EMBnet.journal
Bioinformatics in Action

Donato et al. (2021) EMBnet.journal 26(Suppl A), e960
http://dx.doi.org/10.14806/ej.26.A.960

Precision Epidemiology of Multi-drug resistant bacteria: bioinformatics tools

John Donato[1]✉, Luis Lugo[1], Hermes Perez[1], Harold Ballen[1], Diego Talero[1], Sebastian Prada[1], François Brion[2], Veronica Rincon[1], Laurent Falquet[3], Maria Teresa Reguero[1], Emiliano Barreto-Hernandez[1]

[1] Bioinformatics Center, Biotechnology Institute, Universidad Nacional de Colombia, Bogota, Colombia
[2] Haute Ecole en Hainaut – Department of Sciences and technologies, Mons, Belgium
[3] Biochemistry/Bioinformatics Unit, Université de Fribourg, and Swiss Institute of Bioinformatics, Fribourg, Switzerland
Competing interests: JD none; LL none; HP none; HB none; DT none; SP none; FB none; VR none; LF none; MTR none; EBH none

Precision epidemiology is a tool that allows researchers and the health community to understand, following, and controlling the infections, overall when the microorganisms are antibiotic-resistant. Next-Generation sequencing (NGS) techniques have become an important tool for the precise identification and genomic characterisation of these microorganisms, facilitating their accurate epidemiological monitoring. The World Health Organization (WHO) has prioritised the development of tools such as the Global Antimicrobial Resistance Surveillance System (GLASS) (WHO, 2015) to collect clinical, epidemiological, and laboratory data as a global system for antimicrobial resistance surveillance and to serve as a repository for accurate epidemiological monitoring.

Considering the WHO guidelines, we are developing the Genomic Information Management System (SGIG) (Donato, 2018) to integrate clinical, epidemiological, laboratory, and genomic data obtained by NGS. It consists of several modules: 1) the module enabling the entry of patient, microbiological, and molecular data; 2) the module for processing of NGS data, from its quality assessment to obtaining the annotation and comparison of the assembled genomes, using tools such as FastQC, Spades, Prokka, RGI CARDdb, and Roary; 3) the module for the identification and typification of bacteria, that uses a bidirectional recurrent neural network architecture (Lugo, 2018), along with the inclusion of standard methods such as rMLST (Jolley *et al.*, 2012); 4) the module that uses a script in Python for the prediction of the antibiotic resistance profile, which searches for resistance genes in each genome stored in the local database and relates them with the antibiotic resistance data associated to each gene, through the presence-absence rules of genomic determinants obtained and cured of the information reported in the literature (Perez, 2017). And module 5, which creates reports where it is possible to have access to the clinical and demographic statistics, isolate phenotypic resistance profiles, isolate resistance profile predictions based on the genomics data and isolate comparative genomic results.

The system is built using the Groovy and Grails v3.1.9 frameworks. The system backend consumes services built with Biopython v1.70 and BioPerl v5.8. The database is implemented in MySQL v5.6.30. The system is installed on a DELL server with the operating system openSUSE Leap 42.1.4. Funded by Univesridad Nacional de Colombia and Colciencias (project Código:66234).

References

1. Donato JJ (2018) Sistema de gestión de información de genomas completos de cepas de Acinetobacter baumannii para la identificación, tipificación y seguimiento de resistencia a antibióticos. Master thesis, Universidad Nacional de Colombia - Sede Bogotá. https://repositorio.unal.edu.co/handle/unal/62957 (accessed 02 February 2021)
2. Jolley KA, Bliss CM, Bennett JS, Bratcher HB, Brehony C *et al.* (2012) Ribosomal multilocus sequence typing universal characterisation of bacteria from domain to strain. Microbiology **158**(Pt 4):1005-1015. https://doi.org/10.1099/mic.0.055459-0
3. Lugo LE (2018) A Recurrent Neural Network approach for whole-genome bacteria classification. Master thesis, Universidad Nacional de Colombia - Sede Bogotá. https://repositorio.unal.edu.co/handle/unal/68663 (accessed 05 February 2021)
4. Perez H. (2017) Modelo bioinformático para predecir la resistencia a antibióticos a partir del genoma de una Bacteria. Master thesis, Universidad Nacional de Colombia - Sede Bogotá. https://repositorio.unal.edu.co/handle/unal/62924 (accessed 05 February 2021)
5. WHO - World Health Organization (2015) Global Antimicrobial Resistance Surveillance System (GLASS). Retrieved from: https://www.who.int/glass/en/ (accessed 02 February 2021)