Original Article

Subtractive Genomics Based Analysis For In Silico Recognition And Characterization Of New Drug Targets In Yersinia Pestis

Faisal A Al Zahrani*, Alam Jamal

Department of Biochemistry, King Abdulaziz University, Jeddah 21589, Saudi Arabia
*Corresponding author. E-mail: faahalzahrani@kau.edu.sa
DOI: 10.47750/pnr.2023.14.03.372

Abstract

Yersinia pestis (**Y. pestis**) is a gram-negative bacteriam.it cause plague disease. Between 30% and 100% of untreated patients die, with pneumonic plague being the deadliest. Chloramphenicol Streptomycin and tetracycline are the three main antibiotics that are used for plague treatment. Even if antibiotics are given right away, 16% of people with bubonic plague and more than 50% of people with pneumonic plague die. The number of plague cases and epidemics over the past 15 years has increased to the point where the disease is now considered reemerging. Researchers have discovered that some strains of the plague are immune to the most effective medications now in use to treat the disease. Y. pestis is worrisome since there is no effective vaccination, it must be treated with antibiotics, and it has a history of being used as a bioweapon. Antibiotic-resistant bacteria are a worldwide problem. It has been revealed that two strains of Y. pestis are highly resistant to eight antimicrobial medicines used for cure. Additionally, it is resistant to a number of commonly used antibiotics, including ampicillin, kanamycin, and spectinomycin. In subtractive genomics approach the bacterial total proteome is gently nethermost to a few numbers of probable targets for drug. The steps which are applied in this method to detect targets proteins which are non-similar to human and are vital for bacteria and contribution of the sorted out proteins in bacterial pathways of metabolism that are vital for continuity of pathogen. In this study the approach of subtractive genomics is applied on proteins of Y. pestis and culminated with 4 proteins that may be powerful drug targets and novel zestful molecules can be designed against them to cure the bacteria associated infection.

Key words: Bacteria, Subtractive genome analysis, Y. pestis, Drug targets, Proteins

1. Introduction

Yersinia pestis (**Y. pestis**) is a gram-negative bacterium. Humans can become infected with this bacterium, which is facultatively anaerobic. The zoonotic bacterium Y. pestis causes the devastating disease known as plague. The plague can take on a number of different forms, but there are mostly three major types. The bubonic plague is the most common and widespread of these. The spread of infection can take several distinct forms. Bubonic plague and primary septicemic plague are both caused by infected flea or animal bite. Pneumonic plague, which is caused by inhaling aerosolized Y. pestis, is the well-known type of this infection. Untreated bubonic plague may lead to subsequent septicemic and pneumonic forms. [1]. The death ratio is 100% in untreated Pneumonic plague patients. Pneumonic plague is an extremely contagious type of this infection which is transfer via respiratory droplets of infected person among humans. When a flea carries Y. pestis and bites a person, the bacteria travels via the bloodstream and into the lymph nodes, where it multiplies and causes swelling and inflammation. Incubation of the bacterium can take anything from one to ten days.[2]. Pustules might form on the bubo's skin as the infection progresses. [3]. The infection might lead to severe complications when reaches the lungs. Bubonic plague

have various symptoms like weakness, high body temperature, malaise, chills, headaches and muscle pain. [2]. In cases where bubonic plague is not treated, mortality rates range from 40% to 70%. Between eighty and ninety-five percent of all plague cases are due to this. An incubation period of one to seven days is typical for Y. pestis infection.

Pneumonic plague is another form of the infection, which causes an infection of the lungs and can develop from bubonic plague or be transmitted from an infected person. This is the worst strain of the plague, and its incubation period can be as brief as 24 hours or as long as four days. Pneumonic plague is characterized by weakness, headache and high body temperature and pneumonia that develops quickly and is marked by respiratory problems, cough, chest pain and in rare cases, bloody or watery mucus. [2]. Pneumonic plague is much more dangerous than bubonic plague. It is fatal when untreated. Its death ratio is higher than 50% since patients have less than 24 hours to get to a hospital and start treatment.[2].

Septicemic plague is the 3rd type of the disease in which blood infection occur. Ten percent to fifteen percent of all plague cases are primary septicemic plague. [4]. Bubonic plague can cause a subsequent infection known as septicemia. Although anyone at any age can contract this epidemic, the elderly are most vulnerable. [5]. The quickly replicating Y. pestis causes an immune cascade that feeds on itself, causing illness [6] . Minor symptoms consist of stomach pain, fever, extreme weakness chills and shock; major symptoms include disseminated intravascular coagulopathy, acute respiratory distress syndrome, multiple organ failure, skin and serosal surface bleeding, and acral gangrene. If neglected, septicemic plague can be lethal. The treatment-affected case fatality ratio ranges from 30-50%. Plague is transmitted to humans through the infected flea bite on a rodent or via infected animal contact.

Plague treatment needs to be started immediately [1]. Antibiotics, oxygen therapy, intravenous fluids, and respiratory support are the present standard of cure for plague patients. [7]. Streptomycin, tetracycline, and chloramphenicol are the three most prescribed antibiotics for treating plague. In addition to penicillin, gentamicin, levofloxacin, ciprofloxacin, doxycycline, moxifloxacin, and chloramphenicol are effective antibiotics against the plague. [6]. Patients with bubonic plague have a 16% mortality rate and those with pneumonic plague have a mortality rate of over 50% even with rapid antibiotic therapy. The number of plague cases and epidemics over the past 15 years has increased to the point where the disease is now considered reemerging. Researchers have discovered that some strains of the plague are immune to the most effective medications now in use to treat the disease. The lack of an effective vaccination, the requirement of antibiotics for treatment, and the pathogen's previous employment as a bioweapon all combine to make Y. Pestis a worrisome potential cause of another catastrophic disaster.

Antibiotic-resistance is a global issue [8]. The bacterial potential to conjugally transmit its antibiotic resistant plasmids to other non-resistant strains of Y. pestis or to Escherichia coli (E. coli) is the main cause of antibiotic resistance [9]. Previous studies have examined the sensitivity of Madagascar's isolated strains to various antimicrobials. Two different Yersinia pestis strains were recently reported to be resistant to antibiotics. High-level resistance to eight antimicrobial medicines for treatment use and several prophylactic treatments has been identified in Y. pestis, which is carried on a plasmid of 150 kb named pIP1202. Furthermore, it has shown resistance to some commonly used antibiotics. Both of these types are currently only found in Madagascar. A Chinese strain with a point mutation in the rpsL gene, that codes for the 30S subunit of the ribosomal protein is found to be Streptomycin resistant [9]. A Y. pestis strain in Madagascar has been reported to be resistant to doxycycline due to plasmid-mediated resistance (isolated from a rat in 1998). Furthermore, in 2000, a marmot-borne strain of multidrug-resistant Y. pestis was discovered in Mongolia, but its genetic origins and potential for transmission were not studied. [10].

Genes for antibiotic resistance could spread to the local plague population if the travelers make it to a country where the disease is endemic. It could possibly affect in the same manner E. coli and other bacteria. Additionally, antibiotic resistance may be spread if certain E. coli possess the plasmid harboring the resistance genes from Y. pestis. [9]. The huge number of cases across Madagascar, leading to its designation as a re-emerging disease, implies that this

bacterium may leads to major issues for a region if it is allowed to spread freely or if environmental conditions are favorable. If new antibiotics or therapies are not found, the two resistant natural stains of Y. pestis emergence could pose a serious threat to public health in the future.[11].

Since its origin in Hong Kong in 1894, the last worldwide pandemic has established several endemic foci all over the world. The death and morbidity rates from the disease were drastically reduced because to antibiotics and the execution of public health measures, but the disease was not eradicated. [12]. There are at least three reasons why the epidemic is being discussed again. First, the number of reported cases has raised according to the Organization of World Health. Two, in 1994, the plague made an epidemic return to nations like Malawi, Mozambique, and India where it had been dormant for 15 to 30 years. Finally, the number of focal points is growing in some nations. For example the number of human plague cases has increased in America. [12]

Five biovars of Y. pestis, distinguished by their unique biochemical profiles, have been identified to date. These are Y. pestis Antiqua, mediaevalist, orientalis, pesto ides (microtus), and intermedium.

The alarming spread of drug-resistant forms of bacteria like Y. pestis has created an urgent need to find a cure for numerous bacterial infections for which no current medications are effective.[13]. To overcome drug-resistant infections, researchers are turning to full bacterial protein sequences to identify novel therapeutic targets. [14]. Several strategies for discovering potential novel therapeutic targets are described in the published literature. [15] [16] [17]. In such a scenario, computational subtractive genomics analysis is the most reliable method for discovering potential novel therapeutic targets. [14]. To uncover several novel therapeutic targets against Y. pestis, we applied a computational subtractive genomes analysis strategy in this study. As part of the current research, the entire proteome of Y. pestis strain FDAARGOS 601 was screened in a high-throughput manner using the BLAST tool in comparison to the human genome to identify non-homologous sequences. The Database of Essential Genes (DEG) is a useful tool for identifying possible therapeutic targets that may be used in conjunction with subtractive genome research. As a next stage, the discovered potential therapeutic candidates have been subjected to differential pathway analysis and subcellular localization. Non-similar proteins which are crucial to the Y. pestis strain FDAARGOS 601 survivals were identified using a subtractive genomics approach and may be targets for future treatment development.

2. Material and methods

The FDAARGOS 601 strain of Y. pestis was utilized to identify essential proteins using a subtractive genomics technique and these proteins were subsequently examined to identify therapeutic targets. After identifying medication targets, we ran them through the Drug Bank database to see how broadly they may be use as drug targets. Figure 1 depicts the overall process flow.

2.1 Protein sequence retrieval of pathogen

Using NCBI, we were able to obtain the whole set of proteins from the Y. pestis strain FDAARGOS 601. (www.ncbi.nlm.nih.gov/) [18].

2.2. Search for non-duplicate proteins

CD-HIT was utilized with a sequence identity criterion of 0.8 (i.e., 80%) to find paralogous and duplicate protein successions [19]. Complete proteins from Y. pestis strain FDAARGOS 601 were filtered for duplications to isolate only non-duplicate sequences.

2.3. Searching for Protein Sequences Not Found in the Human Proteome

Using a threshold expectation rate (e value 10–3) BLASTp was used to identify Y. pestis sequences of protein which are different to human proteins. In the end, we got a mix of similar and dissimilar sequences. Protein sequences that did not share enough similarity with the host were omitted from the subsequent study.

2.4. Identification of Yersinia pestis non-identical proteins needed for viability

Sequences that did not share any similarities were put through a BLASTp search using DEG at e-value cutoff of 10^{-5}. Proteins having high levels of similarity to the DEG database are identified as likely being crucial to the bacteria's continued existence. [20].

2.5. An examination of metabolic pathways

The metabolic pathways of all known living organisms are catalogued in KEGG [21]. which stands for the Kyoto Encyclopedia of Genes and Genomes. The KAAS service was utilized for predicting sequences of protein which are engaged in several processes of pathogen metabolism.

2.6. Predicting Subcellular Location

Using PSORTb version 3.0, we were able to predict the subcellular localization of functionally distinct but equally important proteins [22]. The fundamental idea behind Sub Cellular Location BLAST is to run BLASTp against a database of proteins with known subcellular locations and retrieve all non-identical sequences of required proteins. PSORTb offers prediction results for diverse subcellular locations, including as the cytoplasmic membrane, the cytoplasm and the extracellular space, the cell wall and the unknown.

2.7. Predicting the function of a family of necessary, hypothetical, and unrelated proteins

There is currently no actionable data that can be used for predicting the functional family division of proteins, particularly proteins whose function is speculative. We used the Interproscan server. Once the non-similar hypothetical protein sequences were generated, they were sent to the interproscan system for functional family prediction.

2.8. Draggability potency of short-listed proteins

All the protein sequences were screened using BLASTp against the Drug Bank [23] database of protein targets for FDA-approved drugs. BLASTp search with e value 10–3 was used to uncover novel drug targets**.**

2.9. Examining Human Gut Microbiomes

In human body there many organisms which are beneficial for them. This relationship is mutually important for both of them. These microorganisms play very important roles in human body like the immune system training, production of hormones, inhibit harmful species growth and vitamins synthesis. So, if we inhibit these proteins then it will be very dangerous for the host. So to remove these similar proteins, the targeted proteins of Y. Pestis were matched to the gut flora proteins of human by BLASTp with an E-value cutoff score of 1 at the Human Microbiome Project database server [24].

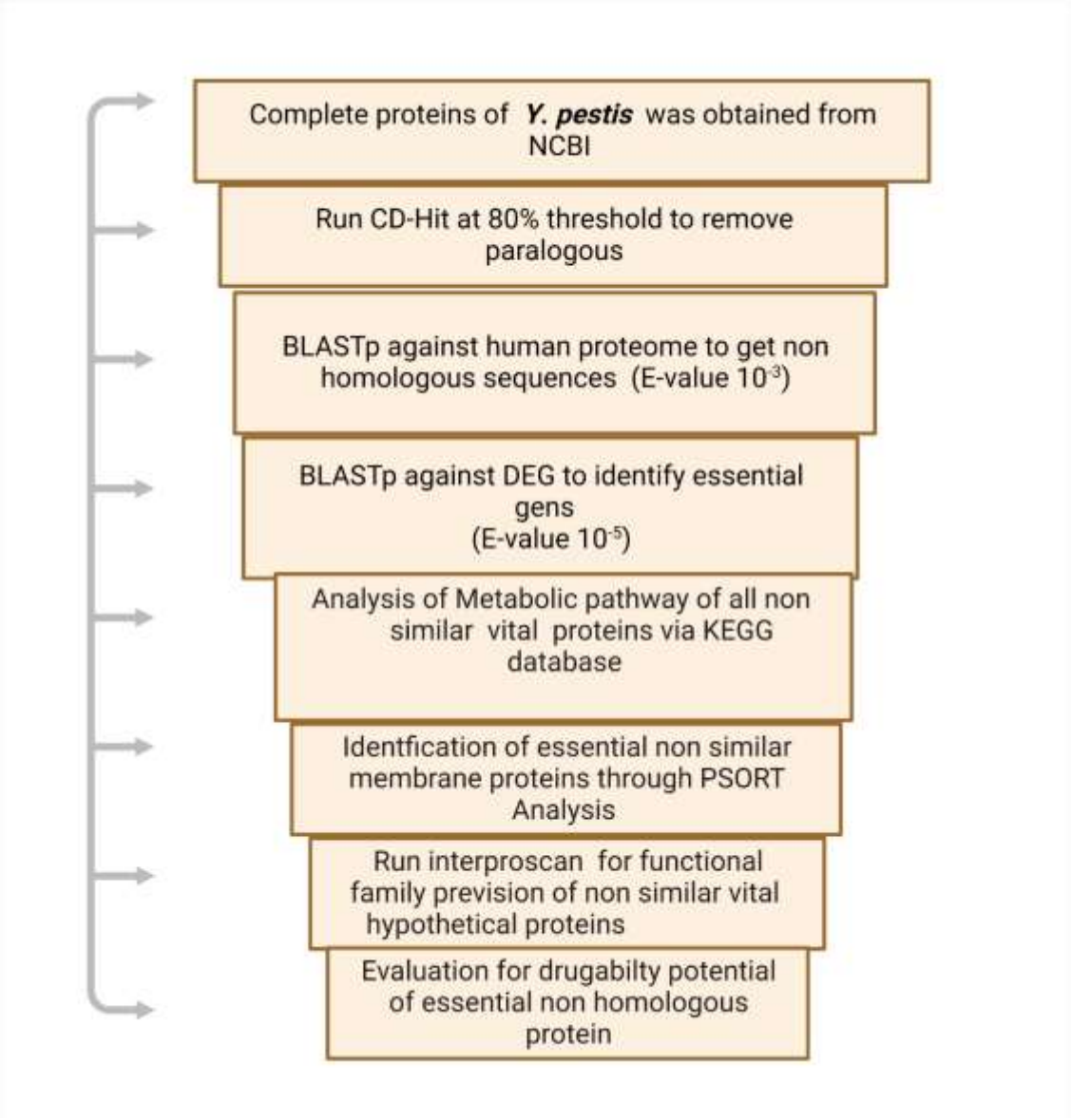

Fig. 1 Overall workflow chart

3. Results and discussion

To search the potent drug targets against Y. pestis strain FDAARGOS_601 was the major objective of this study. The drug targets that are predicted must fulfill the drug potentiality criterion which enable that targets of drug should host non-homologous, vital for pathogen and perform essential major pathogen metabolic pathway. The fig.1 representing the plan of the research work and different outcomes of each step is shows by Table 1. The total proteome of Y. pestis strain FDAARGOS_601 consists of 4104 number of protein sequences which was obtained from NCBI. Then these 4104 protein sequences were then screened via CD-HIT with 0.8 threshold. To remove paralogue and less than 100 amino acids proteins this step was done. This step results in the identification of 277 duplicate and 414 less than 100 amino acids sequences of protein which were removed and so 3413 proteins were remained. Then non-duplicate 3413 proteins were subjected in BLASTp against genome of human. This outcomes displayed 2538 sequences which were non similar with genome of host. Proteins that are similar to genome of human when used as targets for drug in the patients may cause some unpleasant side effects and toxic reactions. Then these proteins of Y. pestis were passed via

BLASTp against DEG with an 10-5 E-value threshold to obtained set of 230 number of proteins. Overall summary of the project is depicted in Table 1. Genes vital for the pathogen survival also with reality that these genes should be different from genome of host; clench huge pledge to be use targets for drugs in specific species.

S. No.	Steps	Y. pestis
1	Complete set of proteins	4104
2	Extracted paralogous (>80% identity) in CD-HIT	3413
3	No. of proteins non similar to host using BLASTp (E-value 10^{-3})	2538
4	Vital proteins sequences in DEG (E-value 10^{-5})	230
5	Vital proteins involved in metabolic pathways (KEGG)	103
6	No. of vital membrane proteins (PSORT)	63
7	No. hypothetical protein as vital proteins (Interproscan)	02
8	Essential drug target proteins (DBD)	04

Table 1. The various steps of subtractive genomics and their results**.**

3.1. Subcellular location of essential and non-homologous proteins

Protein location is important requirement to show its activity. To do its function properly protein needed to be in a specific place. There are methods available for foretelling of proteins subcellular localization by investigating the proteins. PSORTb is the best approach for foretelling of subcellular location. The vital and non-similar sequences were introduced to PSORTb that results that the most of proteins sequences (49 %) are present in the cytoplasm region (Fig. 2). The 2nd high number of the proteins were present in cytoplasm membrane. proteins of Cytoplasm membrane may be the dynamic vaccine targets.

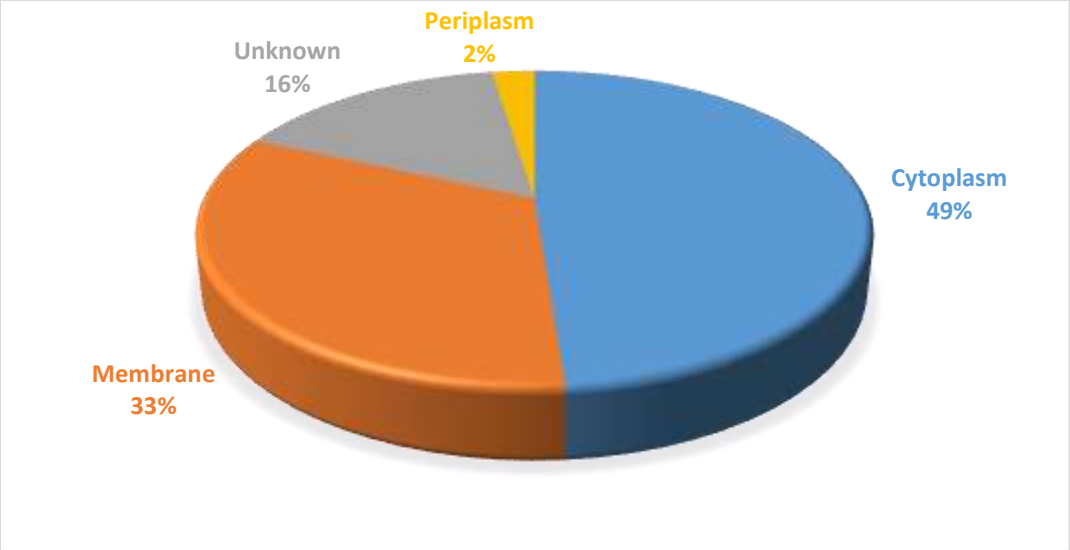

Fig. 2. Subcellular localization of vital non-similar proteins of **Y. pestis**.

3.2. Vital, non-similar and hypothetical proteins family prediction

Hypothetical proteins have known sequences but have unknown function. There are numbers of sequences that are hypothetical in bacteria which causes infection. For prediction of function of hypothetical sequences, the good approach is to classify proteins according to homogeneity in functional groups sequence. One of the best approaches to categorize sequences in functional groups/families is Interproscan. In Y. pestis 2 hypothetical proteins sequences were present. For this foretelling we used the interproscan method. The most common families were ubiquitin-protein transferase family

3.3. KEGG metabolic pathways analysis

Then vital non similar proteins were introduced in KASS server. For detection of the contribution of the subjected sequences in various crucial metabolic pathways which are involved in bacteria this server was used. Total 230 sequences of proteins were subjected to this server. The different pathways of metabolism in which the Y. pestis take part are. Glycolysis, Gluconeogenesis, Citrate cycle, Galactose and mannose metabolism, metabolism of sulfur, Oxidative phosphorylation, Fructose metabolism, metabolism of purine, Aldar ate and Ascorbate metabolism, Propanoate metabolism, metabolism of Sucrose and Starch, Nitrogen metabolism, metabolism of methane, Dicarboxylate and Glyoxylate metabolism, metabolism of Butanoate, metabolism of Pyruvate, Carbon fixation pathways in prokaryotes, Pyrimidine metabolism, Methionine and Cysteine metabolism, biosynthesis of Lysine, Arginine biosynthesis, D-Amino acid metabolism-Antigen repeat unit biosynthesis, metabolism of Threonine, Glycine and serine, biosynthesis of Phenylalanine, Peptidoglycan biosynthesis, biosynthesis of Fatty acid, Arabinogalactan biosynthesis– Mycobacterium, Thiamine metabolism, Biotin metabolism, Biosynthesis of Ubiquinone and other terpenoid-quinone, ABC transporters, Terpenoid backbone biosynthesis, Siderophore group non ribosomal peptides biosynthesis, Nucleotide excision repair, Monobactam biosynthesis, Atrazine degradation, metabolism of Drug, biosynthesis of Ribosome and Aminoacyl-tRNA, RNA degradation, Homologous recombination, DNA replication, Mismatch repair, Phosphotransferase system (PTS),Two-component system, Flagellar assembly, Plant hormone signal transduction, Cell cycle – Caulobacter, Quorum sensing, Biofilm formation - Vibrio cholera, Bacterial chemotaxis, Biofilm formation - Pseudomonas aeruginosa, Biofilm formation - Escherichia coli, Pertussis, Epithelial cell signaling in Helicobacter pylori infection, resistance to beta Lactam, resistance to Cationic antimicrobial peptide (CAMP), resistance to Vancomycin.

In above mentioned metabolic pathways some are associated to pathogen unique metabolic pathways. figure 3 shows distinctive pathways of metabolism. Those proteins which are associated to these pathways are the most powerful drug targets due to non-availability of antagonist pathways in human so there is because in human there are no antagonist pathways so the chances of side effects are extremely rare.

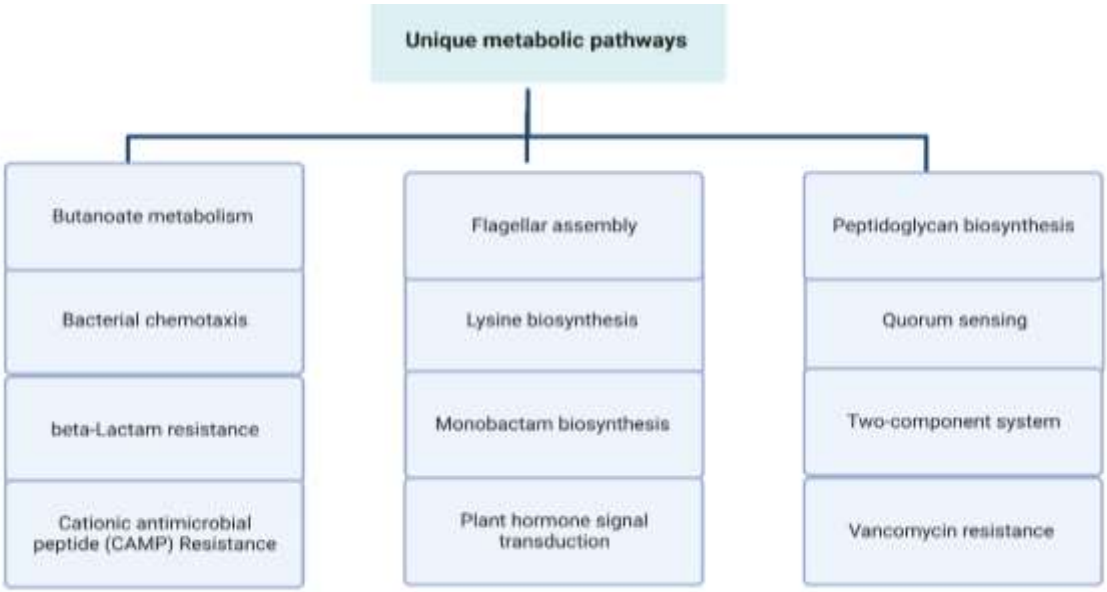

Figure 3. Unique metabolic pathways of **Y. pestis** from KEGG

3.4. Shortlisted Sequences Draggability capacity

By screening out the draggability capacity of all the shortlisted vital proteins this study was further augmented. In the database of Drug Bank, It results the recognition of 4 proteins that have resemblance with the FDA sanctioned drugs targets. Information's of choose proteins with IDs of target were shown in Table No. 2. The Four targets of drugs which shows resemblance to Drug Bank homologs are corelated to Probable pyruvate-flavodoxin Oxidoreductase, Pyruvate synthase, Antizyme inhibitor 2 and Ornithine decarboxylase. These Four proteins can be dynamic drug targets since in Drug Bank all have one analogue with at least 40 % homogeneity in sequence.

Protein ID	Protein name	Drug bank ID	Drug bank organism
P52647	Probable pyruvate-flavodoxin Oxidoreductase	DB00698	E.coli, Salmonella enteritidis PT4 Citrobacter rodentium
P94692	Pyruvate synthase	DB00507	E.coli, Homo sapiens
Q96A70	Antizyme inhibitor 2	DB00125	E.coli, Mus musculus, Homo sapiens
P11926	Ornithine decarboxylase	DB00127; DB06243	E.coli,Homo sapiens,Trypanosoma brucei brucei,Mus musculus

Table 2. non similar vital proteins of **Y. pestis** strain FDAARGOS_601 similar to FDA approved drugs against database of Drug Bank using BLASTP

3.5. Human gut-metagenomes screening

So, to target only those proteins which are only present in the Y. Pestis we perform blastp against human microbiome data base which results the above four proteins sequences have no similarity to human microbiome data base so these four proteins can be a valuable drug targets.

4. Conclusion

Via various tools of bioinformatics, the process of discovery of drug have revolutionized by the idea of analysis and elucidation of available sequences of proteins and genome of different pathogens at various databases. The case of raise in resistance to drug backing the use of in silico approaches to recognize probable targets for drug which should not show any similarity to human proteome. This process is made possible by approach of Subtractive genomics in the characterization of vital, non-similar proteins that can be targeted for finding of therapeutic compounds against Y. pestis strain FDAARGOS_601. Drugs that target Strain specific vital non homologous pathogen proteins guarantee the elimination of infection with little host side effects

5. Acknowledgement

This research was funded by Deanship of Scientific Research (DSR) at King Abdul-Aziz University, Jeddah, grant number G: 335-662-1440" and "The APC was funded by". Faisal A Al Zahrani.

6. References

1. Anisimov, A.P.; Amoako, K.K.J.J.o.m.m. Treatment of plague: promising alternatives to antibiotics. 2006, 55, 1461-1475.
2. Ditchburn, J.-L.; Hodgkins, R.J.B.; Health. Yersinia pestis, a problem of the past and a re-emerging threat. 2019, 1, 65-70.
3. Malloy, G.S.; Brandeau, M.L.; Goldhaber-Fiebert, J.D.J.T.m.; disease, i. Modeling the cost-effectiveness of interventions to prevent plague in Madagascar. 2021, 6, 101.
4. Riedel, S. Plague: from natural disease to bioterrorism. In Proceedings of the Baylor University Medical Center Proceedings, 2005; pp. 116-124.
5. Hull, H.F.; Montes, J.M.; Mann, J.M.J.J.o.I.D. Septicemic plague in New Mexico. 1987, 155, 113-118.
6. van Roon, A.; Maas, M.; Toale, D.; Tafro, N.; van der Giessen, J.J.P.o. Live exotic animals legally and illegally imported via the main Dutch airport and considerations for public health. 2019, 14, e0220122.
7. Poland, J.D.; Dennis, D.T.J.P.m.e., distribution, surveillance; control. Treatment of plague. 1999, 55-62.
8. Fair, R.J.; Tor, Y.J.P.i.m.c. Antibiotics and bacterial resistance in the 21st century. 2014, 6, PMC. S14459.
9. Dai, R.; He, J.; Zha, X.; Wang, Y.; Zhang, X.; Gao, H.; Yang, X.; Li, J.; Xin, Y.; Wang, Y.J.P.n.t.d. A novel mechanism of streptomycin resistance in Yersinia pestis: Mutation in the rpsL gene. 2021, 15, e0009324.
10. Cabanel, N.; Bouchier, C.; Rajerison, M.; Carniel, E.J.I.J.o.A.A. Plasmid-mediated doxycycline resistance in a Yersinia pestis strain isolated from a rat. 2018, 51, 249-254.
11. Galimand, M.; Carniel, E.; Courvalin, P.J.A.a.; chemotherapy. Resistance of Yersinia pestis to antimicrobial agents. 2006, 50, 3233-3236.
12. Control, C.f.D.; Prevention. Human plague--United States, 1993-1994. 1994, 43, 242-246.
13. Matsuoka, M.; Suzuki, Y.; Garcia, I.E.; Fafotis-Morris, M.; Vargas-Gonzalez, A.; Carreno-Martinez, C.; Fukushima, Y.; Nakajima, C.J.J.j.o.i.d. Possible mode of emergence for drug-resistant leprosy is revealed by an analysis of samples from Mexico. 2010, 63, 412-416.
14. Wadood, A.; Jamal, A.; Riaz, M.; Khan, A.; Uddin, R.; Jelani, M.; Azam, S.S.J.M.p. Subtractive genome analysis for in silico identification and characterization of novel drug targets in Streptococcus pneumonia strain JJA. 2018, 115, 194-198.
15. Barh, D.; Tiwari, S.; Jain, N.; Ali, A.; Santos, A.R.; Misra, A.N.; Azevedo, V.; Kumar, A.J.D.D.R. In silico subtractive genomics for target identification in human bacterial pathogens. 2011, 72, 162-177.
16. Uddin, R.; Saeed, K.J.C.b.; chemistry. Identification and characterization of potential drug targets by subtractive genome analyses of methicillin resistant Staphylococcus aureus. 2014, 48, 55-63.
17. Uddin, R.; Saeed, K.; Khan, W.; Azam, S.S.; Wadood, A.J.G. Metabolic pathway analysis approach: identification of novel therapeutic target against methicillin resistant Staphylococcus aureus. 2015, 556, 213-226.
18. Tatusova, T.A.; Karsch-Mizrachi, I.; Ostell, J.A.J.B. Complete genomes in WWW Entrez: data representation and analysis. 1999, 15, 536-543.

19. Li, W.; Godzik, A.J.B. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. 2006, 22, 1658-1659.
20. Zhang, R.; Ou, H.Y.; Zhang, C.T.J.N.a.r. DEG: a database of essential genes. 2004, 32, D271-D272.
21. Moriya, Y.; Itoh, M.; Okuda, S.; Yoshizawa, A.C.; Kanehisa, M.J.N.a.r. KAAS: an automatic genome annotation and pathway reconstruction server. 2007, 35, W182-W185.
22. 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.J.B. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. 2010, 26, 1608-1615.
23. Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.J.N.a.r. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. 2010, 39, D1035-D1041.
24. Wu, G.D.; Lewis, J.D.J.C.G.; Hepatology. Analysis of the human gut microbiome and association with disease. 2013, 11, 774-777.