Application Note

Food Authenticity

Agilent

Trusted Answers

Workflow for Authenticity Testing of Plant Extract Using Revident LC/Q-TOF and MassHunter Explorer

Authors

Megan Laryea-Akrong and Siheng Li
doTERRA International, LLC, Pleasant Grove, UT, USA

Hui Zhao and Jim Lau
Agilent Technologies, Inc.

Abstract

Interest in food authenticity testing is growing rapidly across the food manufacturing industry. Food manufacturers often encounter adulteration and false labeling in the complex food supply chain and consequently, there is a high demand for powerful, convenient, and easy-to-use analytical tools. High-resolution mass spectrometry (HRMS) used in a non-targeted approach is gaining popularity in detecting food fraud and adulteration. This study used lavender essential oil to demonstrate a novel authenticity workflow. This workflow utilized Agilent Revident liquid chromatography/quadrupole time-of-flight mass spectrometer (LC/Q-TOF) with non-targeted data collection to profile and fingerprint complex lavender extracts, and Agilent MassHunter Explorer software for feature extraction, statistical analysis, and compound identification. Besides the integrated Agilent library search, Explorer provides direct access to the NIST LCMS (MS/MS) Search along with complementary, direct access to SIRIUS and enhances accuracy and confidence in identifying unknown compounds.

Introduction

Lavender essential oil (LEO) is often analyzed using gas chromatography/triple quadrupole or gas chromatography/quadrupole time-of-flight mass spectrometry (GC/TQ or GC/Q-TOF) for volatile compounds. The chemical composition of lavender extracts is also rich in secondary metabolites, such as phenolic acids and coumarins, which can produce unique fingerprints using LC/MS.[1] A comprehensive non-targeted approach was applied to study composition differences and characterize unknown chemical components of LEO and its common adulterants using high-resolution accurate mass LC/Q-TOF. Agilent MassHunter Explorer software[2] was used to efficiently process LC/MS and LC/MS/MS data by multivariate analysis using tools that perform feature extraction and alignment to find and align compounds across samples. MassHunter Explorer analyzed the samples by comparing compound similarities and differences between samples or groups of samples. Statistically valid differentiation of lavender essential oil originating from different batches from the same geographical region, different geographical regions, and/or detecting adulteration was accomplished. Convenient visual review of critical compounds that are significantly different or common between groups was provided by statistical tools, such as two-way analysis of variants (ANOVA), hierarchical clustering, fold change, volcano plot, and Venn diagrams. Once the interested unknown compounds were selected, identification followed. Extracted MS and MS/MS spectra were matched to curated Agilent compound databases and ChemVista which allows access to public domain libraries such as MassBank and MoNA. Furthermore, direct export from Explorer to NIST MS Search assisted with compound identification by providing reference mass spectra for LC/MS/MS. Similarly, direct, complementary data export from Explorer to SIRIUS CSI:Finger ID generated molecular formulas by combining isotope pattern analysis and fragmentation analysis which resulted in a hypothetical fragmentation tree. The straightforward process dramatically increased the accuracy of unknown identification. In summary, the accurate mass LC/Q-TOF detection technique, combined with the advanced MassHunter Explorer differential analysis workflow, successfully analyzed and interpreted lavender essential oil profiling results, confirming adulteration, and identifying unknown compounds. This workflow provided an integrative quality control method for relevant plant extracts.

Experimental

Equipment

All experiments in this study were performed using an Agilent 1290 Infinity II LC consisting of an Agilent 1290 Infinity II multisampler (G7167B), an Agilent 1290 Infinity II high speed pump (G7120A), and an Agilent 1290 Infinity II multicolumn thermostat (G7116B) coupled to an Agilent Revident Q-TOF (G6575AA), see Figure 1. The system was controlled by Agilent MassHunter Acquisition software, version 12.1. Data processing was performed with MassHunter Explorer (version 2.0) and MassHunter Qualitative Analysis software (version 12.0).

Chromatographic conditions

Parameter	Setting
Analytical Column	Agilent InfinityLab Poroshell 120 StableBond-Aqueous, 2.1 × 150 mm, 2.7 µm (part number 683775-914)
Column Oven	40 ± 2 °C
Injection Volume	2 µL
Run Time	30 min
Autosampler	5 ± 2 °C
Mobile Phase A	0.1% Formic acid in water
Mobile Phase B	0.1% Formic acid in methanol
Needle Wash	0.1% Formic acid in acetonitrile
Gradient	Time (min) Flow (mL/min) %A %B 0 0. 25 90 10 10.0 0. 25 40 60 15.0 0. 25 20 80 22.0 0.25 0 100 25.0 0. 25 0 100 26.0 0.25 90 10 30.0 0.25 90 10

Revident Q-TOF parameters

Parameter	Setting
Drying Gas Temperature	275 ºC
Drying Gas Flow	12 L/min
Nebulizer Gas	25 psi
Sheath Gas Temperature	385 °C
Sheath Gas Flow	12 L/min
Capillary Voltage	ESI+ 4,000 V
Nozzle Voltage	ESI+ 1,000 V
Ion Mode	AJS ESI Positive
Fragmentor	125 V
Skimmer	45 V
MS Tune	*m/z* 50–1,700
MS Acquisition	MS Scan and Auto MS/MS MS/MS: 40V CE
MS Range	*m/z* 70–1,100 for MS and *m/z* 25–1,100 for MS/MS
Reference Mass	922.0098 ($[M+H]^+$ for protonated HP-0921)

Figure 1. Agilent 1290 Infinity II LC with Revident LC/Q-TOF.

Samples

Lavender essential oil was used in this study to test a novel authenticity workflow that allows the differentiation of lavender essential oil from different geographical regions, detecting adulteration, and identifying the critical compounds which are significantly different or common between groups. The authentic lavender oil was obtained from the steam distillation of the flowering tops of *Lavandula angustifolia*. Ten different lots from three geographical regions and four subtypes of *Lavandula angustifolia* were provided by doTERRA (Pleasant Grove, UT, USA). Adulterated lavender oil including seven commonly used less expensive plant extracts such as ho leaf oil, clove oil, and one oil prepared by synthetic fragrance compounds were also provided by doTERRA. The information on authentic and adulterated lavender oil samples was shown in Table 1. The samples were stored at room temperature.

Table 1. Information on authentic lavender oil samples and adulterated lavender oil samples

Authentic Lavender Essential Oil
Lavender Bulgarian (LB)
Lavender France (LF)
Lavender, China Blue Flower (LCB)
Lavender, China White Flower (LCW)
Commonly Used Lavender Oil Adulterants
A: Ho wood crude oil
B: Ho leaf crude oil
C: Raw clove bud oil
D: Eucalyptus oil
E: Rosewood oil
F: Potentially adulterated lavender essential oil from different vendor
G: Adulterated lavender oil by synthetic constituents

Workflow

Sample preparation and analysis: All the samples, including authentic and adulterated plant extracts, were diluted in methanol (0.4%, v/v). The authentic finished product was also blended with 1, 5, and 20% selected fraudulent ingredients to simulate adulterated samples. Table 2 lists the detailed preparation. Quality control (QC) samples (authentic extracts or authentic finish product) were prepared at replicates of three to six. Selected adulterated sample B was prepared at replicates of three to further verify the reproducibility.

Table 2. Preparation of authentic finished product as control and adulterated samples blended in with different % of fraud plant extracts.

Sample Groups	Fraudulent Product Blended in %	Replicates
Authentic Control	0	5
Adulterated sample B-1%	1	3
Adulterated sample B-5%	5	3
Adulterated sample B-20%	20	3
Adulterated sample D-1%	1	1
Adulterated sample D-5%	5	1
Adulterated sample D-20%	20	1
Adulterated sample E-1%	1	1
Adulterated sample E-5%	5	1
Adulterated sample E-20%	20	1
Adulterated sample F-1%	1	1
Adulterated sample F-5%	5	1
Adulterated sample F-20%	20	1
Adulterated sample G-1%	1	1
Adulterated sample G-5%	5	1
Adulterated sample G-20%	20	1

Statistical analysis: A comprehensive non-targeted approach was applied to study composition differences and characterize unknown chemical components of lavender essential oil and its common adulterants using high-resolution accurate mass LC/Q-TOF. Raw Q-TOF data were imported into the Agilent SW, MassHunter Explorer. Explorer efficiently processed LC/MS and LC/MS/MS data by multivariate analysis using tools that perform feature extraction and alignment to find and align compounds across samples. Explorer investigated the samples by comparing compound similarities and differences between samples or groups of samples. Statistically valid differentiation of lavender essential oil from different batches from the same geographical region, different geographical regions, and/or detecting adulteration was accomplished. Convenient visual review of critical compounds, which are significantly different or common between groups was provided by statistical tools, such as hierarchical clustering and volcano plot.

Unknown identification: Once the interested unknown compounds were selected, identification followed. Extracted MS and MS/MS spectra were matched to curated Agilent compound databases and public domain libraries such as MassBank and MoNA. Furthermore, direct export from Explorer to NIST MS Search assisted with compound identification by providing reference mass spectra for LC/MS/MS. Similarly, direct, complementary data export from Explorer to SIRIUS generated molecular formulas especially for compounds without reference spectra by combining isotope pattern analysis and fragmentation analysis which resulted in a hypothetical fragmentation tree. The process dramatically increased the accuracy of unknown identification.

Results and discussion

Sample elution profile and Q-TOF LC/MS detection

LEO is a highly complex matrix, with thousands of compounds eluting from the selected StableBond aqueous column, followed by detection using electrospray ionization (ESI) Q-TOF in full MS scan and AutoMSMS modes. This setup provided satisfactory sensitivity and acceptable analysis time. The total ion chromatogram (TIC) in Figure 2A illustrated the elution profile of the Bulgarian-origin LEO sample. Over 2,500 compounds were extracted using Find and Align, as shown in Figure 2B, highlighting the high separation efficiency achieved. Compounds were resolved based on retention time and/or accurate mass-to-charge ratio (*m/z*), with online calibration using reference ions maintaining mass accuracy within 1 ppm.

Features finding and alignment

In the Find and Align segment of MassHunter Explorer, parameters are set based on sample chemistry; ion features (e.g., isotopes, adducts, and charge states, RT tolerance, etc.) are extracted from chromatographic data using nontargeted algorithms. The related ion features are grouped into compound features using accurate mass and retention time, Isotopic distribution patterns, and coelution behavior. These features are then aligned across multiple samples based on retention time and accurate mass (*m/z*) into compound groups, allowing for consistent compound identification and comparison. Quick and efficient review of extracted ion chromatograms and spectra are achieved by extensive graphical displays. Through the previously described treatment, over 6,000 features in 17 authentic and fraudulent plant extract samples were obtained, which were subjected to further data filtering and normalizing. The data acquired through the Q-TOF MS scan mode were also verified on data reproducibility in both retention time (RT) and mass-to-charge ratio (*m/z*).

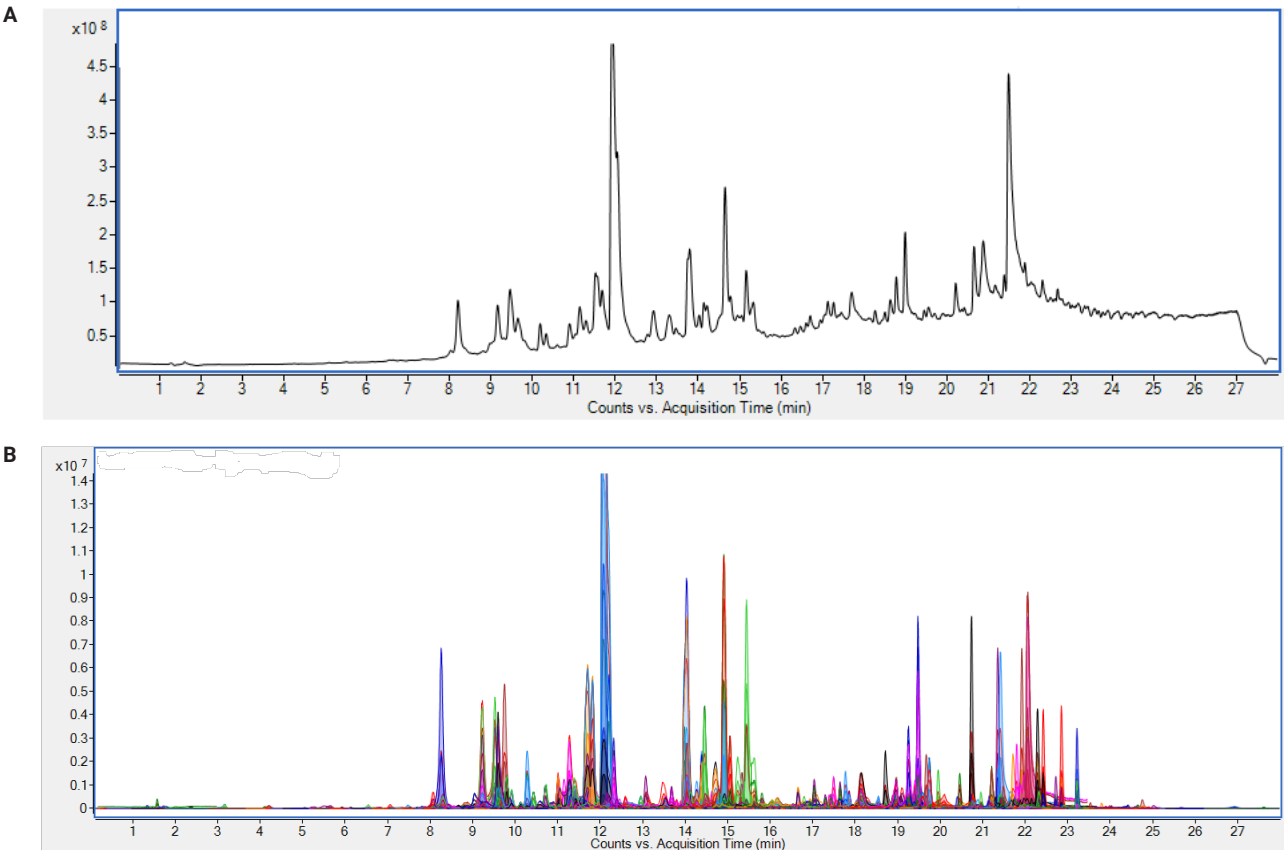


Figure 2. (A) TIC of LEO Bulgarian origin, (B) Find and Align elution profile of LEO Bulgarian origin.

Statistical analysis: principal component analysis (PCA)

PCA is a dimension reduction technique that detects major trends and brings out strong patterns from large and multivariate data by capturing a few Principal Components (PC 1, PC 2, etc.). These Principal Components convey the most variation in the dataset. The PCA plot displays one data point per sample and the samples with similar expression profiles will cluster together. Strong visualization of the separations between sample groups and replicates by different colors are offered. In a 2D PCA plot, PC1 reveals the most variation, while PC2 reveals the second most variation, thus PC1 and PC2 are plotted first to capture the maximum variation of the dataset. In a 3D PCA plot, PC1, PC2 and PC3 are plotted. The compound features from Find and Align step were normalized, filtered, and reviewed with PCA. Unwanted systematic errors due to sample preparation or instrument process can be reduced by setting up the normalization parameters. Certain features (compounds) can be removed by specifying abundance, variability, and frequency of occurrence within the sample groups.

Statistical analysis: differential analysis using Volcano Plot and Hierarchical Clustering

MassHunter Explorer offers a comprehensive set of statistical and differential analysis tools, along with advanced visualization techniques, to support the interpretation of mass spectrometry data. These capabilities support the detection of meaningful patterns, identification of key molecular features, and generation of insights to guide next steps. The platform enables accurate, efficient, and reproducible identification of significant compounds from complex datasets, streamlining data analysis and enhancing confidence in results. PCA plots were generated to evaluate authentic lavender essential oil samples from different batches within the same geographical region, across different geographical regions, and detection of adulteration. Volcano Plot and Hierarchical Clustering analysis were then employed to illustrate differential profiling across four distinct models.

Evaluation 1: two batches of authentic LEO from same origin

Extraction of the features from the same origin of two different Lavender Bulgarian batches (LB1 and LB2) showed same or varied levels for features among the two groups. Example extracted ion chromatograms (EICs) are shown in Figure 3. Then, Figure 4 of the resulting PCA showed tight clustering of the replicates within each batch while also demonstrated clear differences between the two different batches. Despite sharing the same origin, LB1 cluster and the LB2 cluster were clearly separated along PC1, which accounted for most of the variance (84.51%). PC1 (84.51%) plus PC2 (6.46%) explained over 90% of the total variance. It means that the two-dimensional PCA plot captured nearly all meaningful variations in the dataset. This typically indicates a strong, dominant trend (e.g., concentration differences of key compounds), and low noise or minimal irrelevant variation in the dataset. The example EIC data showing that the same compounds were present in both batches but at different levels, which supported that PC1 reflected measurable differences in chemical composition, not compositional class alterations. These differences may reflect batch-to-batch variability, harvest timing, processing conditions, or storage effects, even within a consistent geographical source. To better illustrate the abundance change of the features across the two batches, hierarchical cluster analysis (HCA) was applied to cluster the data sets based on the similarity in feature abundance, as shown in Figure 5. The HCA heatmap showed consistent compound presence across both batches, with variation in color intensity (blue to red), indicating differences in abundance, not identity. The sample clustering map grouped the two batches into separate clusters, but not with extreme discrepancy, suggesting related but quantitatively distinct features. The compound clustering map showed that the same set of compounds was being measured across all samples, reinforcing that the difference was in relative levels, not in compound presence/absence.

Combining HCA, PCA, and EIC compound-level data, it can be confidently concluded that the two batches from the same origin differed only in the levels of certain compounds rather than in overall composition or the presence/absence of unique compounds. These results were likely due to natural batch-to-batch variability, harvest timing, or processing conditions and not due to adulteration or compositional inconsistency.

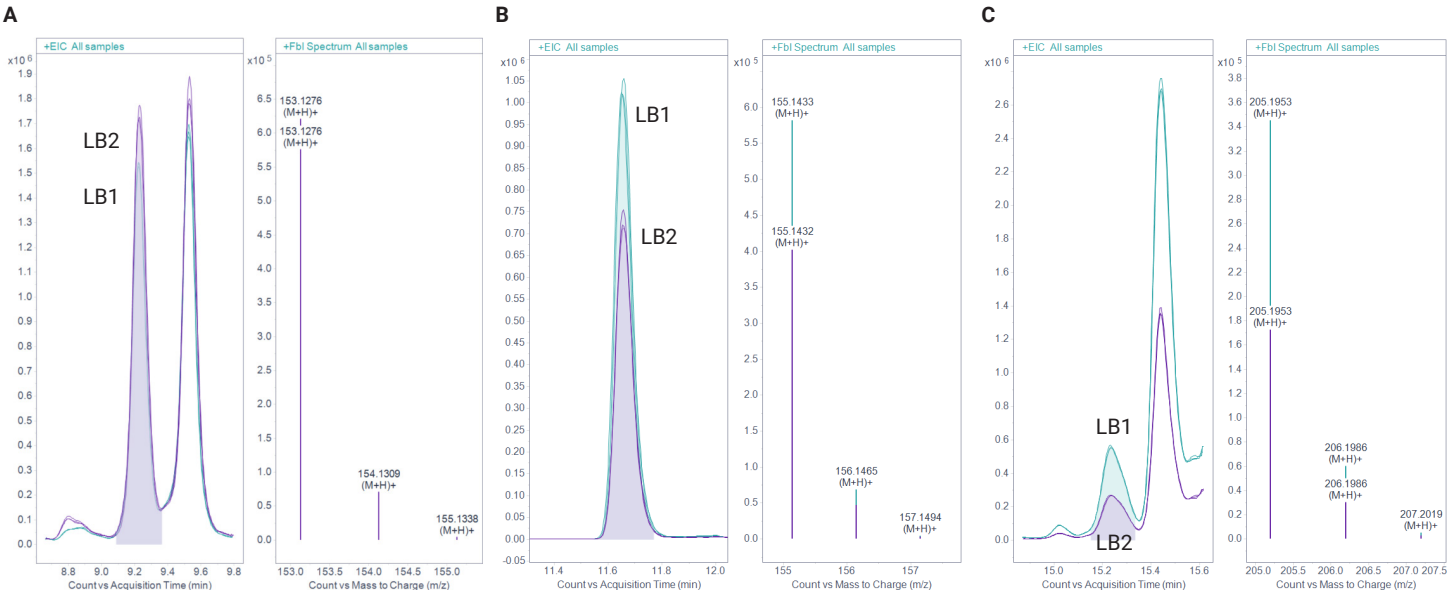


Figure 3. Extracted chromatograms demonstrating the similar abundance for the selected compound (*m/z* of 153.1276) or clear variations in abundance for the selected compounds (*m/z* of 155.1433 and *m/z* of 205.1953) in two batches of same origin authentic LEO from Bulgaria.

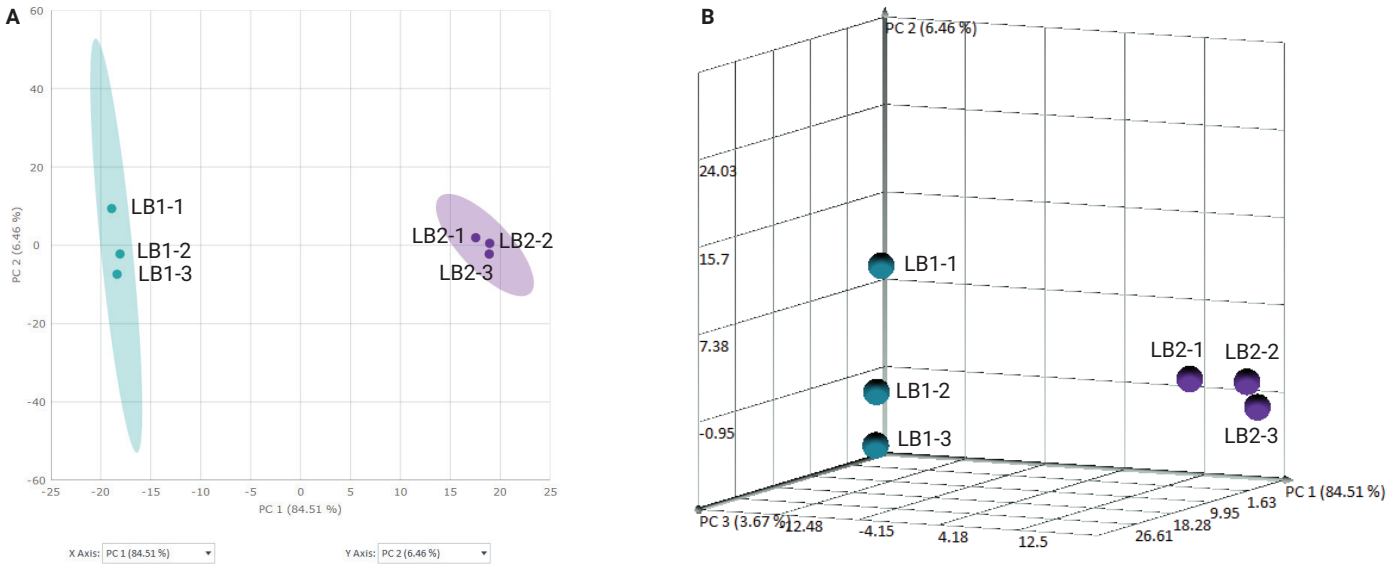


Figure 4. PCA plot of two batches of same origin authentic LEO from Bulgaria. (A) 2-D PCA plot, (B) 3-D PCA plot.

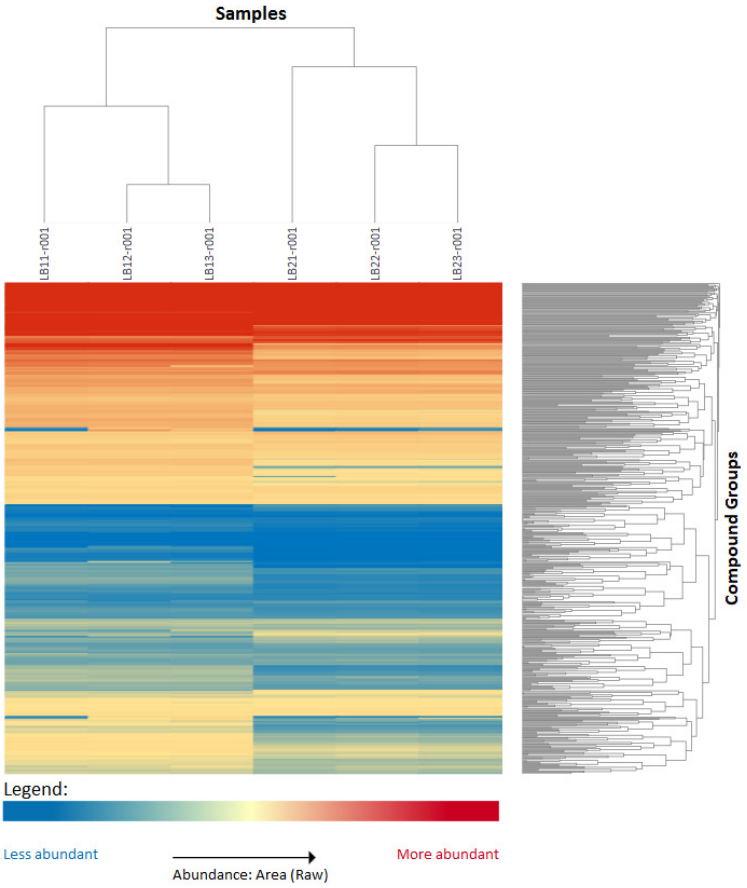


Figure 5. Hierarchical clustering analysis of the two batches of same origin authentic LEO from Bulgaria, showing level differences within two batches.

Evaluation 2: authentic ingredients versus common adulterants

To evaluate the chemical distinction between authentic lavender essential oil samples and known adulterants, a comparative project was constructed with a combination of PCA, volcano plot differential analysis, and examples of features variation analysis. The PCA score plot in Figure 6 revealed a clear separation between the pooled authentic samples (LB1 and LB2) and three common adulterated samples, A (Ho wood crude oil), B (Ho leaf crude oil), and C (raw clove bud oil). The first three principal components, PC1, PC2, and PC3, accounted for 58.10, 21.51, and 9.85% of the total variance, respectively, capturing over 90% of the dataset variability. The LB samples clustered tightly on the left side of the plot, indicating high compositional consistency across batches from the same geographical origin. The adulterants (A, B, C) separated along both PC1 and PC2 on the right side, suggesting significant differences in chemical composition compared to the authentic oils. A volcano plot statistical analysis in Figure 7 was generated to further investigate the chemical differences between authentic lavender essential oils and common adulterants. The plot displayed the distribution of compounds based on their $\log_2$ fold change and corrected p-values, comparing two sample groups: authentic products (two lots of same origin) and adulterants (three common types). Compounds significantly more abundant in authentic samples are highlighted in blue, while those more abundant in adulterants are shown in red. Non-significant compounds are represented in gray. The vertical green line at $\log_2$(fold change) = 0 separates the two groups, and the horizontal green line at $-\log_{10}$(corrected p-value) = 1.25 marked the threshold for statistical significance. The volcano plot analysis revealed distinct chemical markers with statistically significant differences and differentiated authentic lavender essential oil samples from common adulterants. Features more abundant in adulterated samples suggested the presence of foreign components. One example of EIC in Figure 8A with the peak eluting at ~14.0 minutes with $(M+H)^+$ ion at *m/z* 197.1539, was detected exclusively in authentic samples (LB1 and LB2) but absent in all adulterant samples (A, B, and C). This validated its presence and reproducibility in authentic oils. The absence of this feature in adulterants, combined with its statistical significance in the volcano plot, supported its role as an authenticity marker. Figure 8 presented three additional EIC examples: one feature absent in a single adulterant sample, one feature present only in a single adulterant sample, and one feature shared by all samples.

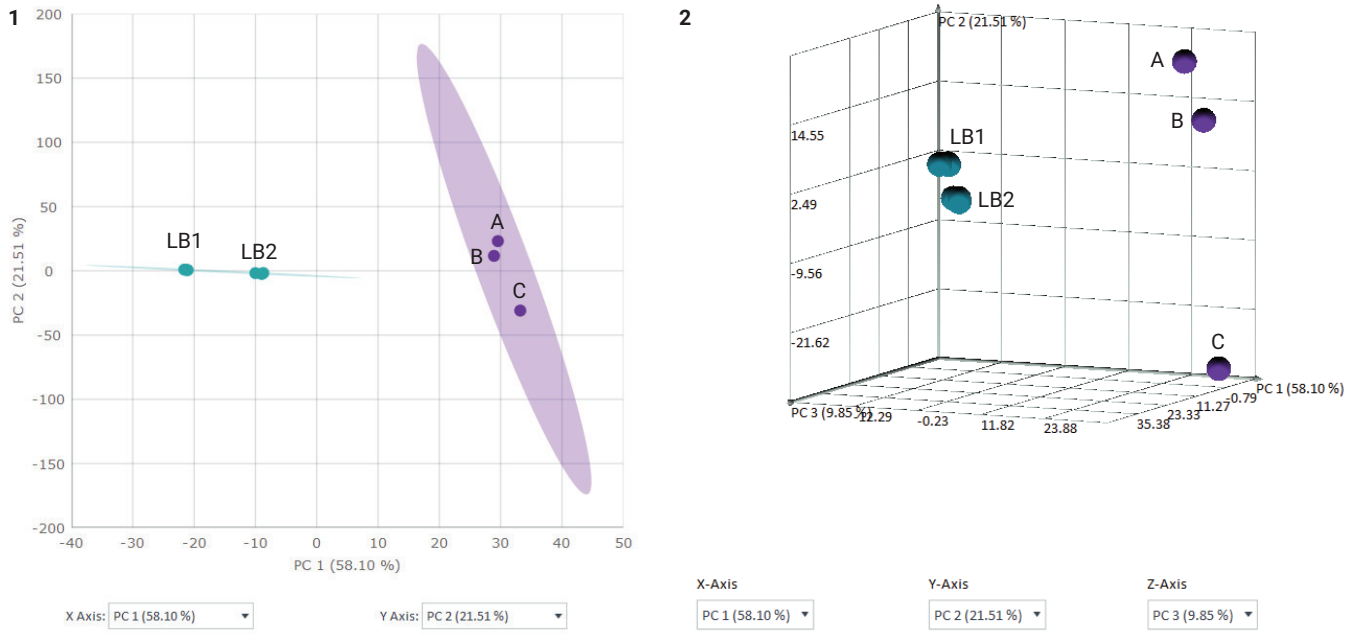


Figure 6. PCA plot of pooled two batches of same origin authentic LEO from Bulgaria versus three common adulterant samples, A (Ho wood crude oil), B (Ho leaf crude oil), C (raw clove bud oil). (1) 2-D PCA plot, (2) 3-D PCA plot.

The integration of PCA, volcano plot, and EIC analysis provided a robust and effective tool for distinguishing authentic lavender essential oils from adulterants. PCA revealed overall compositional differences, volcano plots highlighted statistically significant features, and EIC analysis confirmed the presence of specific markers.

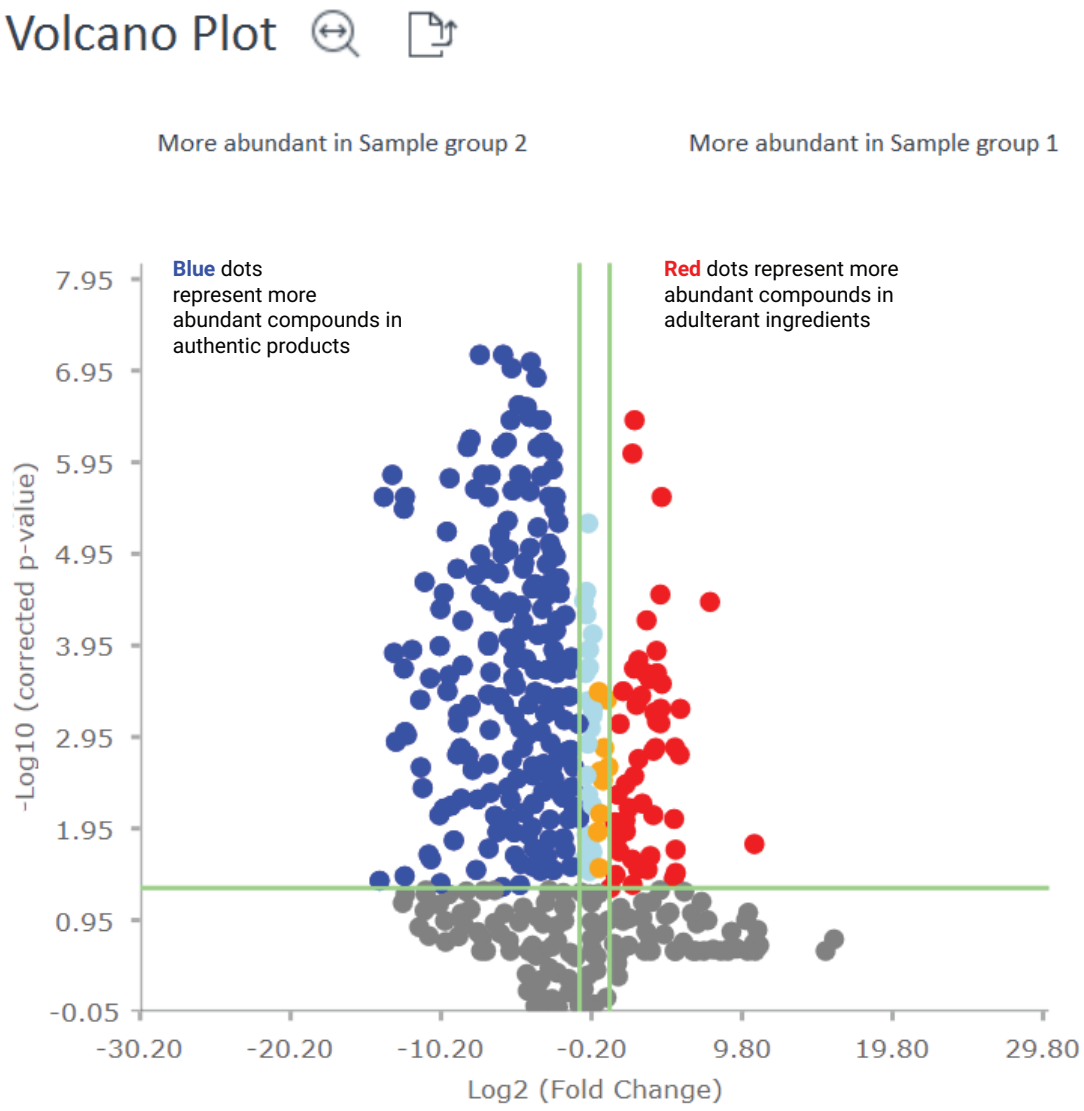
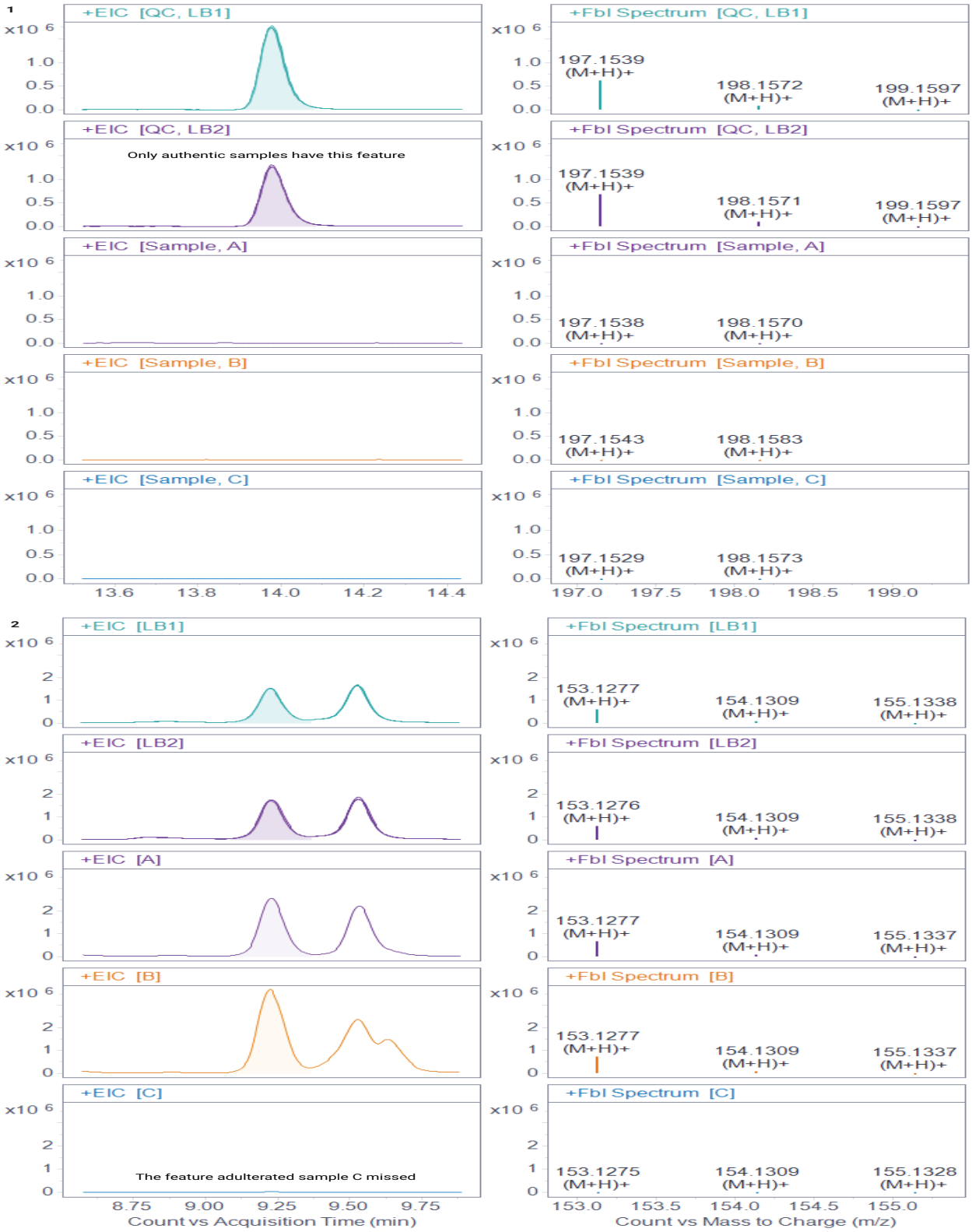


Figure 7. Volcano analysis of authentic LEO from Bulgaria versus three common adulterant samples A, B, and C.


1
+EIC [QC, LB1]
+Fbl Spectrum [QC, LB1]
197.1539 (M+H)+
198.1572 (M+H)+
199.1597 (M+H)+
+EIC [QC, LB2]
Only authentic samples have this feature
+Fbl Spectrum [QC, LB2]
197.1539 (M+H)+
198.1571 (M+H)+
199.1597 (M+H)+
+EIC [Sample, A]
+Fbl Spectrum [Sample, A]
197.1538 (M+H)+
198.1570 (M+H)+
+EIC [Sample, B]
+Fbl Spectrum [Sample, B]
197.1543 (M+H)+
198.1583 (M+H)+
+EIC [Sample, C]
+Fbl Spectrum [Sample, C]
197.1529 (M+H)+
198.1573 (M+H)+
x10 6
13.6 13.8 14.0 14.2 14.4
197.0 197.5 198.0 198.5 199.0
2
+EIC [LB1]
+Fbl Spectrum [LB1]
153.1277 (M+H)+
154.1309 (M+H)+
155.1338 (M+H)+
+EIC [LB2]
+Fbl Spectrum [LB2]
153.1276 (M+H)+
154.1309 (M+H)+
155.1338 (M+H)+
+EIC [A]
+Fbl Spectrum [A]
153.1277 (M+H)+
154.1309 (M+H)+
155.1337 (M+H)+
+EIC [B]
+Fbl Spectrum [B]
153.1277 (M+H)+
154.1309 (M+H)+
155.1337 (M+H)+
+EIC [C]
+Fbl Spectrum [C]
The feature adulterated sample C missed
153.1275 (M+H)+
154.1309 (M+H)+
155.1328 (M+H)+
8.75 9.00 9.25 9.50 9.75
153.0 153.5 154.0 154.5 155.0
Count vs Acquisition Time (min)
Count vs Mass to Charge (m/z)

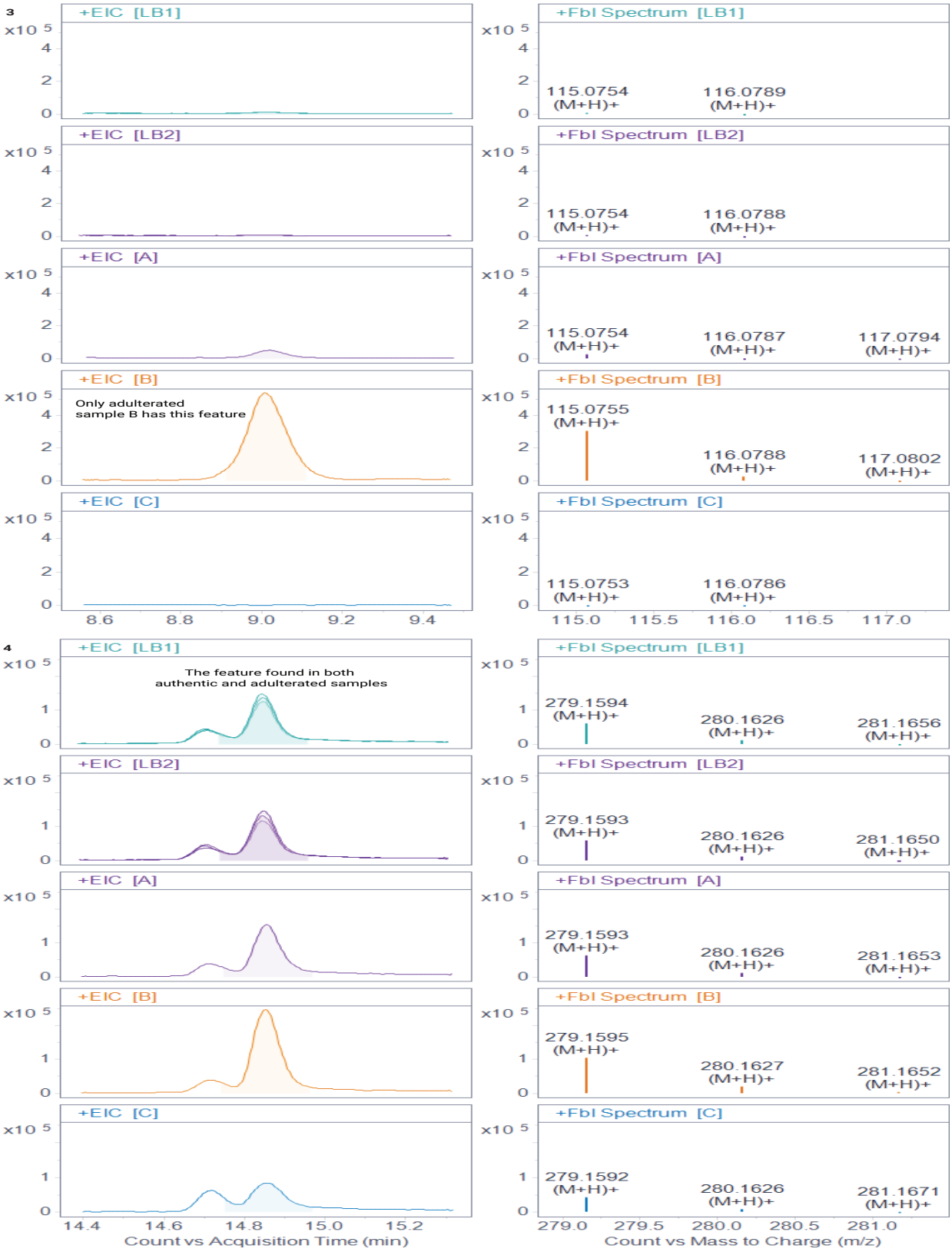


Figure 8. Extracted chromatograms demonstrating the variations (1 through 4) for the selected compounds in authentic Lavender Bulgarian ingredients versus three common adulterant samples A, B, and C.

Evaluation 3: authentic ingredients from different geographical origins

PCA was employed to investigate the chemical variability among authentic samples originating from different geographical sources. The resulting two-dimensional PCA score plot (Figure 9) displayed the distribution of samples along the first two principal components, PC1 and PC2, which accounted for 43.65 and 19.62% of the total variance, respectively. Distinct clustering patterns were observed, with samples from each origin forming well-defined groups enclosed within confidence boundaries. Samples labeled LB (Lavender Bulgarian), LCB (Lavender, China Blue Flower), LCW (Lavender, China White Flower), and LF (Lavender France) exhibited clear intra-group alignment while maintaining inter-group separation. These clustering patterns suggested that geographical origin-specific factors (e.g., soil, climate, processing) played a significant influence on the chemical profile and possibly agronomic factors, even among samples classified as authentic. The separation along PC1 and PC2 highlighted the power of multivariate analysis in distinguishing origin-specific features.

Evaluation 4: authentic lavender oil versus authentic lavender oil blended with 1, 5, and 20% adulterants

Although blending adulterants at levels below 20% is generally not financially favorable for fraudulent practices, adulterated samples were deliberately prepared by fortifying the authentic oil with 1, 5, and 20% of selected fraudulent products, B (Ho Leaf Crude Oil), D (Eucalyptus oil), E (Rosewood oil), F (Potentially adulterated lavender essential oil from different vendor), and G (Adulterated lavender oil by synthetic constituents). This was done to evaluate the sensitivity and effectiveness of PCA in detecting varying degrees of adulteration in lavender essential oil.

The PCA revealed clear and robust separation between the authentic control samples and the adulterated samples across all levels of adulteration. Control samples clustered tightly, indicating high reproducibility in authentic samples. The distinct clustering of different adulteration levels (1, 5, 20%) showed that PCA can capture and differentiate the chemical fingerprint changes caused by adulterants. The inclusion of replicates for each level of adulterant B added statistical robustness, confirming that the observed

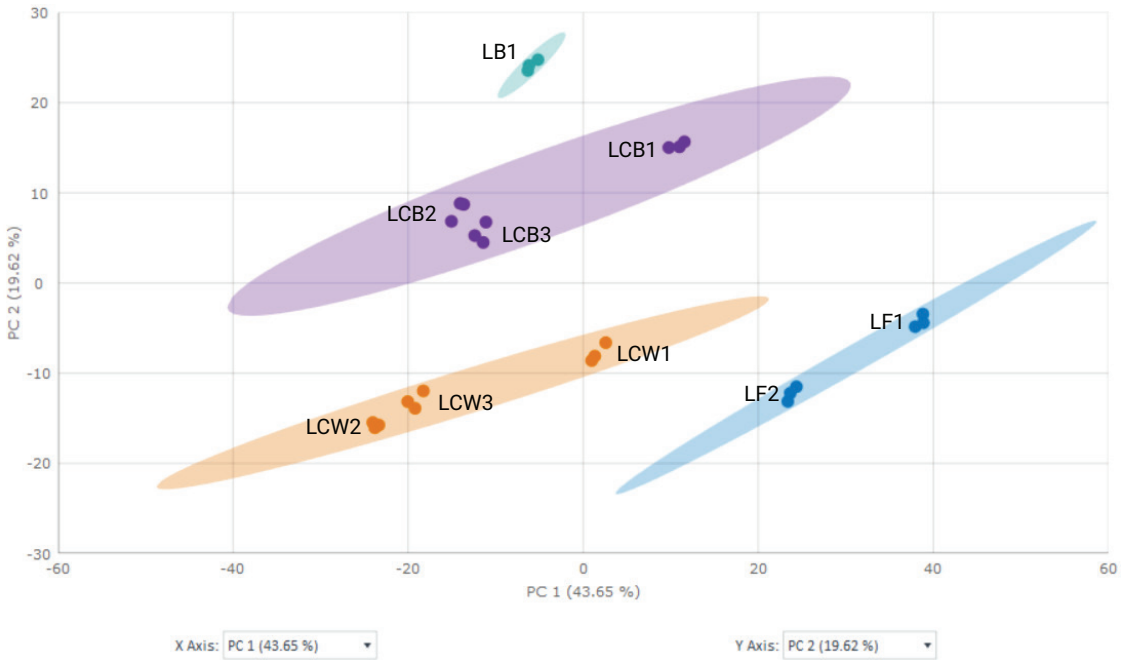


Figure 9. PCA plot of authentic ingredients from different geographical origins.

separation is consistent and reproducible. The PCA plot provided a visual confirmation of adulteration. Different colors and labels (e.g., B-1%, D-1%, etc.) showed how each adulterant affected the product differently, even at the same concentration.

A detailed analysis of adulterant behavior is also included, using sample D (Eucalyptus oil) as an example. The chemical profile of D-1% shared a largely similar chemical profile to the authentic oil at position (–20, 0) on PCA plot. PCA showed minimal, but observable separation, indicating that 1% adulteration is not easily detectable for sample D. D-5% (–10, –10) showed noticeable deviation along both PC 1 and PC 2. This suggested that 5% adulteration impacted the oil's chemical composition enough for PCA to detect a moderate shift in profile. D-20% (20, –10) exhibited a distinct chemical signature, clearly separated from both the control and lower concentration samples. Its position far along PC 1 and PC 2 indicated strong alteration due to high-level adulteration, and PCA effectively captured this change. Another example is sample G (containing synthetic constituents). G-1% located near the origin, around (–10, –20). This cluster is close to the control group, indicating that at 1% synthetic adulteration the chemical profile was altered. PCA still managed to distinguish it, showing its sensitivity to even minimal synthetic content. G-5% located slightly right and downward, around (0, –35). The shift from G-1% to G-5% was more pronounced, suggesting that synthetic components begin to dominate certain chemical signals. This movement away from the control cluster reflects an increasing deviation in the chemical fingerprint. G-20% located farther right and downward around (50, –70). This cluster was well separated from both the control and lower adulteration levels indicating a strong synthetic signature, likely due to dominant artificial compounds that differed significantly from natural lavender oil constituents. The large separation between G-1% and G-20% confirmed that PCA can effectively track and quantify the progression of adulteration. These distances can be used to set thresholds for detection or to train classification models for quality control.

PCA is further proved to be a powerful tool for detecting and quantifying adulteration in essential oils. Its ability to distinguish even 1% synthetic adulteration makes it ideal for quality control and authentication in the fragrance and natural product industries.

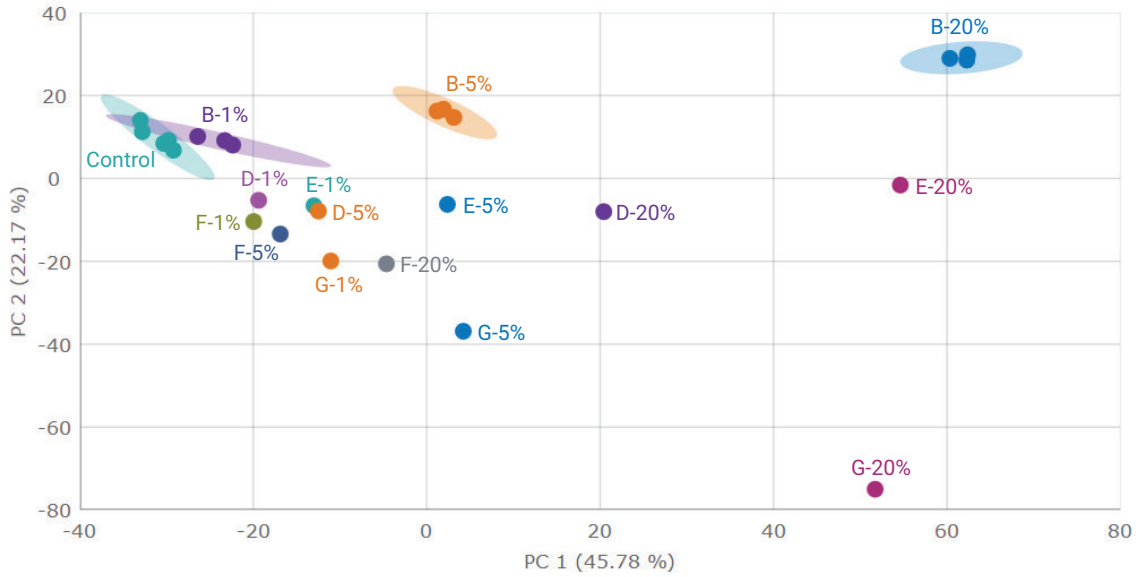


Figure 10. PCA plot of authentic lavender oil samples blended with 1, 5, and 20% adulterants with high sensitivity and effectiveness of chemometric discrimination.

Unknown identification

Agilent personal database search: Following PCA and statistical analysis, compound identification was performed on selected individual compound. The neutral mass calculated from ion evidence, charge states, and adduct annotations, was matched against theoretical neutral masses in a compound database. Additionally, the observed isotope pattern was compared with the expected distribution based on the compound elemental formula. When available, the measured retention time (RT) is also matched with the expected RT stored in the database to enhance identification confidence.

In Figure 11, an unknown feature with *m/z* of 147.0443 was consistently detected across six replicate samples, demonstrating excellent reproducibility and supporting the validity of this feature for further investigation. Figure 12 presented the result of compound identification for an unknown feature, which was putatively assigned as coumarin based on MS and MS/MS spectrum matching and isotope pattern analysis with database entries. Multiple statistical metrics, database match score, mass accuracy, isotope pattern, peak intensity, and isotope distribution were presented in Table 3. Excellent database match score, mass accuracy, and isotopic fidelity strongly validated the identification of the unknown compound as coumarin.

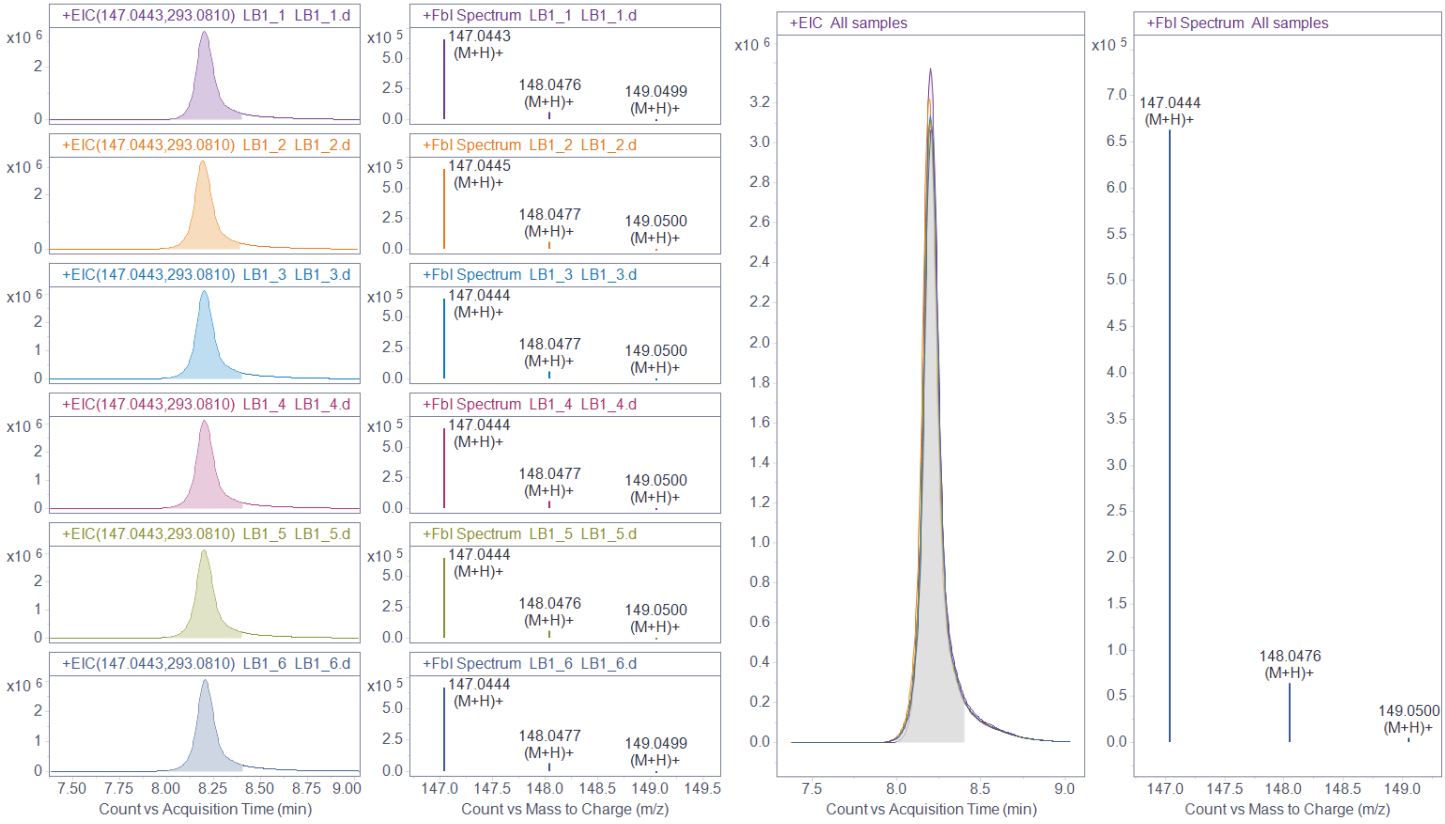


Figure 11. Reproducibility of an unknown feature at *m/z* 147.0443 across six replicates.

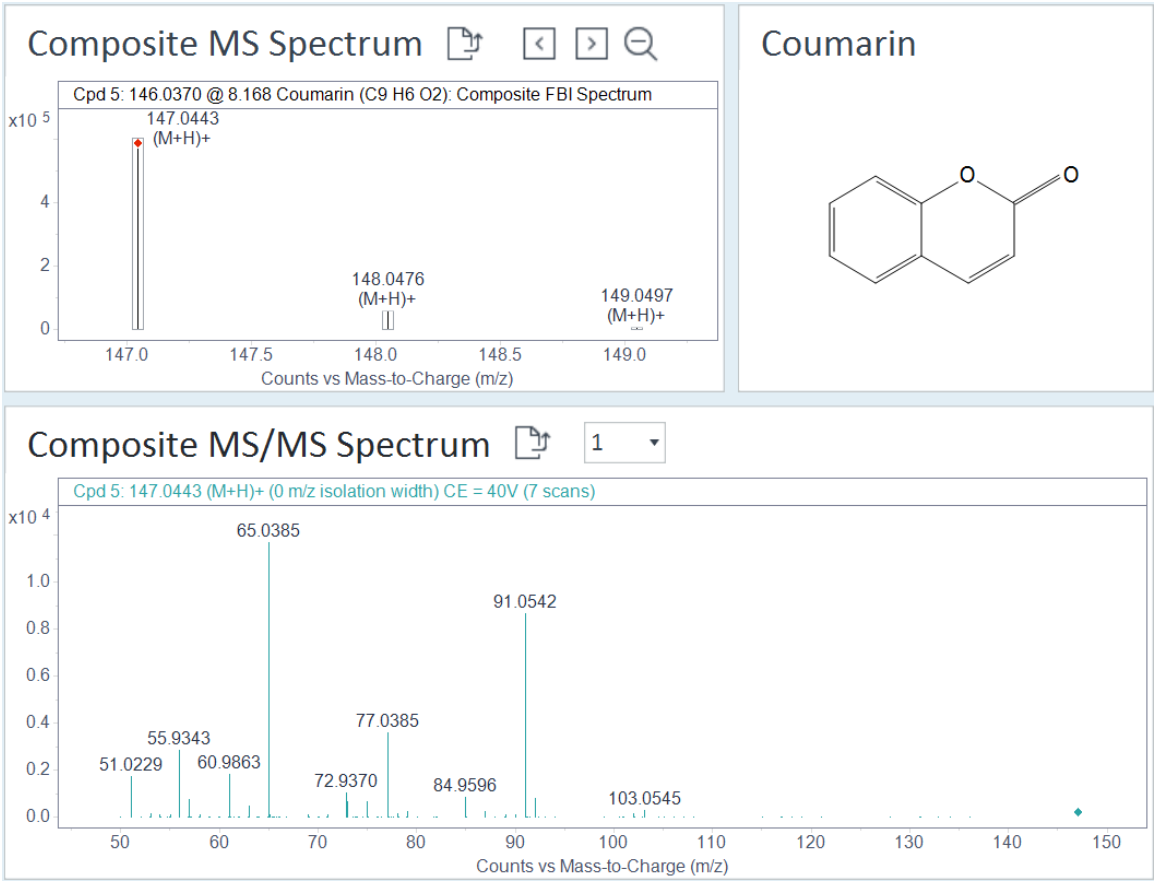


Figure 12. Putative identification of an unknown compound assigned as coumarin.

Table 3. Statistical summary of putative identification: coumarin.

Name	Formula	Diff (ppm)	Diff (mDa)	Score (DB)	DB Source
Coumarin	$C_9H_6O_2$	1.70	0.25	99.61	LEO01.cdb
Species	*m/z*	Score (mass)	Score (isotope abundance)	Score (isotope spacing)	–
$(M+H)^+$	147.0443	99.23%	99.99%	99.89%	–
m/z	***m/z* (calc)**	**Diff (ppm)**	**Diff (mDa)**	**Height**	**Height%**
147.0443	147.0441	1.76	0.26	605210.31	100.00
148.0476	148.0474	1.05	0.15	59696.56	9.86
149.0497	149.0496	0.52	0.08	4796.07	0.79

NIST Library Search: Using the MassHunter Explorer direct link to the NIST 2023 and MS Search 2.4 , the unknown compound with *m/z* 147.0443 $(M+H)^+$, was searched against the spectral database. In Figure 13 the search returned coumarin as the top hit, supported by High Dot Product and Reverse Dot Product scores, indicating strong spectral similarity between the unknown and the library entry.

The sample MS/MS spectrum showed excellent alignment with the reference spectrum for coumarin, including key fragment ions such as *m/z* 65.0385 and 91.0542. This NIST spectral match confirmed the putative identification of the unknown compound as coumarin, with high confidence based on both mass accuracy and fragmentation pattern similarity.

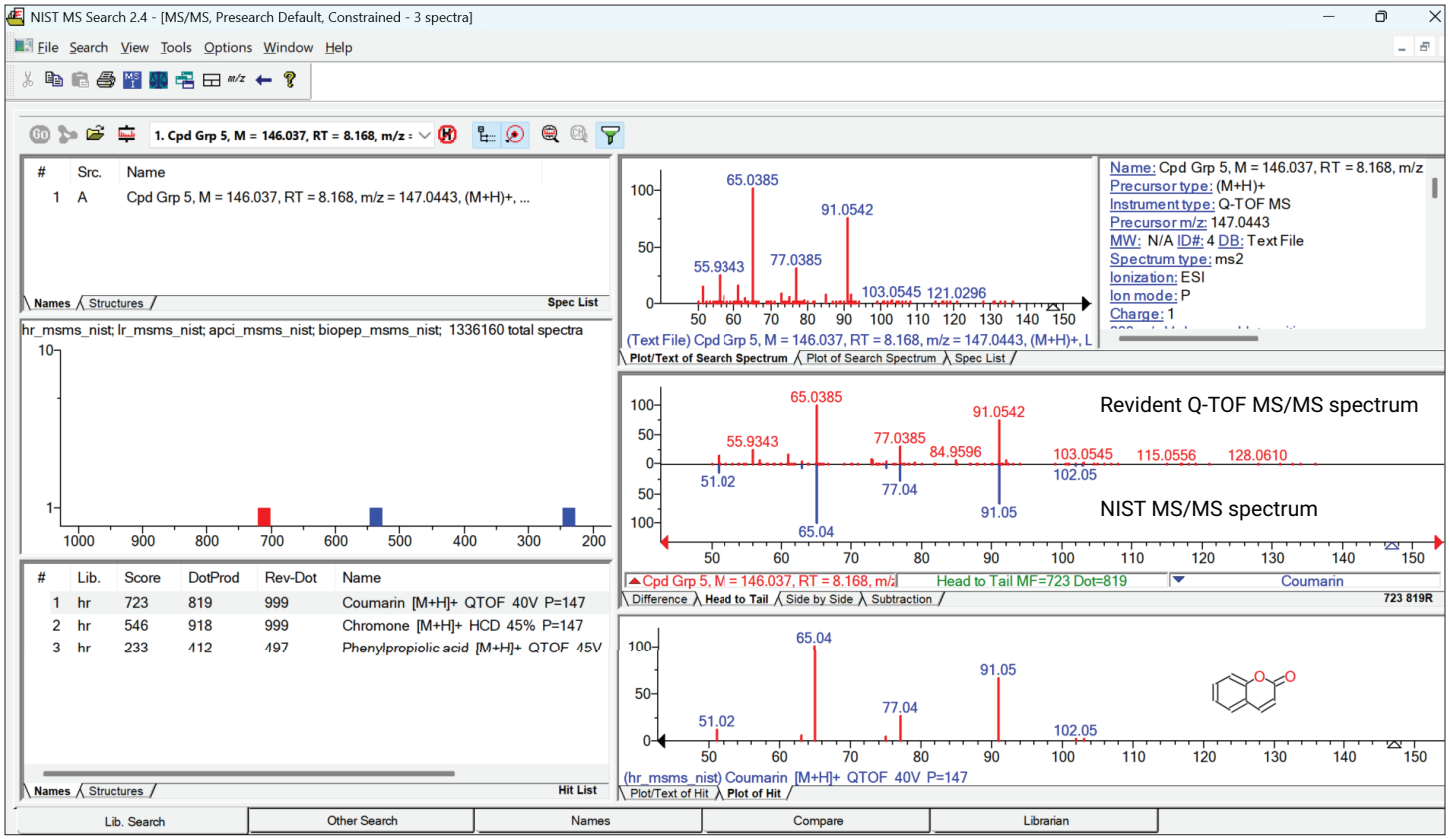


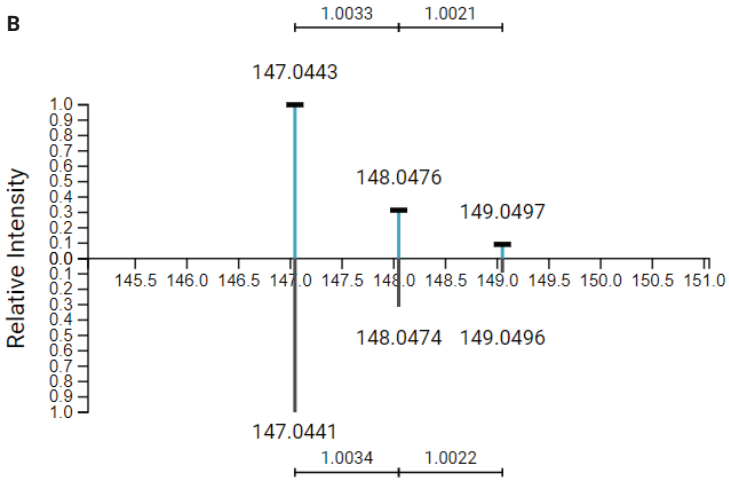
Figure 13. NIST Library Search confirmation for coumarin.

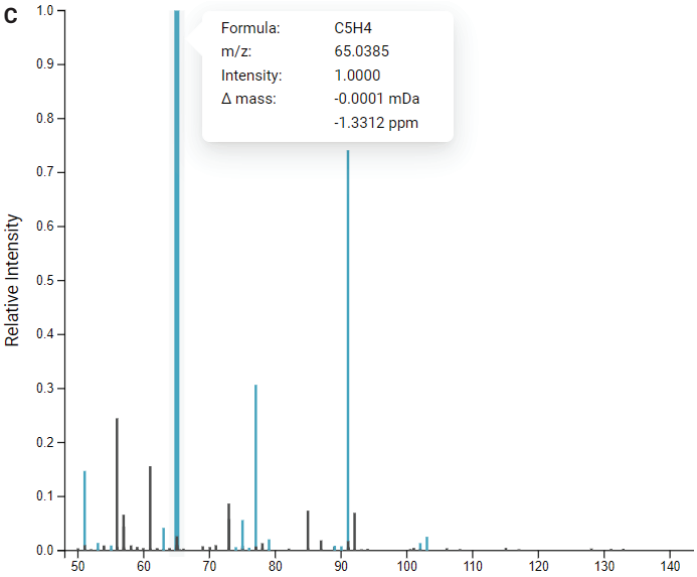
SIRIUS[3,4]: While extensive and curated libraries are available, often there are cases where an analyte of interest is not present in the libraries, and identification may be limited to chemical formula. SIRIUS, an AI supported software, aids identification in these cases. SIRIUS robust machine learning framework predicts molecular formulas, chemical classes, and candidate structures from MS and MS/MS data. The process integrated isotope pattern analysis, fragmentation tree construction, and database matching to achieve high-confidence compound identification and structure elucidation. Explorer provides direct link and complimentary access to SIRIUS, even for commercial users.

An AutoMSMS file containing MS and MS/MS spectra of an unknown compound was loaded into SIRIUS. As shown in Figure 14, SIRIUS computed theoretical isotope distributions for candidate molecular formulas and compared them to the experimental spectrum. The top-ranked formula, $C_9H_6O_2$, achieved a normalized SIRIUS score of 100%, supported by high isotope and tree scores. Using the MS/MS spectrum, SIRIUS generated a fragmentation tree that models the sequential breakdown of the precursor ion. This tree explained the observed fragment peaks and intensities, contributing to the overall confidence in the molecular formula.

A

Rank	Molecular Formula	Adduct	Zodiac Score	Sirius Score (normalized)	Isotope Score	Tree Score	Explained Peaks	Total Explained Intensity	Median Mass Error (ppm)	Median Mass Error (mDa)
1	C9H6O2	[M + H]+	NaN	100.000%	3.189	73.793	15/15	70.071%	-0.864	-0.067
2	C7H4N3O	[M + H]+	NaN	0.000%	1.491	25.541	9/15	46.233%	-1.108	-0.067





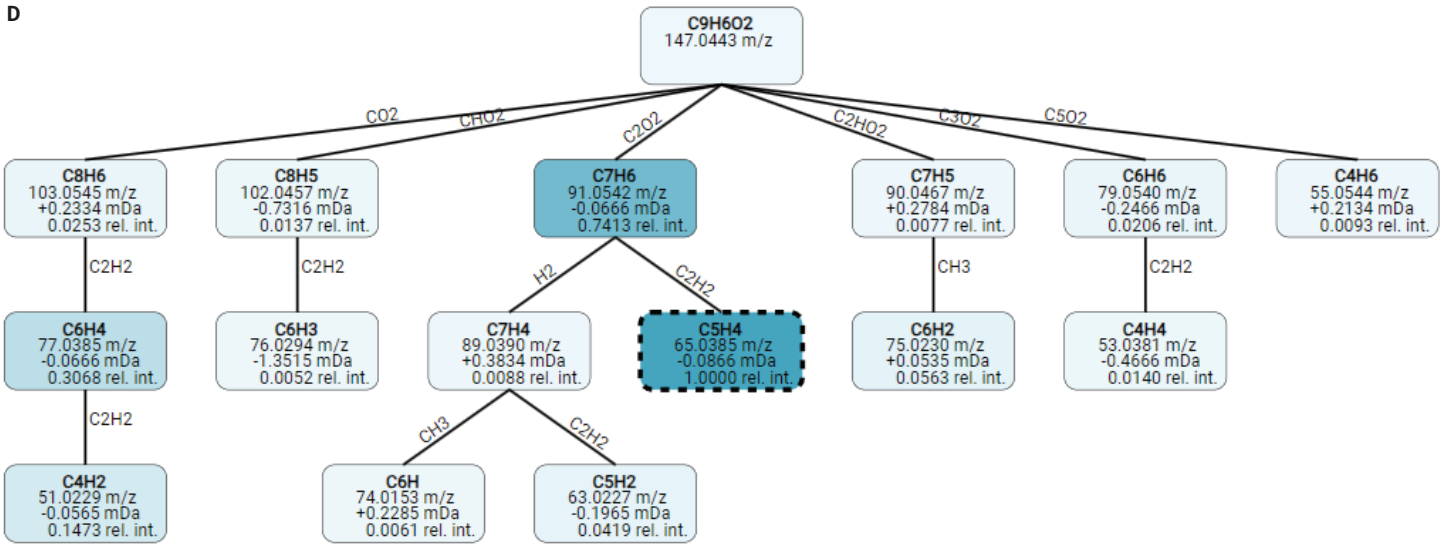


Figure 14. (A) Predicated formula and molecular formula ranking based on (B) MS isotope pattern, (C) MS/MS spectrum and (D) fragment tree analysis.

In Figure 15, CANOPUS, a module within SIRIUS, predicted the compound chemical class based on fragmentation features. The compound was classified as a coumarin derivative, belonging to the broader categories of phenylpropanoids and polyketides. The predicted formula was searched against structural databases. Three top candidate structures were retrieved, with similarity scores ranging from −16.975 to −49.650. These structures provided identities for the unknown compound and guidance for further validation. The integrative approach led to a confident molecular formula ($C_9H_6O_2$) and classified the compound within a relevant chemical family. The fragmentation tree explained a significant portion of the MS/MS spectrum, and the structure matching provided solid compound identification. By combining spectral data with computational prediction for unknown compound identification, SIRIUS has demonstrated its irreplaceable capability in mass spectrometry-based compound discovery.

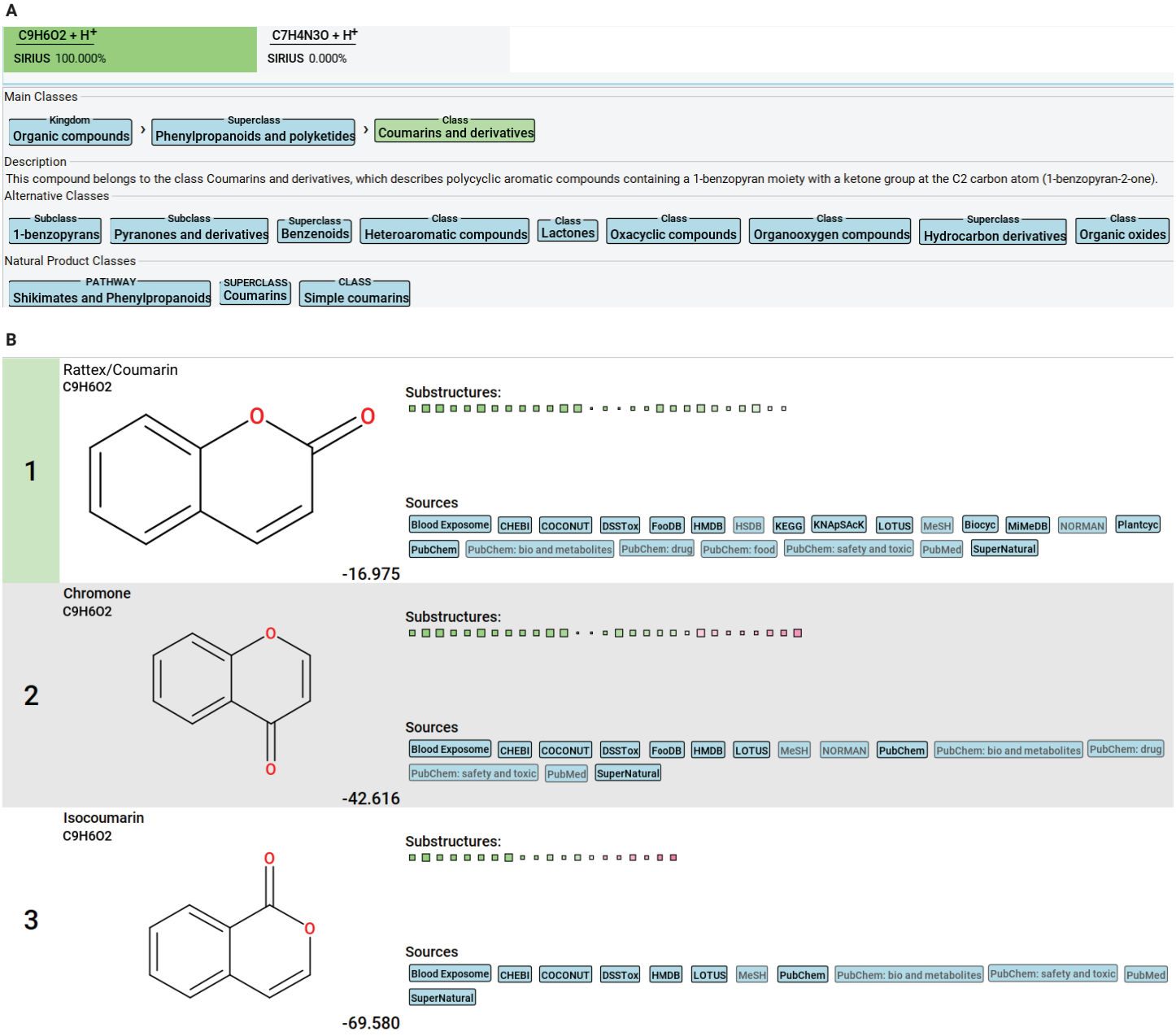


Figure 15. (A) Predicted chemical classes and (B) structure matched for the top-ranked formula.

Conclusion

In summary, the accurate mass LC/Q-TOF technique, combined with the advanced MassHunter Explorer differential analysis workflow, successfully analyzed and interpreted lavender essential oil profiling results, confirming adulteration, and identifying unknown compounds. This workflow provided a reliable integrative quality control and authenticity verification method for relevant plant extracts.

References

1. Aprotosoaie, A. C.; Gille, E.; Trifan, A.; Luca, V. S.; Miron, A. Essential Oils of Lavandula Genus: a Systematic Review of Their Chemistry. *Phytochem. Rev.* **2017**, *16*, 761–799.
2. Agilent MassHunter Explorer Overview, Agilent Technologies, Inc, **2023**.
3. Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Pieter Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS4: a Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16*, 299–302.
4. Kim, H. W; Wang, M. X.; Leber, C. A.; Nothias, L. F.; Reher, R.; Kang, K. B.; van der Hooft, J' J. J.; Dorrestein, P. C.; Gerwick, W. H.; Cottrell, G. W. NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. *Journal of Natural Products* **2021**, *84*, 2795–2807.

www.agilent.com

DE-011126

Printed in the USA, January 16, 2026
5994-8855EN

Agilent
Trusted Answers