BioNumerics®

RELEASE NOTE

MYCOBACTERIUM TUBERCULOSIS SCHEMA

for whole genome sequence typing

We are proud to present a schema for true whole genome multi-locus sequence typing (wgMLST) of M. tuberculosis in BioNumerics. When used in combination with our cloud-based Calculation Engine, typing M. tuberculosis isolates up to strain level using whole genome sequencing is now easily accessible to everyone.

What is the schema exactly?

Based on the core genome MLST definition published by *Kohl et al.*[1], a pan-genomic schema has been defined together with international co-workers. A set of 46 publicly available reference sequences were included, among which *M. bovis, M. africanum* and *M. canettii* strains. To echo the known diversity of *M. tuberculosis*, prototuberculosis strains were included as well. By also capturing the accessory loci, we increased the discriminatory power of the schema, permitting the detection of subtype- or outbreak-specific markers, for powerful classification and outbreak definition tools.

Which loci are present?

Starting from the 46 annotated reference genomes, our in-house developed schema creation procedure uses a sampling-based multi-reciprocal BLAST procedure to determine those sets of alleles that make up the stable loci in the accessory genome. A per-locus allele assessment procedure then determines the central prototype allele, and thus the definition of the locus. The accessory schema, including 1141 loci, is then complemented with the 2891 core loci to obtain maximal consistency with classical and novel multi-locus sequence typing initiatives for *M. tuberculosis*.

How will it help you?

By using BioNumerics and the integrated powerful calculation infrastructure, analyzing whole genome sequencing data for *M. tuberculosis* has become a lot more straightforward. Our cloud-based Calculation Engine offers a high-throughput environment for all your sample processing needs. Its quality-controlled de novo assembly possibilities allow you to easily assemble whole genome sequencing data without the need of local computing power. The two allele detection procedures (assembly-based and assembly-free) allow you to perform fast and reliable allele calling for e.g. cluster detection which can be combined with whole genome SNP analysis to obtain the utmost resolution within your sample comparisons.

The BioNumerics wgMLST schema for *M. tuberculosis* has been tested, validated and approved by our microbiologists.

Great care has been taken to create an analysis procedure that minimizes sample artifacts, while maintaining a huge discriminatory power that supersedes the core genome schema.

With turnaround times of less than 30 minutes per sample, the ability to include both wgMLST and wgSNP analyses and the capability to process many samples simultaneously, the power of high-performance computing will be brought to your desktop with a few clicks.

Interested?

Click on this link to request a calculation engine project or scan the QR code:

References:

(1) Kohl, TA, Diel, R, Harmsen, D, Rothgänger, J, Walter, KM, Merker, M, Weniger, T, and Niemann, S. Whole-genome-based Mycobacterium tuberculosis surveillance: a standardized, portable, and expandable approach. J. Clin. Microbiol. 2014, 52: 2479-86: 2479-86

APPLIED MATHS

A BIOMÉRIEUX COMPANY