Journal of
Health & Medical Informatics

Chordia et al., J Health Med Informat 2019, 10:2

Research Article Open Access

Interactome Analysis for Identification of Common Drug Targets in *Salmonella* Species

Nikita Chordia*, Priyesh Hardia and Priya Jain

Bioinformatics Sub-Center, School of Biotechnology, Takshashila Campus, Devi Ahilya University, Indore, India

Abstract

A wide variety of human population is infected with different species of *Salmonella* causing salmonellosis. At present, the major hurdle in treating infection is the development of resistance to existing antibiotics. Therefore, there is a need to identify new drug targets so that new drugs can be designed that can cure the infection efficiently. Here, we have design an interactome for all species of *Salmonella* using its essential and non-homologous protein to humans. So that, all types of infections of *Salmonella* can be treated with single antibiotic. 1399 essential proteins of *Salmonella* have been analyzed using interactome studies. We found 09 proteins as putative drug targets that can be used to treat all species of *Salmonella*.

Keywords: Drug target; Interactome; *Salmonella*; Bacteria; Non-homologous; Infection; Antibiotic resistance; Treatment

Introduction

Salmonella species are gram-negative, flagellated, facultative anaerobic bacilli of the family Enterobacteriaceae [1]. *Salmonella* infection (salmonellosis) is a common bacterial disease that affects the intestinal tract. Clinically it ranges from the common *Salmonella* gastroenteritis (abdominal cramps, diarrhea and fever) to enteric fevers (typhoid) that maybe sometimes life threatening and therefore, requires prompt antibiotic therapy [2]. But, today the main hurdle in treatment of the bacterial infection is the development of resistance to existing antibiotics [3]. Therefore, there is a need to identify novel bacterial targets for drug development [4].

Salmonella species are one of the major causes of human gastroenteritis and thousands of cases were reported annually for salmonellosis [5]. *Salmonella* typically live in animal and human intestines. Humans become infected most frequently through contaminated water or food. The most commonly identified food sources include meat products, eggs, dairy products and raw fruits and vegetables. Faecal/intestinal contamination of carcasses is the principal source of human food-borne infections [6].

Brenner et al. reported 2,463 serotypes (serovars) of *Salmonella* which includes several species such as *S. enterica*, *S. bongori*, *S. typhimurium*, *S. typhi*, *S. enteritidis*, *S. heidelberg*, *S. subterranean* and many more [7]. Serotypes of *Salmonella* can be divided into two main groups-typhoidal and non-typhoidal. Typhoidal serotypes are strictly adapted to humans and include *Salmonella typhi*, *S Paratyphi A*, *S Paratyphi B*, and *S Paratyphi C*. Nontyphoidal serotypes are more common and generally results in food poisoning. They are zoonotic i.e. they can be transferred between humans and other animals. Most infections are caused by the *S. typhimurium*, *S. enteric* or *S. enteritidis* [8].

Parry and Threlfall reported antimicrobial resistance in typhoidal and non-typhoidal salmonellae. They reported variable rates of resistance in *Salmonella* with particular reference to quinolones and extended spectrum cephalosporins [9,10]. Increasing occurrence of drug resistance for salmonellosis is a major public health problem. To overcome drug resistance to existing antibiotic therapy for salmonellosis, there is a need to identify common drug targets for all the pathogens of *Salmonella* infection.

Here, in this study we have taken all the essential proteins available for *Salmonella* species. Undesirable proteins which are non-homologous to human are removed from the study and an interactome was created for rest of the proteins. The interactome analysis was performed and impactful proteins in the network are considered to be the common drug target for all the species of *Salmonella*.

Methodology

Dataset of essential proteins

Essential proteins of all *Salmonella* species were downloaded from DEG (Database of Essential Genes) database available online at http://tubic.org/deg_bak/ [11]. Essential proteins are absolutely necessary for the survival of an organism and are considered as foundation of life. Therefore, we have taken essential proteins of *Salmonella* to find the dug target.

Non-host homologous proteins

All essential proteins were subjected to similarity searching to find the non-homologous proteins with the human host. Proteins are identified which were not homologous to the human host to eliminate the chances of cross reactivity with the human genome and to minimize the risk of drug toxicity. This was done using BLAST-P (Basic Local Assignment Search Tool) available online at http://www.ncbi.nlm.gov/blast/ [12].

All essential and non-homologous human proteins were then subjected for interactome construction. Interactome is the whole set of interactions between and among proteins. This was done using STRING database available online at http://string-db.org/ [13].

***Corresponding author:** Nikita Chordia, Bioinformatics Sub-Center, School of Biotechnology, Takshashila Campus, Devi Ahilya University, Khandwa Road, Indore-452001, India, Tel: 0091-731-2470373; 0091-9685281822; E-mail: nikita.chordia25@gmail.com

Received April 03, 2019; **Accepted** May 08, 2019; **Published** May 20, 2019

Citation: Chordia N, Hardia P, Jain P (2019) Interactome Analysis for Identification of Common Drug Targets in *Salmonella* Species. J Health Med Informat 10: 331.

Metabolic functional interaction can be determined by using the methods viz. neighborhood, gene fusion, co-occurrence, co-expression and experiment at medium confidence. Then score was calculated by using the following formula given by Kushwaha and Shakya [14].

$$\text{Confidence score of a target} = \frac{\text{Number of interactants of target by used methods}}{\text{Total number of methods used}}$$

Protein with score value till 0.400 was taken as the metabolically functional protein at medium confidence level.

Interactome analysis

Obtained Interactome was analyzed using Cytoscape 3.7 software [15]. Different parameters were considered namely clustering coefficient, characteristics path length, betweenness centrality and neighborhood connectivity. The values of the whole network were compared with the values of the targeted node (protein). It was done by taking the values of all the parameters of whole network then values are calculated after deleting each node one by one. Difference in values is directly related with the impact of respective protein into interactome.

Clustering coefficient is defined as the degree to which nodes in a network tend to cluster together. It generally shows greater value for the entire protein network and is decreased when node (protein) is deleted. The characteristic path length is defined as the expected distance between the nodes. Betweenness centrality is the measure of centrality, for node "x" it is calculated by summing the number of shortest paths between pairs of nodes that pass through node "x" divided by the total number of shortest paths between pairs of nodes. Neighborhood connectivity is average connectivity of neighbors of given vertex. On the basis of the above mentioned parameters, most impactful/potent protein was selected as drug targets.

Results and Discussion

A total of 1399 essential proteins of *Salmonella* species were downloaded from DEG database. A molecule can be drug target only if it is essential for the survival of the pathogen. Another requirement for drug target is that it must not be homologous to host. So, out of 1339, a total 1220 non-homologous proteins to human were identified using BLAST. All these 1220 proteins were submitted in STRING database for network construction. As we have submitted the sequences of different species of *Salmonella* only therefore, similar sequences are omitted from the interactome. Finally, we got the interactome of 927 proteins as shown in Figure 1.

Interactome was analyzed using cytoscape 3.7. In the first step of analysis, values of clustering coefficient, characteristics path length, betweenness centrality and neighborhood connectivity for the whole network was calculated. Then value was calculated to see the effect of every node in the network. This was done by deleting each node of the network and then values of all the parameter (clustering coefficient, characteristics path length, betweenness centrality and neighborhood connectivity) were noted. It is done for all the 927 nodes of the network. After getting all the values of the parameter for each node, difference in the value was calculated by using the values of the whole network. Difference was noted in almost all the nodes of the network that shows the impact of protein in the network. This is because we have taken all the essential proteins of the pathogen. But noticeable differences in the values of all the four parameters were observed in 9 nodes that are very important for the network. Here, we have taken the cut off value that must be fulfilled by the protein. For clustering coefficient, the difference

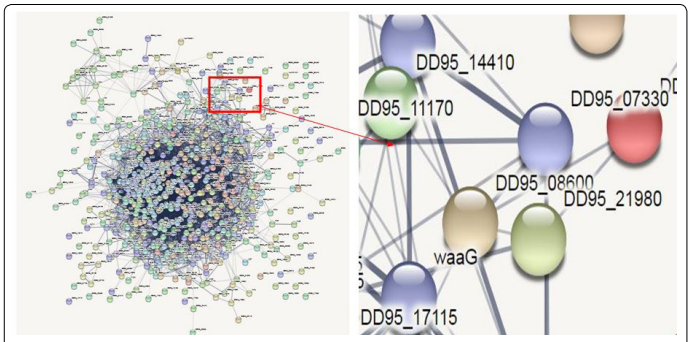

Figure 1: Interactome created by STRING. Different colour shows different interactions like gene neighbourhood, gene fusion, gene co-expression, co-occurrence and text mining.

in values can be positive and negative and any difference in the value is taken as many nodes are showing no difference. For characteristic path length, the difference of greater than 3 was taken for consideration. The cut-off of 1.0E-02 and 160.00 was taken for betweenness centrality and neighborhood connectivity respectively. Table 1 shows the results of those nodes of the network that are showing differences in all the parameters according to the mentioned criteria.

Identified nine proteins from the interactome are having major influence on interaction network. Therefore, targeting any of these protein effect the pathogen metabolism and alterations in these proteins result in killing of the pathogen. The identified potential drug target was also reported by other researchers as drug target in other pathogenic diseases. This shows that our strategy in identifying drug targets from interactome is correct and can be apply for other pathogens also.

waaG codes for Glycosyltransferase and is reported as drug target by Sun [16]. Similarly, *rfbG* is CDP-glucose 4,6-dehydratase and found to be virulent in *Candida albicans* [17]. Thiamine-monophosphate kinase coded by *thiL* and Chromosome partition protein coded by *mukE* was reported to be drug target for Tuberculosis [18,19]. *uppS*-(Isoprenyl transferase) and *secE* (Protein translocase subunit SecE) was also reported as drug target [20,21]. Node *DD95_14630* (Riboflavin biosynthesis protein) was reported as anti-infective drug target [22]. *Alas* (Aminolevulinate synthase) was also reported as drug target for malaria parasite [23]. *rsgA* (Small ribosomal subunit biogenesis GTPase RsgA) is also pinpointed to be drug target by Maguire [24]. All these reported targets are playing crucial role in pathogen's biological processes that are important for the pathogen survival. Therefore, any of these targets can be selected for the disturbance of pathogen machinery. These putative targets have been predicted using interactome analysis. Drug targets using interactome was also reported for other pathogens like *Listeria monocytogenes* [25]. After *in silico* studies, further wet lab work is required. Computational combinatorial methods can be used to find the inhibitors for these targets that can serve as the drug to cure the disease. In addition to this, comparative genomics can be employed so that same medication can be used for different pathogens.

Conclusion

In this study, we identified 09 potential drug targets for *Salmonella* through interactome of *Salmonella* and are less toxic for the host organism. In addition to this, interactome analysis can be employed for any other pathogens to find new drug targets leading to fast drug discovery process. The limitation with this methodology is that it requires complete and reliable data of pathogen's interactome.

Gene name	Clustering coefficient	Clustering coefficient (after deleting node)	Characteristics path length	Characteristics path length (after deleting node)	Betweenness centrality	Betweenness centrality (after deleting node)	Neighborhood connectivity	Neighborhood connectivity (after deleting node)
waaG	0.78	0.32	2.77	3.12	3.32 E-04	1.20E-06	95.15	80.2
rfbG	0.56	0.8	1.56	2.15	2.2 E-05	8.3E-04	101.56	97.12
thiL	0.96	0.56	3.1	2.89	4.44 E-05	7.36E-02	115.26	101.23
mukE	1	0.6	2.86	2.15	1.89E-06	1.66E-04	98.99	96.45
uppS	0.92	0.79	1.98	2.55	2.56 E-05	2.24E-04	89.46	87.56
secE	0.65	0.36	2.82	2.15	7.58E-05	8.93E-04	82.45	81.2
DD95_14630	0.82	0.59	2.95	2.65	5.28E-05	2.98E-04	109.21	101.01
Alas	0.11	0.79	2.29	2.98	2.91E-04	3.47E-04	93.85	92.88
rsgA	0.72	0.14	2.14	2.76	6.32E-05	4.78E-04	96.84	91.25

Table 1: Results of those nodes of the network that are showing differences in all the parameters according to the mentioned criteria.

Acknowledgement

Authors acknowledge the facilities of the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi (DBT) under the Bioinformatics Sub Centre as well as M.Sc. Biotechnology program used in the present work.

Conflict of Interest

The authors confirm that they have no conflict of interest.

References

1. Black PH, Kunz LJ, Swartz MN (1960) Salmonellosis-A review of some unusual aspects. N Engl J Med 262: 921-927.
2. Christenson JC (2013) Salmonella infections. Pediatr Rev 34: 375-383.
3. Paterson DL (2006) Resistance in gram-negative bacteria: Enterobacteriaceae. Am J Infect Control 34: S20-S28.
4. Norrby SR, Nord CE, Finch R (2005) Lack of development of new antimicrobial drugs: A potential serious threat to public health. Lancet Infect Dis 5: 115-119.
5. Sanchez S, Hofacre CL, Lee MD, Maurer JJ, Doyle MP (2002) Animal sources of salmonellosis in humans. J Amer Vet Med Assoc 221: 492-497.
6. L Plym F, Wierup M (2006) Salmonella contamination: A significant challenge to the global marketing of animal food products. Rev Sci Tech 25: 541-554.
7. Brenner FW, Villar RG, Angulo FJ, Tauxe R, Swaminathan B (2000) Salmonella nomenclature. J Clin Microb 38: 2465-2467.
8. Hohmann EL (2001) Nontyphoidal salmonellosis. Clin Infect Dis 32: 263-269.
9. Parry CM, Threlfall EJ (2008) Antimicrobial resistance in typhoidal and nontyphoidal salmonellae. Curr Opin Infect Dis 21: 531-538.
10. Gupta SK, Medalla F, Omondi MW, Whichard JM, Fields PI, et al. (2008) Laboratory-based surveillance of paratyphoid fever in the United States: Travel and antimicrobial resistance. Clin Infect Dis 46: 1656-1663.
11. Gao F, Luo H, Zhang CT, Zhang R (2015) Gene essentiality analysis based on DEG 10, an updated database of essential genes. Methods Mol Biol 1279: 219-233.
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.
13. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, et al. (2017) The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 4: D362-D368.
14. Kushwaha SK, Shakya M (2010) Protein interaction network analysis-approach for potential drug target identification in *Mycobacterium tuberculosis*. J Theoretical Biol 262: 284-294.
15. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504.
16. Sun XL (2013) Glycosyltransferases as potential drug targets. Med Chem 3: e106.
17. Sen M, Shah B, Rakshit S, Singh V, Padmanabhan B, et al. (2011) UDP-glucose 4,6-dehydratase activity plays an important role in maintaining cell wall integrity and virulence of Candida albicans. PLoS Pathogens 7: e1002384.
18. Khare G, Kar R, Tyagi AK (2011) Identification of inhibitors against *Mycobacterium tuberculosis* thiamin phosphate synthase, an important target for the development of anti-TB drugs. PloS One 6: e22441.
19. Nisa S, Blokpoel MC, Robertson BD, Tyndall JD, Lun S, et al. (2010) Targeting the chromosome partitioning protein ParA in tuberculosis drug discovery. J Antimicrobial Chemother 65: 2347-2358.
20. Wiesner J, Jomaa H (2007) Isoprenoid biosynthesis of the apicoplast as drug target. Curr Drug Targets 8: 3-13.
21. Chaudhary AS, Chen W, Jin J, Tai PC, Wang B (2015) SecA: A potential antimicrobial target. Future Med Chem 7: 989-1007.
22. Long Q, Ji L, Wang H, Xie J (2010) Riboflavin biosynthetic and regulatory factors as potential novel anti-infective drug targets. Chem Biol Drug Des 75: 339-347.
23. Bonday ZQ, Dhanasekaran S, Rangarajan PN, Padmanaban G (2000) Import of host δ-aminolevulinate dehydratase into the malarial parasite: Identification of a new drug target. Nat Med 6: 898-903.
24. Maguire BA (2009) Inhibition of bacterial ribosome assembly: A suitable drug target? Microbiol Mol Biol Rev 73: 22-35.
25. Chordia N, Sharma NK, Kumar A (2015) An interactomic approach for identification of putative drug targets in Listeria monocytogenes. Int J Bioinfo Res Appl 11: 315-325.