Department of Medical Microbiology and Infectious Diseases. Erasmus MC

Contact details: Dr. Nikolaos Strepis   e-mail: n.strepis@erasmusmc.nl
Dr. Maria Suarez Diez maria.suarezdiez@wur.nl

Deployment of NGS for diagnosis of bacterial infections.

With the upraise of NGS methods, thousands of bacterial genomes are constantly being sequenced. Although the method evolves, still specific criteria remain undefined. Usually for the sequencing of a bacterial genome, an arbitrary coverage of 50x is strived at. However, the optimal/minimal coverage needed to obtain accurate and reliable results has not been established.

In the department of Medical Microbiology and Infectious Diseases, we have obtained an Illumina sequencer and we are currently sequencing bacterial genomes. In this laboratory basic research as well diagnostics are being performed using NGS. Especially for diagnostics the accuracy and precision of results is crucial. A connection of coverage and quality of sequencing results should be addressed and studied. This may differ per microbial species, we should identify the optimal sequencing coverage per species. As the demand from diagnostics and the multiple research questions there is a high demand of sequencing. Lowering the coverage for acquiring excellent results may lead to additional bacteria sequenced per run.

Experimental Design
Random data(sub)sets will have to be created with varying levels of coverage. Reads will be generated for achieving low and high coverage in a random distribution. A manual tool may be needed to better assess this issue.

❖ Tweaking of in silico sequence reads will lead to a different quality sequenced genome and assemblies
❖ Genomes will be validated with frequently used criteria such as N50, number of contigs and mismatches, MLST typing as well as SNP differences
❖ A coverage-to-quality correlation will be established and expanded to each species
❖ The defined minimal/optimal coverage will be the criteria for correct performance of NGS in diagnostics and routine in the Medical Microbiology of EMC.

Techniques

❖ Use of in silico software for reads generation
❖ Use of *possible* scripting for designing a tool for random generation and manipulation of sequenced reads
❖ Use of QC tools for assembled genomes
❖ Use of MLST / SNPs analysis with related tools