

## MULTI-LOCUS SEQUENCE TYPING (MLST)

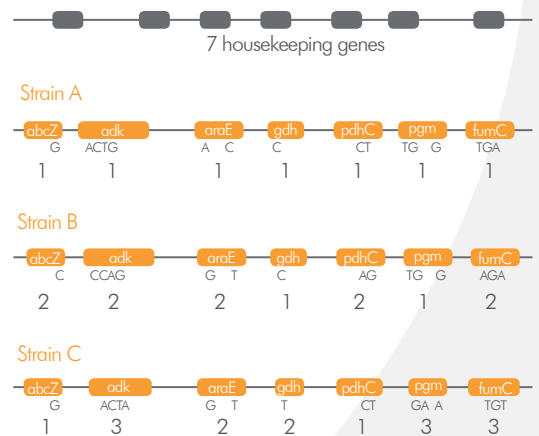
### What is Multi-Locus Sequence Typing?

Often referred to as MLST, Multi-Locus Sequence Typing is a technique where a number of well chosen housekeeping genes (loci) are sequenced, usually in part.

In a typical approach, recombination is expected to occur with a much higher frequency than point mutations. Therefore, one does not look at the total sequence similarity between strains. Instead, each sequence for a given locus is screened for identity with already known sequences for that locus. If the sequence is

different, it is considered to be a new allele and is assigned a unique (arbitrary) allele number. In case seven housekeeping genes are studied, each strain is thus characterized by a profile of seven allele numbers, with each unique profile having its own designated sequence type.

The allelic profiles can be considered as a character set of 7 categorical characters. MLST has been used successfully to study population genetics and reconstruct micro-evolution of epidemic bacteria and other micro-organisms.

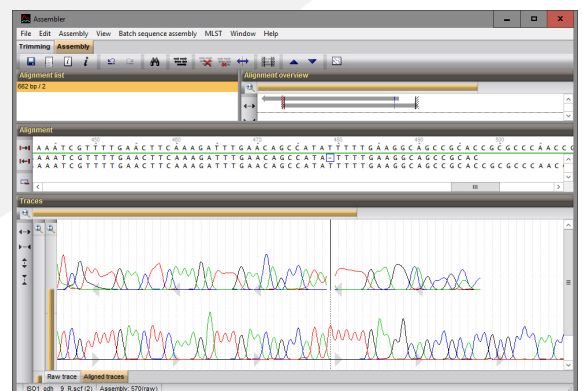
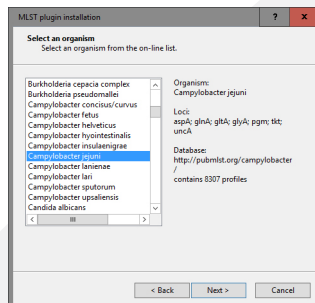


### MLST in BioNumerics

Applied Maths has contributed to the analysis of MLST data through the use of Minimum Spanning Trees. Through the availability of an MLST plugin, the BioNumerics software is widely used for the storage and analysis of MLST sequences. BioNumerics automatically analyses batches of sequence trace files, connects to online MLST databases, retrieves corresponding allele numbers, sequence types as well as available clonal complex information. BioNumerics can process hundreds of isolates in only seconds. Results are stored in the database and are available for statistical and population analysis, clustering, partitioning, and identification using BioNumerics' impressive set of analysis tools.

#### NOTE:

The MLST plugin is a free add-on. The minimal configuration consists of the "Sequence Data", "Character Data" and "Tree and Network Interference" modules. Please contact us for more information.



## YOUR ADVANTAGES



WHY USE **BioNumerics** FOR YOUR MLST ANALYSIS?

- ✓ Easy database setup
- ✓ Automated import/assembly of traces
- ✓ Fast allele and MLST type identification
- ✓ Impressive set of analysis tools

## Easy database setup



An MLST database can be easily set up based on any publicly available MLST scheme definition for e.g. *Escherichia coli*, *Salmonella enterica*, *Klebsiella pneumoniae*, *Campylobacter*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* or any other organism available from an online repository, simply by selecting the organism from a list. Alternatively, you can set up your own MLST schemes.

BioNumerics automatically connects to the selected repository to retrieve the names of the housekeeping genes and creates the database environment to store the information for all the genes. The software will download alleles, trimming positions and MLST types locally and update its database at startup or fetch data for each analysis from web services.

## Automated import/assembly of traces



Automated import and assembly of batches of sequencer trace files from various sources (Applied BioSystems, Beckman, Amersham, FASTA). File names are parsed into strain and gene information using a parsing definition.

Consensus sequences are automatically trimmed using start and stop signatures and placed in the right direction.

When the batch assembly is finished, an overview report is shown, listing processing status and occasional problems for each strain/gene combination.

Double-clicking on a particular issue opens the Assembler with the problem position selected. For each problem position, nearest existing alleles and base differences can be shown.

## Fast allele and MLST type identification

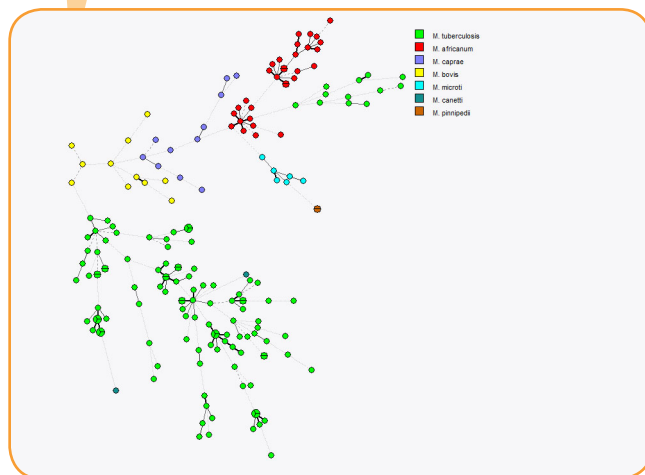
Alleles and MLST types can be identified by automatic on-line requests to MLST web services or by comparing to locally stored allele database, automatically updated at startup.

Allele and MLST type information for own strains is stored in the database and can be updated at any time for a selection of strains.

## Impressive set of analysis tools

Calculate population modelling networks in the finest and most comprehensive cluster analysis application available today, using standard or custom priority rules and with branch significance support indication.

Calculate and display partitioning for clonal complexes and use BioNumerics' rich set of statistics tools.



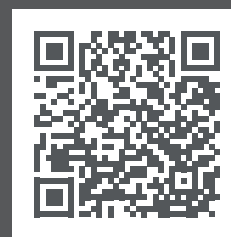
## CONVINCED? INTRIGUED? TRY IT FOR YOURSELF!

[www.bionumerics.com](http://www.bionumerics.com)

Scan the QR codes to access each step



1. Make sure you have a **BioNumerics** license (also see note on first page).



2. Watch tutorial movies or download sample data for use in **BioNumerics**.

### Headquarters

Keistraat 120 • 9830 Sint-Martens-Latem • Belgium  
+32 922 22 100 info@applied-maths.com

### USA and Canada

11940 Jollyville Rd., Suite 115N • Austin, TX 78750 USA  
+1 512 482 9700 info-us@applied-maths.com