BioNumerics®

RELEASE NOTE

CLOSTRIDIUM DIFFICILE

schema for whole genome typing

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

What is the schema exactly?
A core and pangenomic schema has been defined in collaboration with international coworkers. This resulted in 259 selected reference sequences that reflect the known diversity of *C. difficile*. By also capturing the accessory loci, this increased the discriminatory power of the schema. At the same time, the extended schema also allows for the detection of subtype- or outbreak-specific markers, thus enabling more powerful classification and outbreak definition tools.

Which loci are present?
Starting from the 259 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 6713 loci, is then complemented with the 1999 core loci, 7 loci from the MLST schema of the Pasteur institute[1], 7 MLST loci[2], 6 CWP cluster loci and 13 "other" loci, associated with antibiotic resistance and virulence, as featured on pubMLST.org[3] to obtain maximal consistency with classical and novel multi-locus sequence typing initiatives for *C. difficile*.

How will it help you?
By using BioNumerics and the integrated powerful calculation infrastructure, analyzing whole genome sequencing data for *C. difficile* 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 *C. difficile* 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 an enormous discriminatory power that supersedes the core genome schema.

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

Interested?

Simply request a calculation engine project to get started:

References:
(1) Lemee, L., Dhalluin, A., Pestel-Caron, M., Lemeland, J. F., & Pons, J. L. (2004). Multilocus sequence typing analysis of human and animal Clostridium difficile isolates of various toxigenic types. Journal of clinical microbiology, 42(6), 2609-2617.
(2) Griffiths, D., Fawley, W., Kachrimanidou, M., Bowden, R., Crook, D. W., Fung, R., ... & Kirton, R. (2010). Multilocus sequence typing of Clostridium difficile. Journal of clinical microbiology, 48(3), 770-778.
(3) Jolley, K. A., & Maiden, M. C. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC bioinformatics, 11(1), 595.

APPLIED MATHS
A BIOMÉRIEUX COMPANY