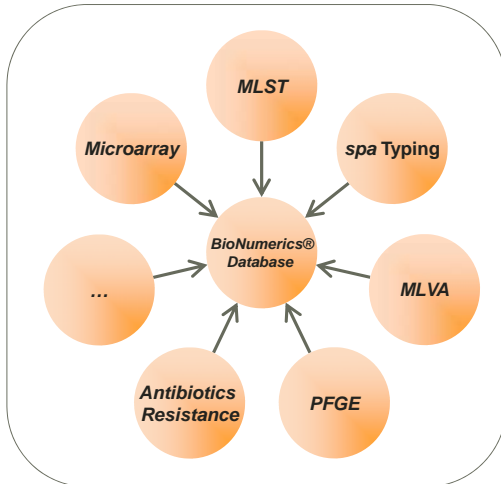


MRSA monitoring in Belgium: improved *spa* typing and analysis using the BioNumerics® software

K. Janssens, L. Vauterin, B. Pot, P. Vauterin
Applied Maths NV, Keistraat 120, B9830 Sint-Martens-Latem, Belgium

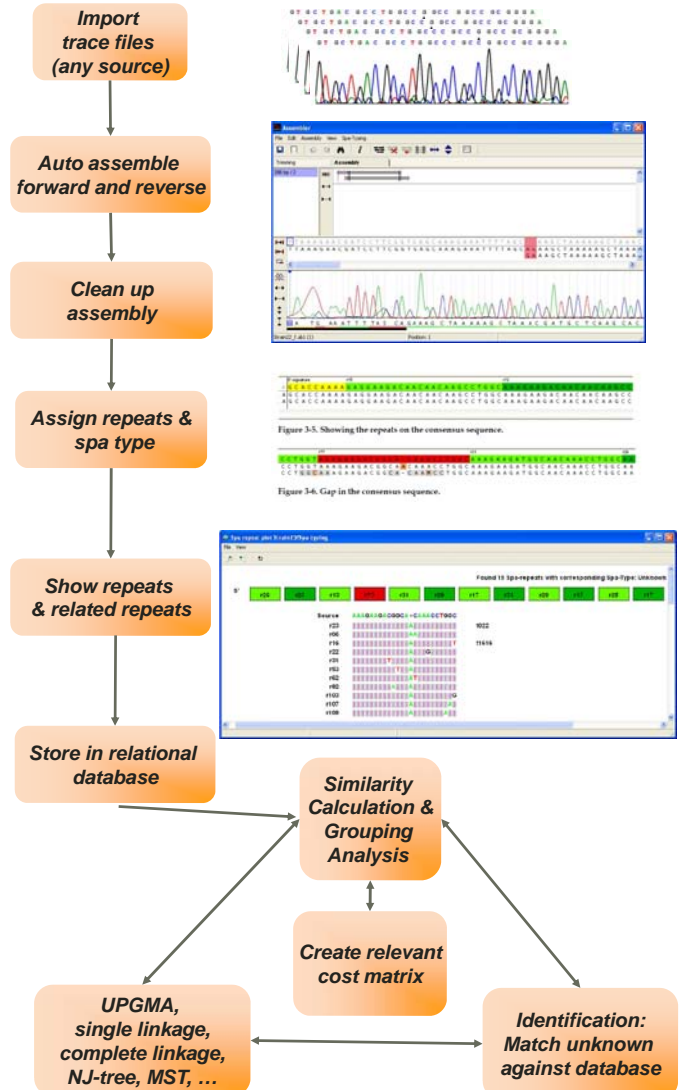
INTRODUCTION: Today, MRSA subtyping and identification of nosocomial infections and transmissions relies mostly on genotypic methods. Sequence-based techniques offer many advantages over traditional typing methods. The BioNumerics® software, widely used for storage and analysis of microbiological data, can combine information from different phenotypic and genotypic methods in a single database. We developed a BioNumerics® database with all known *spa* types (May 31, 2007; data from SeqNet: SpaServer) as a starting point for our analysis. Basic data mining was performed on all 2548 known types.

Overview of the different MRSA typing methods that can be managed and stored by BioNumerics®.



DATA PROCESSING:

Workflow for the use of the **Spa-Typing** functionality in BioNumerics®



DATA ANALYSIS:

Plotting tools

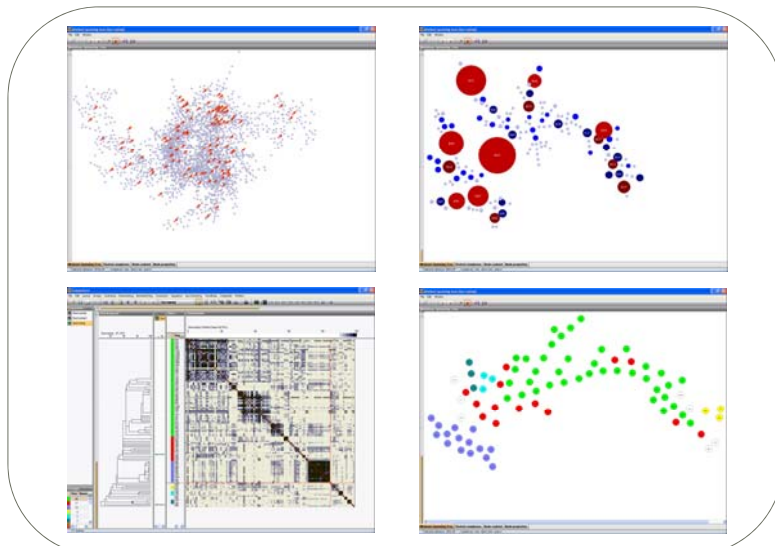
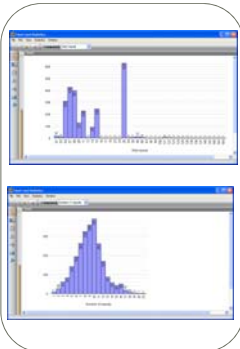
- Show histograms of the starting repeats
- Plot number of repeats present in all *spa* types

Clustering tools

- Similarity calculation after DSI alignment (Benson, 1997) using a cost matrix in which amino acid changes are given a weight factor of two, compared to nucleotide changes.
- Traditional clustering tools
- The Minimal Spanning Tree (MST) calculation, showing a theoretical relationship between available *spa* types

Mining tools

- Use the potential of BioNumerics® to combine multiple data types useful to compare and/or validate different methodological approaches.



EXAMPLE:

Top

Minimum Spanning Tree for all *spa* types found in Belgium (left: detailed view; right: cluster representation).

The 117 types found in Belgium were mapped on the MST of all theoretical types, illustrating the scattered diversity of the Belgian types (red flags). The tree inference settings used were: Gap creation: 250%; Gap extension: 50%; Duplication creation: 25%; Duplication extension: 25%.

Bottom left

Single Linkage dendrogram for the same data. Different groups are indicated by colors (right of the dendrogram). Further right is the similarity matrix calculated after DSI alignment.

Bottom right

Selected result of the Single Linkage analysis was mirrored on the MST.