O104:H4 outbreak and non outbreak *E. coli* strain identification by mass spectrometry: Differentiation of the two groups using ABOS, a new software tool

D.Dressler[1], M. Andrian[2,] O. Petrini [3]
[1] BioTeSys GmbH, Esslingen, Germany; [2] ARS NOVA, Esslingen, Germany; [3] POLE Pharma Consulting, Breganzona, Switzerland

Samples belonging to a large outbreak of O104:H4 Shiga-toxigenic E.coli in 2011 were analyzed by MALDI-TOF mass spectrometry after formic acid extraction (fae) and direct deposition (dsd) of samples. Toxigenic strains were separated from non-toxigenic strains using the recently developed ABOS software and sensitivity, specificity and efficiency of the ABOS identification were compared with those of a conventional analysis.
ABOS analyses the overall peak pattern of every sample. Depending on the degree of similarity with reference samples, unknown samples are classified as outbreak or non-outbreak strains. The performance of ABOS on spectra obtained with fae or dsd data as well as the impact of different threshold levels have been examined using a set of reference data of only 5 O104:H4 and 5 non outbreak strains.
ABOS differentiation between outbreak and non outbreak strains showed a specificity of 92.9%, a sensitivity of 86.3 % and an efficiency of 88.6% at a threshold level of 8, with no differences in power of discrimination for fae and dsd sample data. Similar results have been generated also using a threshold level of 4, with fae and dsd samples having been correctly classified with an even higher specificity (100%) but lower sensitivity (74.2%) and slightly lower efficiency (83.3%). In contrast, the use of a threshold level of 2 has shown much lower reliability of discrimination.
ABOS was able to classify reliably the spectra studied using only 5 strains of each group as references. In addition, and in concordance with earlier results, a number of discriminatory peaks that could be used for strain classification have been identified or confirmed.
ABOS has thus been shown to be a helpful tool to quickly detect O104:H4 *E. coli* during outbreaks. Even unprocessed (dsd) samples are useful for the identification of these strains if a threshold level of at least 4 is used.

Materials and Methods

MALDI-TOF mass spectormetry (MS) data of the 2011 northern Germany outbreak of shigatoxin producing and normal *E. coli* strains were obtained from UKE Hamburg, Germany (294 sample cultures: MALDI-TOF MS spectra from 3000 to 11700). We compared data obtained after formic acid extraction (fae) or from direct sample deposition (dsd). Three replicate measurements were merged and analyzed as single data points preprocessed with different signal to noise threshold limits (2, 4 or 8 by MALDIQuant, Bruker Daltonics). To ensure comparability the identical set of five positive (outbreak) and five negative (non-outbreak) reference strains were used for all calculations. Reference strain allocation was verified by Christner et al. (2014) using PCR genotyping and multilocus sequence typing (MLST).

Results (Table 1)

•Using ABOS, five reference samples of outbreak and non-outbreak strains are sufficient to reliably and correctly assign unknown strains to either group.

•Low signal to noise thresholds (cutoff value of 2) resulted in poorer sensitivity especially when analyzing dsd samples (Table 1). A threshold of 4 resulted in improved sensitivities and specificities. Further increase of the threshold (cut off value 8) stabilizes the sensitivity on a high level.

•ABOS provides an overview of the masses that contribute most to the analysis results as well as of their relative contribution (see figure at bottom). These masses are potential biomarkers to enable the separation of outbreak from non-outbreak strains. Our results are virtually superimposable to those obtained by Christner et al. (2014), who described the same masses as the most important candidate markers for strain classification.

		fae - signal to noise threshold = 2	dsd - signal to noise threshold = 2	fae- signal to noise threshold = 4	dsd - signal to noise threshold = 4	fae-signal to noise threshold = 8	dsd-signal to noise threshold = 8
learning groups/reference strains (sample ID)	Non-outbreak	001; 002; 004; 010; 014					
	outbreak	278; 281; 284; 288; 292					
most important (characteristic) masses		6600.9; 6842.5; 3085.9; 3356.3; 3489.4; 4449.4; 5855.8; 6711.1; 8801; 8814.3	5231.1; 3021.6; 5406.5; 10463.3; 10739.7; 3853.5; 6277.6; 9596.3; 7707.9; 11757	3356.2; 6601.2; 6711.1; 3327.5; 5442.8; 8801; 10889.3; 3854.6; 4449.4; 5855.8	3356.2; 6601.2; 6711.1; 5442.8; 5855.8; 8801; 8814.4; 9252.2; 10889.3; 11710.5	3356.2; 6711.1; 3327.5; 5339.3; 5442.8; 5855.8; 10889.3; 4449.4; 4726.2; 6842.3	3356.2; 6711.1; 3327.5; 5339.3; 5442.8; 5855.8; 10883.9; 4449.4; 4726.2; 6842.3
result quantitative	Positive/Positive	140	9	143	129	158	158
	Negative/Positive (outliers)	50	181	47	61	32	32
	Positive/Negative	104	70	104	104	104	104
	Negative/Negative (outliers)	0	34	0	0	0	0
		294	294	294	294	294	294
statistics	**Sensitivity**	0.7368	0.0474	0.7526	0.9789	0.8316	0.8316
	lower 95% CI	0.6682	0.0219	0.685	0.6075	0.7706	0.7706
	upper 95 % CI	0.7979	0.088	0.8122	0.7447	0.8819	0.8819
	Specificity	1	0.6731	1	1	1	1
	lower 95% CI	0.9652	0.5741	0.9652	0.9652	0.8498	0.8498
	upper 95 % CI		0.7619				
	Efficiency	0.8299	0.2687	0.8401	0.7925	0.8912	0.8912
	lower 95% CI	0.782	0.2189	0.7931	0.7416	0.8498	0.8498
	upper 95 % CI	0.8711	0.3232	0.8801	0.8374	0.9244	0.9244
	PVP	1	0.2093	1	1	1	1
	PVN	0.6753	0.2789	0.6887	0.6303	0.7647	0.7647

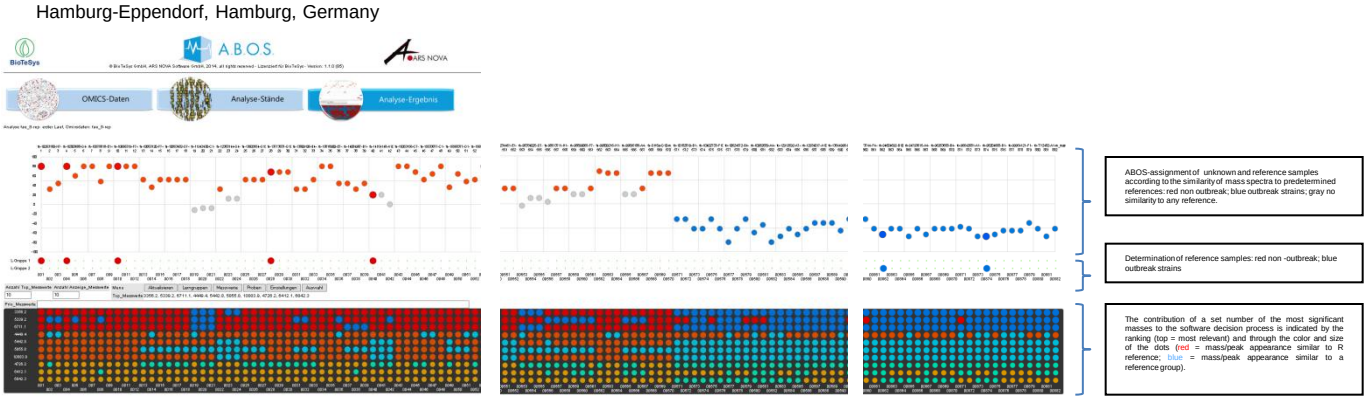
Table 1. Results of classification analysis by ABOS. fae = formic acid extracted; dsd = direct sample deposition; sensitivity = true positive rate; specificity = true negative rate; efficiency = correct classification rate; CI = confidence interval; PVP = predictive value of positive test; PVN = predictive value of negative test.

Summary

ABOS could quickly and reliably identify outbreak and non-outbreak *E. coli* strains on the basis of only ten reference strains (identified by PCR genotyping and MLST). Optimized signal to noise ratios help to strengthen the statistical power of the identification. This study has shown that ABOS is a suitable tool to analyse MALDI-TOF MS data and to provide reliable results within hours.

Acknowledgements

MALDI-TOF data were provided by Dr. Martin Christner, Department of Medical Microbiology, Virology and Hygiene, University Medical Center Hamburg-Eppendorf, Hamburg, Germany



Rapid MALDI-TOF Mass Spectrometry Strain Typing during a Large Outbreak of Shiga-Toxigenic *Escherichia coli.* Christner et al. PLOS One 9(7) 2014