www.sepsolve.com

Uncovering hidden compositional changes in breath profiles using TD–GC×GC–TOF MS and untargeted chemometrics

Laura McGregor *, Ryan Sutherill and Bob Green

SepSolve Analytical, 4 Swan Court, Cygnet Park, Peterborough, PE7 8GX, UK. * Email address: lmcgregor@sepsolve.com

Introduction

Volatile organic compounds (VOCs) emitted in breath have great potential for use in non-invasive disease diagnosis, but in the biomarker discovery phase of large-scale clinical trials, an incorrect identification can compromise the validity of an entire trial, meaning that both robust analytical techniques and confident data mining are required.

Thermal desorption (TD) coupled with GC–MS is known as the 'gold standard' for breath analysis, due to its ability to capture a complete breath profile with high sensitivity. Here, we combine TD with advanced separation and detection by GC×GC–TOF MS to gain greater insight into sample composition.

However, data acquisition is just the beginning – the information-rich chromatograms must then be transformed into meaningful results. In this study, we demonstrate the use of a powerful chemometrics platform to align and compare chromatograms using automated untargeted workflows.

Experimental

Sampling: 1 L breath samples were collected using sampling bags from two groups of participants and transferred to 'Biomonitoring' sorbent tubes (Markes International).

TD: Instrument: Centri® (Markes International) using the tube-based TD module with a 50-tube autosampler.

GC×GC: INSIGHT® flow modulator (SepSolve Analytical); P_M 2.5 s.

TOF MS: BenchTOF2™ (SepSolve Analytical); m/z 45–450 at 100 Hz in Tandem Ionisation® mode.

Software: Full instrument control by ChromSpace®, with data mining and chemometrics in ChromCompare+ (SepSolve Analytical).

Figure 1: Analytical system used in this study.

Results and discussion

Chromatographic alignment

In the untargeted data analysis workflow, the first step was chromatographic alignment (Figure 2) to account for any retention time drift over the course of the study and minimise the risk of false hits.

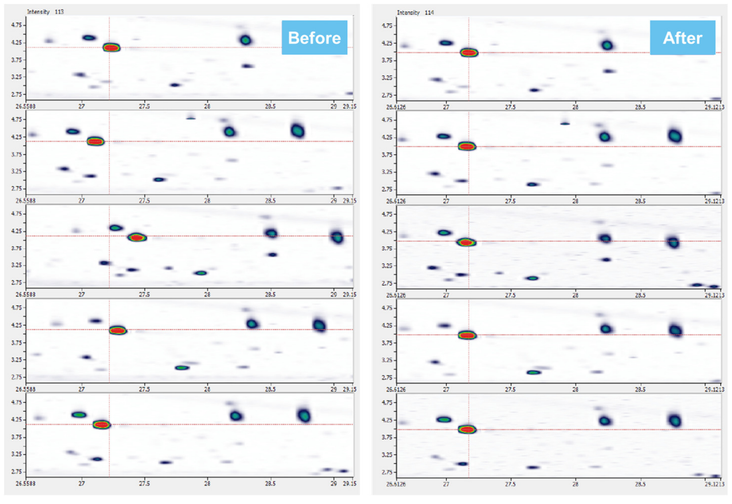


Figure 2: Enhanced region of TD–GC×GC–TOF MS colour plots showing the use of chromatographic alignment in ChromCompare+.

Feature discovery

Next, feature discovery was performed on the raw data to find significant changes across sample classes. In metabolomics matrices, the diagnostic compounds are rarely of high abundance – by adopting a raw data approach, trace peaks are not overlooked (Figure 3). Additionally, the use of raw data enables automated workflows to be adopted, minimising laborious pre-processing steps and accelerating analytical workflows.

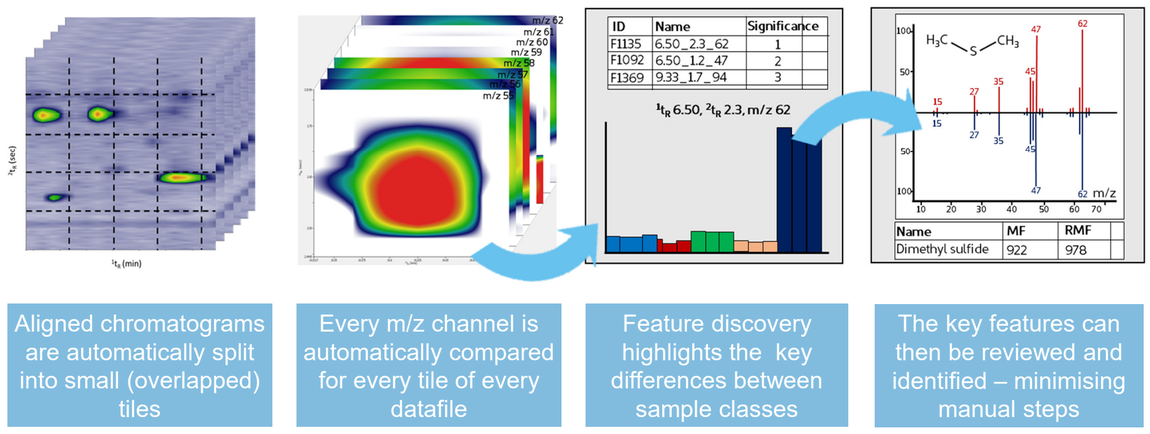


Figure 3: Overview of the automated untargeted workflow in ChromCompare+.

This reduced feature list was then viewed as a volcano plot (Figure 4), which conveniently highlights the features that are statistically increased (red) or decreased (blue) in the Group B samples relative to the Group A samples.

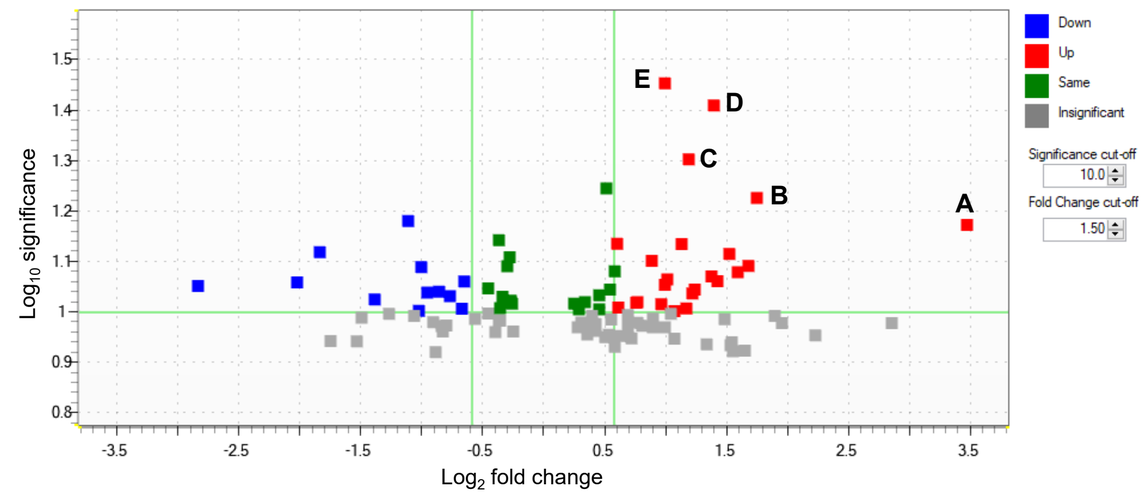
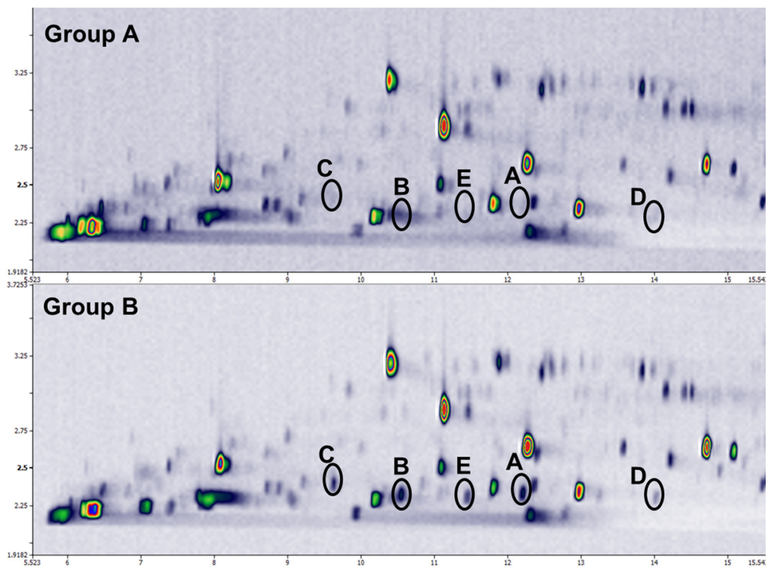


Figure 4: Volcano plot in ChromCompare+ for 100 features selected using feature discovery. Five statistically significant features are annotated (A–E) and identified in Figure 5.

Five statistically significantly features stand out in the volcano plot (labelled A–E). Using the retention time and m/z information provided in the feature list, the analytes represented by these features were identified (Figure 5). This streamlined workflow eliminates the need to integrate and identify hundreds of peaks that may not be relevant.



A. (*Z*)-1-Methylthio-1-propene
B. Allyl methyl sulfide
C. Methyl propyl sulfide
D. Dimethyl disulfide
E. (*E*)-1-Methylthio-1-propene

Figure 5: Enhanced region of the TD–GC×GC–TOF MS colour plots for breath samples from Groups A and B, showing identification of five of the most significant differences between the sample classes (labelled A–E in Figure 4).

Five statistically significantly features stand out in the volcano plot (labelled A–E). Using the retention time and m/z information provided in the feature list, the analytes represented by these features were identified (Figure 5). This streamlined workflow eliminates the need to integrate and identify hundreds of peaks that may not be relevant.

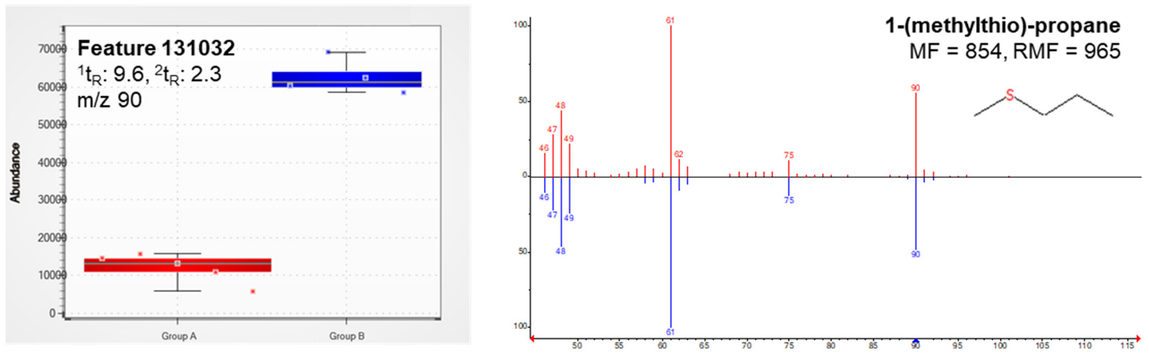


Figure 6: Box and whisker plot for one of the most significant features and its identification using BenchTOF.

Conclusions

- TD–GC×GC–TOF MS captures comprehensive breath profiles with high sensitivity to gain maximum insight into sample composition.
- Fully automated data analysis in ChromCompare+ minimises laborious pre-processing steps and accelerates workflows.
- Importing the entire raw dataset reduces the risk of missing important trace differences, thereby increasing confidence in results.
- Interactive charts, such as PCA plots, volcano plots and box plots, allow easy visualisation of trends and differences between samples.

SepSolve Analytical Ltd
T: +44 (0)1733 669222 (UK) +1 888-379-3835 (USA)
E: hello@sepsolve.com