Journal of Microbiological Methods 101 (2014) 1–8

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Comparative genomics study for the identification of drug and vaccine targets in *Staphylococcus aureus*: MurA ligase enzyme as a proposed candidate

Soma Ghosh [a], Jyoti Prava [b], Himanshu Bhusan Samal [b], Mrutyunjay Suar [a], Rajani Kanta Mahapatra [a,*]

[a] School of Biotechnology, Campus-11, KIIT University, Odisha 751024, India
[b] Bioinformatics Department, BJB Autonomous College, Bhubaneswar, Odisha 751001, India

ARTICLE INFO

Article history:
Received 16 December 2013
Received in revised form 10 March 2014
Accepted 19 March 2014
Available online 29 March 2014

Keywords:
Staphylococcus aureus
Computational comparative genomics
KEGG
Drug targets
MurA enzyme

ABSTRACT

Now-a-days increasing emergence of antibiotic-resistant pathogenic microorganisms is one of the biggest challenges for management of disease. In the present study comparative genomics, metabolic pathways analysis and additional parameters were defined for the identification of 94 non-homologous essential proteins in *Staphylococcus aureus* genome. Further study prioritized 19 proteins as vaccine candidates where as druggability study reports 34 proteins suitable as drug targets. Enzymes from peptidoglycan biosynthesis, folate biosynthesis were identified as candidates for drug development. Furthermore, bacterial secretory proteins and few hypothetical proteins identified in our analysis fulfill the criteria of vaccine candidates. As a case study, we built a homology model of one of the potential drug target, MurA ligase, using MODELLER (9v12) software. The model has been further selected for in silico docking study with inhibitors from the DrugBank database. Results from this study could facilitate selection of proteins for entry into drug design and vaccine production pipelines.

1. Introduction

The emergence of antibiotic resistance is a growing global threat to the effective treatment of infectious diseases. *Staphylococcus aureus* is one such gram positive bacterial pathogen which has been resistant to many antibiotics present in the market. It is responsible for a myriad of diseases, from skin and soft-tissue infections, to more invasive diseases including necrotizing pneumonia and sepsis (Diekema et al., 2001). Methicillin resistant *S. aureus* (MRSA) is resistant to all members of the β-lactam class of antibiotics including penicillin, cephalosporin and carbapenem, thereby disarming all previous mainstay treatments against *S. aureus*. In addition, MRSA is frequently resistant to other common antimicrobial agents. This strongly indicates that there is a continuous need to search for additional drug targets in bacterial genomes that would offer better protection and less long-term resistance (Perumal et al., 2007). The arrival of the post-genomic era has brought with it the possibility of genome-wide application of a rational new drug target and vaccine candidate selection methodology. This study is based on comparative and reductive genomics using computational approaches with integrated data from genomics, proteomics, and metabolomics (Butt et al., 2012a).

The availability of complete genome sequence of two methicillin resistant *S. aureus* strains N315 and NCTC 8325 represent an excellent opportunity to computationally predict non-homologous essential proteins and associated metabolic pathways; thus accelerating drug discovery steps. Here, we report the first computational comparative genomic analysis of the metabolic pathway in two different strains of *S. aureus* with differential numbers of essential genes for the identification of novel drug and vaccine targets. It is expected that the identified targets will expand our understanding of the molecular mechanisms of *S. aureus* pathogenesis, and facilitate the production of novel therapeutic agents.

2. Materials and methods

2.1. Identification of metabolic pathways of host and pathogens

A systematic workflow was defined that involved several bioinformatics tools, databases and drug target prioritization parameters (Fig. 1), with the goal of obtaining information about the drug and vaccine targets in *S. aureus* genome but absent in its host, therefore avoiding any potential side effects. We referred to the previous work on a comparative/subtractive genomics workflow by Butt et al. (2012a,b). The identification of metabolic pathways of host and pathogen was done by using KEGG database (Kyoto Encyclopedia of Genes and Genomes) as a source of metabolic pathway information

* Corresponding author. Tel.: +91 9337627961.
E-mail address: rmahapatra@kiitbiotech.ac.in (R.K. Mahapatra).

http://dx.doi.org/10.1016/j.mimet.2014.03.009
0167-7012/

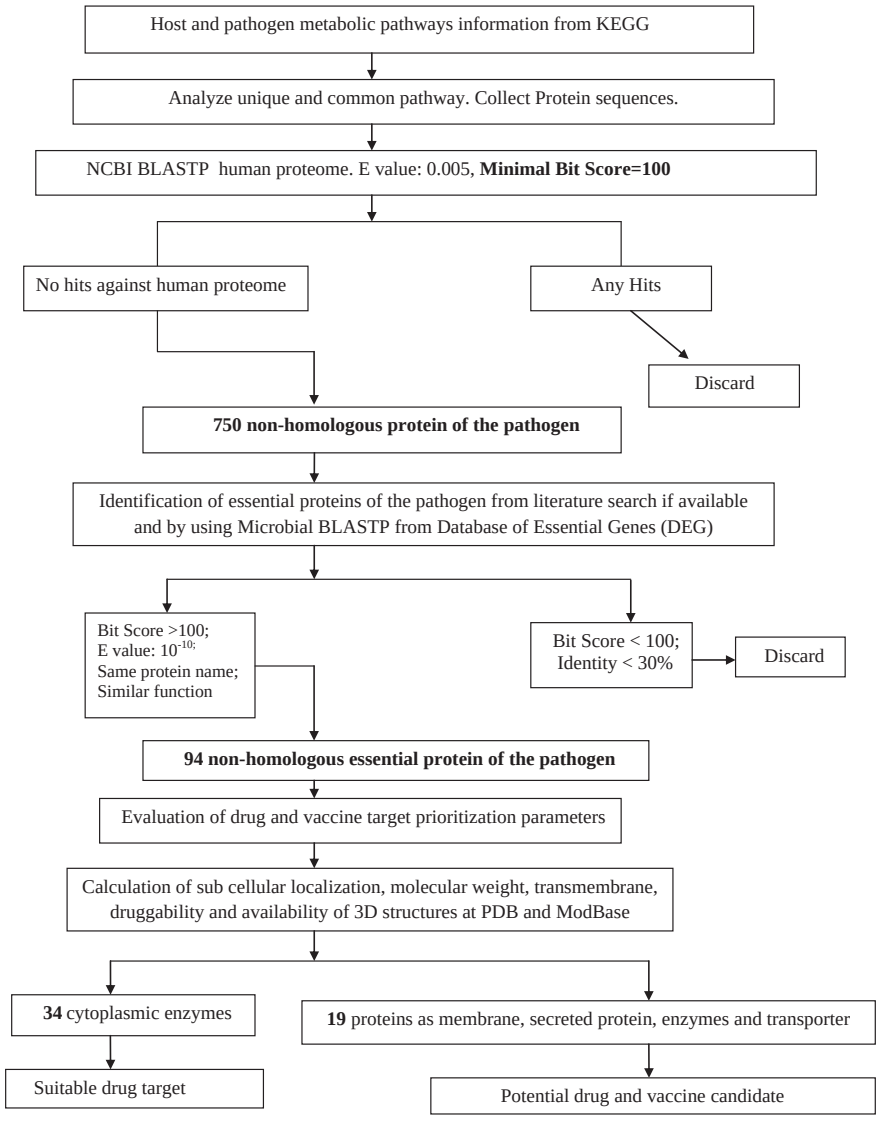

Fig. 1. Schematic representation of steps involved in computational comparative genomics-based target identification in *Staphylococcus aureus*.

(Kanehisa et al., 2006). A preliminary comparison was conducted manually between the two pathogenic strains of *S. aureus* to find out common and unique pathways in between them.

2.2. Identification of non-homologous and essential proteins

The corresponding protein sequences of common and unique pathways were obtained from UniProt database (www.uniprot.org). They were subjected to BLASTP (Altschul et al., 1997) analysis against the non-redundant database with the e-value inclusion threshold set to 0.005 against human proteome. Proteins which did not have hit below the e-value were considered as non-homologous protein. The criterion for selection e-value was based on the previous study by Anishetty et al. (2005), Butt et al. (2012a,b), and Damte et al. (2013).

After BLASTP analysis the non-homologous genes were further analyzed for essentiality to pathogen by the DEG database (http://tubic.tju.edu.cn/deg/) (Zhang et al., 2004). The DEG search was performed to identify the critical genes necessary for the survival of the pathogen *S. aureus*. E-value was set to 10^{-10}; a minimum bit score of 100 and identity percentage more than 30% was kept to screen out the non-homologous essential proteins of *S. aureus* by using the DEG BLASTP. For setting the e-value we have referred to the previous work by Butt et al. (2012a,b).

2.3. Drug target prioritization

There are certain criteria that help in determining suitable drug targets which were evaluated for each of the potential drug targets (Aguero et al., 2008). This involved calculation of molecular weight (MW) using computational tools and drug targets associated literature available at Swiss-Prot database. Trans-membrane predictions were made by TMHMM server (Krogh et al., 2001), and we searched for the presence of solved 3D structures through the Protein Data Bank (PDB) (Bernstein et al., 1977) and Mod-Base (Pieper et al., 2011). Additionally, the membrane proteins were analyzed for MAPPP (MHC-I Antigenic Peptide Processing Prediction) (http://www.mpiib-berlin.mpg.de/MAPPP/) with the SYFPEITHI matrix which combines existing prediction tools for proteasomal processing and MHC class I anchoring.

Druggability is another important target prioritization criterion. The default parameters for BLASTP were used to line up the potential drug targets from *S. aureus* against the list of protein targets of compounds found within the DrugBank. The selection criteria on BLASTP results is based on previous study (Holman et al., 2009), that is alignments with e-values less significant than 1×10^{-25} were removed.