Novel vaccine candidates identified by comprehensive analysis of global collection of *Streptococcus agalactiae* genomes

Tiraput Poonpanichakul[1,2], Adrian Cazares[3], Pisut Pongchaikul[2], Aras Kadioglu[1], Marie Yang[1]

[1]Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom
[2]Chakri Naruebodindra Medical Institute, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Samut Prakan, Thailand
[3]EMBL's European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, United Kingdom

UNIVERSITY OF LIVERPOOL

Background

Group B Streptococcus (GBS: *Streptococcus agalactiae*) is a major cause of life-threatening neonatal diseases globally. There is currently no licensed vaccine for GBS and current treatment and prevention rely on the use of antibiotics. However, the rate of antimicrobial resistance is alarmingly rising hence prophylactic approaches to GBS are highly warranted. Recent advances in whole-genome sequencing (WGS) and its continually decreasing costs have promoted the production of large genome datasets worldwide which can be exploited towards the development of vaccines. In the present study, we identified putative vaccine candidates by applying an analytical pipeline that integrates epidemiological, clinical and genomic data.

Methods

To identify the novel universal vaccine candidates, we performed the comprehensive analysis of publicly available GBS genomes. An unprecedented large collection of 15,693 genomes were assembled and annotated. Antimicrobial resistance genes, sequence typing, and clonal complex grouping were assessed. After quality control, 11,841 samples underwent core genome analysis. Core genes were ranked based on in silico prediction of cellular localisation, transmembrane topology, adhesins, physicochemical property, antigenicity, similarity to host protein, and functional group assessment.

Conclusion

We performed a large-scale genomic analysis to generate an atlas of vaccine candidates that cover the heterogeneous population of GBS. Novel candidates for GBS vaccine were predicted and characterised computationally. Further in vitro and in vivo validation are now required to confirm their immunogenicity and protective efficacy.

Results

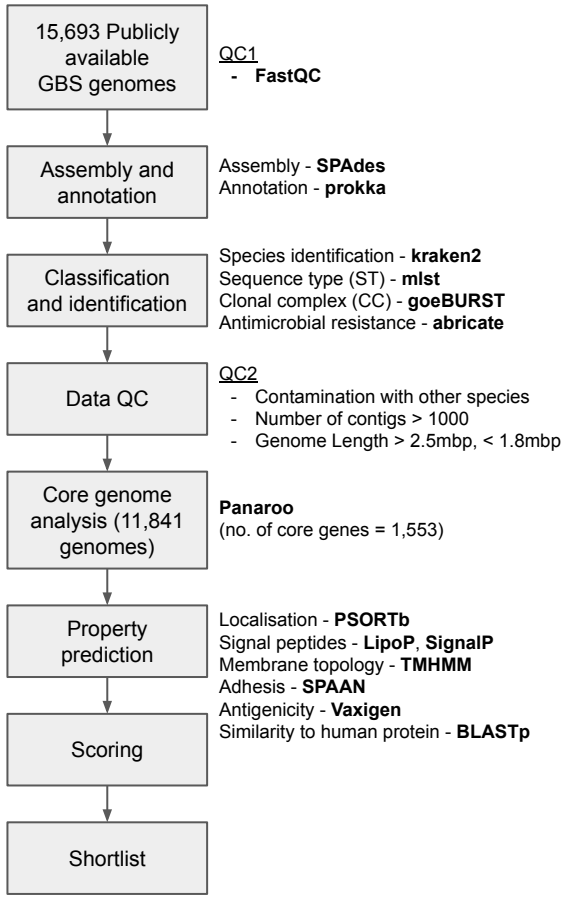

Figure 1: Diagram showing workflow of the analysis. Software names are highlighted in bold text.

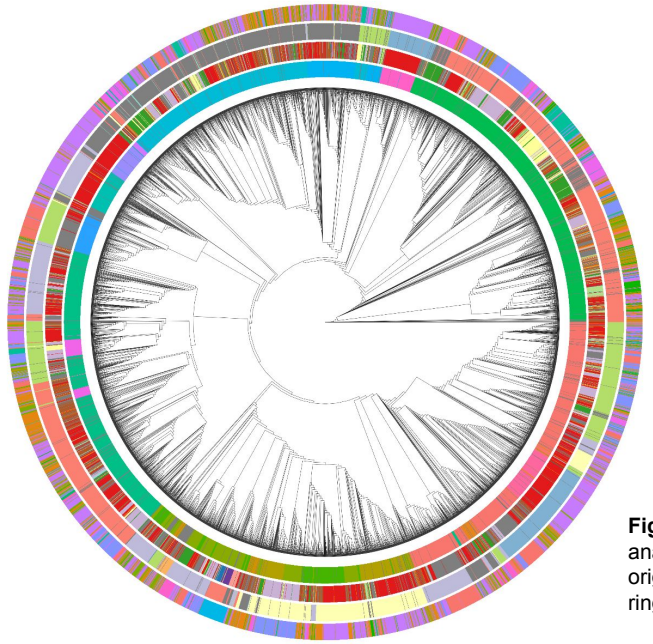

Figure 2: Overview characteristic of GBS genomes in the analysis. Colour bars highlight clonal complex (CC), country of origin, capsular serotype, and bioproject number (from inner ring to outer ring respectively).

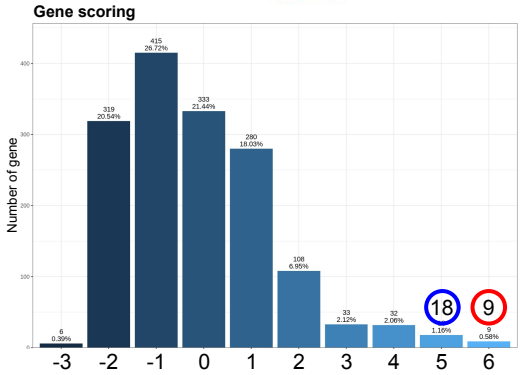

Figure 3: Bar plot shows cumulative score distribution of 1,553 core genes. Core genes was scored using 7 software listed in property prediction step in Figure 1 to shortlist potential candidate proteins. Score was given +1 for positive feature, -1 for negative feature, and 0 for ambiguous result.

Functional Group	Total
Amino acid transport and metabolism	3
Cell wall/membrane/envelope biogenesis	6
Function unknown	10
Inorganic ion transport and metabolism	1
Intracellular trafficking, secretion, and vesicular transport	1
Nucleotide transport and metabolism	3
Post-translational modification, protein turnover, chaperones	3

Figure 4: Functional group of candidates those have the score 5-6 in the analysis were identified through Conserved Domains Database.