *Mil. Med.* 182 (S2), 4–10. <https://doi.org/10.7205/ajam2017.00064>.
- Prabhurajeshwar, C., 2018. Shigellosis: Its Prevention and Management issues. *Int. J. Med. Sci. Diagnosis Res.* 2 (4).
- ProtParam, E. (2017). ExPASy-ProtParam tool. In.
- Qiu, S., Wang, Y., Xu, X., Li, P., Hao, R., Yang, C., Liu, N., Li, Z., Wang, Z., Wang, J., Wu, Z., Su, W., Yang, G., Jin, H., Wang, L., Sun, Y., Yuan, Z., Huang, L., Song, H., 2013. Multidrug-resistant atypical variants of *Shigella flexneri* in China. *Emerg. Infect. Dis.* 19 (7), 1147–1150. <https://doi.org/10.3201/eid1907.111221>.
- Regev, T., Myers, N., Zarivach, R., Fishov, I., Chatterji, D., 2012. Association of the chromosome replication initiator DnaA with the *Escherichia coli* inner membrane in vivo: quantity and mode of binding. *PLoS ONE* 7 (5), e36441. <https://doi.org/10.1371/journal.pone.0036441>. <https://doi.org/10.1371/journal.pone.0036441.g001>. <https://doi.org/10.1371/journal.pone.0036441.g002>. <https://doi.org/10.1371/journal.pone.0036441.g003>. <https://doi.org/10.1371/journal.pone.0036441.g004>. <https://doi.org/10.1371/journal.pone.0036441.g005>. <https://doi.org/10.1371/journal.pone.0036441.g006>. <https://doi.org/10.1371/journal.pone.0036441.g007>. <https://doi.org/10.1371/journal.pone.0036441.g008>. <https://doi.org/10.1371/journal.pone.0036441.g009>. <https://doi.org/10.1371/journal.pone.0036441.g010>. <https://doi.org/10.1371/journal.pone.0036441.g011>. <https://doi.org/10.1371/journal.pone.0036441.g012>. <https://doi.org/10.1371/journal.pone.0036441.t001>. <https://doi.org/10.1371/journal.pone.0036441.t002>. <https://doi.org/10.1371/journal.pone.0036441.t003>. <https://doi.org/10.1371/journal.pone.0036441.t004>. <https://doi.org/10.1371/journal.pone.0036441.t005>. <https://doi.org/10.1371/journal.pone.0036441.t006>. <https://doi.org/10.1371/journal.pone.0036441.t007>. <https://doi.org/10.1371/journal.pone.0036441.t008>. <https://doi.org/10.1371/journal.pone.0036441.t009>. <https://doi.org/10.1371/journal.pone.0036441.t010>. <https://doi.org/10.1371/journal.pone.0036441.t011>. <https://doi.org/10.1371/journal.pone.0036441.s001>. <https://doi.org/10.1371/journal.pone.0036441.s002>. <https://doi.org/10.1371/journal.pone.0036441.s003>. <https://doi.org/10.1371/journal.pone.0036441.s004>. <https://doi.org/10.1371/journal.pone.0036441.s005>.
- Renaux, A., 2018. UniProt: the universal protein knowledgebase (vol 45, pg D158, 2017). *Nucleic Acids Res.* 46 (5), 2699–2699.
- Sakharkar, K.R., Sakharkar, M.K., Chow, V.T., 2004. A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*. In *Silico Biol.* 4 (3), 355–360.
- Sati, H.F., Bruinsma, N., Galas, M., Hsieh, J., Sanhueza, A., Ramon Pardo, P., Espinal, M. A., Smyth, D.S., 2019. Characterizing *Shigella* species distribution and antimicrobial susceptibility to ciprofloxacin and nalidixic acid in Latin America between 2000–2015. *PLoS ONE* 14 (8), e0220445. <https://doi.org/10.1371/journal.pone.0220445>. <https://doi.org/10.1371/journal.pone.0220445.g001>. <https://doi.org/10.1371/journal.pone.0220445.g002>. <https://doi.org/10.1371/journal.pone.0220445.g003>. <https://doi.org/10.1371/journal.pone.0220445.g004>. <https://doi.org/10.1371/journal.pone.0220445.g005>. <https://doi.org/10.1371/journal.pone.0220445.g006>. <https://doi.org/10.1371/journal.pone.0220445.g007>. <https://doi.org/10.1371/journal.pone.0220445.g008>. <https://doi.org/10.1371/journal.pone.0220445.g009>. <https://doi.org/10.1371/journal.pone.0220445.g010>. <https://doi.org/10.1371/journal.pone.0220445.g011>. <https://doi.org/10.1371/journal.pone.0220445.g012>. <https://doi.org/10.1371/journal.pone.0220445.s001>. <https://doi.org/10.1371/journal.pone.0220445.s002>. <https://doi.org/10.1371/journal.pone.0220445.s003>. <https://doi.org/10.1371/journal.pone.0220445.s004>. <https://doi.org/10.1371/journal.pone.0220445.s005>. <https://doi.org/10.1371/journal.pone.0220445.s006>. <https://doi.org/10.1371/journal.pone.0220445.s007>. <https://doi.org/10.1371/journal.pone.0220445.s008>. <https://doi.org/10.1371/journal.pone.0220445.s009>. <https://doi.org/10.1371/journal.pone.0220445.s010>. <https://doi.org/10.1371/journal.pone.0220445.s011>. <https://doi.org/10.1371/journal.pone.0220445.s012>.

- journal.pone.0220445.s00210.1371/journal.pone.0220445.s00310.1371/journal.pone.0220445.s00410.1371/journal.pone.0220445.s00510.1371/journal.pone.0220445.s006.
- Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* 31 (13), 3381–3385.
- Shad, A.A., Shad, W.A., 2021. *Shigella sonnei*: virulence and antibiotic resistance. *Arch. Microbiol.* 203 (1), 45–58.
- Shahid, F., Ashfaq, U.A., Saeed, S., Munir, S., Almatroudi, A., Khurshid, M., 2020. In Silico Subtractive Proteomics Approach for Identification of Potential Drug Targets in *Staphylococcus saprophyticus*. *Int. J. Environ. Res. Public Health* 17 (10), 3644.
- Shiragannavar, S.S., Shettar, A.K., Madagi, S.B., Sarawad, S., 2019. Subtractive genomics approach in identifying polysaccharide biosynthesis protein as novel drug target against *Eubacterium nodatum*. *Asian J. Pharm. Pharmacol.* 5 (2), 382–392.
- Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G., Laub, M.T., Arkin, A., 2005. Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol.* 3 (10), e334. <https://doi.org/10.1371/journal.pbio.0030334>.
<https://doi.org/10.1371/journal.pbio.0030334.g00110.1371/journal.pbio.0030334.g00310.1371/journal.pbio.0030334.g00410.1371/journal.pbio.0030334.g00510.1371/journal.pbio.0030334.g00610.1371/journal.pbio.0030334.g00710.1371/journal.pbio.0030334.t00210.1371/journal.pbio.0030334.t00110.1371/journal.pbio.0030334.t00310.1371/journal.pbio.0030334.t00410.1371/journal.pbio.0030334.sg00110.1371/journal.pbio.0030334.sg00210.1371/journal.pbio.0030334.sg00310.1371/journal.pbio.0030334.sg00410.1371/journal.pbio.0030334.sg00510.1371/journal.pbio.0030334.st00110.1371/journal.pbio.0030334.st00210.1371/journal.pbio.0030334.st003>.
- Sybilski, A.J., Ambroziak, G., & Strzelecka, J. Life-threatening infection with *Shigella flexneri* in fully immunocompetent girl.
- Taneja, N., Mewara, A., 2016. Shigellosis: Epidemiology in India. *Indian J. Med. Res.* 143 (5), 565–576. <https://doi.org/10.4103/0971-5916.187104>.
- Torraca, V., Holt, K., Mostow, S., 2020. *Shigella sonnei*. *Trends Microbiol.* 28 (8), 696–697.
- Vasanthanathan, P., Taboureau, O., Oostenbrink, C., Vermeulen, N.P.E., Olsen, L., Jørgensen, F.S., 2009. Classification of cytochrome P450 1A2 inhibitors and noninhibitors by machine learning techniques. *Drug Metab. Dispos.* 37 (3), 658–664.
- Wei, W., Ning, L.-W., Ye, Y.-N., Guo, F.-B., Lespinet, O., 2013. Geptop: a gene essentiality prediction tool for sequenced bacterial genomes based on orthology and phylogeny. *PLoS ONE* 8 (8), e72343. <https://doi.org/10.1371/journal.pone.0072343>.
<https://doi.org/10.1371/journal.pone.0072343.g00110.1371/journal.pone.0072343.g00210.1371/journal.pone.0072343.g00310.1371/journal.pone.0072343.g00410.1371/journal.pone.0072343.g00510.1371/journal.pone.0072343.g00610.1371/journal.pone.0072343.g00710.1371/journal.pone.0072343.g00810.1371/journal.pone.0072343.t00110.1371/journal.pone.0072343.t00210.1371/journal.pone.0072343.t00310.1371/journal.pone.0072343.s001>.
- Yang, H., Lou, C., Sun, L., Li, J., Cai, Y., Wang, Z., Li, W., Liu, G., Tang, Y., 2019. admetsAR 2.0: web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* 35 (6), 1067–1069.
- Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26 (13), 1608–1615.
- Zakrzewska-Czerwińska, J., Jakimowicz, D., Zawilak-Pawlik, A., Messer, W., 2007. Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiol. Rev.* 31 (4), 378–387.