Jurnal
Teknologi

Full Paper

IDENTIFICATION AND CLASSIFICATION OF BACTERIAL AUTOTRANSPORTERS

Hina Aslam Butt, Syeda Marriam Bakhtiar*

Department of Bioinformatics and Biosciences, Mohammad Ali Jinnah University, Expressway Zone IV, Islamabad, Pakistan

Article history
Received
25 October 2015
Received in revised form
14 December 2015
Accepted
9 February 2016

*Corresponding author
marriam@jinnah.edu.pk

Abstract

Bacterial Autotransporters is a huge and assorted super family of proteins present in external boundary of the gram negative bacteria. Autotransporters consists of C-terminal beta domain, N-terminal passenger domain and a pathogenic function. The beta domain of the bacterial autotransporter is associated with pathogenicity with three types of functions namely adhesions, esterase and proteases. Identification and classification of a bacterial autotransporter is important in order to target it and hinder its ability to cause disease. Identification and classification therefore has become a challenging problem due to its homology sequence. In this paper, we ought to establish a web based tool that could identify and classify bacterial autotransporters on basis of homology with already reported bacterial autotransporters. The input required for tool amino acid sequence in FASTA format. The tool will analyze the query based on similarity with already reported sequence in database and generate output with information in terms of function, organism, class, length, protein ID and symbol. Further, this user friendly and freely accessible tool is integrated with NCBI, InterPro, Pfam, pubmed. Moreover, newly detected bacterial autotransporter can be easily added in the Bacterial Autotransporter Detector tool database with authentic reference.

Keywords: Pathogenicity; bacterial autotransporter; adhesions; esterase,proteases

© 2016 Penerbit UTM Press. All rights reserved

1.0 INTRODUCTION

Bacteria have an extremely varied range of protein secretion systems, involving movement of proteins, toxins, enzymes and other effector molecules from inner to outer of the bacterial cell. Characteristic of Gram Negative Bacteria includes cytoplasmic cell membrane, lipopolysacchride in its outermost cover of membrane, phospholipids in it's inter most leaflet and porins. Gram Negative bacteria have two membranes that makes the secretion complex. Bacterial Protein secretion system includes Type 1 to Type VI classification system. Type 1 secretion system also known as T1SS involves transfer of ions, drugs and discrete molecules to protein of numerous sizes. Type 2 secretion system also called as Secretory dependent system is only found in Gram negative bacteria. It transports proteins via Sec pathway. Type 3 secretion system known as T3SS transport effector proteins surpassing the anchor as well as bacterial membranes into the cytosol of anchor cells includes different functions such as immune response and defense response. Type 4 secretion system commonly called as TFSS is effective in delivering virulence factor proteins directly into host cells and taking up DNA and proteins. In Type V protein secretion system, additional findings related to assembly, structure and function have enlarged the classification and it now includes Type Va "autotransporters", Type Vb "two-partner systems", Type Vc "trimeric autotransporters", Type Vd "Patatin-like proteins" and Type Ve "intimins and invasins" [1]. In this paper we focused on the Type V protein secretion system i.e. Bacterial Autotransporters. Thomas Meyer and colleagues used the terminology for the first time and carryout IgA protease from *Neisseria meningitidis* [2]. The foundation of investigation was the previous findings that amino acid chains of the protein itself anchors the role for surface display and for movement across the external boundary. Bacterial Autotransporters are extensively spread in the external boundary of the gram negative bacteria.

Bacterial autotransporters are mainly connected with pathogenicity. Three domains are found in

autotransporters, Type Va protein secretion system namely signal sequence, passenger domain and the translocation domain. Signal sequence acts as indicator, exist at N-terminus of protein enables the focusing of amino acid chains to the internal boundary that translocate to the periplasm. Traveller or passenger domain named as alpha and secreted domain grant variety of impact functions of various autotransporters. Translocation domain named as helper and beta domain reside at C-terminus end of the amino acid chains. It is made up of small joined region along with alpha helical subordinate formation and beta core that obtain a beta barrel tertiary structure that induced in the outermost boundary opens door for movement of the alpha domain via outermost boundary. The actual model of external boundary protein secretion system anticipate that the prime of the beta strands proceed from the periplasmic area to the outer surface hence mislay the passenger domain interim enlarging into the periplasm [3]. Contemporary findings by Oliver et al. impart the existence of an intramolecular chaperone in BrkA, comprise by remnant 606 to 702. This area correlate to PD002475 in the ProDom database [4,5] and institutes in various bacterial autotransporters but not in other proteins in the database. Figure 1 shows the Ultimate constitution of bacterial autotransporter. Main regions are represented and the scissors indicate the areas of managing

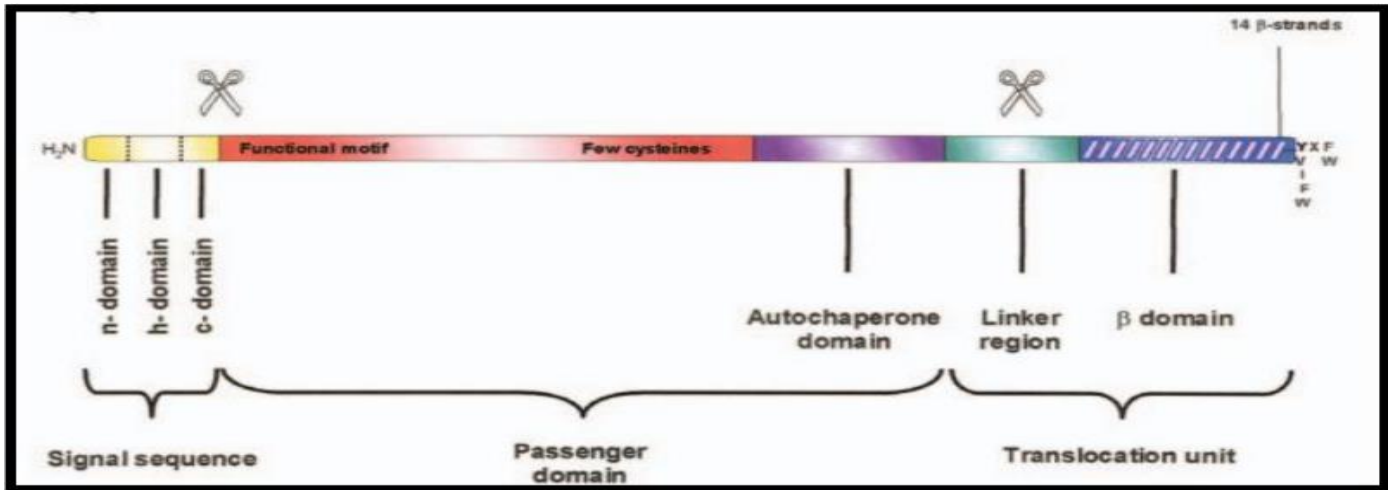

Figure 1 Constitution of bacterial autotransporter where scissors indicate areas involved in regulation

Signal sequence at the start of the polypeptide in Bacterial Autotransporter helps in targeting the polypeptide to the inner membrane by passing periplasm. Passenger domain of Bacterial Autotransporter is associated in causing virulence by encoding effector function. Effector function include three types namely adhesions, esterase and proteases. Besides causing virulence, adhesions are composed of cell surface components of bacteria that facilitate linkage to other cells or to surfaces. Esterase and proteases are enzyme that perform hydrolase activity and proteolysis activity respectively. Species of bacteria contains extend molecular approaches to energetically persuade their approach into focused cells for reproduction and or circulation to other vector tissues. Infringement can progress by undeviating association of surface anchor cell receptors or by undeviating movement of bacterial proteins into the vector cell cytosol that will encourage reposition of the plasma membrane architecture, inducing pathogen engulfment. Bacterial adhesions negotiate the linking of bacteria to their recess such as the tissue of an impair host. Adhesions have to be translocate over the cell cover to become functional and throughout protein secretion process they overlap into their final conformation. Examples of bacterial adhesions include enteropathogenic E. coli planted clearance of the intestinal striated border microvilli that promote to diarrhea. It is responsible for hemorrhagic colitis and sporadically hemolytic uremic syndrome. Translocation domain also known as beta barrel area consists of small alpha helical join fragment and twelve beta strands that accumulated into beta barrel in external boundary. β domain was firstly suggested to function as a route that translocate its own traveler domain over the external boundary [6]. Autotransporter have been identified as the first glycosylated proteins in Gram-negative bacteria [7]. Firstly, it was conceived that bacterial autotransporters are capable to place into the bacterial Outer Membrane without the participation of other factors. During current years, the necessary Outer Membrane protein BamA has been shown to be critical for bacterial autotransporter biogenesis [8]. True model of autotransporter suggested that aperture is produced in the Outer Membrane by the end of polypeptide of the autotransporter and hairpin twists out via aperture. Following this, the transported alpha region begins to fold from end of the polypeptide to the start of polypeptide, all other proteins are pulled out. One of the example of bacterial autotransporter include Bordetella. *Bordetella parapertussis and Bordetella pertussis are adjacent* associated organisms that are accountable for human pertussis. Animal pathogen named as *Bordetella*

bronchiseptica leads to the respiration disease in host. Virulence factors produced by these three includes filamentous hemagglutinin (FHA), adenylate cyclase, pili, pertactin. *Bordetella pertussis* generates pertussis toxin [9]. Traveller domains of the pertactin molecules are represented in *Bordetella pertussis* [10], *Bordetella parapertussis* [11], and *Bordetella bronchiseptica* [12] as proteins of 68 kDa, 69 kDa, and 70 kDa respectively. Organized responsible traveler domain residues non covalently linked with the 30 kDa beta domain in a manner recollect of Ag43 and AIDA-I[13]. *Helicobacter pylori* is pathogen exist in humans that populate the human stomach and remain determined to the gastric mucous layer. The hygienic findings disclose a strong link between *Helicobacter pylori* disease and capability of the bacterium to generate the cytotoxic associated antigen, blood group antigen binding adhesion and vacuolating cytotoxins[14,15]. It shows that bacterial autotransporter are highly associated with pathogenesis. Hence, their identification is necessary to find the root cause of the pathogenecity. These recent studies provide dimensions to further analyze the bacterial autotransporter because it ultimately helps in diagnostics, drug development, disease prevention and other aspects. The sequence analysis provide help in the identification and characterization of bacterial autotransporters. With the advancement in bioinformatics, many tools are developed that store the sequence and provide the analysis on the retrieved sequences. Therefore, we focused on the sequence analysis of the bacterial autotransporters and developed a tool that provides sequence analysis of the bacterial autotransporters.

2.0 METHODOLOGY

Bacteria consists multiple arrangement of protein secretion system which impart important function to the organism. Many autotransporters proteins are reported to be associated with virulence. Similarly cell to cell adhesion as well as biofilm formation also involves autotransporters. With the development of modern bioinformatics and recent advances in bioinformatics methods have been adopted to overcome the problem associated with detection of autotransporter sequences. We also target to automate the classification of identified autotransporter proteins. As bacterial genomes have been sequenced and stored by using bioinformatics databases which also contain amino acid sequences. Researchers identified the specific region which are homologous to the autotransporters. We constructed a database in which all the bacterial autotransporters regions are stored along with their functions and important information. Main objectives of the project was to establish database for already reported autotransporter identification, secondly identification and classification of autotransporter query sequence using database of already reported proteins. Purpose of this web based tool is to collect, store, analyze and retrieve information of the bacterial autotransporters. This web based is able to identify and classify the sequences of the bacterial autotransporters based on the previous knowledge.

3.0 AUTOTRANSPORTER TOOL

Autotransporters are the highly involved in causing the virulence. The sequence analysis helps in identification of the bacterial autotransporter. Therefore, we developed an automated web based tool named as Bacterial Autotransporter Detector Tool that provides help in identification and classification of bacterial autotransporters. Bacterial Autotransporter Detector Tool is a web based tool constructed by using Java Server Pages. Java coding is implemented for the identification and characterization of already reported bacterial autotransporters sequences. Apache Tomcat is used for connecting Java Server Pages. Database is created in XAMPP. Bacterial Autotransporter tool is integrated with Pubmed, InterPro, Pfam, NCBI etc. PubMed is an unoccupied complete collection of literature related to biomedical and life sciences. PubMed is an online tool that provide articles, journals, reviews to related query. NCBI stands for "The National Center for Biotechnology Information" and is component of branch of National Institutes of Health named as United States National Library of Medicine. National Center for Biotechnology Information is most recognized website that provides complete information related to search. Pfam is a database of those regions which are preserved in protein families and are extensively utilized by biologists and researchers to explain and characterize proteins. The database of Pfam consists of two classes named as Pfam A and Pfam B. Family Pfam A comprises Hidden Markov Model, complete alignment and seed alignment with database links, notation and information data references. Family

Pfam B robotically produced alignments of sequence clusters attained from the use of Automatic Domain Decomposition Algorithm database with no notation that surcharge the Pfam A families. InterPro is database of domains, protein families and functional sites. In this recognizable characteristics present in familiar proteins can be applied to upcoming protein sequences in order to functionally characterize them. InterPro areas are constructed about diagnostic marks and proteins that outstandingly match. InterPro is a facility that gives practical and functional investigation of the protein sequences by categorizing them into families and anticipating the existences of domains and important sites. Input of the our tool is amino acid sequence in FASTA format, a text based format for indicating some peptide sequences or nucleotide sequences in which amino acids or nucleotides are indicating using single letter codes. Sequence name and comments are preceded by the sequences in this format. FASTA format can be pasted by the user or can be uploaded from the stored files. This tool will have ability to detect Bacterial Autotransporter from

provided input. Beside detection, it provides complete information of the Bacterial Autotransporter which includes function, organism, class, length, literature results. For authentication of results tool will provide references to the related detected Bacterial Autotransporter. Database for Bacterial Autotransporter Detector tool is shown in figure 2,and graphical user interface of Bacterial Autotransporter Detector tool is shown in figure 3.

ID	SYMBOL	AMINO ACID SEQUENCE	DESCRIPTION	FUNCTION	ORGANISM	LENGTH
1	AoaA	>gi|158331148|dbj|BAF88633.1| uncharacterized protein AZC_2635 [Azorhizobium caulinodans ORS 571] MVFEWWVGIMSARIAETMFRSQGESFRQPLLRKRLLRSSALIPFLALVLLAPPAFAGGGGNGGGNSGSPG GLDNYDAGGTGGTGTAATSIGSGGGAGGGGAGLTGGAGGAANGTNGAGGAGGTHAYVGSGSGVTNSLTGG NGAAGSSGSNNGGGGGAGGTAVVITGNSSGVLNWSSATFTLTGGVGGAGGSGSSAQNGGEGGTGGDGLYF GQDITFTSSNSLLITGGQGGAGSNGRAGGNGIEFATGTNSRATINGQVQGGAGGTSTKTNSNTGNGGNGI LADTLDLTLNAMVSGGAGSSGAGDVGGGAGGDGISALAFTGMINAAVTGGTGGSTGASSSASAGSGGYGL RLNGASTVTVNANVTGGNSGAASVSGGYGGIGIYATPSLSLTIQNNAIVRGGDAGAVTGSAYTTGGSGIY LDYATSDITLLSGQVIGGNASGSDSAATGGYGIESHGAPSVTITLGAGSTISGGLGGNGMSDRSARANAI FIENFDGTPATNTLVLLGSGGTSGSTYATIIGDVVATPADSTTNKMTFGGAGGVFDVGQIDLGAGSSAAT IFRGFTEFTVDTTGTWIMTGASTTGGTALPATWTVAAGTLQLGNASTAGTLLGSVGVNSGGTLANGGRTA TVTNGVTVTSGATLSLTAVSGGPALSITSGNLNLNTGSTLTITLGAPTTQSLISVAAGSAVLGSTLNITD AGTMAAGSYTLLSYSGTLSGPGLTLGTTPSAFQFSVDTSSTAGQVLLLVTGGAPTVIYWNGSTTGGSSGP VAGGTGTWTAASSGVTNWTNSAGTSRVVSDPSLTAIFAGTAGTVTVSAASGAVSAKGLEFDTSGYTITGS ELTLVNGSTMPQVNVVGSSASATISAPLAGSNGLEK	UNCHARACTERIZED PROTEIN	UNKNOWN	Azorhizobium caulinodans	3766
2	Arp	>gi|154091346|gb|ABS57467.1| acidic repeat protein [Bartonella henselae] MSKKILLSYTTTAAIILFNIHSNAYAISLFSGEGENKTAPTQESYENIYALDGGKIHGKDLKIIGPSTQE ESIITDISGVEARKFGSMIELEGDTTIKNVSIGLLAKESGTIKMNDGSIQVKKVHIQTPIGIAAVSNGAI ILNNVEIDASNQEQSIKTIDETGIGDGTGASLKSGGTLSMTGGSIKSNYLGITLEESSSDKNKLENVKIN ITNLANATKESIGIRVIKTSKVTLNQVTIRHARTSIHASDSSEITISGGLIQGNHTGINVEKESVITLKN DVEVLSNDHGLSANGLHSKITMQGGKLITAGLQPAVLAGSGGEINLINVVVHIDDLTDIVMHIDDNETQT QKLELERKEAPLTTQGLQAQYAQSKITMIRGSITTTGLNPAVLAGSGGQIDLTNVPMKVHNVGLQAQAEQ SKIVMTRGSITTTGMNPAVLAGSGGEIDLNDVVIKTKDIALQAQDKQSKITMRGGKLIKTGPRAAIFVTC GGQIDLLDAQLHTDSNGLAVRGRESKITLKDSEVRANILLVGKPNDEDDNGEANVIADHSILEGGARNSE RKPTQIIFSLINGTTWYLKANMQSHKIQKLDLIKKLHSEVFKLNLNNSTIVFRTPREDQYQTLHIGNKSP HLTNNNTTHKTVYNATGDAKIYFNTEWNNEVPKEQQKTDRLLIHGDVSGITTIHFRNLLKGKKTKEKNTG PVNTRGLSLVQVSGKAEENSFKLANGYTTIKGLPYKYTLNAYGPTSSRGKASIEQSFVGEEEDFWDFRLQ NAILDAEETISPDTQETNFFNSEEITPLNTEETASLNDGETTPLYSKKAASPDAEETISSDTQETNFFNS EEITPLNTEETASLNDGETTPLYSKKAASPDAEETISSDTQETNFFNSEEITPLNTEETASL	ACIDIC REPEAT PROTEIN	UNKNOWN	Bartonella henselae	1441

Figure 2 Bacterial autotransporter detector database

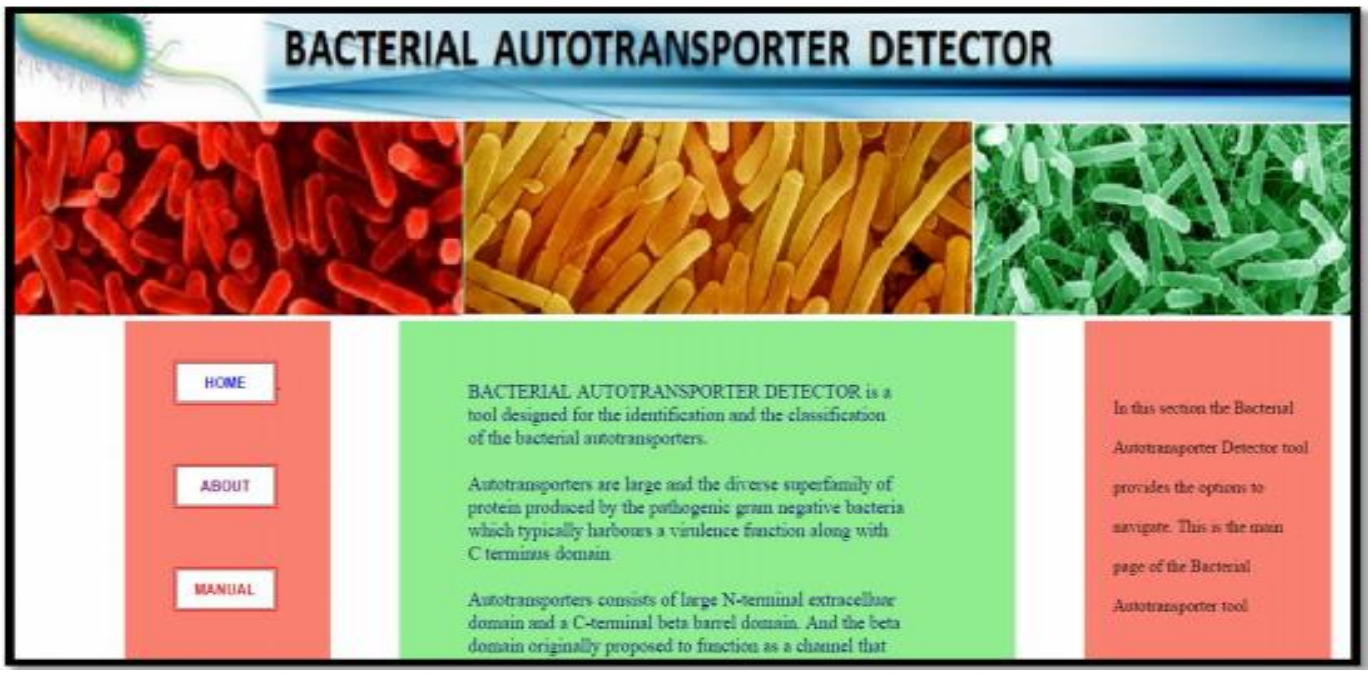

Figure 3 Bacterial autotransporter detector database graphical user interface

This tool is an implementation of web based tools development techniques, for identification and classification of bacterial autotransporters for researchers and students, which can be used by those who are working on bacterial autotransporters. Bacterial autotransporters are responsible for pathogenesis. Therefore a lot of research is focused on "autotransporters". Problem of unavailability for bacterial autotransporter identification and classification is there. Researchers and students working on bacterial autotransporters face a lot of problems in their studies. Solution is the automation of

autotransporter protein identification by a web based tool using sequence alignment with a database of all the reported autotransporter proteins. Researchers and students working on protein analysis and pathogenesis of bacteria usually come across small sequence. Assigning function and identification of these small sequences is important. As these autotransporters are associated with pathogenesis, therefore a tool is required which help in the identification and classification of the bacterial autotransporters. System architecture diagram of tool is shown in Figure 4

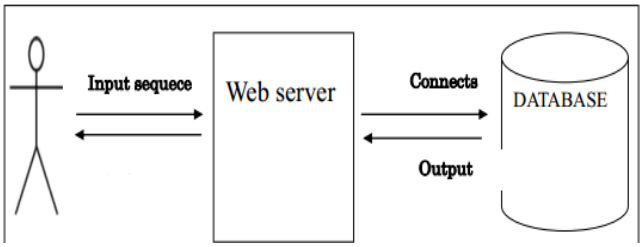

Figure 4 System architecture diagram of system

4.0 CONCLUSION

Bacteria consists of extremely multiple arrangement of protein secretion system which impart important function to the organism. Many autotransporters proteins are reported to be associated with virulence. Similarly cell to cell adhesion as well as biofilm formation also involves autotransporters. With the development of modern bioinformatics and recent advances in bioinformatics methods have been adopted to overcome the problem associated with detection of autotransporter sequences. We also target to automate the classification of identified autotransporter proteins. As regular BLAST skips short sequence such as autotransporter therefore this tool is a help to identify Autotransporter and functionally annotate these small sequence. Development of this web based tool is helpful in identification of autotransporter proteins of bacteria by sequence alignment. Maintenance of the database of all reported autotransporter proteins will further helps to classify the autotransporters and integration of this data with our developed tool that will classify and give access to the basic information. System is able to classify autotransporter into already established classes. Major problem faced by sequence analyst these days is to identify the short sequences which encode for many important molecules such as bacteriocins, miRNAs and autotransporters. These sequences usually skip identification and alignment from BLAST because of their small size. Many tools are now available which target these small molecules e.g. BAGEL for bacteriocins which helps to automate the identification and classification of the small molecules. This tool is definitely a step forward in this way.

References

[1] Celik, N., C. T. Webb, D. L. Leyton, K. E. Holt, E. Heinz, R. I. Gorrel, T. Kwok, T. Naderer, R. A. Strugnell, T. P. Speed, R. D. Teasdale, V. A. Likić, and T. Lithgow. 2012. A Bioinformatic Strategy For The Detection, Classification And Analysis Of Bacterial Autotransporters. *PLoS One*. 7(8): e43245,

[2] Leo, J. C., I. Grin, and D. Linke. 2012. Type V Secretion: Mechanism(S) Of Autotransport Through The Bacterial Outer Membrane. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367(1592): 1088–101.

[3] Henderson, I. R., F. Navarro-Garcia, and J. P. Nataro. 1998. The Great Escape: Structure And Function Of The Autotransporter Proteins. Trends Microbiol. 6: 370-378.

[4] Oliver, D. C., G. Huang, E. Nodel, S. Pleasance, and R. C. Fernandez. 2003. A Conserved Region Within The *Bordetella Pertussis* Autotransporter Brka Is Necessary For Folding Of Its Passengerdomain. *Mol. Microbiol*. 47: 1367-1383.

[5] Corpet, F., F. Servant, J. Gouzy, and D. Kahn. 2000. ProDom and ProDom-CG: Tools For Protein Domain Analysis And Whole Genome Comparisons. *Nucleic Acids Res*. 28: 267-269

[6] Bernstein, H. D. 2007. Are Bacterial "Autotransporters" Really Transporters?. *Trends Microbiol.* 15(10): 441–447.

[7] Benz and M. A. Schmidt. 2011. Structures And Functions Of Autotransporter Proteins In Microbial Pathogens. *Int. J. Med. Microbiol.* 301(6): 461–468.

[8] Leo, J. C., I. Grin, and D. Linke. 2012. Type V Secretion: Mechanism(s) of Autotransport Through The Bacterial Outer Membrane. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367(1592): 1088–1101.

[9] Parton, R. Review Of The Biology of *Bordetella pertussis*. *Biologicals*. 1999. 27: 71 –76.

[10] Charles, I. G., G. Dougan, D. Pickard, S. Chatfield, M. Smith, P. Novotny, P. Morrissey, N. F. Fairweather. 1989. Molecular Cloning And Characterization Of Protective Outer Membrane Protein P.69 From *Bordetella Pertussis*. *Proc Natl Acad Sci USA*. 86(10): 3554–3558.

[11] Li, L. J., G. Dougan, P. Novotny, I. G. Charles. 1991. P.70 Pertactin, An Outer-Membrane Protein From *Bordetella Parapertussis*: Cloning, Nucleotide Sequence And Surface Expression in *Escherichia coli*. *Mol Microbiol*. 5(2): 409–417.

[12] Li, J., N. F. Fairweather, P. Novotny, G. Dougan, I. G. Charles. 1992. Cloning, Nucleotide Sequence And Heterologous Expression Of The Protective Outer-Membrane Protein P.68 Pertactin from *Bordetella bronchiseptica*. *J Gen Microbiol.* 138(8): 1697–1705.

[13] Charles, I., N. Fairweather, D. Pickard, J. Beesley, R. Anderson, G. Dougan, M. Roberts. 1994. Expression of the *Bordetella pertussis* P.69 Pertactin Adhesin In *Escherichia Coli*: Fate Of The Carboxy-Terminal Domain. *Microbiology*. 140(12): 3301–3308.

[14] Figura, N. 1997. *Helicobacter* Pylori Factors Involved In The Development Of Gastroduodenal Mucosal Damage And Ulceration. *J Clin Gastroenterol*. 25: S149–S163.

[15] Gerhard, M, N. Lehn, N. Neumayer, T. Boren, R. Rad, W. Schepp, S. Miehlke, M. Classen, C. Prinz. 1999. Clinical Relevance Of The *Helicobacter Pylori* Gene For Blood-Group Antigen-Binding Adhesin. *Proc Natl Acad Sci USA*. 96(2): 12778-12783