World Journal of Pharmaceutical Science and Research
www.wjpsronline.com
Research Article
ISSN: 2583-6579
SJIF Impact Factor: 3.454
Year - 2023
Volume: 2; Issue: 5
Page: 117-122

POTENTIAL DRUG TARGET IDENTIFICATION IN *BRUCELLA ABORTUS* BY SUBTRACTIVE GENOMICS APPROACH

Manjushree Awari and Vivek Keshri*

Article Received: 20 August 2023 || Article Revised: 08 September 2023 || Article Accepted: 29 September 2023

Corresponding Author: Vivek Keshri
REVA University, School of Applied Sciences, Department of Biotechnology, Bengaluru-560064, Karnataka, India.

ABSTRACT
The discovery of essential proteins in the pathogenic bacterium *Brucella abortus*, distinct from those in its host, *Bos taurus* (cattle), presents a promising avenue for combating bacterial infections. Using a combination of subtractive genomics and the NCBI's BLAST program, researchers have unveiled new antibiotic candidates through an examination of *Brucella abortus*, the causative agent of brucellosis. This study has spotlighted 64 significant protein sequences out of the total 2976 in *Brucella abortus* that have minimal resemblance to cattle proteins. A thorough investigation has unveiled the involvement of these 64 critical proteins in various metabolic processes, offering novel targets for antibacterial therapy against bovine brucellosis. Within this group of 64 proteins, two highly virulent ones have been identified as potential targets for future drug development efforts, paving the way for innovative treatments.

KEYWORDS: Antibacterial intervention; Brucellosis; Cattle (*Bos taurus*); Therapeutic target proteins; Virulent Proteins.

INTRODUCTION

Brucellosis is a bacterial infection caused by *Brucella* species, most notably *Brucella abortus*, *Brucella melitensis, Brucella suis*, and *Brucella canis*.[1] These can infect both animals (cattle, goats and sheep, swine, dogs, wildlife and others) and humans, resulting in a zoonotic infection.[1] The Brucella genus consists of ten species, with eight inhabiting terrestrial environments and two adapted to marine habitats.[2] While each species has the potential to infect various host species, it typically exhibits a preference for a particular host. Bovine brucellosis, primarily caused by *Brucella abortus*, poses a significant economic challenge due to the substantial harm it causes to the commercial livestock industry.[3] Humans are often exposed to the disease through direct or indirect contact with diseased animals or their products. Controlling brucellosis is imperative due to its dual threat as a zoonotic disease and an economic burden. Searching for new drug targets in bacteria is crucial due to the global impact of brucellosis as a significant public health and economic burden. Current treatment options for brucellosis are limited and often associated with challenges such as prolonged therapy, relapses, and the emergence of antibiotic resistance. By identifying novel drug targets, researchers can develop more effective and specific therapies that target essential bacterial processes, improving treatment outcomes, reducing treatment duration, and minimizing the risk of resistance. *Brucella abortus* is a gram-negative

bacteria, the causative agent of brucellosis in cattle (causing reproductive problems such as abortions, stillbirths, and reduced fertility) and also infects other animals and, less commonly, humans. The bioinformatics approach to novel drug target identification represents a transformative paradigm in modern pharmaceutical research, seamlessly integrating cutting-edge computational techniques with vast biological datasets to unravel complex disease mechanisms and expedite drug discovery. Computational approaches in particular comparative and subtractive genomics have been extensively used to identify novel drug targets in infectious pathogens. These approaches are powerful, speedy and cost-effective in drug discovery and development processes compared to conventional methods. Taking this as an advantage, we implemented a subtractive genomics approach[4] to predict potential drug targets in *Brucella abortus* which has fetched two druggable proteins. The current drug target identification will not only enhances patient care and well-being but also contributes to controlling the spread of brucellosis, safeguarding livestock industries, and aligning with broader efforts in combating antimicrobial resistance and promoting One Health initiatives for integrated human, animal, and environmental health.

MATERIAL AND METHODS

Data retrieval

The protein sequences of *Brucella abortus* and cattle (*Bos taurus*) were obtained from the NCBI database. The bacterial essential genes were acquired from the Database of Essential Gene Database (DEG). [5–8] A collective count of 2,976 sequences of *Brucella abortus,* 26,619 essential proteins from DEG, and 63,628 sequences of *Bos taurus* were obtained. The complete data analysis is depicted in Fig.1.

Detection of essential proteins in *Brucella abortus*

Essential genes in bacteria were identified through BLASTp[9] analysis in the DEG (Database of Essential Genes).[5–8] BLAST results were then analyzed, with a focus on identifying significant matches to established essential genes in DEG, using specific criteria including an E-value threshold of 1e-10 and an alignment score exceeding 70%. The host *Bos taurus* lacked these essential genes.

Metabolic Pathway Analysis

The KAAS (KEGG Automatic Annotation Server) at KEGG performed a metabolic pathway study for the metabolic processes[10,11] in different pathways. KAAS performs functional gene annotation by conducting BLAST comparisons against the meticulously curated KEGG GENES database. The outcome comprises KO assignments as well as KEGG pathways that are generated automatically.

Prediction of Virulent proteins

The bi-layer cascade support vector machine (SVM) prediction program VirulentPred[12] was utilized to predict sequences of bacterial virulent proteins. In the initial layer, diverse protein sequence features were employed to train and fine-tune SVM classifiers. These classifiers were subsequently cascaded to the second layer SVM classifier to create the ultimate classifier. The selected prediction methods for the query encompassed amino acid composition, dipeptide composition, similarity searching, higher-order dipeptide composition, position-specific scoring matrix (PSSM), and the cascaded SVM module.

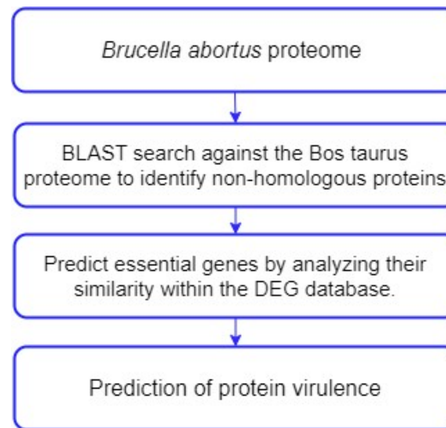

Fig. 1: Flow chart of systematic identification of novel targets in *Brucella abortus*.

RESULT AND DISCUSSION

Biological strategies have been employed for the identification of essential genes within bacterial organisms, while computational methods have been utilized for the same purpose in prokaryotes. In our study, we focused on Brucella abortus, a cattle pathogenic bacterium, and identified potential antibacterial targets. Our approach involved aligning protein sequences from the NCBI with the DEG essential gene database. From the initial pool of 2976 proteins in Brucella abortus, we identified 423 unique sequences that showed no significant similarity to cattle sequences. Subsequent alignment with the essential gene database yielded 64 strongly matched unique sequences. Functional classification based on gene names or descriptions suggested that these proteins might be specific to the pathogen and play a role in a crucial metabolic pathway. The KEGG GENES database further categorized these 64 significant genes based on their involvement in distinct metabolic pathways, using the KEGG orthology (KO) approach. This database serves as a valuable resource for cross-species genome annotation. The annotated sequences included various proteins such as glycyl-tRNA synthetase alpha chain, chorismate synthase, 3-oxoacyl-[acyl-carrier-protein] synthase II and III, and several others, as detailed in Table 1. These protein sequences represent potential novel targets for further investigation.

CONCLUSIONS

In conclusion, the application of the Support Vector Machine (SVM) approach in predicting virulence/non-virulence properties unveiled two virulent proteins (MBT2252796.1 and MBT2253300.1). This prediction/analysis suggests that these essential proteins may play a significant role in the regular functioning of the pathogen within the host. As a result, these proteins hold promise as potential targets for drug development.

Table 1: KEGG orthology of all essential potential antimicrobial targets.

Query Seq.	KO list	(non)Virulent	Pathways
MBT2250995.1	K01581	Non-Virulent	E4.1.1.17, ODC1, speC, speF; ornithine decarboxylase
MBT2251071.1	K01892	Non-Virulent	HARS, hisS; histidyl-tRNA synthetase [EC:6.1.1.21]
MBT2251078.1	K04077	Non-Virulent	groEL, HSPD1; chaperonin GroEL [EC:5.6.1.7]
MBT2251083.1	K01870	Non-Virulent	IARS, ileS; isoleucyl-tRNA synthetase [EC:6.1.1.5]
MBT2251185.1	K00525	Non-Virulent	E1.17.4.1A, nrdA, nrdE; ribonucleoside-diphosphate reductase alpha chain
MBT2251186.1	K00526	Non-Virulent	E1.17.4.1B, nrdB, nrdF; ribonucleoside-diphosphate reductase beta chain
MBT2251413.1	K00382	Non-Virulent	DLD, lpd, pdhD; dihydrolipoamide dehydrogenase [EC:1.8.1.4]
MBT2251417.1	K00658	Non-Virulent	DLST, sucB; 2-oxoglutarate dehydrogenase E2 component

			(dihydrolipoamide succinyltransferase)
MBT2251418.1	K00164	Non-Virulent	OGDH, sucA; 2-oxoglutarate dehydrogenase E1 component [EC:1.2.4.2]
MBT2251419.1	K01902	Non-Virulent	sucD; succinyl-CoA synthetase alpha subunit [EC:6.2.1.5]
MBT2251421.1	K00024	Non-Virulent	mdh; malate dehydrogenase [EC:1.1.1.37]
MBT2251469.1	K01586	Non-Virulent	lysA; diaminopimelate decarboxylase [EC:4.1.1.20]
MBT2251482.1	K09812	Non-Virulent	ftsE; cell division transport system ATP-binding protein
MBT2251505.1	K14572	Non-Virulent	MDN1, REA1; midasin
MBT2251598.1	K01889	Non-Virulent	FARSA, pheS; phenylalanyl-tRNA synthetase alpha chain [EC:6.1.1.20]
MBT2251602.1	K03686	Non-Virulent	dnaJ; molecular chaperone DnaJ
MBT2251642.1	K00647	Non-Virulent	fabB; 3-oxoacyl-[acyl-carrier-protein] synthase I [EC:2.3.1.41]
MBT2251678.1	K02945	Non-Virulent	RP-S1, rpsA; small subunit ribosomal protein S1
MBT2251736.1	K01681	Non-Virulent	ACO, acnA; aconitate hydratase [EC:4.2.1.3]
MBT2251767.1	K02470	Non-Virulent	gyrB; DNA gyrase subunit B [EC:5.6.2.2]
MBT2251894.1	K00812	Non-Virulent	aspB; aspartate aminotransferase [EC:2.6.1.1]
MBT2251900.1	K01955	Non-Virulent	carB, CPA2; carbamoyl-phosphate synthase large subunit [EC:6.3.5.5]
MBT2252132.1	K02622	Non-Virulent	parE; topoisomerase IV subunit B [EC:5.6.2.2]
MBT2252134.1	K02433	Non-Virulent	gatA, QRSL1; aspartyl-tRNA(Asn)/glutamyl-tRNA(Gln) amidotransferase subunit A [EC:6.3.5.6 6.3.5.7]
MBT2252144.1	K03168	Non-Virulent	topA; DNA topoisomerase I [EC:5.6.2.1]
MBT2252380.1	K00821	Non-Virulent	argD; acetylornithine/N-succinyldiaminopimelate aminotransferase [EC:2.6.1.11 2.6.1.17]
MBT2252480.1	K00208	Non-Virulent	fabI; enoyl-[acyl-carrier protein] reductase I [EC:1.3.1.9 1.3.1.10]
MBT2252494.1	K01662	Non-Virulent	dxs; 1-deoxy-D-xylulose-5-phosphate synthase [EC:2.2.1.7]
MBT2252517.1	K09458	Non-Virulent	fabF, OXSM, CEM1; 3-oxoacyl-[acyl-carrier-protein] synthase II [EC:2.3.1.179]
MBT2252796.1	**K03106**	**Virulent**	**SRP54, ffh; signal recognition particle subunit SRP54 [EC:3.6.5.4]**
MBT2252819.1	K02111	Non-Virulent	ATPF1A, atpA; F-type H+/Na+-transporting ATPase subunit alpha [EC:7.1.2.2 7.2.2.1]
MBT2252821.1	K02112	Non-Virulent	ATPF1B, atpD; F-type H+/Na+-transporting ATPase subunit beta [EC:7.1.2.2 7.2.2.1]
MBT2252878.1	K00615	Non-Virulent	tktA, tktB; transketolase [EC:2.2.1.1]
MBT2252909.1	K03798	Non-Virulent	ftsH, hflB; cell division protease FtsH [EC:3.4.24.-]
MBT2252978.1	K00940	Non-Virulent	ndk, NME; nucleoside-diphosphate kinase [EC:2.7.4.6]
MBT2253022.1	K01714	Non-Virulent	dapA; 4-hydroxy-tetrahydrodipicolinate synthase [EC:4.3.3.7]
MBT2253129.1	K02469	Non-Virulent	gyrA; DNA gyrase subunit A [EC:5.6.2.2]
MBT2253157.1	K00382	Non-Virulent	DLD, lpd, pdhD; dihydrolipoamide dehydrogenase [EC:1.8.1.4]
MBT2253158.1	K00627	Non-Virulent	DLAT, aceF, pdhC; pyruvate dehydrogenase E2 component (dihydrolipoamide acetyltransferase) [EC:2.3.1.12]
MBT2253159.1	K00162	Non-Virulent	PDHB, pdhB; pyruvate dehydrogenase E1 component beta subunit [EC:1.2.4.1]
MBT2253160.1	K00161	Non-Virulent	PDHA, pdhA; pyruvate dehydrogenase E1 component alpha subunit [EC:1.2.4.1]
MBT2253163.1	K01689	Non-Virulent	ENO, eno; enolase [EC:4.2.1.11]
MBT2253179.1	K01647	Non-Virulent	CS, gltA; citrate synthase [EC:2.3.3.1]
MBT2253236.1	K00939	Non-Virulent	adk, AK; adenylate kinase [EC:2.7.4.3]
MBT2253240.1	K02988	Non-Virulent	RP-S5, MRPS5, rpsE; small subunit ribosomal protein S5
MBT2253259.1	K02358	Non-Virulent	tuf, TUFM; elongation factor Tu
MBT2253260.1	K02355	Non-Virulent	fusA, GFM, EFG; elongation factor G
MBT2253264.1	K03046	Non-Virulent	rpoC; DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6]
MBT2253265.1	K03043	Non-Virulent	rpoB; DNA-directed RNA polymerase subunit beta [EC:2.7.7.6]
MBT2253272.1	K02358	Non-Virulent	tuf, TUFM; elongation factor Tu
MBT2253300.1	**K06861**	**Virulent**	**lptB; lipopolysaccharide export system ATP-binding protein**
MBT2253457.1	K00525	Non-Virulent	E1.17.4.1A, nrdA, nrdE; ribonucleoside-diphosphate reductase alpha chain

MBT2253480.1	K00029	Non-Virulent	maeB; malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+) [EC:1.1.1.40]
MBT2253486.1	K00383	Non-Virulent	GSR, gor; glutathione reductase (NADPH) [EC:1.8.1.7]
MBT2253533.1	K00342	Non-Virulent	nuoM; NADH-quinone oxidoreductase subunit M [EC:7.1.1.2]
MBT2253541.1	K09810	Non-Virulent	lolD; lipoprotein-releasing system ATP-binding protein [EC:7.6.2.-]
MBT2253562.1	K01756	Non-Virulent	purB, ADSL; adenylosuccinate lyase [EC:4.3.2.2]
MBT2253589.1	K01887	Non-Virulent	RARS, argS; arginyl-tRNA synthetase [EC:6.1.1.19]
MBT2253636.1	K01876	Non-Virulent	DARS2, aspS; aspartyl-tRNA synthetase [EC:6.1.1.12]
MBT2253698.1	K11717	Non-Virulent	sufS; cysteine desulfurase / selenocysteine lyase [EC:2.8.1.7 4.4.1.16]
MBT2253700.1	K09013	Non-Virulent	sufC; Fe-S cluster assembly ATP-binding protein
MBT2253706.1	K01866	Non-Virulent	YARS, tyrS; tyrosyl-tRNA synthetase [EC:6.1.1.1]
MBT2253767.1	K01961	Non-Virulent	accC; acetyl-CoA carboxylase, biotin carboxylase subunit [EC:6.4.1.2 6.3.4.14]

Data Availability: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. All data generated or analyzed during this study are included in this published article.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Research funding: None declared

Competing interests: Authors state no conflict of interest.

REFERENCES

1. De Massis F, Zilli K, Donato G Di, et al (2019) Distribution of Brucella field strains isolated from livestock, wildlife populations, and humans in Italy from 2007 to 2015. PLoS One 14. https://doi.org/10.1371/journal.pone.0213689.
2. Yu WL, Nielsen K (2010) Review of detection of brucella spp. by polymerase chain reaction. Croat. Med. J., 51: 306–313.
3. Khurana SK, Sehrawat A, Tiwari R, et al (2021) Bovine brucellosis–a comprehensive review. Vet. Q., 41: 61–88.
4. Keshri V, Singh DP, Prabha R, et al (2014) Genome subtraction for the identification of potential antimicrobial targets in Xanthomonas oryzae pv. oryzae PXO99A pathogenic to rice. 3 Biotech, 4: 91–95. https://doi.org/10.1007/s13205-013-0131-7.
5. Luo H, Lin Y, Gao F, et al (2014) DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic Acids Res, 42. https://doi.org/10.1093/nar/gkt1131.
6. Zhang R, Lin Y (2009) DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. Nucleic Acids Res, 37: 455–458. https://doi.org/10.1093/nar/gkn858.
7. Zhang R, Ou HY, Zhang CT (2004) DEG: a database of essential genes. Nucleic Acids Res, 32. https://doi.org/10.1093/NAR/GKH024.
8. Zhang R, Ou HY, Zhang CT (2004) DEG: A database of essential genes. Nucleic Acids Res, 32. https://doi.org/10.1093/nar/gkh024.
9. Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. J Mol Biol, 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

10. Moriya Y, Itoh M, Okuda S, et al (2007) KAAS: An automatic genome annotation and pathway reconstruction server. Nucleic Acids Res, 35. https://doi.org/10.1093/nar/gkm321.
11. Kanehisa M, Furumichi M, Tanabe M, et al (2017) KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res, 45: D353–D361. https://doi.org/10.1093/nar/gkw1092.
12. Garg A, Gupta D (2008) VirulentPred: A SVM based prediction method for virulent proteins in bacterial pathogens. BMC Bioinformatics, 9: 62. https://doi.org/10.1186/1471-2105-9-62.