FindR-TB: A cloud–based tool for antibiotic resistance prediction in *Mycobacterium tuberculosis*

Author Names Paul Walsh[a], Micheál Mac Aogáin[b], Brendan Lawler[a]
[a]Department of Computer Science
CIT – Cork Institute of Technology, Ireland
e-mail:
[b]Department of Clinical Microbiology, School of Medicine,
TCD – Trinity College Dublin.
e-mail: m.macaogain@tcd.ie

Keywords: Cloud computing, Microbiology, Antibiotic resistance, Tuberculosis

The emergence of affordable DNA sequencing technology holds promise for infectious disease diagnostics. Access to whole-genome sequence (WGS) data of pathogenic micro-organisms is now within reaching distance of hospital diagnostic laboratories and promises to bring about positive disruptive changes in diagnostic microbiology. However, without 'point and click' applications addressing computational bottlenecks in the analysis of complex bacterial DNA sequences, clinical adoption of genome-based molecular diagnostics will not be realised.

Mycobacterium tuberculosis (TB) is a global infectious disease that affects approximately 9 million people every year leading to over a 1 million deaths. An estimated 450,000 TB infections exhibit drug resistance. The basis of drug resistance is underpinned by genetic mutations in the TB genome and is therefore predictable given access genomic sequence data. The application of WGS analysis to TB has gained traction as an aid in the diagnosis of TB promising gains in antibiotic resistance detection times and more accurate disease transmission mapping of infected individuals.

Here we report a new development in our cloud-based bacterial DNA management software; the FindR-TB resistance predictor. This tool is integrated into our cloud-based microbial DNA analysis platform (SimplicityTM) and functions in the prediction of antibiotic resistance in *M. tuberculosis* from in-putted user sequence data. In an initial characterisation of our tool we have tested it on genomic data from 100 *M. tuberculosis* isolate genomes allowing us to document the sensitivity and specificity of the tool by using WGS data from phenotyped strains with characterised antibiotic sensitivities.

Diagnostic test statistics were generated for 100 strains allowing us to gauge the potential of this tool as a diagnostic aid for clinical application; i.e. prediction of antibiotic resistance profiles of TB strains and subsequent guidance of therapeutic choices. Resistance to first line antibiotic agents for the treatment of TB including isoniazid, rifampicin, ethambutol and streptomycin was detected with a sensitivity of 91%, 98%, 96% and 100% respectively, and corresponding specificity of 97%, 90%, 73% and 100%. The first line agent pyrazinamide had a significantly lower sensitivity of 34% with a specificity of 100%. These results highlight the ability of our *in silico* predictions to be comparible with currently available diagnostic tests for antibiotic resistant TB while highlighting areas of discord where further characterisation of resistance mechanisms is required.

We plan to test our FindR-TB on a broader collation of TB isolates in order to hone the diagnostic test statistics and improve results. Future tools are currently in development for other pathogenic bacteria including *Clostridium difficile* and *Staphylococcus aureus.* Ultimately these tools will become cornerstones of the clinical microbiology lab as WGS technologies become increasingly integrated into the laboratory work-flows of modern medical diagnostics.