ChromMine™ Application Note 25-002

Dried herbs – demonstrating the power of ChromMine™ clustering

Abstract

Dried culinary herbs are widely used in domestic and commercial cooking, yet their complex aroma profiles vary considerably between species and blended products. In this study, six common kitchen herbs and herb mixtures - oregano, rosemary, thyme, basil, mixed herbs, and Italian herbs - were analysed in duplicate using thermal desorption gas chromatography–mass spectrometry (TD-GC-MS) headspace sampling. Compound tables exported from **Agilent MassHunter Unknowns Analysis™** were processed and visualised in **ChromMine™**, enabling rapid comparison of volatile fingerprints.

ChromMine™ automatically grouped replicates, highlighted inconsistencies in compound identification, and facilitated reproducible data correction prior to re-analysis. Distinct chemical signatures were observed for each herb type, with the blended products forming intermediate clusters consistent with their mixed composition. This study demonstrates how **ChromMine™** streamlines non-targeted GC-MS data handling, providing transparent, metadata-driven insight into complex flavour mixtures in seconds.

Introduction

In this study, six common kitchen herbs and herb mixtures - oregano, rosemary, thyme, basil, mixed herbs, and Italian herbs - were analysed in duplicate using thermal desorption gas chromatography–mass spectrometry (TD-GC-MS) headspace sampling. The resulting chromatograms were first processed in **Agilent MassHunter Unknowns Analysis™**, and the exported results tables were imported into **ChromMine™** for processing, visualisation, and comparison of volatile fingerprints.

Initial import and clustering revealed that some compounds had been identified differently between replicate samples - highlighting one of **ChromMine's** key strengths: the ability to detect and flag inconsistencies at the data-structure level. Once these discrepancies were resolved, the dataset was re-imported and re-processed, producing a clean, reproducible foundation for comparative analysis.

Experimental

Sampling and GC-MS analysis

Headspace sampling was carried out directly from small quantities (~2.5 g) of each dried herb or herb mixture placed in 250 mL wide-neck glass Schott bottles (Fisher Scientific, UK).

Samples were equilibrated at room temperature before collecting the headspace vapour onto glass Tenax-TA sorbent tubes (CAMSCO Inc., US) and analysed using thermal desorption gas chromatography–mass spectrometry (TD-GC-MS) under standard non-targeted conditions (Markes International TD100-xr, Agilent 8890–5977B GC/MS).

The resulting chromatograms were processed using Agilent MassHunter Unknowns Analysis™ (UA) software (NIST23™ library, minimum 75% match factor, 1 ng toluene equivalent minimum area).

Compound identifications and integrations were then exported as per-compound, per-sample CSV files for ChromMine™ processing.

ChromMine import and processing

The raw data exported from UA consisted of a per-compound, per-sample CSV file with ten columns (including file name, sample name, retention time, compound name, component area, CAS number, molecular formula, and match factor) and 254 rows—amounting to approximately 2 500 data points. ChromMine™ ingested and parsed this dataset within seconds, automatically structuring the UA output into an analysis-ready table for downstream filtering and clustering.

Cluster comparison - ChromMine's "superpower"

Figure 1 shows the ChromMine™ cluster-plot projection for the herb and blank samples, with the common siloxane contaminants removed by filtering, but prior to any data or identification review.

Unlike Principal Component Analysis (PCA), the relative positioning of samples in the ChromMine Clustering Algorithm (CCA) plot carries direct meaning. With the chosen axes - **sample similarity** along the x-axis and **concentration similarity** along the y-axis - replicate analyses would be expected to align both vertically and horizontally. Vertically, because replicates should contain the same VOC composition; and horizontally, because they should exhibit comparable overall VOC concentrations.

Reviewing the initial cluster plot made it immediately clear that something was inconsistent. The thyme replicates (highlighted with a red box) displayed divergent VOC profiles, as did the mixed herbs analyses. Similarly, the two Italian herb replicates showed comparable VOC patterns but markedly different overall concentrations.

Looking at the reconstituted two-dimensional line chromatograms of the thyme replicates (**Figure 2**), the concentration disparity is evident. The blue sample was analysed before the orange sample, suggesting that insufficient time was allowed for headspace re-equilibration between runs. However, the chromatographic profiles appeared far more similar than the clustering suggested.

Investigating the large peak at ~18 minutes (by hovering over it with the mouse in ChromMine) it became apparent that two different isomers of thymol had been identified in the two different runs. Using the sample profile comparison tool in ChromMine (which sorts by retention time order) this difference is even more obvious as demonstrated in **Figure 3**.

Returning to UA and checking the match factors for these two thymol isomer identifications - it could be seen that they were 98.5% and 98.3% - this is a very common occurrence for substituted structural isomers. To ensure consistency, thymol was attributed to that peak in both runs.

The same procedure was used to identify discrepancies in the mixed herbs – where structural isomers had also caused a smaller discrepancy in sample similarity – in that case between *o*-cymene and *p*-cymene.

Looking at the Italian herb replicates it was evident that they had the same overall profile - but the second run had lower concentrations for each component across the board. As with the thyme analysis, this was put down to not leaving the sample long enough to fully recover the headspace before taking the replicate.

Having observed that misidentification of structural isomers could cause such large discrepancies in replicate samples, the whole data set was looked at in the bubble plot view (figure 4) to determine if there were any other, sample-set-wide compounds that required investigation.

Figure 1: Cluster plot of herb and blank samples before data review – thyme replicates highlighted with a red box, Italian herb replicates highlighted with a blue box and mixed herbs with the green box. Horizontal differences indicate different VOC profiles, vertical differences point at compound concentration.

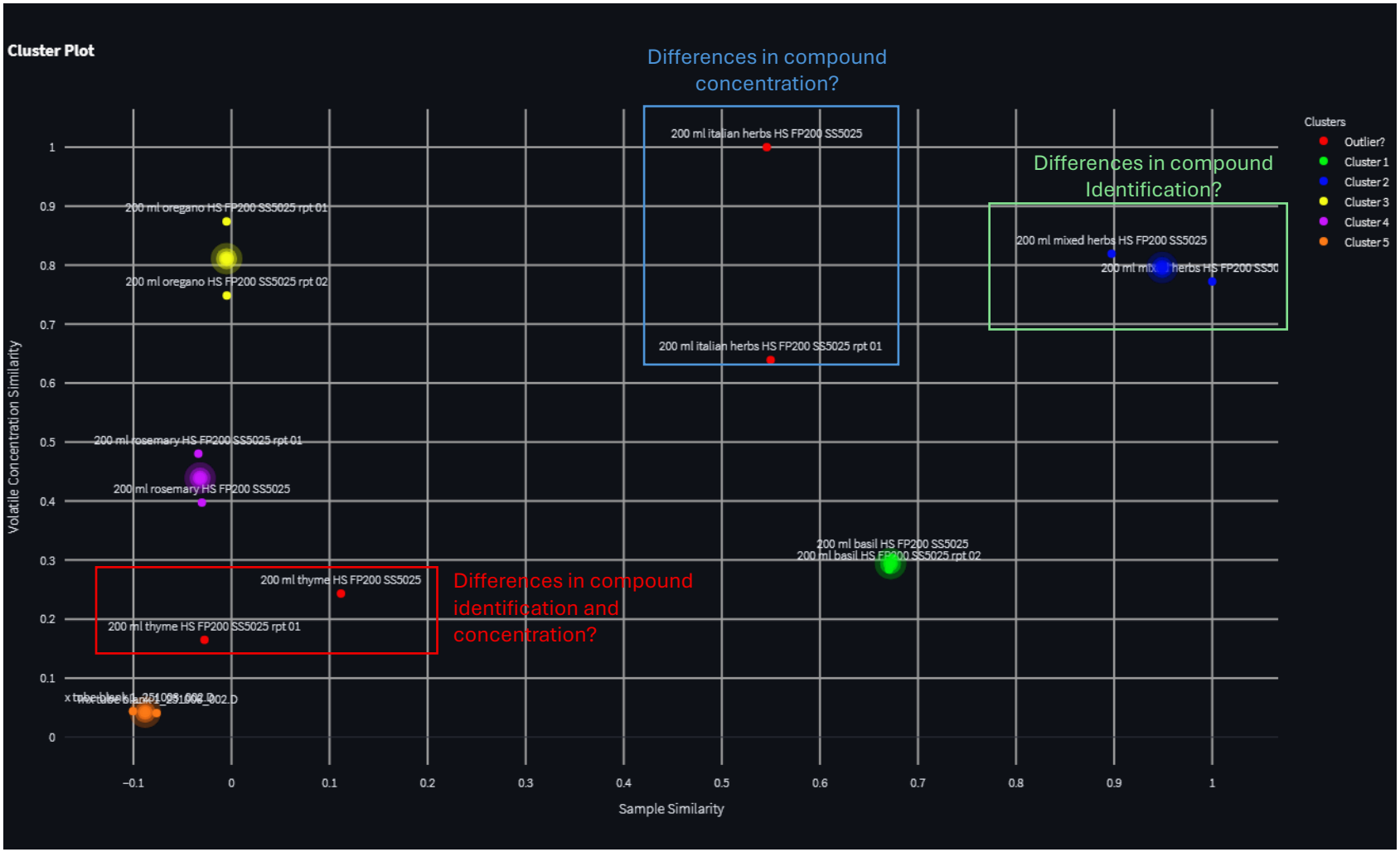


Figure 2: ChromMine reconstituted line chromatogram comparing the two runs of thyme headspace

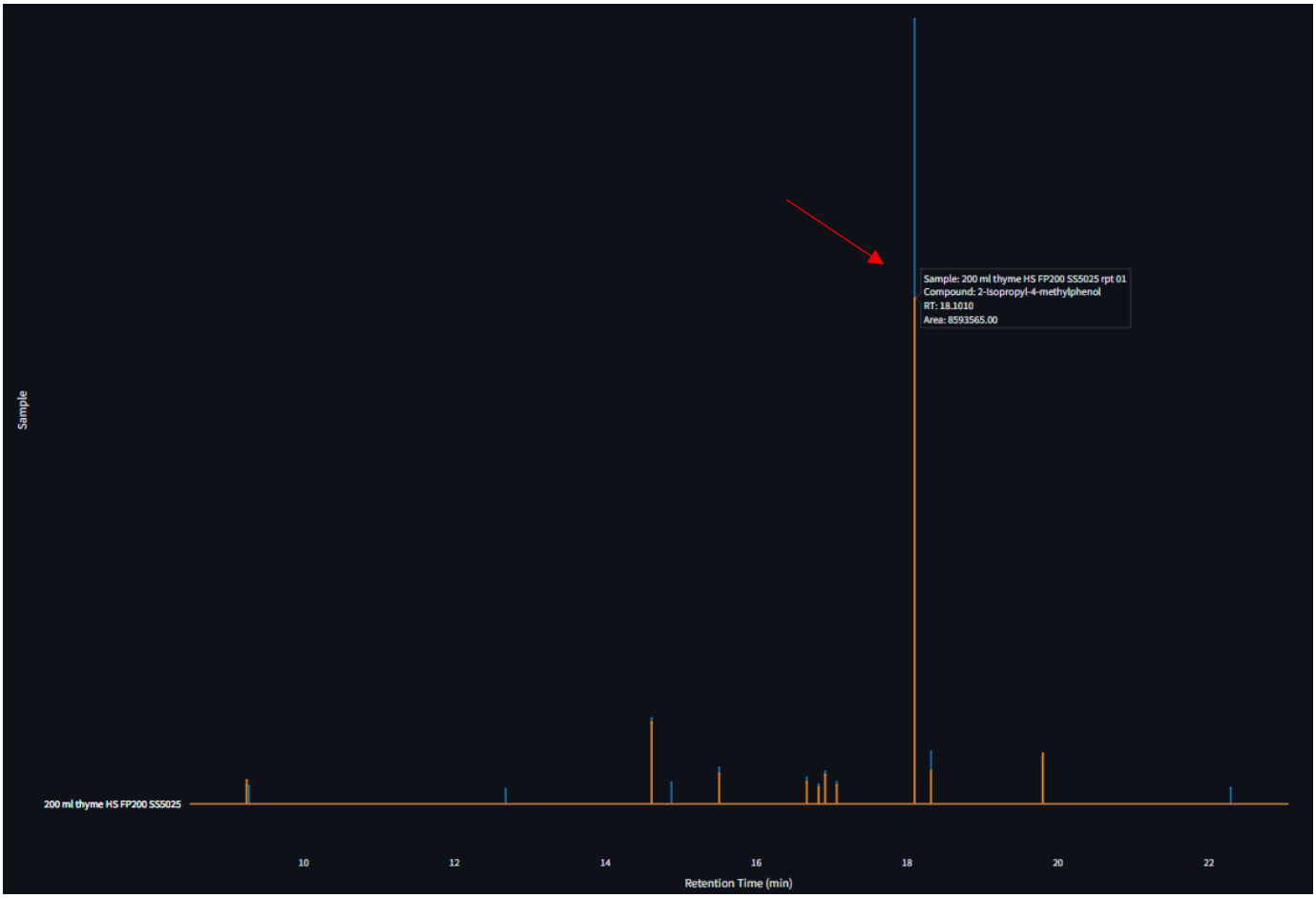
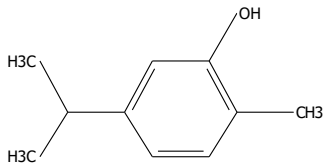


Figure 3: Sample profile comparison tool (sorted by retention time) comparing the two thyme headspace replicates with the two structural isomers of thymol highlighted in the red box

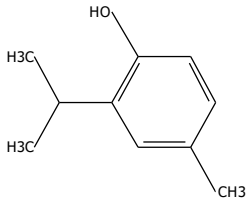
Figure 4: Excerpt from the all-sample Bubble Plot view whilst determining if there were set-wide misidentified isomers from one sample to another

The Bubble Plot shows that a dominant compound in the Italian herbs at 18.317 minutes was potentially also found at 18.318 minutes in the oregano samples, with the same molecular formula (see **Figure 4**). Again, returning to UA, it could be seen that these were structural isomers.

Found in Italian herbs:



Found in oregano sample:



Given that the main ingredient of Italian herbs is oregano, and that the match factor to identify as either of these isomers was 97.8% versus 97.7%, it was decided that these compounds should also be attributed the same structure for consistency. Making these, and the mixed herb changes, the cluster plot was redone and gives what we see in **Figure 5**.

Reviewing **Figure 5** we can now see that the oregano and Italian herb samples group closer together, as do the thyme and mixed herbs samples.

Conclusions

This application note demonstrates that the volatile organic compounds responsible for the aroma of different dried herbs vary appreciably from one type to another: an observation familiar to any chef.

More importantly, it highlights the power of **ChromMine™**, and particularly its clustering algorithm, to identify when compounds have been misidentified between replicate analyses or across a whole sample set.

By distinguishing true analytical differences from artefacts of compound identification, **ChromMine™** ensures reproducible, data-driven insight into complex chemical mixtures.

Figure 5: Cluster plot redone with corrected identifications and improved clustering