

RESEARCH ARTICLE

Subtractive genomic analysis for computational identification of putative immunogenic targets against clinical *Enterobacter cloacae* complex

Negin Bolourchi¹ , Sepideh Fereshteh¹ , Narjes Noori Goodarzi², Farzad Badmasti^{1,3*} 

1 Department of Bacteriology, Pasteur Institute of Iran, Tehran, Iran, **2** Department of Pathobiology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran, **3** Microbiology Research Center (MRC), Pasteur Institute of Iran, Tehran, Iran

 These authors contributed equally to this work.

* fbadmasti2008@gmail.com



OPEN ACCESS

Citation: Bolourchi N, Fereshteh S, Noori Goodarzi N, Badmasti F (2022) Subtractive genomic analysis for computational identification of putative immunogenic targets against clinical *Enterobacter cloacae* complex. PLoS ONE 17(10): e0275749. <https://doi.org/10.1371/journal.pone.0275749>

Editor: René Massimiliano Marsano, University of Bari: Università degli Studi di Bari Aldo Moro, ITALY

Received: May 20, 2022

Accepted: September 22, 2022

Published: October 13, 2022

Copyright: © 2022 Bolourchi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information](#) files.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: CDD, conserved domain database; CDSs, coding sequences; cg-MLST, core genome

Abstract

Background

Enterobacter is a major nosocomial genus of *Enterobacteriaceae* responsible for a variety of nosocomial infections, particularly in prolonged hospitalized patients in the intensive care units. Since current antibiotics have failed treating colistin- and carbapenem-resistant *Enterobacteriaceae*, efforts are underway to find suitable alternative strategies. Therefore, this study conducted a reverse vaccinology (RV) to identify novel and putative immunogenic targets using core proteome of 20 different sequence types (STs) of clinical *Enterobacter* spp. Moreover, we introduced a structural-based approach for exploration of potential vaccine candidates against the *Enterobacteriaceae* family using their conserved domain analysis.

Results

A number of 2616 core coding sequences (CDSs) were retrieved from 20 clinical strains of *Enterobacter* spp. with a similarity of $\geq 50\%$. Nine proteins with a score of ≥ 20 considered as the shortlisted proteins based on the quartile scoring method, including three TonB-dependent receptors, WP_008500981.1, WP_058690971.1 and WP_058679571.1; one YjbH domain-containing protein, WP_110108068.1; three flagellar proteins, WP_088207510.1, WP_033145204.1 and WP_058679632.1; one spore-coat U domain-containing protein, WP_039266612.1; and one DD-metalloendopeptidase family protein, WP_025912449.1. In this study, proteins WP_058690971.1 and WP_110108068.1 were detected as the top candidates with regard to immune stimulation and interactions with TLRs. However, their efficacy is remaining to be evaluated experimentally.

Conclusions

Our investigation introduced common ferrichrome porins with high sequence similarity as potential vaccine candidates against the *Enterobacteriaceae* family. These proteins belong to the iron acquisition system and possess all criteria of suitable vaccine targets. Therefore,

multi-locus sequence typing; ECC, *Enterobacter cloacae* complex; NJ, neighbor joining; RV, reverse vaccinology; ST, sequence type; UTIs, urinary tract infections.

they need to be specifically paid attention for vaccine development against clinically important members of *Enterobacteriaceae* family.

Introduction

Enterobacter is a genus of common Gram-negative, facultatively anaerobic, rod-shaped, non-spore-forming, and motile bacteria belonging to the *Enterobacteriaceae* family. There are 22 highly similar species within this genus, comprising *Enterobacter cloacae* complex (ECC) [1, 2]. Some ECC strains originate from soil and water, and some are natural commensals of the animal and human gut. However, they are capable of causing a variety of infections such as respiratory tract infections, urinary tract infections (UTIs), soft-tissue infections, septicemia, and meningitis with severe medical sequela [3]. Among members of this complex, *E. cloacae*, *E. hormaechei*, *E. kobei*, and *E. roggenkampii* are frequently isolated from clinical specimens. *E. hormaechei* and *E. kobei* account for more than 70% of community-acquired infections. *E. cloacae* is responsible for 10% of post-surgical peritonitis and four to five percent of nosocomial sepsis, pneumonia cases, and UTIs [4]. These species have wide geographical distribution all over the world, frequently reported from Australia, the United States, Germany, and China [5].

The treatment of ECC infections is problematic due to the increasing resistance to various antimicrobial agents [6]. Recent studies have shown that *Enterobacter* spp. are often the second or the third most common nosocomial *Enterobacteriaceae* harboring carbapenemase enzymes [7]. Therefore, there is an urgent need to eliminate such highly resistant bacteria via new options other than antibiotics. With this regard, various effective strategies can be introduced. Among them, vaccination is promising as it represents microorganism-specific prevention, confining the spread of infection and reducing clinical manifestations, drug side effects, patient hospitalization as well as average medical expenses [8, 9]. However, despite the highly frequent and extended antimicrobial resistance of ECC, no effective vaccine has been developed so far.

The ability of conventional methods to discover immunological compounds of microorganisms is limited [10]. The application of whole genomic data along with computational bioinformatics enables us to explore proteins for vaccine design. Reverse vaccinology (RV) is a desirable method because it computationally identifies vaccine candidates, the majority of which cannot be detected through wet-lab experiments [11]. A few studies have been conducted on *in silico* identification of novel putative vaccine candidates against ECC up to now. In 2020, three outer-membrane porin proteins, including LpfC, OmpA, and FimD as well as an arginine transporter were introduced using a subtractive study of *E. cloacae* reference proteome [12]. More recently, Alshammari *et al.*, used bioinformatics to design a multi-chimeric vaccine against *E. xiangfangensis* using the ferrichrome porin (FhuA) and peptidoglycan-associated lipoprotein (Pal) [13]. Also, Phosphoporin E (PhoE) and a putative outer-membrane porin protein presented appropriate characteristics for vaccine design [14].

In the present study, we aimed to introduce novel putative immunogenic candidates against different sequence types (STs) of clinical ECC strains using their core proteome. In addition, we investigated consensus protein structures with promising immunogenicity against other clinically important *Enterobacteriaceae*.

Materials and methods

Initial protein screening

Selection of *Enterobacter* spp. Strains. The dataset included almost all *E. cloacae* complex causing human clinical infections isolated from various clinical samples (not animal/

environmental sources), different geographical regions and different periods of time (from 2008 to 2020). Please see [S1 Table](#). We picked up 20 *Enterobacter* spp. belonging to seven species including *E. cloacae*, *E. hormaechei*, *E. kobei*, *E. roggenkampii*, *E. cancerogenus*, *E. asburiae*, and *E. bugandensis*, considering all the above criteria for our analysis. The complete genomes of selected ECC strains were retrieved from the NCBI database (<https://www.ncbi.nlm.nih.gov/>). The genetic characteristics and clinical information of strains have been provided in [S1 Table](#).

Genetic and phylogenetic comparison of ECC strains. PubMLST (<https://pubmlst.org/>) was used to determine strains' STs based on the allele numbers of six housekeeping genes (*rpoB*, *fusA*, *gyrB*, *leuS*, *pyrG*, and *rplB*). PubMLST is a collection of integrated population sequence data, containing the provenance, phylogenetic and phenotypic information for over 100 different microbial species and genera [15]. To compare the strains on the whole genomic level, multiple circular alignment was performed using the BLAST Ring Image Generator (BRIG) software version 0.95. BRIG is a free application that can visualize similarities and differences of genomes and compare their genetic features [16]. In this study, *E. cloacae* subsp. *cloacae* ATCC 13047 (accession number: NC_014121.1) was used as the reference strain. To elucidate their phylogenetic distances, a Neighbor Joining (NJ) dendrogram based on the core genome multi-locus sequence typing (cg-MLST) was performed using cano-wgMLST_Bac-Compare (<http://baccompare.imst.nsysu.edu.tw>). This server enables users to determine the evolutionary relationship of bacteria using a whole genome (wg-) MLST and canonical MLST [17]. The number, distribution, and functional classification of pan, core, accessory and unique coding sequences (CDSs) of 20 ECC strains were achieved by the Bacterial Pan Genome Analysis (BPGA) tool [18]. Finally, the core proteome was obtained using BPGA with an identity cut-off value of $\geq 50\%$.

Prediction of subcellular localization of proteins. All proteins were uploaded to PSORTb (www.psort.org/psortb/) as a database to predict the subcellular localization of proteins [19]. Only cell wall, extracellular, secreted, and surface-exposed proteins were selected. The topology of the proteins was predicted using the HMMTOP (<http://www.enzim.hu/hmmtop/>) database which is an automatic server for predicting transmembrane helices and the topology of proteins [20].

Antigenicity and allergenicity determination of proteins. The antigenicity of the proteins was predicted with the VaxiJen online tool (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) using a cut-off value of ≥ 0.5 . VaxiJen is a server for alignment-independent prediction of protective antigens [21]. Subsequently, the allergenicity of the antigenic proteins was determined using the AlgPred 2.0 tool (<https://webs.iitd.edu.in/raghava/algpred2/batch.html>) with a cut-off value of ≥ 0.5 . This web tool has been developed for prediction of amino acid residues with allergenic characteristics in a protein sequence [22].

Sequence similarity of putative immunogenic targets with the human proteome. PSI-BLAST (<https://blast.ncbi.nlm.nih.gov>) with the defined threshold (coverage $\geq 30\%$ and identity $\geq 25\%$) was used to assess the homology of selected proteins versus human proteome (*Homo sapiens*, taxid: 9606) [23, 24]. PSI-BLAST provides a BLASTp search with a custom, position-specific, scoring matrix which can help to find distant evolutionary relationships [25].

Comparative analysis

Characterization of physicochemical properties, functional class and adhesin probability. The molecular weight and other physicochemical properties of the selected proteins such as theoretical pI, half-life, instability and hydropathy indexes were estimated using the ProtParam tool (<https://web.expasy.org/protparam/>). This web tool computes various physical and

chemical parameters for a particular protein [26]. The adhesion probability and functional class of the proteins were determined using the Vaxign (<http://www.violinet.org/vaxign2>) and VICMpred (<http://www.imtech.res.in/raghava/vicmpred/>) databases, respectively. Vaxign predicts adhesion probability of proteins using an optimized SPAAN. It has been approved that the SPAAN prediction has a sensitivity of 89% and specificity of 100% [27]. VICMpred provides a SVM-based method to stimulate functional classification of bacterial proteins and classifies them into virulence factors, information molecules, cellular processes and metabolic molecules [28].

Determination of linear B-cell epitopes and human MHC II binding sites. The BepiPred v2.0 tool (<http://www.cbs.dtu.dk/services/BepiPred/>) was used to identify linear B-cell epitopes of all previously selected proteins with a threshold of ≥ 0.61 . This tool uses a random forest algorithm trained on epitopes of antibody-antigen protein structures [29]. For each protein, the B-cell epitopes' ratio was calculated (the proportion of the number of amino acids in B-cell epitopes to the total number of amino acids of each protein). Next, TepiTool (<http://tools.iedb.org/tepitool/>) from the Immune Epitope Database (IEDB) resource, was used to predict human MHC II binding sites (T-cell epitopes) with a cut-off value of the top 5% of peptides. The ratio of MHC II binding sites was calculated for each protein (the proportion of binding sites to the total amino acids of each protein). This database uses experimental data on antibody and T-cell epitopes studied in humans, non-human primates, and other animal species in relation to infectious diseases, autoimmunity, allergies, and transplantation [30].

Prediction of conformational B-cell epitopes. In this section, the tertiary structures of the putative immunogenic candidates were first predicted using the Robetta web tool (<https://rosetta.bakerlab>) [31]. The quality of each tertiary structure prediction was checked using QMEAN (<https://swissmodel.expasy.org/qmean/>) and the ProSA-web server (<https://prosa.services.came.sbg.ac.at/prosa.php>). The QMEAN server provides the quality estimation of protein structure models by ranking potentially unreliable regions within them [32]. ProSA-web server is useful for the recognition of errors in three-dimensional (3D) structures of proteins [33].

Next, the conformational B-cell epitopes of selected proteins were identified using the IEDB analysis resource (<http://tools.iedb.org/ellipro/>) with a threshold value of ≥ 0.8 [34]. The tertiary structures of proteins were visualized using PyMOL version 2.3.4 (Schrödinger, LLC). PyMOL is a user-sponsored molecular visualization system on an open-source foundation.

Shortlist selection based on the quartile scoring method

The quartile method is based on the evaluation of eight properties including adhesion probability, antigenicity index, hydropathy index, instability index, functional class (virulence, cellular process, information and storage, and metabolic molecule), B-cell epitopes ratio, T-cell epitopes ratio, and the number of conformational B-cell epitopes. The sum of all scores for each protein was considered in the final score. Proteins with a score of ≥ 20 were considered putative immunogenic proteins and underwent further analysis. The quartile method measures the dispersion of values by dividing their distribution into lower, median, and upper quartiles to form four intervals. The quartile method provides us a rational target selection by comparing proteins based on several unweighted criteria at the same time [35].

Immune simulation and molecular docking

The immune-reactivity of the shortlisted proteins were simulated using C-ImmSim (<https://kraken.iac.rm.cnr.it/C-IMMSIM/index.php>). C-ImmSim utilizes a position-specific scoring matrices (PSSM) originated from machine learning techniques to simulate immunological

reactions [36]. C-ImmSim also predicts the anatomical regions where crucial events of immunity occur [37]. The number of injections was considered 1 time without LPS for each immunogenic candidate. The protein-protein rigid docking was also performed for the prediction of their binding affinity to surface-exposed human TLR-1, 2, and 4 using pyDockWEB (<https://life.bsc.es/pid/pydockweb/default/index>). This is a web-server for the prediction of protein-protein rigid interactions [38].

Conserved domains analysis

The conserved domain database, CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>), was used for prediction of taxonomic relationships of proteins and their functions based on their conserved domains [39]. In addition, the best putative candidates were investigated for their similar proteins in other genera of the *Enterobacteriaceae* family using this web tool.

Results

Comparative genomic analysis of ECC strains and their phylogenetic relationship

The multiple sequence alignment of selected clinical *Enterobacter* spp. showed high similarity of strains at the whole genomic level. See Fig 1A. The strains belonged to different STs including 513, 484, 595, 795, 764, 54, 520, 78, 806, 1, 23 191, 1140 and 1794. See Fig 1B. The phylogenetic dendrogram based on 1680 genes with $\geq 95\%$ sequence similarity indicated a close ancestral relationship between different members of ECC. Strains belonging to each single *Enterobacter* spp. were located close in consensus nodes except for *E. cloacae* complex sp. 35734.

During core proteome analysis, a number of 2616 core CDSs were found among 20 clinical strains of *Enterobacter* spp. with $\geq 50\%$ sequence similarity. The lowest and highest number of accessory CDSs were for *E. hormaechei* strain AKB48 (1253) and *E. cloacae* strain PIM-B10EC27 (1776), respectively. In addition, *E. asburiae* strain A2563 (27 CDSs) and *E. cancerogenus* strain 1428 (362 CDSs) had the minimum and maximum numbers of unique CDSs, respectively. See Fig 1B.

The pan/core gene profiling of ECC strains indicated that the number of pan gene families differs meaningfully among the strains. While, the distribution of core gene families was almost close. See Fig 2A. The core, accessory and unique orthologous genes were clustered in 19 known functional classes and one unknown class. See Fig 2B. The majority of core CDSs were involved in general functions as well as amino acid metabolism/transportation while of them very few were associated with intercellular trafficking and defense mechanisms. The majority of accessory genes were related to general functions and transcription. The majority of unique CDSs were involved in general functions, replication/repair as well as transcription. The fewest percentage of accessory and unique genes were involved in nucleotide transport/metabolism, translation and ribosomal biogenesis.

Subtracted proteins

Among 2616 core proteins, only 48 proteins were outer membrane or extracellular. The number of the transmembrane helices in all proteins were ≤ 1 and no proteins weighed >110 kDa. Among them, 42 proteins were antigenic and five were allergenic. Overall, 37 antigenic, non-allergen proteins with no similarity to human proteins were subtracted for comparative analysis. The flowchart summarizing step by step subtraction of vaccine candidates has been shown in Fig 3.

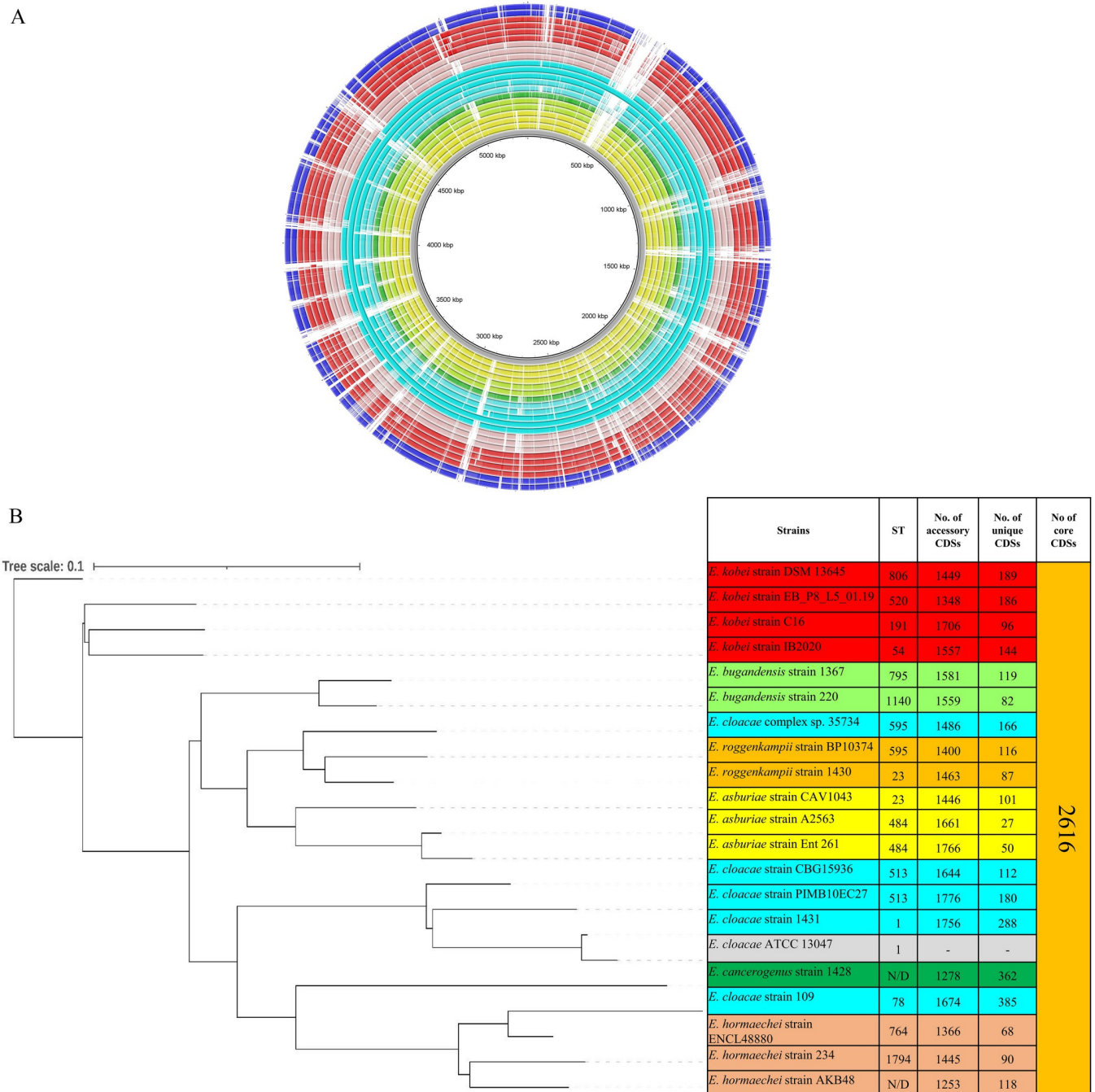


Fig 1. A. The whole genomic multiple sequence alignment of 20 different clinical *Enterobacter* spp. isolates. The results showed their high similarity. B. The phylogenetic dendrogram based on cg-MLST using 1680 genes with $\geq 95\%$ sequence similarity. The results indicated a close ancestral relationship between different members of *Enterobacter* spp. A number of 2616 core coding sequences were retrieved from their genome with $\geq 50\%$ similarity.

<https://doi.org/10.1371/journal.pone.0275749.g001>

Shortlisted proteins based on the quartile scoring method

Nine proteins with a score of ≥ 20 were considered as the shortlisted proteins based on eight properties, including adhesion probability, antigenicity index, hydrophathy index, instability index, functional class (virulence, cellular process, information and storage, and metabolic

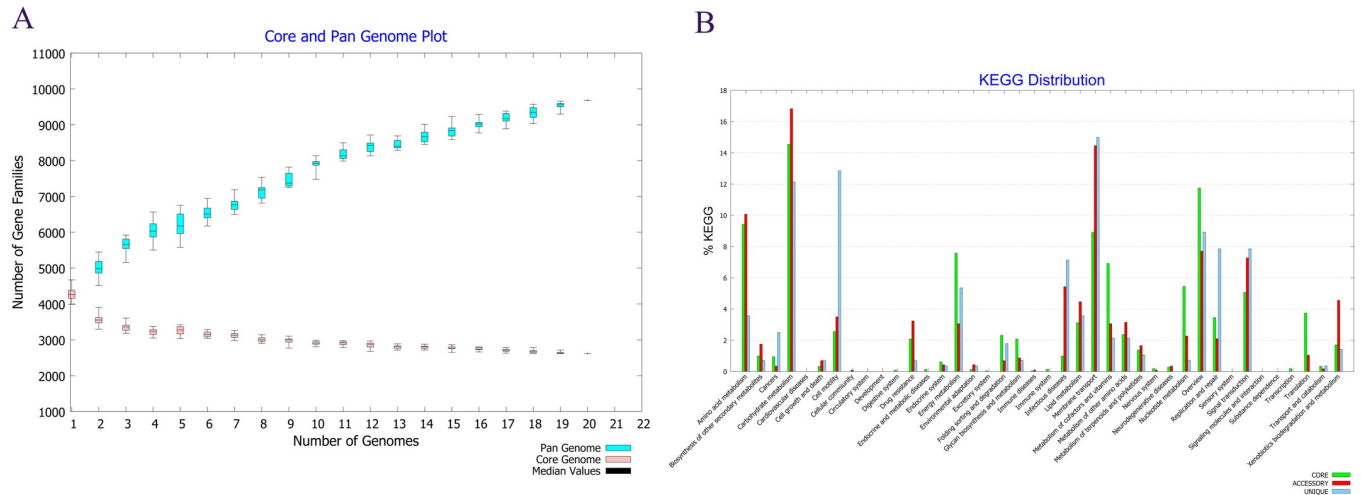


Fig 2. A. The distribution of pan/core gene families of 20 clinical *Enterobacter* species. The pan/core gene profiling of ECC strains indicated that the number of pan gene families differs meaningfully among the strains. While, the distribution of core gene families was almost close. B. The functional classification of core, accessory and unique genes among 20 clinical *Enterobacter* species. The orthologous core, accessory and unique genes were clustered in 19 known functional classes and one unknown class.

<https://doi.org/10.1371/journal.pone.0275749.g002>

molecule), B-cell epitopes ratio, T-cell epitopes ratio, and the number of conformational B-cell epitopes. See Fig 4A. The shortlisted putative candidates were as follows: five outer membrane proteins including WP_008500981.1 (TonB-dependent siderophore receptor), WP_058690971.1 (TonB-dependent siderophore receptor), WP_058679571.1 (TonB-dependent vitamin B12 receptor BtuB) and WP_110108068.1 (YjbH domain-containing protein); WP_025912449.1 (peptidoglycan DD-metalloendopeptidase family protein); and four

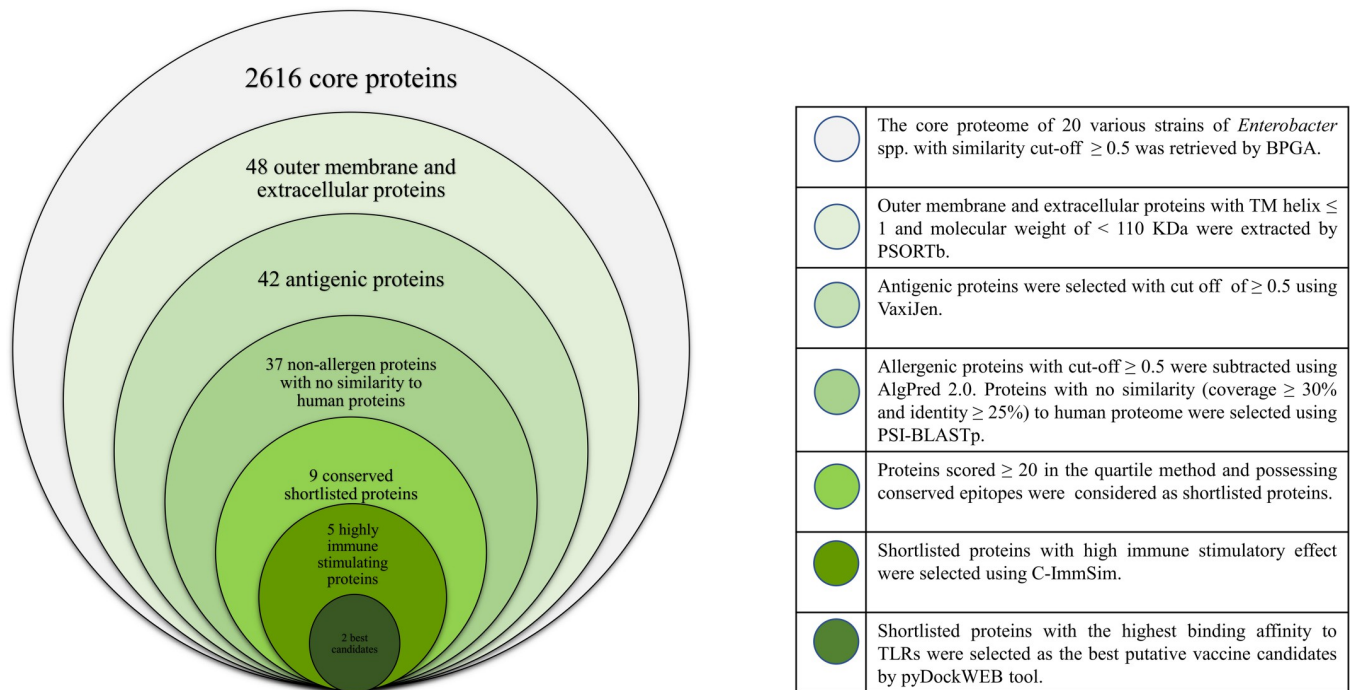


Fig 3. The workflow used for the identification of putative immunogenic candidates against *E. cloacae* complex.

<https://doi.org/10.1371/journal.pone.0275749.g003>

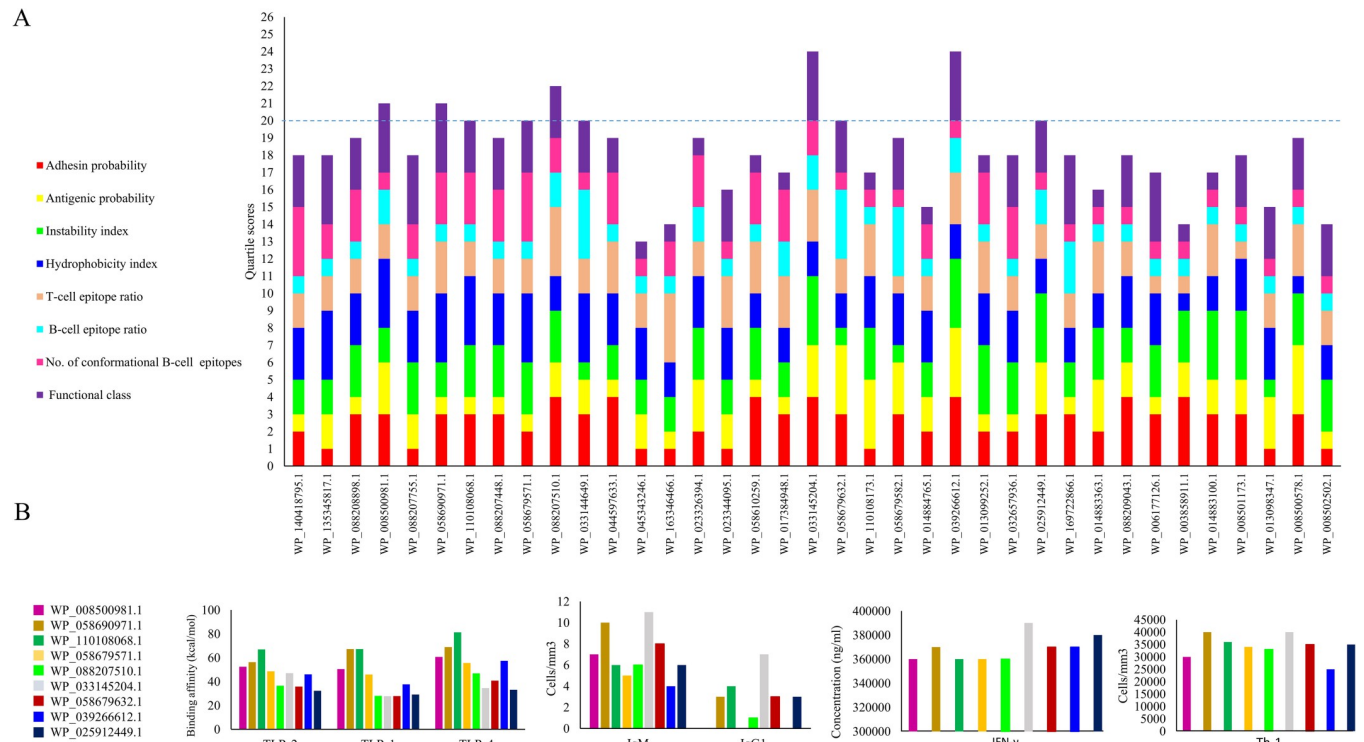


Fig 4. A. The comparison of 37 proteins of *E. cloacae* complex based on the quartile method. Nine proteins with a score of ≥ 20 were considered as the shortlisted proteins based on the following eight properties: adhesion probability, antigenicity index, hydropathy index, instability index, functional class (virulence, cellular process, information and storage, and metabolic molecule), B-cell epitopes ratio, T-cell epitopes ratio, and the number of conformational B-cell epitopes. B. The comparison of immunological responses induced by nine shortlisted vaccine candidates against *E. cloacae* complex as well as their binding affinity to TLRs. The results showed that five proteins including WP_058690971.1, WP_110108068.1, WP_033145204.1, WP_058679632.1 and WP_025912449.1 could stimulate the production of IFN- γ , IgM, IgG1. Among these, WP_058690971.1 and WP_110108068.1 had the highest binding affinity to TLR-1, 2 and 4.

<https://doi.org/10.1371/journal.pone.0275749.g004>

extracellular proteins including WP_088207510.1 (flagellar hook-associated protein FlgK), WP_033145204.1 (flagellar hook protein FlgE), WP_058679632.1 (flagellar hook length control protein FliK), and WP_039266612.1 (spore-coat U domain-containing protein). See Fig 4A.

Comparison of the shortlisted proteins

Molecular weights of proteins ranged from 26.13 to 82.87 kDa. All of them were predicted to be stable. WP_033145204.1 has the lowest instability index (14.73). TonB-dependent siderophore receptors, FlgE and FliK were virulence factors while other remaining proteins were involved in cellular processes. WP_008500981.1 and WP_025912449.1 had the maximum and minimum number of linear B-cell epitopes, respectively. The B-cell epitope ratio ranged from 0.10 (for WP_058690971.1) to 0.51 (for WP_058679632.1). WP_058690971.1 and WP_025912449.1 had the highest and lowest number of T-cell epitopes, respectively. WP_088207510.1 had the maximum T-cell epitope ratio (0.15). The number of conformational B-cell epitopes ranged from two (for WP_025912449.1) to eight (for WP_058679571.1). Physicochemical properties, functional class, number of linear and conformational B-cell epitopes, MHC II binding sites of shortlisted proteins have been presented in S2 Table. The tertiary structure of shortlisted proteins as well as their conformational B-cell epitopes have been depicted in Fig 5.

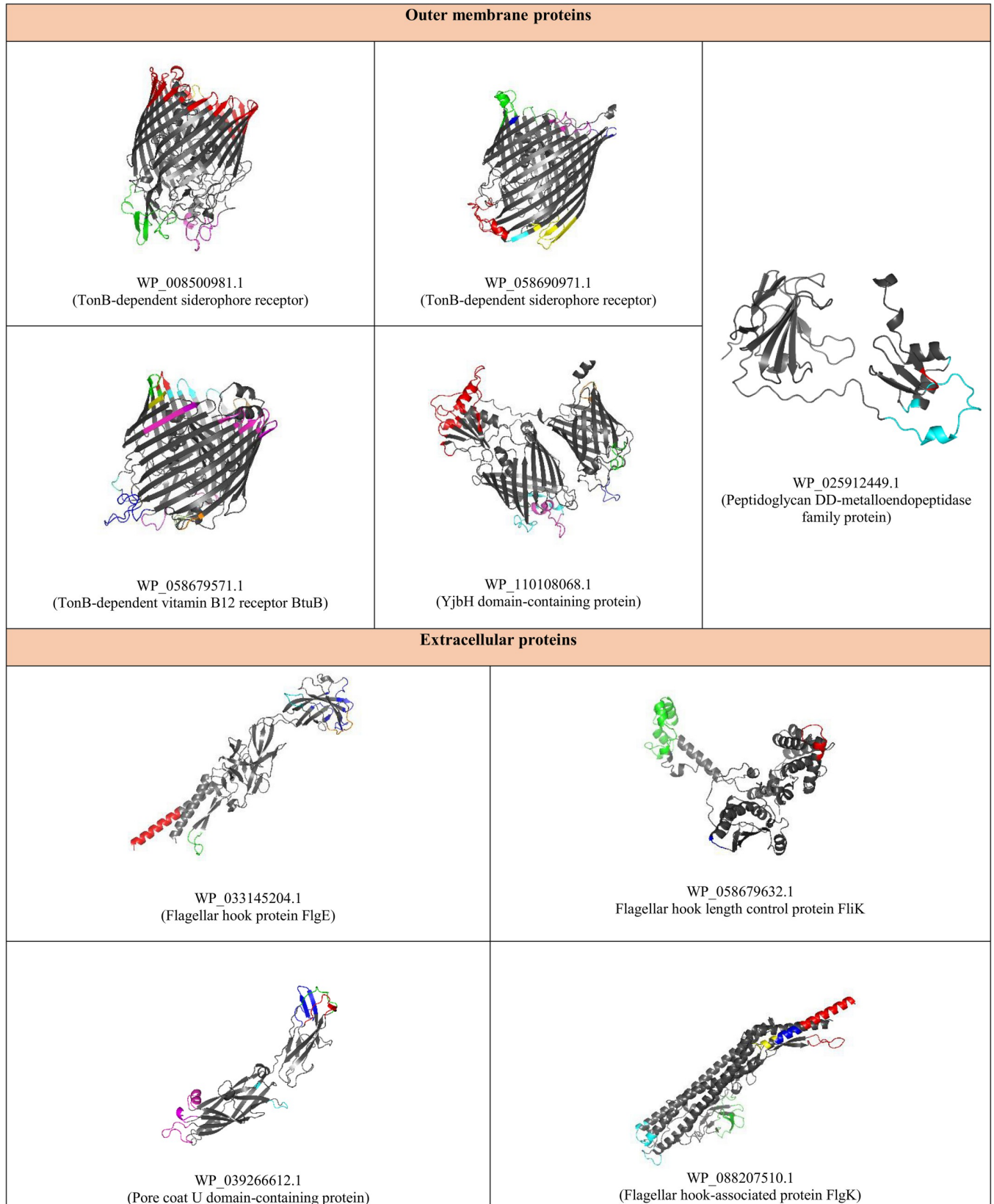


Fig 5. The tertiary structure and conformational B-cell epitopes of nine shortlisted proteins against *E. cloacae* complex. Nine proteins with a score of ≥ 20 were subtracted. The shortlisted putative candidates were as follows: five outer membrane proteins including WP_008500981.1 (TonB-dependent siderophore receptor), WP_058690971.1 (TonB-dependent siderophore receptor), WP_058679571.1 (TonB-dependent vitamin B12 receptor BtuB) and WP_110108068.1 (YjbH domain-containing protein); WP_025912449.1 (peptidoglycan DD-metalloendopeptidase family protein); and four extracellular proteins including WP_088207510.1 (flagellar hook-associated protein FlgK), WP_033145204.1 (flagellar hook protein FlgE), WP_058679632.1 (flagellar hook length control protein FliK), WP_039266612.1 (spore-coat U domain-containing protein).

<https://doi.org/10.1371/journal.pone.0275749.g005>

Immune simulation and molecular docking

The results of the immune simulation showed that five proteins including WP_058690971.1, WP_110108068.1, WP_033145204.1, WP_058679632.1 and WP_025912449.1 could stimulate the production of IFN- γ , IgM and IgG1. See Fig 4B and Table 1. Among them, WP_058690971.1 and WP_110108068.1 had the highest binding affinity to TLR-1, 2 and 4. The interaction of these two proteins with TLR-1, 2 and 4 has been presented in Fig 6. Considering both criteria, WP_058690971.1 and WP_110108068.1 were selected as the best vaccine candidates against the clinical *Enterobacter* spp. The results of immune simulation and molecular docking have been summarized in Table 1.

Conserved domains

The results obtained from CDD demonstrated that nine shortlisted proteins have superfamilies with three main functions: iron uptake (WP_008500981.1, WP_058690971.1, and WP_058679571.1), exopolysaccharide production (WP_110108068.1) and flagellar assembly (WP_088207510.1, WP_033145204.1, and WP_058679632.1). WP_039266612.1 had a spore-coat U domain involved in spore coating. WP_025912449.1 had domains responsible for

Table 1. Results of immune simulations and molecular dockings of nine putative vaccine candidates against clinical *Enterobacter* spp.

Immune response prediction	Immune simulations				Binding affinities		
	IFN- γ (ng/ml)	IgM (cells/mm ³)	IgG1 (cells/mm ³)	Th1 (cells/mm ³)	TLR-2 (kcal/mol)	TLR-1 (kcal/mol)	TLR-4 (kcal/mol)
Vaccine candidates							
WP_008500981.1 (TonB-dependent siderophore receptor)	360000	7	0	30000	-52.427	-50.506	-60.703
WP_058690971.1 (TonB-dependent siderophore receptor)	370000	10	3	40000	-56.349	-67.264	-68.984
WP_110108068.1 (YjbH domain-containing protein)	360000	6	4	36000	-66.994	-67.264	-81.345
WP_058679571.1 (TonB-dependent vitamin B12 receptor BtuB)	360000	5	0	34000	-48.639	-45.921	-55.599
WP_088207510.1 (Flagellar hook-associated protein FlgK)	360000	6	1	33000	-36.503	-28.029	-46.974
WP_033145204.1 (Flagellar hook protein FlgE)	390000	11	7	40000	-47.016	-27.741	-34.604
WP_058679632.1 (Flagellar hook length control protein FliK)	370000	8	3	35000	-35.844	-27.842	-40.801
WP_039266612.1 (Pore coat U domain-containing protein)	370000	4	0	25000	-46.123	-37.685	-57.411
WP_025912449.1 (Peptidoglycan DD-metalloendopeptidase family protein)	380000	6	3	35000	-32.314	-29.089	-33.082

*Bold parameters indicate values above the mean.

<https://doi.org/10.1371/journal.pone.0275749.t001>

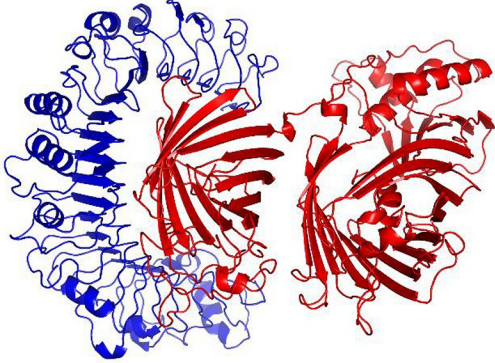


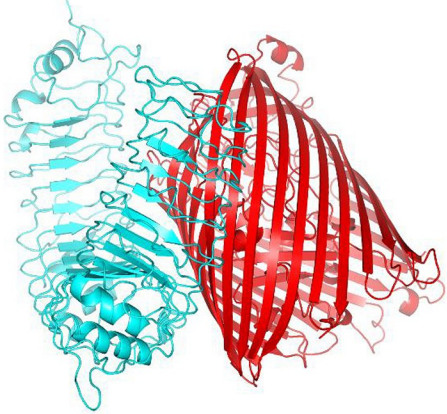
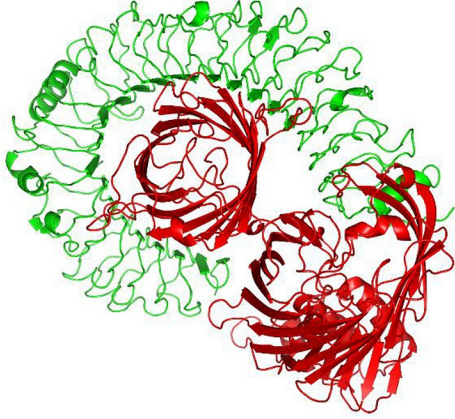

Protein Human TLRs	WP_110108068.1 (YjbH domain-containing protein)	WP_058690971.1 (TonB-dependent siderophore receptor)
TLR-1		
TLR-2		
TLR-4		

Fig 6. The interaction of WP_058690971.1 (TonB-dependent siderophore receptor) and WP_110108068.1 (YjbH domain-containing protein) with TLR-1, 2 and 4.

<https://doi.org/10.1371/journal.pone.0275749.g006>

peptidoglycan hydrolase (NlpD, M23 peptidase and LysM). The full results of the CDD search have been summarized in [Table 2](#).

Evaluation of TonB-dependent siderophore receptors in other members of *Enterobacteriaceae* as putative vaccine candidates

We identified the conserved domain of the ferrichrome outer membrane transporter superfamily in WP_058690971.1 (TonB-dependent siderophore receptor). This superfamily exists in ferrichrome porin FhuA of all members of *Enterobacteriaceae*. We detected this ferrichrome porin in major clinically important *Enterobacteriaceae* including *Escherichia coli* (WP_000124438.1 and WP_000124388.1) *Klebsiella pneumoniae* (WP_004178624.1 and WP_012542816.1), *Salmonella enterica* (WP_000113211.1), *Shigella flexneri* (WP_011587185.1) and *Citrobacter koseri* (WP_012134008.1). All selected proteins scored ≥ 20 based on the quartile method. See [Fig 7A](#). The results of immune simulation and molecular docking for all seven proteins were similar to that of our candidate. See [S3 Table](#). The tertiary structures of *Enterobacteriaceae* ferrichrome porins have been presented in [Fig 7B](#). In addition, all information regarding these proteins have been shown in [S3 Table](#).

Discussion

This study utilized the core proteome of various *Enterobacter* spp. to screen several immunogenic candidates in the first place. The results of cg-MLST indicated a close phylogenetic relationship among selected strains on the core genomic level and the number of their core CDSs were also in the same range. Therefore, by excluding the accessory genetic content; the core proteome represents the most prevalent proteins with the high conservancy [40]. Keeping this in mind, the extraction of appropriate candidates for vaccination against ECC using their core genomic content seems logical. As, it provides proteins with high distribution and conservancy among all ECC.

An ideal or promising immunogenic candidate has several criteria. It should be exposed to extracellular space to effectively elicit protective immune responses. Besides, it has to be highly antigenic without any allergenicity. It needs to be highly conserved among widely distributed strains. In addition, an ideal immunogenic candidate should play an important role in the pathogenesis of bacteria and ideally be expressed during bacterial infection [35]. We used the quartile scoring method to subtract our candidates as it provides an option to consider all the above criteria at the same time.

In this study, we introduced nine putative immunogenic proteins as vaccine candidates against ECC, three of which were TonB-dependent proteins. In Gram-negative bacteria, these receptors are involved in the transport and uptake of large substrates such as iron siderophore complexes and vitamin B12 [3]. TonB-dependent receptors are considered excellent candidates for vaccine development due to their critical role in bacterial virulence and vast extracellular exposure. TonB-dependent receptors have been investigated in various Gram-negative pathogens [41]. The TonB-dependent siderophore receptor and the siderophore enterobactin receptor FepA of *Klebsiella* spp. fulfilled all vaccine parameters [42]. In *Acinetobacter baumannii*, two TonB-dependent receptors (BauA and BfnH) showed a partial protective effect on animal models. Aghajani *et al.* indicated that the mortality rate and bacterial load of *A. baumannii* in immunized mice was lower compared to that of the control group [43]. YncD of *Salmonella enterica*, an *in vivo*-induced antigen, elicited a significant immune-protection against the lethal wild-type challenge [44]. Furthermore, TbpA of *Neisseria meningitidis* conferred protection against serogroup B [45].

Table 2. Subcellular localization, conserved domains, related taxonomy and functions of nine putative vaccine candidates against clinical *Enterobacter* spp.

Protein Accession Number	Subcellular Localization	Conserved Domain	Domain taxonomy	Function
WP_008500981.1 (TonB-dependent siderophore receptor)	Outer Membrane	FepA family TonB-dependent siderophore receptor	<i>Enterobacteriaceae</i>	TonB-dependent siderophore receptor acts as a channel to allow import of iron-siderophore complexes, such as <i>Escherichia coli</i> ferrienterobactin receptor, which is involved in the initial step of iron uptake by binding ferrienterobactin.
		FepA	Proteobacteria	FepA is the outer membrane receptor for ferrienterochelin and colicins [Inorganic ion transport and metabolism].
		Ligand_gated_channel	Bacteria	TonB dependent/Ligand-Gated channels are created by a monomeric 22 strand (22, 24) anti-parallel beta-barrel. Ligands apparently bind to the large extracellular loops. The N-terminal 150–200 residues form a plug from the periplasmic end of barrel. Energy (proton-motive force) and TonB-dependent conformational alteration of channel (parts of plug, and loops 7 and 8) allow passage of ligand. FepA residues 12–18 form the TonB box, which mediates the interaction with the TonB-containing inner membrane complex. TonB preferentially interacts with ligand-bound receptors. Transport thru the channel may resemble passage thru an air lock. In this model, ligand binding leads to closure of the extracellular end of pore, then a TonB-mediated signal facilitates opening of the interior side of pore, deforming the N-terminal plug and allowing passage of the ligand to the periplasm. Such a mechanism would prevent the free diffusion of small molecules.
		TonB_dep_Rec	Bacteria	TonB-dependent siderophore receptor; This subfamily model encompasses a wide variety of TonB-dependent outer membrane siderophore receptors. It has no overlap with TonB receptors known to transport other substances, but is likely incomplete due to lack of characterizations. It is likely that genuine siderophore receptors will be identified which score below the noise cutoff to this model at which point the model should be updated. [Transport and binding proteins, Cations and iron carrying compounds, Transport and binding proteins, Porins].
WP_058690971.1 (TonB-dependent siderophore receptor)	Outer Membrane	Ferrichrome outer membrane transporter	<i>Enterobacteriaceae</i>	-
		Ligand_gated_channel	Bacteria	-
		TonB-siderophore	Bacteria	TonB-dependent siderophore receptor; This subfamily model encompasses a wide variety of TonB-dependent outer membrane siderophore receptors. It has no overlap with TonB receptors known to transport other substances, but is likely incomplete due to lack of characterizations. It is likely that genuine siderophore receptors will be identified which score below the noise cutoff to this model at which point the model should be updated. [Transport and binding proteins, Cations and iron carrying compounds, Transport and binding proteins, Porins].
		CirA	Bacteria	Outer membrane receptor proteins, mostly Fe transport [Inorganic ion transport and metabolism].
		TonB_dep_Rec	Bacteria	-
WP_110108068.1 (YjbH domain-containing protein)	Outer Membrane	YjbH	Bacteria	YjbH domain-containing protein, similar to <i>Escherichia coli</i> K-12 YjbH which is a putative lipoprotein and/or porin involved in exopolysaccharide production. It is an exopolysaccharide biosynthesis protein. YjbH is a family of Gram-negative beta-barrel.
WP_058679571.1 (TonB-dependent vitamin B12 receptor BtuB)	Outer Membrane	BtuB	<i>Enterobacteriaceae</i>	TonB-dependent vitamin B12 receptor BtuB is involved in the active translocation of vitamin B12 (cyanocobalamin) across the outer membrane to the periplasmic space.
		TonB-B12	Gammaproteobacteria	TonB-dependent vitamin B12 receptor; This model represents the TonB-dependent outer membrane receptor found in Gammaproteobacteria responsible for translocating the cobalt-containing vitamin B12 (cobalamin). [Transport and binding proteins, Other, Transport and binding proteins, Porin].
		TonB_sider_MxcH	Bacteria	TonB-dependent siderophore myxochelin receptor MxcH.
		Ligand_gated_channel	Bacteria	-
		TonB_dep_Rec	Bacteria	-

(Continued)

Table 2. (Continued)

Protein Accession Number	Subcellular Localization	Conserved Domain	Domain taxonomy	Function
WP_088207510.1 (Flagellar hook-associated protein FlgK)	Extracellular	FlgK	Proteobacteria	Flagellar hook-associated protein FlgK forms the junction between the hook and the filament in the flagellum together with FlgL and provides a structural base where flagellin, a filament-forming protein, is inserted for the initiation of filament elongation.
		FlgK_ends	Bacteria	Flagellar hook-associated protein FlgK; The flagellar hook-associated protein FlgK of bacterial flagella has conserved N- and C-terminal domains. The central region is highly variable in length and sequence, and often contains substantial runs of low-complexity sequence. This model is built from an alignment of FlgK sequences with the central region excised.
		Flagellar basal body rod FlgEFG protein C-terminal	Bacteria	This family consists of a number of C-terminal domains of unknown function. This domain seems to be specific to flagellar basal-body rod and flagellar hook proteins in which pfam00460 is often present at the extreme N-terminus.
WP_033145204.1 (Flagellar hook protein FlgE)	Extracellular	FlgE	Bacteria	Flagellar hook protein FlgE functions as a nano-sized universal joint, which is essential for dynamic and efficient bacterial motility and taxis.
		Flagellar basal body rod FlgEFG protein C-terminal	Bacteria	This family consists of a number of C-terminal domains of unknown function. This domain seems to be specific to Flagellar basal-body rod and flagellar hook proteins in which pfam00460 is often present at the extreme N-terminus.
WP_058679632.1 (Flagellar hook length control protein FliK)	Extracellular	PRK10118	<i>Enterobacteriaceae</i>	Flagellar hook-length control protein FliK controls elongation by determining hook length and by stopping the supply of hook protein to the filament protein.
		C-terminal domain of flagellar hook-length control protein FliK and similar domains;	Cellular organisms	The flagellar hook-length control protein FliK is a soluble cytoplasmic protein that is secreted during Flagellar formation. It controls hook elongation by two successive events: by determining hook length and by stopping the supply of hook protein. It contains an N-terminal domain that determines hook length and a C-terminal domain that is responsible for switching secretion from the hook protein to that of the filament protein, by interacting with FlhB, the switchable secretion gate.
		Flg_hook	Cellular organisms	It controls the length of the Flagellar hook by directly measuring the hook length as a molecular ruler. This family also includes YscP of the <i>Yersinia</i> type III secretion system, and equivalent proteins in other pathogenic bacterial type III secretion systems.
WP_039266612.1 (Spore coat U domain-containing protein)	Extracellular	SCPU	Proteobacteria	This domain is found in a bacterial family of spore coat proteins. Spore coat U (SCPU) domain-containing protein may act as a bacterial spore coat protein or as a secreted pili protein subunit involved in motility and biofilm formation; the family is distantly related to fimbrial proteins
WP_025912449.1 (Peptidoglycan DD-metalloendopeptidase family protein)	Outer Membrane	NlpD	<i>Enterobacteriaceae</i>	Murein hydrolase activator NlpD is the activator of the cell wall hydrolase AmiC. It is Required for septal murein cleavage and daughter cell separation during cell division.
		Peptidase_M23	Bacteria	Members of this family are zinc metallopeptidases with a range of specificities. The peptidase family M23 is included in this family, these are Gly-Gly endopeptidases. Peptidase family M23 are also endopeptidases. This family also includes some bacterial lipoproteins for which no proteolytic activity has been demonstrated. This family also includes leukocyte cell-derived chemotaxin 2 (LECT2) proteins. LECT2 is a liver-specific protein which is thought to be linked to hepatocyte growth although the exact function of this protein is unknown.
		LysM	Bacteria	-

<https://doi.org/10.1371/journal.pone.0275749.t002>

We realized that the TonB-dependent siderophore receptor (WP_058690971.1) has the conserved domain of the ferrichrome outer membrane transporter superfamily belonging to the *Enterobacteriaceae* family. In this study, the *in silico* investigation on ferrichrome porins in different members of *Enterobacteriaceae* demonstrates that they all have the criteria of being vaccine candidates. Despite diverse amino acid compositions of these proteins, our data

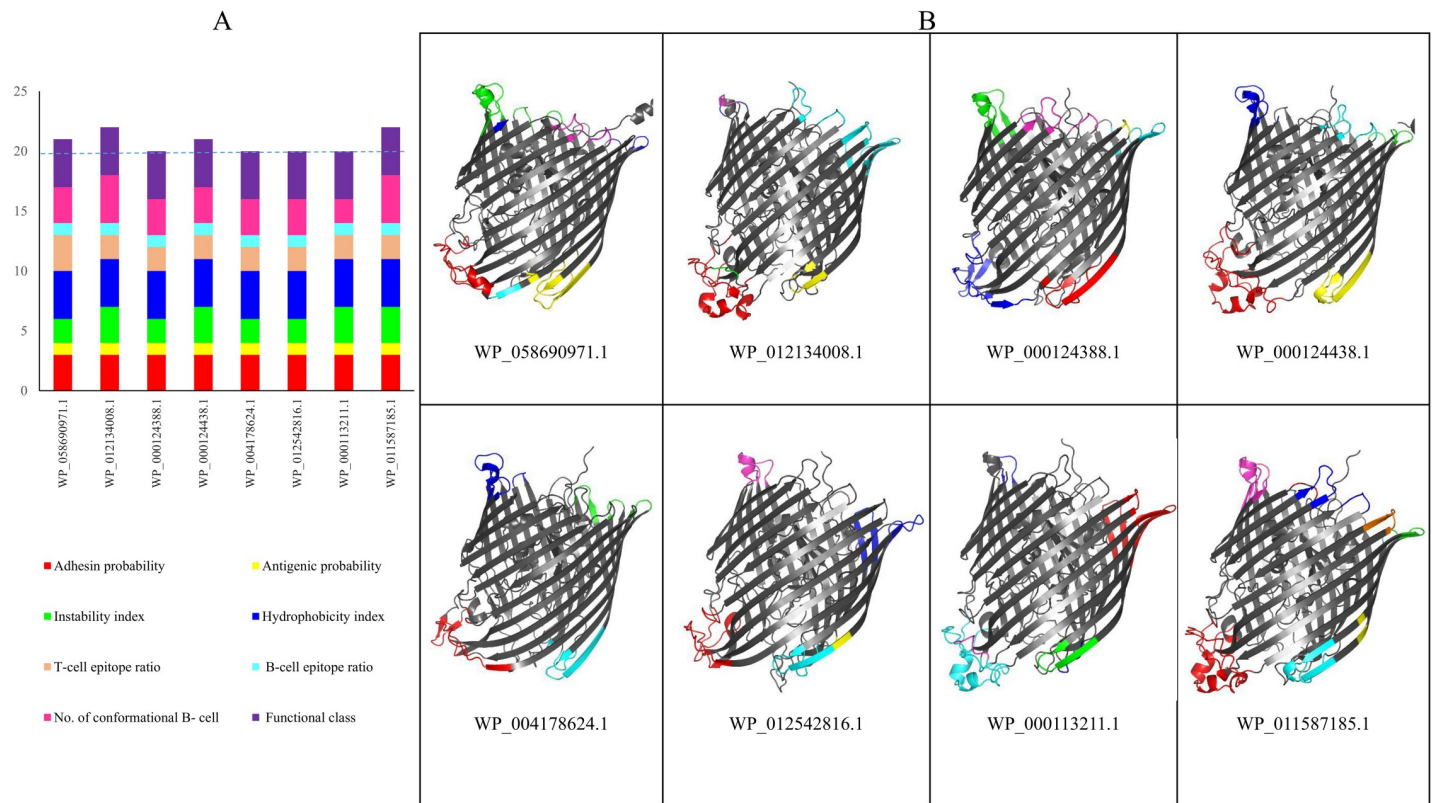


Fig 7. A. The comparison of eight proteins belonging to ferrichrome outer membrane transporter conserved superfamily in eight clinically important genera of *Enterobacteriaceae*. All selected proteins scored ≥ 20 based on the quartile method with physicochemical properties, functional class, number of linear and conformational B/T-cell epitopes similar to our candidate, TonB-dependent siderophore receptor (WP_058690971.1). B. The tertiary structure and conformational B-cell epitopes of ferrichrome porins in major clinically important genera of *Enterobacteriaceae* family.

<https://doi.org/10.1371/journal.pone.0275749.g007>

showed that they are all similar to WP_058690971.1 with regard to physicochemical properties as well as the number of T-cell and conformational B-cell epitopes. The results of immune simulation and molecular docking for these proteins were also close to WP_058690971.1. See [S3 Table](#).

From the other side, proteins belonging to the ferrichrome outer membrane transporter superfamily also have conserved domain of the ligand-gated channel superfamily. This domain was found in four of our candidates including WP_008500981.1, WP_058690971.1, WP_058679571.1 and WP_110108068.1, which had the strongest interaction with TLRs. Their beta-strands are connected with loops on their extracellular side which are required for substrate attachment and transport [46]. It seems that multiple exposed conformational epitopes on the beta-barrel structure of these proteins can interact with the innate immune system effectively. In our previous *in silico* analysis, ligand-gated channel proteins FhuA, BfnH, PapC, DcaP, FatA and IutA were introduced as the best vaccine candidates against *A. baumannii*. Therefore, it seems that proteins with beta-barrel structure belonging to ferrichrome outer membrane transporter and ligand-gated channel superfamilies have potential for vaccine development against bacterial pathogens. Accordingly, structure-dependent exploration of vaccine candidates in bacterial genomes could serve as a method for *in silico* vaccine design itself. This method reduces our need for the comprehensive analysis of whole genomes to approach an appropriate vaccine candidate.

Similarly, NlpD conserve domain seem to be an important structure in proteins with vaccine potential. NlpD of WP_025912449.1 has an important role in cell stability. *Yersinia pestis*

and *Haemophilus influenzae* lacking NlpD showed less virulence in murine models and failed in growth, respectively [47, 48]. Proteins involved in bacterial flagella assembly were among our shortlists. These filamentous structures drive cell locomotion in fluids (swimming process) or on surfaces (swarming), allowing cells to move into favorable environments [49]. In this context, an *in silico* study on *Pseudomonas aeruginosa*, *Morganella morganii* and *Clostridioides difficile* introduced flagellar assembly proteins as valuable vaccine candidates [35, 50, 51]. Therefore, the identification of specific conserved domains in proteins provides essential information regarding their protective potential and possibly estimate their suitability for vaccine design.

Conclusion

This study introduced nine putative immunogenic candidates against clinical members of ECC using a pan/core-genomic analysis and RV approach. Considering the immune-simulation results, WP_058690971.1, TonB-dependent siderophore receptor, and WP_110108068, YjbH domain-containing protein, were the best immunogenic candidates against ECC with respect to TLRs interactions. However, experimental assays are remaining to determine the efficacy of these candidates. Our investigation introduced ferrichrome porins as ideal vaccine candidates against members of *Enterobacteriaceae*. Proteins belonging to the ferrichrome outer membrane transporter superfamily possess all criteria of suitable vaccine candidates. Therefore, they need to be seriously paid attention to eliminate clinically important *Enterobacteriaceae* by vaccine development. Based on our study, structural-based exploration of vaccine candidates can be considered as a fast and more convenient method for *in silico* vaccine development.

Supporting information

S1 Table. The genetic characteristics and clinical information of twenty *Enterobacter* spp.
(DOCX)

S2 Table. Physicochemical properties and number of linear B/T-cell and conformational B-cell epitopes for nine putative vaccine candidates against clinical *Enterobacter* spp.
(DOCX)

S3 Table. Physicochemical properties, number of linear and conformational B/T-cell epitopes, results of immune simulation and TLR binding affinities for eight ferrichrome outer membrane transporters in different members of *Enterobacteriaceae*.
(DOCX)

Acknowledgments

The authors would like to thank the personnel of Pasteur Institute of Iran for their support.

Author Contributions

Conceptualization: Farzad Badmasti.

Investigation: Negin Bolourchi.

Methodology: Farzad Badmasti.

Supervision: Farzad Badmasti.

Visualization: Negin Bolourchi.

Writing – original draft: Negin Bolourchi, Sepideh Fereshteh, Narjes Noori Goodarzi.

Writing – review & editing: Farzad Badmasti.

References

1. Davin-Regli A, Lavigne JP, Pagès JM. *Enterobacter* spp.: Update on Taxonomy, Clinical Aspects, and Emerging Antimicrobial Resistance. *Clin Microbiol Rev.* 2019; 32(4). Epub 20190717. <https://doi.org/10.1128/cmr.00002-19> PMID: 31315895; PubMed Central PMCID: PMC6750132.
2. Davin-Regli A, Pagès JM. *Enterobacter aerogenes* and *Enterobacter cloacae*; versatile bacterial pathogens confronting antibiotic treatment. *Front Microbiol.* 2015; 6:392. Epub 20150518. <https://doi.org/10.3389/fmicb.2015.00392> PMID: 26042091; PubMed Central PMCID: PMC4435039.
3. Mishra M, Panda S, Barik S, Sarkar A, Singh DV, Mohapatra H. Antibiotic Resistance Profile, Outer Membrane Proteins, Virulence Factors and Genome Sequence Analysis Reveal Clinical Isolates of *Enterobacter* Are Potential Pathogens Compared to Environmental Isolates. *Front Cell Infect Microbiol.* 2020; 10:54. Epub 20200221. <https://doi.org/10.3389/fcimb.2020.00054> PMID: 32154188; PubMed Central PMCID: PMC7047878.
4. Ji Y, Wang P, Xu T, Zhou Y, Chen R, Zhu H, et al. Development of a One-Step Multiplex PCR Assay for Differential Detection of Four species (*Enterobacter cloacae*, *Enterobacter hormaechei*, *Enterobacter roggkampii*, and *Enterobacter kobei*) Belonging to *Enterobacter cloacae* Complex With Clinical Significance. *Front Cell Infect Microbiol.* 2021; 11:677089. Epub 20210518. <https://doi.org/10.3389/fcimb.2021.677089> PMID: 34095000; PubMed Central PMCID: PMC8169972.
5. Xu T, Zhang C, Ji Y, Song J, Liu Y, Guo Y, et al. Identification of *mcr-10* carried by self-transmissible plasmids and chromosome in *Enterobacter roggkampii* strains isolated from hospital sewage water. *Environ Pollut.* 2021; 268(Pt B):115706. Epub 20201012. <https://doi.org/10.1016/j.envpol.2020.115706> PMID: 33069047.
6. Dimitrova D, Stoeva T, Markovska R, Stankova P, Mihova K, Kaneva R, et al. Molecular Epidemiology of Multidrug Resistant *Enterobacter cloacae* blood isolates from a University Hospital. *Journal of IMAB—Annual Proceeding Scientific Papers.* 2019; 25(2):2457–64.
7. Jean SS, Hsueh PR. Distribution of ESBLs, AmpC β -lactamases and carbapenemases among *Enterobacteriaceae* isolates causing intra-abdominal and urinary tract infections in the Asia-Pacific region during 2008–14: results from the Study for Monitoring Antimicrobial Resistance Trends (SMART). *J Antimicrob Chemother.* 2017; 72(1):166–71. Epub 20161004. <https://doi.org/10.1093/jac/dkw398> PMID: 27703058.
8. Tilahun M, Kassa Y, Gedefie A, Ashagire M. Emerging Carbapenem-Resistant *Enterobacteriaceae* Infection, Its Epidemiology and Novel Treatment Options: A Review. *Infect Drug Resist.* 2021; 14:4363–74. Epub 20211021. <https://doi.org/10.2147/IDR.S337611> PMID: 34707380; PubMed Central PMCID: PMC8544126.
9. Jansen KU, Knirsch C, Anderson AS. The role of vaccines in preventing bacterial antimicrobial resistance. *Nat Med.* 2018; 24(1):10–9. <https://doi.org/10.1038/nm.4465> PMID: 29315295.
10. Milshteyn A, Schneider JS, Brady SF. Mining the metabiome: identifying novel natural products from microbial communities. *Chem Biol.* 2014; 21(9):1211–23. <https://doi.org/10.1016/j.chembiol.2014.08.006> PMID: 25237864; PubMed Central PMCID: PMC4171686.
11. Rappuoli R, Aderem A. A 2020 vision for vaccines against HIV, tuberculosis and malaria. *Nature.* 2011; 473(7348):463–9. <https://doi.org/10.1038/nature10124> PMID: 21614073.
12. Ismail S, Ahmad S, Azam SS. Vaccinomics to design a novel single chimeric subunit vaccine for broad-spectrum immunological applications targeting nosocomial *Enterobacteriaceae* pathogens. *Eur J Pharm Sci.* 2020; 146:105258. Epub 20200205. <https://doi.org/10.1016/j.ejps.2020.105258> PMID: 32035109.
13. Alshammari A, Alharbi M, Alghamdi A, Alharbi SA, Ashfaq UA, Tahir UI Qamar M, et al. Computer-Aided Multi-Epitope Vaccine Design against *Enterobacter xiangfangensis*. *Int J Environ Res Public Health.* 2022; 19(13). Epub 20220623. <https://doi.org/10.3390/ijerph1913723> PMID: 35805383; PubMed Central PMCID: PMC9265868.
14. Al-Megrin WAI, Karkashan A, Alnuqaydan AM, Aba Alkhayl FF, Alrumaihi F, Almatroudi A, et al. Design of a Multi-Epitopes Based Chimeric Vaccine against *Enterobacter cloacae* Using Pan-Genome and Reverse Vaccinology Approaches. *Vaccines (Basel).* 2022; 10(6). Epub 20220601. <https://doi.org/10.3390/vaccines10060886> PMID: 35746494; PubMed Central PMCID: PMC9227637.
15. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* 2018; 3:124. Epub 20180924.

- <https://doi.org/10.12688/wellcomeopenres.14826.1> PMID: 30345391; PubMed Central PMCID: PMC6192448.
16. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*. 2011; 12:402. Epub 20110808. <https://doi.org/10.1186/1471-2164-12-402> PMID: 21824423; PubMed Central PMCID: PMC3163573.
 17. Liu YY, Lin JW, Chen CC. cano-wgMLST_BacCompare: A Bacterial Genome Analysis Platform for Epidemiological Investigation and Comparative Genomic Analysis. *Front Microbiol*. 2019; 10:1687. Epub 20190724. <https://doi.org/10.3389/fmicb.2019.01687> PMID: 31396192; PubMed Central PMCID: PMC6668299.
 18. Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep*. 2016; 6:24373. Epub 20160413. <https://doi.org/10.1038/srep24373> PMID: 27071527; PubMed Central PMCID: PMC4829868.
 19. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*. 2010; 26(13):1608–15. Epub 20100513. <https://doi.org/10.1093/bioinformatics/btq249> PMID: 20472543; PubMed Central PMCID: PMC2887053.
 20. Tusnády GE, Kalmár L, Simon I. TOPDB: topology data bank of transmembrane proteins. *Nucleic Acids Res*. 2008; 36(Database issue):D234–9. Epub 20071005. <https://doi.org/10.1093/nar/gkm751> PMID: 17921502; PubMed Central PMCID: PMC2238857.
 21. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007; 8:4. Epub 20070105. <https://doi.org/10.1186/1471-2105-8-4> PMID: 17207271; PubMed Central PMCID: PMC1780059.
 22. Saha S, Raghava GP. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res*. 2006; 34(Web Server issue):W202–9. <https://doi.org/10.1093/nar/gkl343> PMID: 16844994; PubMed Central PMCID: PMC1538830.
 23. Pearson WR. An introduction to sequence similarity ("homology") searching. *Curr Protoc Bioinformatics*. 2013;Chapter 3:Unit3.1. <https://doi.org/10.1002/0471250953.bi0301s42> PMID: 23749753; PubMed Central PMCID: PMC3820096.
 24. Rawal K, Sinha R, Abbasi BA, Chaudhary A, Nath SK, Kumari P, et al. Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. *Sci Rep*. 2021; 11(1):17626. Epub 20210902. <https://doi.org/10.1038/s41598-021-96863-x> PMID: 34475453; PubMed Central PMCID: PMC8413327.
 25. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res*. 2022; 50(W1):W276–9. Epub 20220412. <https://doi.org/10.1093/nar/gkac240> PMID: 35412617; PubMed Central PMCID: PMC9252731.
 26. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*. 1999; 112:531–52. <https://doi.org/10.1385/1-59259-584-7:531> PMID: 10027275.
 27. He Y, Xiang Z, Mobley HL. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol*. 2010; 2010:297505. Epub 20100704. <https://doi.org/10.1155/2010/297505> PMID: 20671958; PubMed Central PMCID: PMC2910479.
 28. Saha S, Raghava GP. VICMpred: an SVM-based method for the prediction of functional proteins of Gram-negative bacteria using amino acid patterns and composition. *Genomics Proteomics Bioinformatics*. 2006; 4(1):42–7. [https://doi.org/10.1016/S1672-0229\(06\)60015-6](https://doi.org/10.1016/S1672-0229(06)60015-6) PMID: 16689701; PubMed Central PMCID: PMC5054027.
 29. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017; 45(W1):W24–w9. <https://doi.org/10.1093/nar/gkx346> PMID: 28472356; PubMed Central PMCID: PMC5570230.
 30. Paul S, Sidney J, Sette A, Peters B. TepiTool: A Pipeline for Computational Prediction of T Cell Epitope Candidates. *Curr Protoc Immunol*. 2016; 114:18.9.1–9.24. Epub 20160801. <https://doi.org/10.1002/cpim.12> PMID: 27479659; PubMed Central PMCID: PMC4981331.
 31. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*. 2004; 32(Web Server issue):W526–31. <https://doi.org/10.1093/nar/gkh468> PMID: 15215442; PubMed Central PMCID: PMC441606.
 32. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 2011; 27(3):343–50. Epub 20101205. <https://doi.org/10.1093/bioinformatics/btq662> PMID: 21134891; PubMed Central PMCID: PMC3031035.
 33. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007; 35(Web Server issue):W407–10. Epub

20070521. <https://doi.org/10.1093/nar/gkm290> PMID: 17517781; PubMed Central PMCID: PMC1933241.
34. Ponomarenko J, Bui HH, Li W, Füsseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*. 2008; 9:514. Epub 20081202. <https://doi.org/10.1186/1471-2105-9-514> PMID: 19055730; PubMed Central PMCID: PMC2607291.
 35. Noori Goodarzi N, Fereshteh S, Azizi O, Rahimi H, Bolourchi N, Badmasti F. Subtractive genomic approach toward introduction of novel immunogenic targets against *Clostridioides difficile*: Thinking out of the box. *Microb Pathog*. 2022; 162:105372. Epub 20211222. <https://doi.org/10.1016/j.micpath.2021.105372> PMID: 34954046.
 36. Stolfi P, Castiglione F. Emulating complex simulations by machine learning methods. *BMC Bioinformatics*. 2021; 22(Suppl 14):483. Epub 20211112. <https://doi.org/10.1186/s12859-021-04354-7> PMID: 34772335; PubMed Central PMCID: PMC8588594.
 37. Gomes LGR, Rodrigues TCV, Jaiswal AK, Santos RG, Kato RB, Barh D, et al. In Silico Designed Multi-Epitope Immunogen "Tpme-VAC/LGCM-2022" May Induce Both Cellular and Humoral Immunity against *Treponema pallidum* Infection. *Vaccines (Basel)*. 2022; 10(7). Epub 20220625. <https://doi.org/10.3390/vaccines10071019> PMID: 35891183; PubMed Central PMCID: PMC9320004.
 38. Jiménez-García B, Pons C, Fernández-Recio J. pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. *Bioinformatics*. 2013; 29(13):1698–9. Epub 20130509. <https://doi.org/10.1093/bioinformatics/btt262> PMID: 23661696.
 39. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res*. 2020; 48(D1):D265–d8. <https://doi.org/10.1093/nar/gkz991> PMID: 31777944; PubMed Central PMCID: PMC6943070.
 40. Saralegui C, Ponce-Alonso M, Pérez-Viso B, Moles Alegre L, Escribano E, Lázaro-Perona F, et al. Genomics of *Serratia marcescens* Isolates Causing Outbreaks in the Same Pediatric Unit 47 Years Apart: Position in an Updated Phylogeny of the Species. *Front Microbiol*. 2020; 11:451. Epub 20200331. <https://doi.org/10.3389/fmicb.2020.00451> PMID: 32296400; PubMed Central PMCID: PMC7136904.
 41. Wang J, Xiong K, Pan Q, He W, Cong Y. Application of TonB-Dependent Transporters in Vaccine Development of Gram-Negative Bacteria. *Front Cell Infect Microbiol*. 2020; 10:589115. Epub 20210127. <https://doi.org/10.3389/fcimb.2020.589115> PMID: 33585268; PubMed Central PMCID: PMC7873555.
 42. Allemailem KS. A Comprehensive Computer Aided Vaccine Design Approach to Propose a Multi-Epitopes Subunit Vaccine against Genus *Klebsiella* Using Pan-Genomics, Reverse Vaccinology, and Biophysical Techniques. *Vaccines (Basel)*. 2021; 9(10). Epub 20210927. <https://doi.org/10.3390/vaccines9101087> PMID: 34696195; PubMed Central PMCID: PMC8540426.
 43. Aghajani Z, Rasooli I, Mousavi Gargari SL. Exploitation of two siderophore receptors, BauA and BfnH, for protection against *Acinetobacter baumannii* infection. *Apmis*. 2019; 127(12):753–63. Epub 20191021. <https://doi.org/10.1111/apm.12992> PMID: 31512768.
 44. Xiong K, Chen Z, Xiang G, Wang J, Rao X, Hu F, et al. Deletion of yncD gene in *Salmonella enterica* ssp. *enterica* serovar Typhi leads to attenuation in mouse model. *FEMS Microbiol Lett*. 2012; 328(1):70–7. Epub 2011/12/14. <https://doi.org/10.1111/j.1574-6968.2011.02481.x> PMID: 22150228.
 45. West D, Reddin K, Matheson M, Heath R, Funnell S, Hudson M, et al. Recombinant *Neisseria meningitidis* transferrin binding protein A protects against experimental meningococcal infection. *Infect Immun*. 2001; 69(3):1561–7. <https://doi.org/10.1128/IAI.69.3.1561-1567.2001> PMID: 11179327; PubMed Central PMCID: PMC98056.
 46. Pieńko T, Trylska J. Extracellular loops of BtuB facilitate transport of vitamin B12 through the outer membrane of *E. coli*. *PLoS Comput Biol*. 2020; 16(7):e1008024. Epub 20200701. <https://doi.org/10.1371/journal.pcbi.1008024> PMID: 32609716; PubMed Central PMCID: PMC7360065.
 47. Ercoli G, Tani C, Pezzicoli A, Vacca I, Martinelli M, Pecetta S, et al. LytM proteins play a crucial role in cell separation, outer membrane composition, and pathogenesis in nontypeable *Haemophilus influenzae*. *MBio*. 2015; 6(2):e02575–14. <https://doi.org/10.1128/mBio.02575-14> PMID: 25714719
 48. Tidhar A, Flashner Y, Cohen S, Levi Y, Zauberman A, Gur D, et al. The NlpD lipoprotein is a novel *Yersinia pestis* virulence factor essential for the development of plague. *PLoS one*. 2009; 4(9):e7023. <https://doi.org/10.1371/journal.pone.0007023> PMID: 19759820
 49. Terashima H, Kojima S, Homma M. Flagellar motility in bacteria structure and function of flagellar motor. *Int Rev Cell Mol Biol*. 2008; 270:39–85. [https://doi.org/10.1016/S1937-6448\(08\)01402-0](https://doi.org/10.1016/S1937-6448(08)01402-0) PMID: 19081534.
 50. Ullah A, Ahmad S, Ismail S, Afsheen Z, Khurram M, Tahir Ul Qamar M, et al. Towards A Novel Multi-Epitopes Chimeric Vaccine for Simulating Strong Immune Responses and Protection against *Morganella*

morganii. *Int J Environ Res Public Health*. 2021; 18(20). Epub 20211019. <https://doi.org/10.3390/ijerph182010961> PMID: 34682706; PubMed Central PMCID: PMC8535705.

51. Wan C, Gao C, Xie Q, Wang Y, Cheng X, Fang Y, et al. Flagella hook protein FlgE is a novel vaccine candidate of *Pseudomonas aeruginosa* identified by a genomic approach. *Vaccine*. 2021; 39(17):2386–95. <https://doi.org/10.1016/j.vaccine.2021.03.051> PMID: 33775439