DOI 10.18699/SBB-2022-493

Computational pipeline and database for the epidemiological typing and antimicrobial resistance profiling of important bacterial pathogens

Shelenkov A., Mikhaylova Y., Akimkin V.
Central Research Institute of Epidemiology of the Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing, Moscow, Russia
** fallandar@gmail.com*

Key words: epidemiological surveillance, isolate typing database, whole genome sequencing, antimicrobial resistance, genomic markers, ESKAPE pathogens

Motivation and Aim: Monitoring the spread of particular clones of pathogenic bacteria and associated antimicrobial resistance determinants within a particular healthcare facility, country or across the world represents an important task for national and international public health institutions. During recent years, the antibiotic resistance within different species of nosocomial and community-acquired bacterial pathogens has increased to dangerous levels [1]. This problem cannot be solved without comprehensive investigations of resistance transmission mechanisms and global epidemiological surveillance. Currently, such a surveillance is often dependent on whole genome sequencing (NGS) of bacterial genomes and corresponding bioinformatics analysis pipelines [2]. Although the number of available genomic sequences for bacterial isolates increases very rapidly (more than 100,000 bacterial genomes are already available in NCBI Genbank, https://www.ncbi.nlm.nih.gov/genome), public genomic databases usually lack the essential epidemiological information for the isolates, for example, the results of multilocus sequence typing (MLST) or typing using other schemes, as well as the antimicrobial gene content. Thus, it is very difficult to use the information available in such databases for epidemiological surveillance of resistance spreading. The pipeline developed and the corresponding database under development aim to assist the users to retrieve such information from the genomes of important pathogenic bacteria.
Methods and Algorithms: Earlier we have developed a computational pipeline for seamless integration of various software tools for annotation of newly sequenced clinical bacterial isolates, including typing using various schemes [3]. In this work, we have added additional typing tools (cgMLST and hybrid profiles, SPA types, serovars etc) and performed an analysis of all genomes available in Genbank for important nosocomial pathogens belonging to ESKAPE group (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii, Pseudomonas aeruginosa*, and *Enterobacter* spp). We used Resfinder 4.0 database for antimicrobial gene detection (https://cge.cbs.dtu.dk/services/ResFinder/) and MentaList (https://github.com/WGS-TB/MentaLiST, version 0.2.4) [4] to obtain the cgMLST profiles. The database of MLST-based sequence types (ST) for the isolates available in Genbank is currently available in local network of our institution and contains the information regarding the STs obtained and the assignment of them to known international or global clones of high risk.

Results: We have applied the computational tools developed to investigate the epidemiological characteristics of important nosocomial pathogens. Here we present the results for one of them, *Acinetobacter baumannii,* for which more than 12,000 genomes are available in Genbank. Currently, there are 10 known international clones of high risk for this important pathogen, and they are no longer specific to particular regions of the world and may co-exist even within a single hospital [5]. Most part of the isolates (about 70 %) was found to belong to well-established international clone (IC) 2, while IC1 genomes constituted only about 10 %. These data represents the skew of IC distribution towards well-known clones, while the epidemiological data report different clonal variability [1]. However, the researchers are usually interested in making comparison between the isolates obtained in particular location and the ones available in public databases to infer the possible ways of clone spreading and resistance transmission via, e.g., phylogenetic investigations. For this purpose, the data for comparatively rare occurring STs could be useful, and our database provides such data. At the same time, even the studies of ubiquitous clones like IC1 and IC2 can provide important epidemiological information since the antibiotic gene content cannot be always associated with particular clones, and resistance gene distribution could vary within particular clonal group.

Conclusion: We have developed a user-friendly computational pipeline for typing and annotation of pathogenic bacteria genomes, including antibiotic resistance prediction, and applied this pipeline and other tools available to collect the epidemiologically important data from public genome databases. The obtained data will be available online soon and will help in performing various genomic epidemiology investigations and prepare scientifically valid manuscripts.

Acknowledgements: The study was supported by the industry research program of Rospotrebnadzor (2021–2025) “Scientific support of epidemiological monitoring and public health measures of the Russian Federation territory” and Russian Federation Government Resolution No. 3116-r.

References

1. Kuzmenkov A. et al. AMRmap: An interactive web platform for analysis of antimicrobial resistance surveillance data in Russia. *Front Microbiol*. 2021;12:620002.
2. Maljkovic B. et al. Next generation sequencing and bioinformatics methodologies for infectious disease research and public health: approaches, applications, and considerations for development of laboratory capacity. *J Infect Dis.* 2020;221:S292-S307.
3. Shelenkov A. et al. Characterization of antimicrobial resistance, virulence profiling and analysis of whole-genome sequence of clinical *Klebsiella pneumoniae* isolates. *Antibiotics (Basel).* 2020;9(5):261.
4. Feijao P. et al. MentaLiST – A fast MLST caller for large MLST schemes. *Microbial Genomics*. 2018;4(2):146.
5. Shelenkov A. et al. Diversity of international high-risk clones of *Acinetobacter baumannii* revealed in a Russian multidisciplinary medical center during 2017–2019. *Antibiotics (Basel).* 2021;10(8):1009.