Systematic determination of biomarkers by automated processing of gas chromatography-mass spectrometry data

Saïd Abou El Karam, Nathalie Kondjoyan, Jérémy Ratel, Erwan Engel

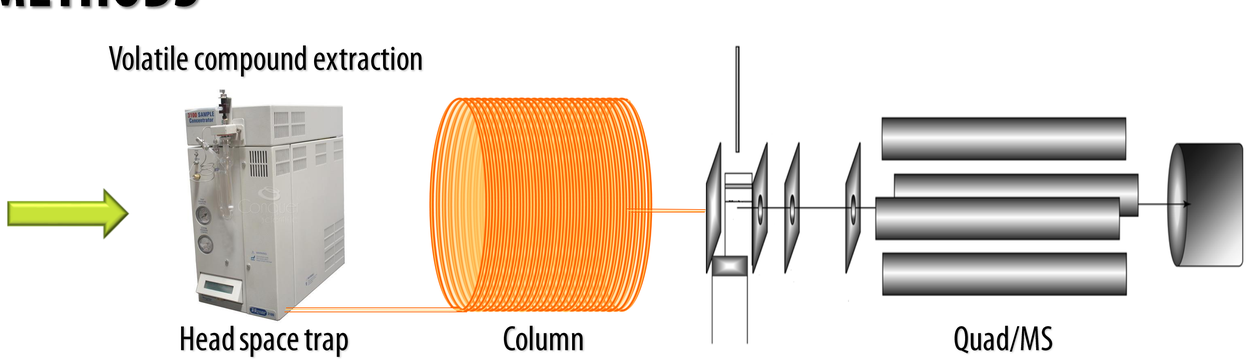
SYSTEMATIC DETERMINATION OF BIOMARKERS BY AUTOMATED PROCESSING OF GAS CHROMATOGRAPHY-MASS SPECTROMETRY DATA

SAÏD ABOU EL KARAM, NATHALIE KONDJOYAN, JÉRÉMY RATEL, ERWAN ENGEL*
INRA, UR370 QuaPA, MASS Group, F-63122 Saint-Genès-Champanelle, France. *e-mail : erwan.engel@clermont.inra.fr

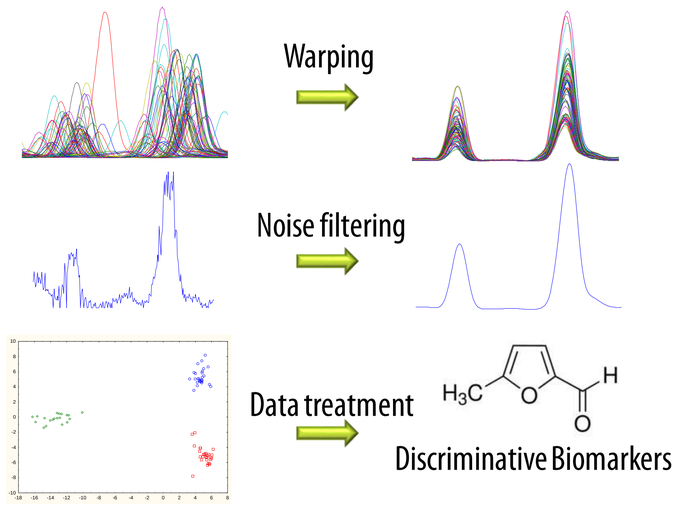
CONTEXT and OBJECTIVES The development of automated tools for systematic search of markers in the gas chromatography-mass spectrometry signal requires breaking down the barriers posed by the signal treatment and more particularly by the alignment. The present study proposes a toolkit including a novel alignment approach inspired by the COW method [3] which was upgraded for correcting local distortions. Beyond alignment problems, the determination of the discriminative markers [2] was based on an original data normalisation method [1]. In order to assess its performance, the toolkit was applied to a data set giving the composition in volatile compounds of three walnut oils from different origins. Each walnut oil was analysed with 20 different analytical setups of the same GC-MS system in order to mimic instrumental drifts.

MATERIAL and METHODS

Three walnut oils [1]



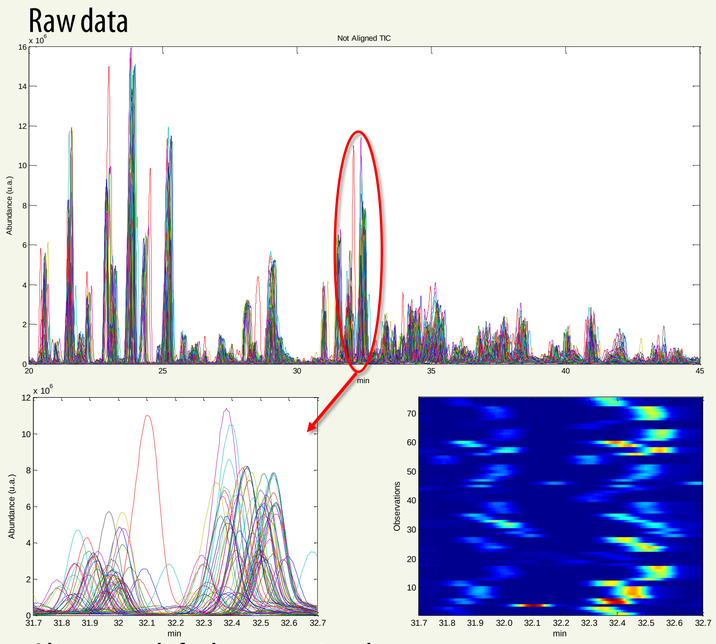
New & worn traps × New & worn Columns × 5 autotuning settings = 20 GC-MS setups [1]



RESULTS

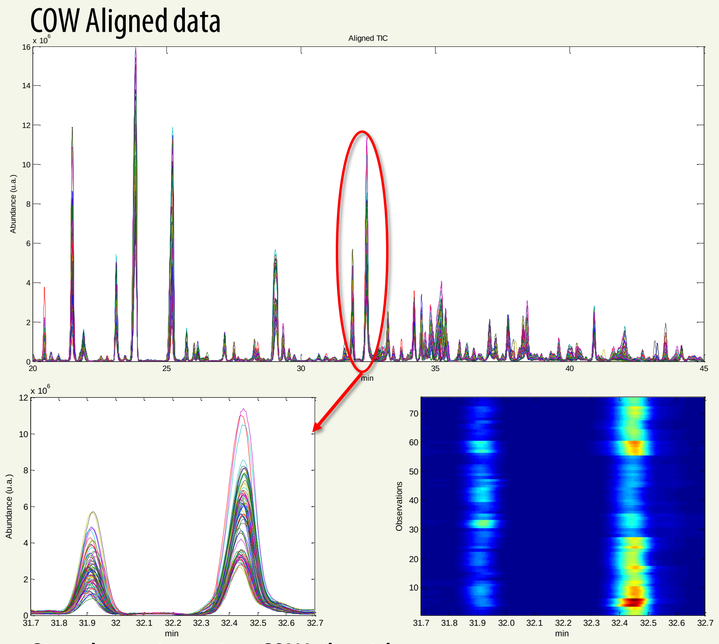
Warping of GC-MS signals

Raw data



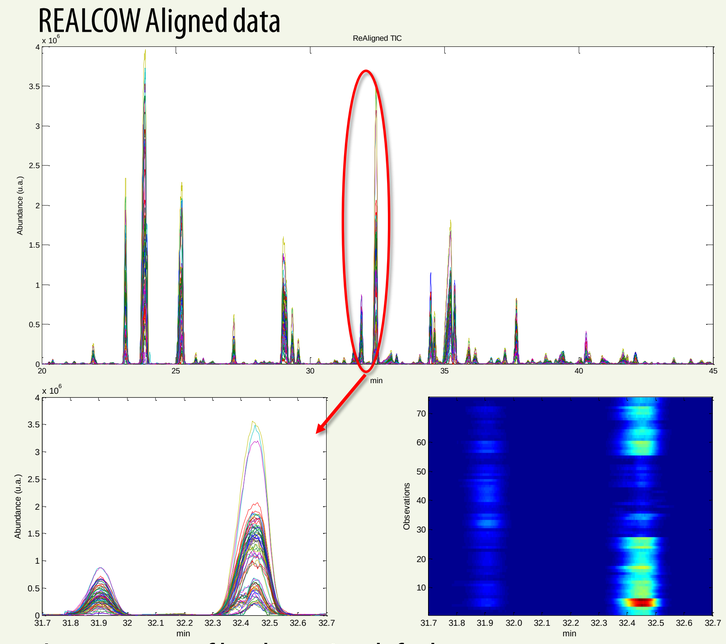
Alignemnt default on raw signals

COW Aligned data



Signal warping using COW algorithm

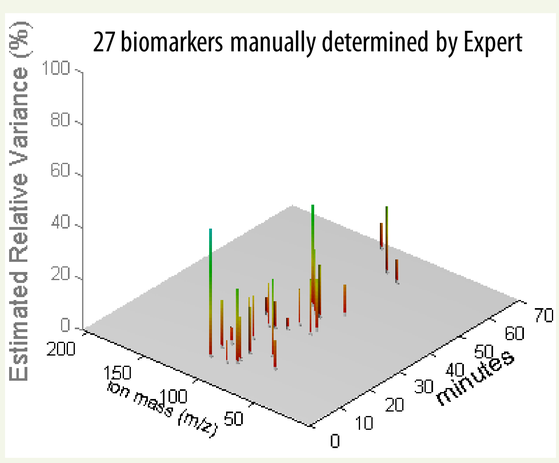
REALCOW Aligned data



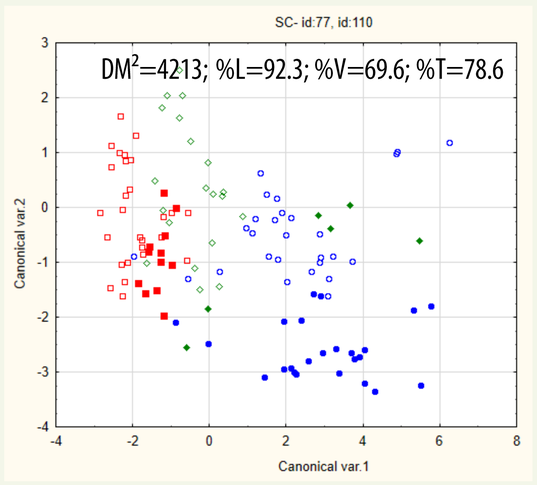
Improvement of local warping defaults

Determination of biomarkers

Expert method (manual approach)

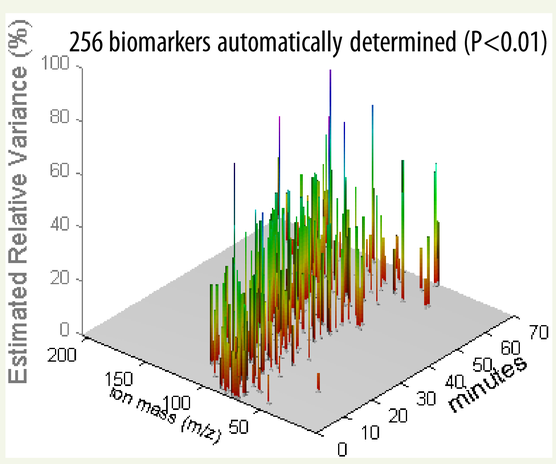


Anovagram

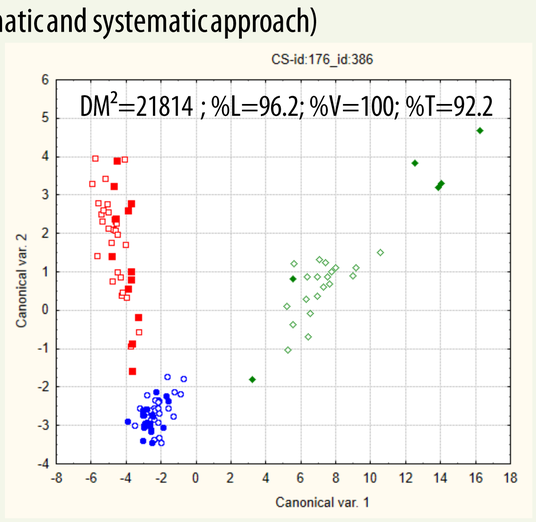


Discriminant analysis
2 best variables among 27

Systematic method using REALCOW algorithm (automatic and systematic approach)



Anovagram



Discriminant analysis
2 best variables among 256

CONCLUSIONS The biomarkers identified by expert were also found by the systematic method which in turn offered much more pertinent and discriminative markers. This method was validated on our data set with important instrumental drifts but its efficiency must be confirmed on other data sets.

REFERENCES [1] Deport, C., Ratel, J., Berdague, J. L., & Engel, E. (2006). Comprehensive combinatory standard correction: A calibration method for handling instrumental drifts of gas chromatography-mass spectrometry systems. *Journal of Chromatography A, 1116*(1-2), 248-258.
[2] Mass Spectra Separation Software (MS3), IDDN.FR.001.270012.000.R.P.2010.000.10800. INRA - UR QuaPA - F-63122 Saint Genès Champanelle, France.
[3] Tomasi, G., van den Berg, F., & Andersson, C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. Journal of Chemometrics, 18(5), 231-241.

6th International Symposium on Recent Advances in Food Analysis (RAFA), 2013, Prague, Czech Republic.

Centre de Clermont-Theix-Lyon
63122 Saint Genès Champanelle – France,
www.clermont.inra.fr