# A Novel Algorithm of Core Genome Phylogeny for Outbreak Investigation of Multidrug Resistant Organisms in Hospital and Community Settings

**Jake Siu-Lun Leung**
Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, HKSAR, People's Republic of China

**Gilman Kit-Hang Siu**
Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, HKSAR, People's Republic of China

## Abstract

In the past few years, multi-locus sequence typing (MLST) is still widely applied in bacterial outbreak investigation in Hong Kong, but it only provided sketchy information due to its poor resolution power. Emerging of WGS technique helps to resolve such issue, but also comes with some new challenges, demand of specialist who is capable to manage, analyse and interpret the WGS data. Hence, my study is not only to propose a new algorithm to conduct core genome SNP (cgSNP) phylogenetic analysis, but also to solve the specialist shortage issue through scripting the pipeline to become a friendly-use tool. The algorithm is firstly to obtain a core gene set as a reference. Afterwards, samples will be extracted core genes by following the defined gene set, followed by phylogenetic analysis. This algorithm was successfully validated by revealing similar results through repeating analysis with same set of *K.pneumoniae* genomic data related to outbreak investigation. Secondly, it was applied to analyse genetic relatedness of vancomycin-resistant *E. faecium* (VRE). The results demonstrated, in the recent years, a ST80 VRE strain keeps causing outbreak in different hospital. The proposed algorithm could produce a highly similar results compared to the other methods, and able to be applied on outbreak investigation.

## Keywords

Core genome phylogeny, outbreak investigation, vancomycin-resistant, enterococcus faecium, VRE.