The reproducibility of matrix-assisted laser desorption inonization-time of flight mass spectrometry (MALDI-TOF) based typing: employment of bioinformatics

M. Oberle[1], N. Wohlwend[2], D. Jonas[3], F. Maurer[4], G. Jost[5], K. Vranckx[6], A. Egli[7,8]

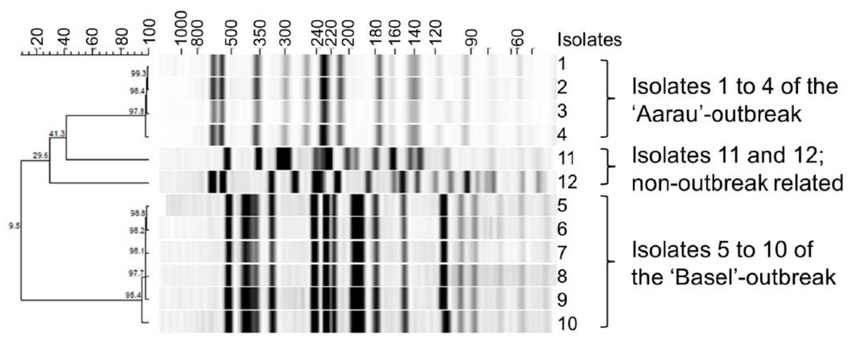
[1] Clinical Microbiology, Cantonal Hospital Aarau, Aarau; [2] Labormedizinisches Zentrum Dr Risch, Berne, [3] Clinical Microbiology, University of Freiburg, Freiburg, Germany; [4] Institute of Medical Microbiology, University of Zurich, Zurich, [5] dianalabs, Geneva; [6] Applied Maths NV, Belgium, 7 Clinical Microbiology, University Hospital Basel, Basel, [8] Applied Microbiology Research, University of Basel, Basel, Switzerland

Purpose

- MALDI-TOF holds promise for an easy access typing method compared to conventional approaches (Egli et al. Plos one 2014).
- However, the technical, biological and center reproducibility of such data has not yet been explored.
- The aim of this study was to compare typing data from multiple centers employing bioinformatics.

Material, Methods and Datasets

- Seven different centers tested 12 extended spectrum beta-lactamase (ESBL)-producing *Escherichia coli* isolates – including two outbreak and one non outbreak-related clusters (Fig. 1).
- Every center used a Microflex MALDI-TOF (Bruker, Bremen, Germany) and used the same standard operating procedure (Egli et al. Plos one 2014).
- Each center recorded four spectra from each isolate (technical repeat) and repeated three times the procedure (biological repeat).
- Raw data was used to calculate the technical and biological reproducibility in/between each center using the Bionumerics software (Applied Maths, Belgium).

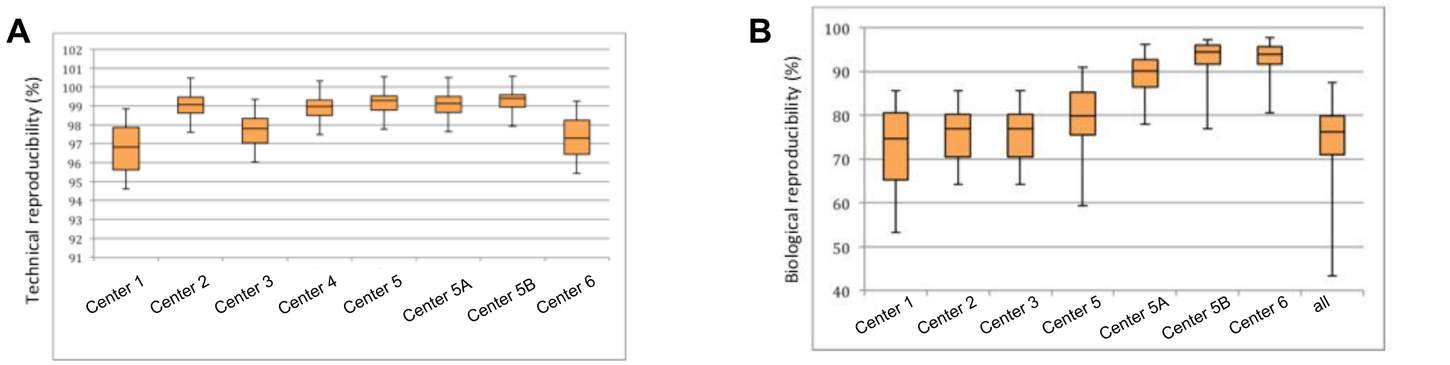


- **Fig 1.** Pulsed field gel electrophoresis (PFGE) of two outbreaks and non-related isolated used for MALDI-TOF based typing.

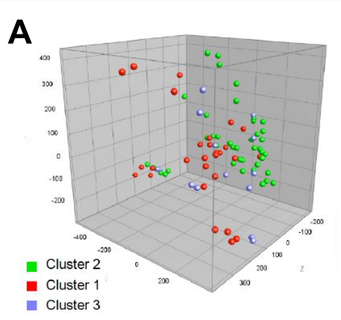
Corresponding author

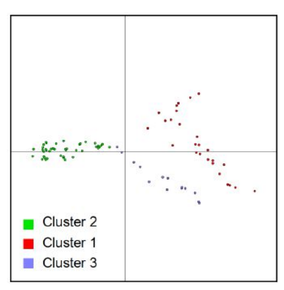
Adrian Egli, MD, PhD; Clinical Microbiology (adrian.egli@usb.ch)
University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland

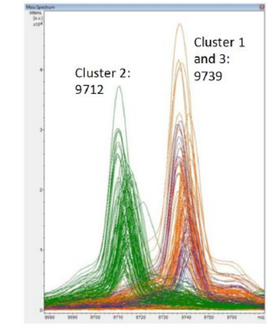
Results



- **Fig 2.** (**A**) Technical reproducibility of MALDI-TOF mass spectra in all centers was above 95%. (**B**) The biological reproducibility between different days is lower, but was sufficient to reliably differentiate clusters (see below).



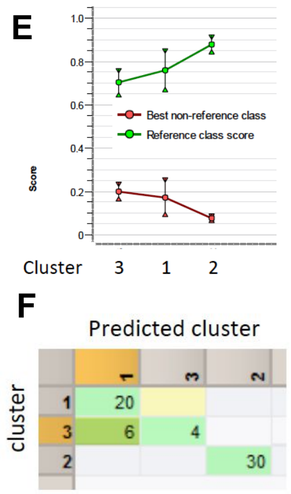




- **Fig 3.** (**A**) A principal component analysis of mass spectra showed a strong center specific effect. (**B**) However, the linear discriminant analysis indicates, that clusters can be separated. (**C**) The identification of separating peaks between all isolates **(D)** and resulted in a list of discriminant peaks to allocate each cluster. (**E**) A classifier data analysisusing a support vector machine indicates a separating capacity per reference center, as well as (**F**) a separation for all centers.

D

Peak Position	Cluster 1	Cluster 3	Cluster 2	Possible proteins (from TagIdent)
3444	Yes	Yes	No	Protamine-like protein
5873	Yes	Yes	No	Regulatory protein MokB
6539	Yes	No	No	50S ribosomal protein L30
7173	Yes	No	No	Pilin, Protein CopA/IncA
7650	No	No	Yes	Response regulator inhibitor for tor operon Protein KleB Protein IscX Cold shock-like protein CspH
7708	Yes	Yes	No	Response regulator inhibitor for tor operon Protein KleB Protein IscX Cold shock-like protein CspH
8326	Yes	Yes	No	Tautomerase PptA Dihydrofolate reductase type 2 Ferrous iron transport protein A
8350	No	No	yes	Tautomerase PptA Dihydrofolate reductase type 2 Ferrous iron transport protein A
9712	No	No	yes	30S ribosomal protein S17 Regulatory protein AriR UPF0386 protein YjhX Acid stress chaperone HdeA
9739	Yes	Yes	No	30S ribosomal protein S Regulatory protein AriR 17 UPF0386 protein YjhX Acid stress chaperone HdeA
10463	Yes	Yes	No	30S ribosomal protein S19 Sugar fermentation stimulation protein B
10489	No	No	Yes	30S ribosomal protein S19 Sugar fermentation stimulation protein B



Summary and Conclusions

- The reproduction of MALDI-TOF mass spectra for typing purposes is strongly dependent on the techinical aspect.
- Specific protocol for a reproducible sample preparation and data quality is a key element.
- Detection of specific clusters is possible even between different centers, however bioinformatic may be required.