Received: 18 January 2023 | Revised: 17 April 2023 | Accepted: 12 June 2023

DOI: 10.1002/pca.3258

RESEARCH ARTICLE

Phytochemical Analysis WILEY

A comparison of high- and low-resolution gas chromatography–mass spectrometry for herbal product classification: A case study with *Ocimum* essential oils

Evelyn J. Abraham[1] | E. Diane Wallace[2] | Joshua J. Kellogg[1,3]

[1]Intercollege Graduate Degree Program in Plant Biology, Pennsylvania State University, University Park, Pennsylvania, USA

[2]Mass Spectrometry Lab, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[3]Department of Veterinary and Biomedical Sciences, Pennsylvania State University, University Park, Pennsylvania, USA

Correspondence
Joshua J. Kellogg, Department of Veterinary and Biomedical Sciences, Pennsylvania State University, University Park, PA 16802, USA.
Email: jjk6146@psu.edu

Funding information
Garden Club of America, Grant/Award Number: Anne S. Chatham Fellowship in Medicinal Botany; National Institute of Food and Agriculture, Grant/Award Number: PEN04772

Abstract

Introduction: Selection of marker compounds for targeted chemical analysis is complicated when considering varying instrumentation and closely related plant species. High-resolution gas chromatography–mass spectrometry (GC–MS), via orbitrap detection, has yet to be evaluated for improved marker compound selection.

Objective: This study directly compares high- and low-resolution GC–MS for botanical maker compound selection using *Ocimum tenuiflorum* L. (OT) and *Ocimum gratissimum* L. (OG) for botanical ingredient authentication.

Methods: The essential oils of OT and OG were collected via hydrodistillation before untargeted chemical analysis with gas chromatography coupled to single-quadrupole (GC-SQ) and orbitrap (GC-Orbitrap) detectors. The Global Natural Products Social Molecular Networking (GNPS) software was used for compound annotation, and a manual search was used to find the 41 most common *Ocimum* essential oil metabolites.

Results: The GC-Orbitrap resulted in 1.7-fold more metabolite detection and increased dynamic range compared to the GC-SQ. Spectral matching and manual searching were improved with GC-Orbitrap data. Each instrument had differing known compound concentrations; however, there was an overlap of six compounds with higher abundance in OG than OT and three compounds with a higher abundance in OT than OG, suggesting consistent detection of the most variable compounds. An unsupervised principal component analysis (PCA) could not discern the two species with either dataset.

Conclusion: GC-Orbitrap instrumentation improves compound detection, dynamic range, and feature annotation in essential oil analysis. However, considering both high- and low-resolution data may improve reliable marker compound selection, as GC-Orbitrap analysis alone did not improve unsupervised separation of two *Ocimum* species compared to GC-SQ data.

KEYWORDS
botanical identification, essential oil, GC–MS, mass spectrometry, metabolomics

1 | INTRODUCTION

Herbal products are increasing in popularity among American consumers; in 2017, 35% of Americans reported using herbal products, and in 2019 there was an 8.9% increase in use.[1] Herbal and botanical supplements are classified as food products under the Dietary Supplement Health and Education Act and thus fall under the jurisdiction of the Food and Drug Administration. The Food, Drug, and Cosmetic Act (FD&C) states that products must be evaluated for their identity, purity, and efficacy, but there is no set testing requirement.[2] The lack of federal and global regulations has resulted in the adulteration of up to 27% of herbal products.[3] While traditional approaches to botanical authentication, including morphology and genetics, help confirm product identity, complications arise when dealing with processed, dried, powdered, or extracted products.[4]

Instead of inconsistent traditional authentication, targeted analytical approaches, such as liquid chromatography coupled to mass spectrometry (LC–MS), are the predominant methods for confirming botanical identity. Targeted chemical analyses focus on one or a small group of compounds specific to the evaluated botanical.[5] Thus, the compounds in the product/species must be known prior to evaluation, and analytical standards must be readily available, complicating the selection of optimal analytes.[6] Additionally, plant metabolite concentrations differ based on geographic origin and environmental conditions,[7] so reliable, consistent targeted analyte selection is tedious. Current investigations into the potential of untargeted, full-profile analysis for authentication schemes to bypass targeted analyte selection rarely appear in industry settings.[5] However, untargeted approaches paired with multivariate statistical analysis provide the opportunity to distinguish between different species based on their chemical profile while identifying the compounds most responsible for the variation. Therefore, untargeted metabolomics studies can simultaneously perform authentication and compound selection for future targeted analysis of dried and processed botanical materials.

Gas chromatography coupled to mass spectrometry (GC–MS) commonly complements LC–MS in herbal product investigations and has been used for product confirmation and analysis for over 50 years. Electron impact (EI) ionisation, a hard ionisation technique, generates reproducible fragments and allows for the creation of multiple databases, including the National Institute of Standards and Technology (NIST) database[8] and has the advantage of possessing fast duty cycles, yielding a rapid, sensitive platform designed for targeted, quantitative analysis.[9] In recent years, GC has been coupled with high-resolution mass analysers, such as a time-of-flight (ToF) or Orbitrap (GC-Orbitrap). High-resolution mass spectrometry (HRMS) provides the ability to further resolve compounds by providing accurate mass measurements (< 5 ppm), which allows for putative identification and chemical formula determination.[10] However, relatively few reports exist investigating the potential of GC–HRMS for improved approaches to authentication and marker compound discovery compared to quadrupole instruments.[11–13] Lacalle-Bergeron *et al.* argue that GC–HRMS might not be necessary and adds excess processing time and expenses since fragmentation libraries and compound resolution is already acceptable with quadrupole instruments.[11] Despite the efforts to introduce high-resolution GC data to herbal authentication, there is no indication that this technology provides authentication benefits over single-quadrupole and other low-resolution instrumentations.

The present study directly compares the ability of a GC-Orbitrap and GC-single quadrupole (GC-SQ) to distinguish two *Ocimum* (basil) species based on their essential oil profiles. *Ocimum gratissimum* L. and *Ocimum tenuiflorum* L. are two species of Holy Basil (Tulsi) with very similar chemical constituents but dynamic differences in chemical concentrations and biological activity.[14] Within these basil species, multiple cultivars exist with varied chemical profiles, complicating species discrimination via traditional approaches, making it an ideal herbal material to demonstrate the potential of GC-Orbitrap instrumentation for essential oil analysis and species classification using real-world market products.

2 | MATERIALS AND METHODS

2.1 | Reagents and chemicals

LC–MS grade ethyl acetate (purity ≥ 99.5%; VWR International, Radnor, PA, USA), methanol (purity ≥ 99.99%; Fisher Scientific International, Inc., Hampton, NH, USA), and *n*-hexane (purity ≥ 99.99%; Fisher Scientific International, Inc.) were used for this study.

2.2 | Basil sampling

To evaluate the potential of GC–HRMS analysis for classification of realistic market samples, we ordered *Ocimum* products from four separate, reputable botanical suppliers, two located in the United States and two located in India. Two species of *Ocimum* were evaluated, *O. gratissimum* (four samples) and *O. tenuiflorum* (five samples). Table 1 provides sample information, including species, supplier (coded for anonymity), and country of origin. Samples were dried prior to shipment and stored in air-tight containers at room temperature in the dark until processing. Certificates of Analysis were obtained from three suppliers for the specific lots of dried products, and documentation verifying plant identity was obtained from the fourth.

2.3 | Hydrodistillation of essential oil

Essential oils were collected from each sample in triplicate using a modified hydrodistillation procedure.[15] Briefly, 30 g of leaf tissue was transferred to a two-necked round-bottom flask with 450 mL distilled water (1:15 tissue/water) with one neck plugged and the other connected to a condenser. Distillate was collected for 1.5 h (until no more oil was collected), and oils were separated using ethyl acetate. Solvent was removed with a Buchi Rotavapor R-300 with a low vacuum and no additional heat to prevent loss of volatile compounds. Dried

TABLE 1 *Ocimum* sample information.

Sample ID	Species	Supplier	Country of origin	Essential oil yield (%)
B1	*Ocimum tenuiflorum*	A	India	0.31 ± 0.15
B2	*Ocimum tenuiflorum*	B	USA	0.45 ± 0.15
B3	*Ocimum tenuiflorum*	C	India	0.17 ± 0.12
B4	*Ocimum tenuiflorum*	C	India	0.20 ± 0.11
B5	*Ocimum tenuiflorum*	D	USA	0.29 ± 0.20
B6	*Ocimum gratissimum*	A	India	0.32 ± 0.15
B7	*Ocimum gratissimum*	B	USA	0.55 ± 0.26
B8	*Ocimum gratissimum*	C	India	0.19 ± 0.17
B9	*Ocimum gratissimum*	D	USA	0.42 ± 0.15

samples were stored at 4°C in the dark until GC–MS analysis. Essential oil yield (from the dried plant material) was calculated with the following formula:

% yield = (Weight extracted oil / Dry weight of sample) × 100

Yield information for each sample is provided in Table 1.

2.4 | Gas chromatography–mass spectrometry

2.4.1 | GC-Orbitrap–HRMS

All triplicate extracts for each sample were analysed. GC separation was performed using a Thermo TraceGOLD TG-5SilMS column. Samples were injected splitless with a volume of 1 µL. GC oven was initialised at 40°C. At 1 min, the GC oven started ramping at a rate of 25°C/min until 300°C (11.5 min). The temperature was held at 300°C for 2.5 min, for a total of 14 min. HRMS acquisition was performed on a ThermoFisher (Waltham, MA, USA) GC Exactive system at 1 mg/mL in methanol via an EI ionisation source. An electron energy of 70 eV was utilised. Samples were acquired in positive mode with a scan range of 50–600 *m/z* and a resolution of 60,000. The ion source was set to 280°C, and the MS transfer line was set to 250°C. The automatic gain control (AGC) target was set to 1×10^6. Raw spectral data was deposited in the MASSive database (ID: MSV000091070, https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=870fe9654fbc43cd8223e1b68375a3bc).

2.4.2 | GC-SQ–MS

The triplicate extracts for each sample were analysed. Samples were analysed on an Agilent 7890A gas chromatograph coupled to an Agilent 5975C inert XL EI/CI MSD (Agilent Technologies, Santa Clara, CA, USA). Separation was performed using a Rxi-5 ms, 30 m, 0.25 mm inner diameter, 0.25 µm df column (Restek, PA, USA). The injection volume was 1 µL. The inlet temperature was set to 250°C with a split ratio of 20:1 and a helium carrier flow rate of 1.0 mL/min. At 1 min, the GC oven started ramping at a rate of 25°C/min until 300°C (11.5 min). The temperature was held at 300°C for 2.5 min, for a total of 14 min. The mass spectrometer detector was operated in positive mode with a full scan range of 50–600 *m/z* using electron ionisation. The transfer line temperature was set to 250°C and the ion source temperature at 230°C. Raw spectral data was deposited in the MASSive database, as earlier.

2.5 | Data preprocessing in MZmine2

Data preprocessing was performed in the open-access MZmine2 software.[16] Steps included peak detection, chromatogram deconvolution, and decomposition, as well as isotope and duplicate peak filtering, followed by sample peak alignment and filtering. Although the same workflow and steps were followed for both untargeted datasets, different thresholds were used to maximise features and minimise noise in each set. Final feature lists were gap-filled prior to export to a .csv file for data analysis.

Specifically, the noise levels for peak detection were 5.0E5 and 2.0E2 for the GC-Orbitrap and GC-SQ, respectively. Peak filtering included removing features not present in all three triplicates of a given sample. Additionally, all features not present in at least one sample at a concentration five-fold higher than the blank were removed from the dataset, and the peak area of the triplicates was averaged. See Supporting Information Table S1 for data preprocessing parameters.

2.6 | Library search using the Global Natural Products Social Molecular Networking (GNPS)

Raw data files were converted to mzXML format using MSconvert[17] and used to generate a molecular network using the "Library Search/Molecular networking GC workflow" from GNPS.[18] Table S2 provides the specific search parameters used for library matching – the same parameters were used for both datasets.

2.7 | Evaluation of known essential oil compound ratios

A list of 50 compounds commonly found in *O. gratissimum* and *O. tenuiflorum* was compiled with a limited metasearch using

"*Ocimum*, *Ocimum gratissimum*, *Ocimum tenuiflorum*, essential oil, chemistry, metabolites, small molecules, Tulsi, Holy Basil" and other similar search keywords in the search engines Google Scholar and PubMed. Papers were chosen that had compounds identified in *O. gratissimum*, *O. tenuiflorum*, or both, and the compounds listed were compiled from eight papers.[14,19–25] Initially, a list of 50 compounds was generated, but it was narrowed to 41 due to nomenclature overlap and availability of MS^2 library data. The library *m/z*, fragmentation patterns, and structural information were used to manually search the GC-SQ and GC-Orbitrap raw data for potential matches. The peak area for the most likely match was averaged in *O. gratissimum* and *O. tenuiflorum* and used to create an abundance ratio between the species. The *P*-values (students' *t*-test, $P = 0.05$) were calculated in Excel. Venn diagrams were created to compare the compounds with the highest abundance in either species using the data from both instruments. Raw peak area information for each compound in each dataset is located in Tables S3–S6.

2.8 | Unsupervised data analysis – principal component analysis (PCA)

Principal component analysis (PCA) was performed in R 4.1.1. Data were log-transformed, Pareto-scaled, and centred before PCA with the MetabolAnalyze (version 1.3.1) and stats (version 3.6.2) packages. Principal components (PCs) were calculated using the autoplot function within the stats package. Hotelling's 95% confidence intervals were used to search for outliers, and none were detected in either PCA model. Data was coloured and plotted based on sample species using the ggplot2 package. Key variables for separation along PC1 and PC2, which contained the highest variance for both PCA models, were identified via the associated loadings plots and the fviz_conrtib tool within the factoextra package. The *m/z* and retention time of the key compounds identified in the loading plots were used for feature annotation using the GNPS spectral library matching results.

TABLE 2 Peak information from gas chromatography-Orbitrap (GC-Orbitrap) and gas chromatography-single-quadrupole (GC-SQ) detectors.

Instrument	Number peaks in raw chromatogram	Number of peaks post processing and filtering
GC-Orbitrap	800–900	737
GC-SQ	150–250	417

Note: Processing and filtering information can be found in Supporting Information Table S1.

3 | RESULTS AND DISCUSSION

3.1 | GC-Orbitrap data yields increased metabolite coverage

To our knowledge, this is the first study evaluating the use of GC-Orbitrap instrumentation for plant essential oil analysis; we compared the detection of *Ocimum* essential oil metabolites using GC-Orbitrap and GC-SQ instruments. Hydrodistilled essential oils of nine *Ocimum* samples belonging to two species, *O. gratissimum* and *O. tenuiflorum*, were analysed at the same concentration using the same chromatography method; the GC-Orbitrap detected almost four times more metabolites in the *Ocimum* essential oil compared to the GC-SQ (Table 2). Specifically, there were 800–900 peaks in the resolved chromatograms of the combined *O. gratissimum* and *O. tenuiflorum* samples from the GC-Orbitrap versus 150–250 peaks in the chromatograms of the GC-SQ datasets (Table 2).

Comparing each instrument's base peak chromatograms (BPCs) confirms differences in peak coverage (Figure 1). The relative intensity of each peak has been scaled based on the highest intensity peak in each data set so that the most intense peak has a relative intensity of 100%. In the blank (Figure 1A), which was in the same queue position between the two instruments, there is similar peak detection resulting from the GC-Orbitrap and GC-SQ instruments. However, since signal intensity is higher in the GC-Orbitrap data due to the trapping functionality, there is also increased high-intensity peak detection

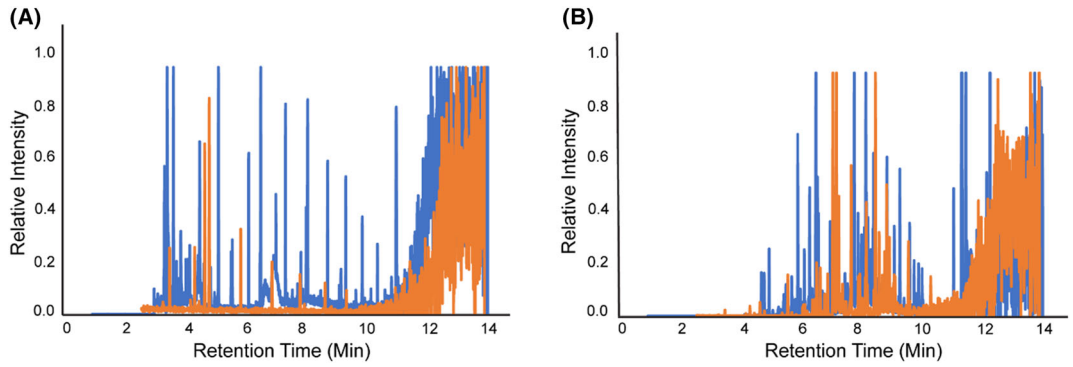


FIGURE 1 Comparison of base peak chromatograms (BPCs) from the gas chromatography-orbitrap (GC-Orbitrap) (blue) and gas chromatography-single quadrupole (GC-SQ) (orange) raw data in a representative blank (A) and sample (B).

throughout the blank. In the blank BPC overlay (Figure 1A), the GC-Orbitrap (blue) peaks have a more consistent distribution throughout the method and a higher overall relative intensity compared to the GC-SQ distribution (orange). Furthermore, more low-intensity peaks are detected in the GC-Orbitrap blank, highlighting increased sensitivity for a range of ion intensities. Improved metabolite detection is an expected output of a higher resolution instrument; however, the variability of peak detection in the blanks highlights that increased sensitivity increases the risk of false peak detection.

When considering a representative sample's BPCs from both instruments, there are apparent differences in peak abundance and distribution (Figure 1B). Using the same chromatography method, more peaks with a more even distribution between minutes 4 and 12 resulted from the GC-Orbitrap (blue) data. This observation aligns with previous findings that Orbitrap systems feature highly efficient ion transmission, improving the number of ions that reach the detector.[26] The two instruments resulted in similar amounts of low-intensity peaks. At the same time, the GC-Orbitrap detected more middle and high-intensity peaks, suggesting that the Orbitrap detector has a better dynamic range in full scan mode. Typically, there is concern that low abundance peaks will be overlooked with untargeted studies, however optimal methods provide a full peak spectrum, not only focusing on low vs. high abundance peaks.[27] This confirms previous reports that GC-Orbitrap analysis provides improved metabolite coverage and sensitivity compared to low-resolution analysers.[28]

Figure 2 compares the BPC of representative *O. gratissimum* (Figure 2A) and *O. tenuiflorium* (Figure 2B) samples from the GC-Orbitrap (blue) and GC-SQ data (orange). There are clear differences in peak abundance and range between the instruments. Both species have more peaks resulting from GC-Orbitrap detection compared to the GC-SQ. There is an overlap of a few peaks from the two species, however the BPC from both instruments demonstrates variable chemical profiles. The appearance of more high and middle intensity peaks in the GC-Orbitrap confirms that the higher resolution instrument provides increased dynamic range. Overall, Figures 1 and 2 demonstrate that the GC-Orbitrap instrument provides increased metabolite detection in *Ocimum* essential oils compared to the GC-SQ instrument.

Following peak processing, which included mass detection, chromatogram building, chromatogram deconvolution, and various filtering steps (Table S1), the GC-Orbitrap and GC-SQ datasets contained 737 and 417 unique compounds, respectively. This confirms that Orbitrap technology leads to increased peak detection, even after filtering noise.

3.2 | Open-source spectral matching is more robust with GC-Orbitrap data

A bottleneck in untargeted metabolomics studies is streamlined, reliable library spectral matching for feature searches and annotation. Once a compound of interest is found, confidently identifying it is not trivial. Algorithmic library searches are the most common approach to identification – algorithms match mass spectra from the experimental data to library spectra of known compounds. Those with the most fragment overlap and similar masses are given a high likelihood score and can be tentatively identified.[29,30] NIST is the most common database used for GC–MS analysis and contains over 70,000 unique compounds.[8] While proprietary software exists with company-specific libraries, we decided to use the open-source GNPS library matching tools to directly compare compound annotation with data from the two instruments.[18] Libraries used for matching included the NIST14, MassBank of North America (MONA) GC library, and FiehnLib.

The parameters for library matching can be found in Table S2; identical parameters were used for both searches. Using GNPS, 210 potential compound hits were found in the GC-SQ data, and 365 potential compound hits were found in the GC-Orbitrap data. This increase of 155 peaks between the two datasets (1.7-fold increase) is proportional to the 1.7-fold increase of peaks in the filtered GC-Orbitrap vs. GC-SQ data. Thus, it is unknown if the improved number of hits is due to better spectral matching or simply because there were more peaks to be matched.

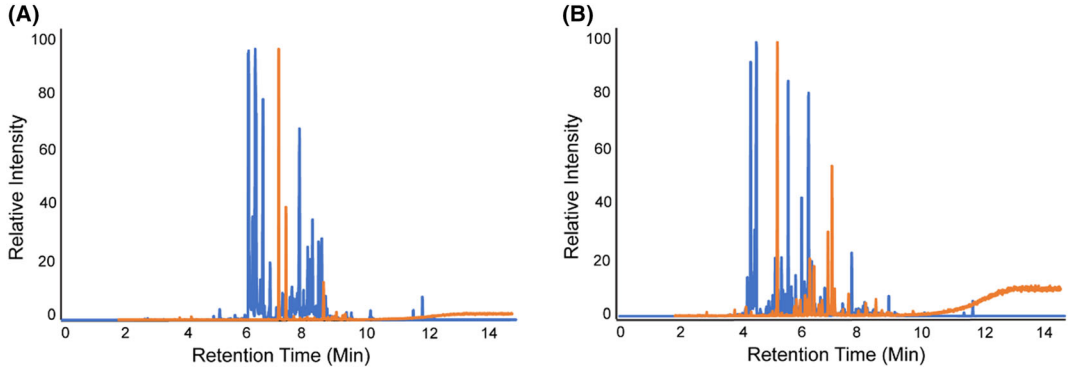


FIGURE 2 Comparisons of base peak chromatograms (BPCs) from representative *Ocimum gratissimum* (A) and *Ocimum tenuiflorum* (B) samples from the gas chromatography-orbitrap (GC-Orbitrap) (blue) and gas chromatography-single quadrupole (GC-SQ) (orange) data.

3.3 | Different marker compounds determined for *O. gratissimum* and *O. tenuiflorum* with GC-Orbitrap and GC-SQ data

In addition to the GNPS library search, we compiled a list of 41 compounds previously identified in *O. gratissimum* and *O. tenuiflorum*. This list was generated with a limited metasearch from Google Scholar and PubMed using the following keywords: *Ocimum*, *Ocimum gratissimum*, *Ocimum tenuiflorum*, essential oil, chemistry, metabolites, small molecules, Tulsi, Holy Basil. The 50 most common *Ocimum* metabolites from this search were narrowed down to 41 due to nomenclature and structural redundancy.[14,19–25] Tables 3 and 4 outline the different known essential oil compounds and each compound's average *m/z* and retention time in the GC-SQ and GC-Orbitrap data, respectively. Compounds were found in the raw data with a manual search, using the fragmentation patterns from published, publicly available spectra to search for the best match. Since there was no retention index (RI) or isolation and follow-up analysis in this study, each compound is assigned a level 3 identification (tentative candidate) according to the level system proposed by Schymanski *et al*.[31] Of the 41 compounds, 29 were found in the GC-SQ data (Table 3, 70%) and 34 in the GC-Orbitrap data (Table 4, 83%). A-farnese, bicyclogermacrene, and copaene were not found in either data set, potentially due to low natural abundance. This demonstrates that using GC-Orbitrap instrumentation for essential oil analysis provides improved coverage of known metabolites. In other words, it may be a better option for semi-targeted analysis where particular compounds are of interest, but a holistic look at the metabolite profile is desired. This may prove especially useful for essential oil analysis in a range of botanicals because many volatile compounds have highly similar structures and masses producing overlapping spectra; GC-Orbitrap data allows separation of these compounds without extensive gradient development.[32]

To meet the goal of an unbiased comparison of the GC-Orbitrap and GC-SQ for compound identification we only considered the direct data outputs instead of adding additional annotation tools. Of course, there are steps to take that are not included in this study that may improve annotation reliability and ease. For example, RI libraries can greatly aid compound annotation and have benefits over spectral and retention time matching alone.[33,34] Additionally, analytical standards for specific compounds of interest allow direct comparison of retention time, fragmentation patterns, and mass detection for absolute annotations for a semi-targeted approach.[35] Future studies may also benefit from gas chromatography-flame ionisation detector (GC-FID) analysis, which provides more accurate quantitative data, but provides little information about compound structure and is unsuitable for untargeted analysis or a direct comparison to Orbitrap data.

3.4 | Implications of species-specific marker variation with differing instrumental analysis

Currently, the most common approach to botanical product authentication is targeted analysis. Marker compounds unique to a species are quantified and used to confirm a sample's identity. In some cases, this is a single compound; in others, it is a unique ratio of compounds.[36] For example, the Association of Official Analytical Chemists (AOAC) official method for identifying ashwagandha (*Withania somnifera* WS) utilises 10 different compounds.[37] Herbal chemotypes are differentiated based on proportions of their major constituents – oregano (*Origanum vulgare* L.) is analysed based on concentration of carvacrol, thymol, and linalool.[38] However, selecting marker analytes is not a simple task, especially since many secondary metabolites in herbs can change drastically based on individual chemotypes, extraction techniques, and instrumentation.[39,40] Furthermore, herbs in the same genus but different species are often grouped tor analytical analysis, despite differences in chemical and bioactive profiles.[41]

To determine if GC-SQ and GC-Orbitrap instruments identify different marker compounds for the same samples, we compared the ratio of the 41 known compounds between *O. tenuiflorum* and *O. gratissimum* (Figures 3 and 4). Once a tentative candidate was identified in the raw data, the peak area was averaged for all *O. tenuiflorum* and *O. gratissimum* samples. The ratio of each compound in *O. tenuiflorum* and *O. gratissimum* was calculated in both datasets and the relative abundance compared between the GC-SQ (Table 3) and GC-Orbitrap (Table 4). Additionally, a student's *t*-test (two-tailed, with unequal variance) was used to determine if any compounds have a significantly different abundance between the two species ($P = 0.05$). Ratios allowed comparisons of compound quantity between instruments, whereas the innate difference in baselines restricts a direct comparison of peak areas. From this evaluation, no compounds were significantly different between *O. tenuiflorum* and *O. gratissimum* in the GC-Orbitrap data (Table 4). Only one, chavicol, was significantly different between the species in the GC-SQ data (Table 3). This is not surprising since the two species are very closely related and literature reports of the species' constitutes are inconsistent. To draw conclusions about compound ratios and potential marker compounds, we used a *P*-value of 0.1 for comparisons in Figures 3 and 4.

Tentative marker compounds were proposed using the ratio of each compound's peak abundance in the two species. Figure 3 compares the top 10 compounds with a larger peak area in *O. gratissimum* than *O. tenuiflorum*; the two instruments provide unique combinations of high abundance compounds. For example, camphor was the most distinct between the two species with the GC-Orbitrap analysis, but eugenol was the most distinct with the GC-SQ analysis. Notably, camphor was not detected in the GC-SQ dataset. However, eugenol was found in both datasets and is significantly different ($P = 0.1$) in the GC-Orbitrap data. Other overlapping compounds with a higher peak area ratio in *O. gratissimum* vs. *O. tenuiflorum* are τ-cadinol, caryophyllene, β-elemene, terpinene, and elemene. Previous studies have reported camphor higher in *Ocimum kilimandscharicum* than other *Ocimum* species, justifying its exclusion as a marker compound.[42] Another study found that τ-cadinol and eugenol have higher concentrations in *O. tenuiflorum* than *O. gratissimum* with a controlled growth study (not market available products).[43] Combined, the six overlapping compounds formulate a unique chemical fingerprint for

TABLE 3 Evaluation of known essential oil compounds found in gas chromatography-single quadrupole (GC-SQ) data.

Compound	Average *m/z*	Retention time (min)	*P*-Value	Average peak area in OG	Average peak area in OT	Peak area ratio (OG/OT)
3-Hexanol	73.1	4.34	0.92	4356.09	4381.57	1.00
α-Bergamotene	204.2	8.08	0.96	3158.92	3240.14	0.97
α-Bisabolene	119.1	8.13	0.05	1676.27	4228.90	0.40
α-Bulnesene	93.1	7.63	0.84	10376.94	8786.31	1.18
α-Cadinol	204.2	8.79	0.18	6420.14	3504.07	1.02
α-Farnese						
Alloocimene						
α-Pinene						
Aromadendrene	105.1	8.04	0.79	3531.87	3961.24	0.89
α-Terpineol	93.1	6.25	0.16	485.83	1123.96	0.43
β-Elemene	81.1	7.32	0.56	16308.68	12868.77	1.27
Biscyclogermacrene						
Borneol						
β-Pinene	69.1	8.09	0.06	4050.14	11462.41	0.35
β-Selinene	93.1	8.83	0.86	23846.91	18045.69	1.23
Camphene	121.1	7.98	0.24	2763.49	4672.99	0.59
Camphor						
Carene	93.1	8.13	0.08	10936.55	3770.39	0.96
Carvone						
Caryophyllene	91.1	8.47	0.18	38114.29	20718.40	1.84
Chavicol	134.1	6.53	**0.01**	4147.71	26489.45	0.16
Copaene						
Elemene	81.1	7.32	0.91	14835.67	13521.51	1.15
Elemol	189.2	8.78	0.72	3335.09	3930.22	0.85
Estragole	148.1	6.29	0.18	748.32	4982.93	0.15
Eucalyptol (1,8-cineole)						
Eugenol	164.1	7.20	0.16	812707.78	494355.96	1.64
Fenchone						
Geraniol	69.1	6.69	0.09	4893.23	16246.93	0.30
Germacrene-D	161.1	8.84	0.51	4000.65	5187.02	0.77
Limonene	136.1	8.29	0.07	1051.12	430.91	1.28
Linalool	71.1	5.57	0.16	3967.91	40329.70	0.10
Menthol						
Methyl cinnamate						
Myrcene	93.1	6.40	0.14	515.09	1458.49	0.35
Ocimene	93.1	8.05	0.11	5263.12	12064.05	0.44
Sabinene						
τ-Cadinol	161.1	8.84	0.51	4000.65	5187.02	0.77
Terpinene	93.1	8.13	0.76	2461.14	2301.60	1.07
Terpinolene						
γ-Cadinene						
γ-Muurolene	105.1	8.79	0.70	10443.30	11800.23	0.89

Note: *P*-value calculated via a two-tailed student's *t*-test between the peak area in *Ocimum gratissumum* (OG) and *Ocimum tenuiflorium* (OT) samples. Bold *P*-values are significant ($P = 0.1$).

TABLE 4 Evaluation of known essential oil compounds found in the gas chromatography-Orbitrap (GC-Orbitrap) data.

Compound	Average *m/z*	Retention time (min)	*P*-Value	Average peak area in OG	Average peak area in OT	Peak area ratio (OG/OT)
3-Hexanol						
α-Bergamotene	93.0700	7.40	0.27	1.17E+08	1.92E+08	0.61
α-Bisabolene	93.07699	7.40	0.27	1.17E+08	1.92E+08	0.61
α-Bulnesene	107.0731	4.63	0.43	2.41E+06	1.07E+07	0.23
α-Cadinol						
α-Farnese						
Alloocimene	121.0682	6.57	0.13	1.05E+09	5.53E+08	1.91
α-Pinene	93.0700	5.03	0.17	4.97E+07	3.45E+08	0.14
Aromadendrene	105.0699	6.87	0.39	3.37E+07	8.96E+07	0.38
α-Terpineol	136.0889	7.92	0.69	7.08E+07	9.16E+07	0.77
β-Elemene	95.0807	5.02	0.48	2.31E+07	9.57E+06	2.41
Biscyclogermacrene						
Borneol	95.0807	5.02	0.19	2.82E+06	2.19E+07	0.13
β-Pinene	121.0882	5.09	0.53	1.03E+07	3.32E+07	0.31
β-Selinene	107.0827	5.14	0.29	6.92E+06	2.48E+07	0.28
Camphene	121.0882	5.09	0.53	1.03E+07	3.32E+07	0.31
Camphor	95.0669	5.76	0.09	1.01E+08	2.83E+07	3.59
Carene						
Carvone	108.0717	4.92	0.10	6.65E+05	1.10E+07	0.06
Caryophyllene	133.0155	3.13	0.32	5.26E+05	2.82E+05	1.87
Chavicol	167.1010	6.07	0.40	5.35E+04	7.57E+04	0.71
Copaene						
Elemene	81.0699	9.96	0.51	2.92E+07	1.61E+07	1.81
Elemol	59.0458	7.33	0.32	2.28E+06	1.23E+06	1.85
Estragole	148.0874	5.74	0.54	1.68E+07	3.96E+07	0.43
Eucalyptol (1,8-cineole)	81.0669	9.87	0.89	2.16E+07	2.37E+07	0.91
Eugenol	164.0827	6.39	0.07	2.00E+09	9.55E+08	2.10
Fenchone	81.0699	9.87	0.89	2.16E+07	2.37E+07	0.91
Geraniol	139.0827	9.96	0.62	2.22E+06	3.72E+06	0.60
Germacrene-D	119.0856	7.61	0.51	1.11E+08	1.48E+08	0.75
Limonene	67.0542	5.98	0.45	3.42E+07	5.39E+07	0.63
Linalool	93.0699	4.30	0.37	1.49E+06	2.62E+06	0.57
Menthol	71.0545	9.88	0.84	8.84E+06	8.02E+06	1.10
Methyl cinnamate	94.0716	5.60	0.37	8.17E+06	1.81E+07	0.45
Myrcene	93.0699	4.30	0.37	1.49E+06	2.62E+06	0.57
Ocimene	93.0700	5.03	0.17	4.96E+07	3.46E+08	0.14
Sabinene	93.0700	4.90	0.17	3.80E+07	2.62E+08	0.15
τ-Cadinol	161.1162	7.02	0.34	6.13E+07	2.63E+07	2.33
Terpinene	136.0888	8.01	0.32	1.79E+08	1.22E+08	1.47
Terpinolene	149.0726	6.58	0.09	1.50E+09	6.78E+08	2.22
γ-Cadinene	188.0845	9.46	0.33	1.38E+06	2.90E+06	0.48
γ-Muurolene						

Note: *P*-value calculated via a two-tailed student's *t*-test between the peak area in *Ocimum gratissumum* (OG) and *Ocimum tenuiflorium* (OT) samples. Bold *P*-values are significant ($P = 0.1$).

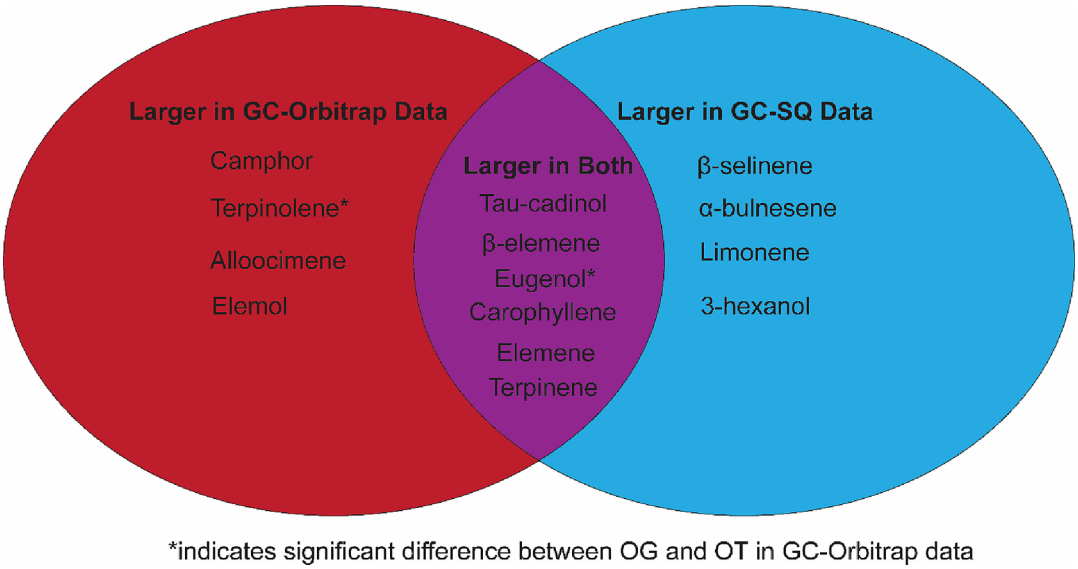


FIGURE 3 Comparison of 10 discriminating compounds between *Ocimum gratissimum* (OG) and *Ocimum tenuiflorum* (OT). These were selected due to their greater peak area in *O. gratissimum*. The selected putative biomarkers differed between the two mass analysers, Orbitrap vs. single quadrupole (SQ), yet there were six that were found in both systems that can discriminate between the two species.

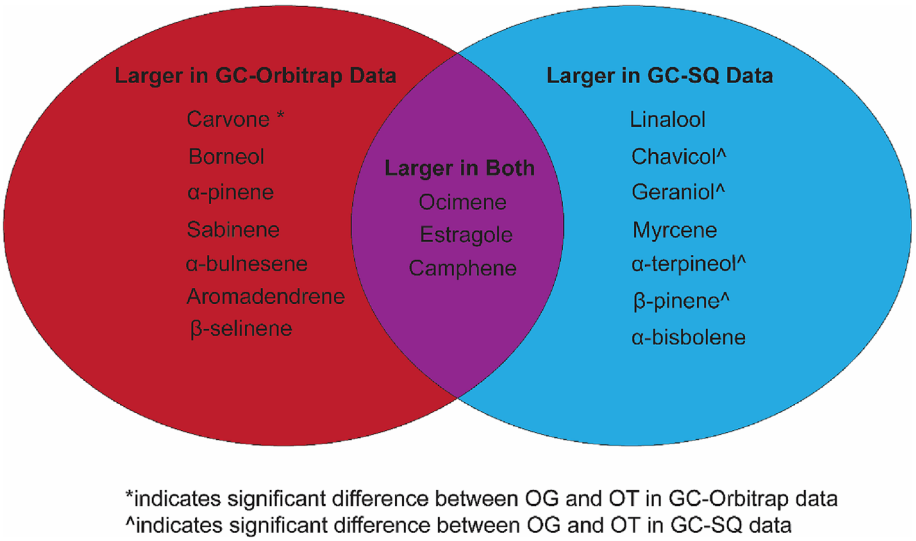


FIGURE 4 Comparison of 10 discriminating compounds between *Ocimum tenuiflorum* (OT) and *Ocimum gratissimum* (OG). These were selected due to their greater peak area in *O. tenuiflorum*. The selected putative biomarkers differed between the two mass analysers, Orbitrap vs. single quadrupole (SQ), yet there were three that were found in both systems that can discriminate between the two species.

O. gratissimum, with consistent detection across analytical instruments.

Similarly, Figure 4 compares the top 10 compounds with a larger peak area in *O. tenuiflorum* than *O. gratissimum*. There is much less overlap between the GC-Orbitrap and GC-SQ in this comparison, with only three compounds, ocimene, estragole, and camphene, similar between the two. There is no significant difference in these three compounds between the two species but together act as potential markers of *O. tenuiflorum*. A review of major essential oil constituents found that estragole and ocimene are often major constituents of *O. tenuiflorum* and *O. gratissimum*,[14] so investigating the ratio of the two in combination with camphene and other discriminatory metabolites would be a robust approach to authentication. The GC-SQ has four compounds with significant differences ($P = 0.1$) between the species with a higher abundance in *O. tenuiflorum*, suggesting that the data may be more suitable for chemotype analysis. Interestingly, α-bulnesene had a higher abundance in *O. tenuiflorum* in the GC-Orbitrap data but a higher abundance in *O. gratissimum* in the GC-SQ data.

Together, this leads to two suggestions for marker compound determination for herbal product studies. First, multiple instruments should be considered when evaluating potential markers and selecting analytes for future authentication. If only the GC-Orbitrap was employed, compounds not detected by other instruments could have

been selected as markers, leading to specious species identification when considering future samples in other studies. This is especially important when considering compounds like aromadendrene, which had a higher peak area in *O. gratissimum* in the GC-SQ data but a higher peak area in *O. tenuiflorum* in the GC-Orbitrap data. The second suggestion is in the case where the use of two instruments is infeasible, monographs and methods should be instrument-specific, with a separate validated workflow and compound list for each analytical approach.

3.5 | Unsupervised classification of *Ocimum* species is inconclusive

An unsupervised PCA evaluated the ability of essential oil analysis to differentiate between *O. gratissimum* and *O. tenuiflorum* (Figure 5). PCA reduces the dimensionality of data to show spatial distribution of samples based on the most variable metabolites, so samples closer together in the scores plot are more chemically similar than those farther apart. The associated loadings plot allows visualisation of the compounds most responsible for variation in each PC.

Unsupervised PCA analysis has been used to distinguish species based on their essential oil profiles. However, it is unknown if GC-Orbitrap data, which provides greater metabolite coverage and increased detection of mid-abundance compounds (Table 2, Figure 1), improves species separation compared to lower resolution data. Figure 5 demonstrates that in the context of this study, neither instrument produces data suitable for separating the species, and the same samples cluster together with both data inputs. This suggests essential oil profiles are capable of distinguishing *Ocimum* characteristics, but without more metadata (which was unavailable from product suppliers) we cannot determine the defining traits of each cluster, however subsequent analysis showed that the clusters are not based on country of origin or supplier (Supporting Information Figures S1 and S2).

Previous research has shown that GC-Orbitrap data can distinguish samples based on geographic origin,[12] but limited evidence for species distinction exists. This is the first study comparing the classification potential of GC-Orbitrap data to lower resolution data; however, previous studies have shown that using chemometric approaches to pattern visualisation and sample classification with LC-Orbitrap data provides similar clustering as LC-SQ analysis, with no clear distinction in their identification performance.[44]

While the scores plots do not provide species-level distinction of *Ocimum* samples, the loadings plots offer insight into the key features responsible for innate variation between the samples. It should

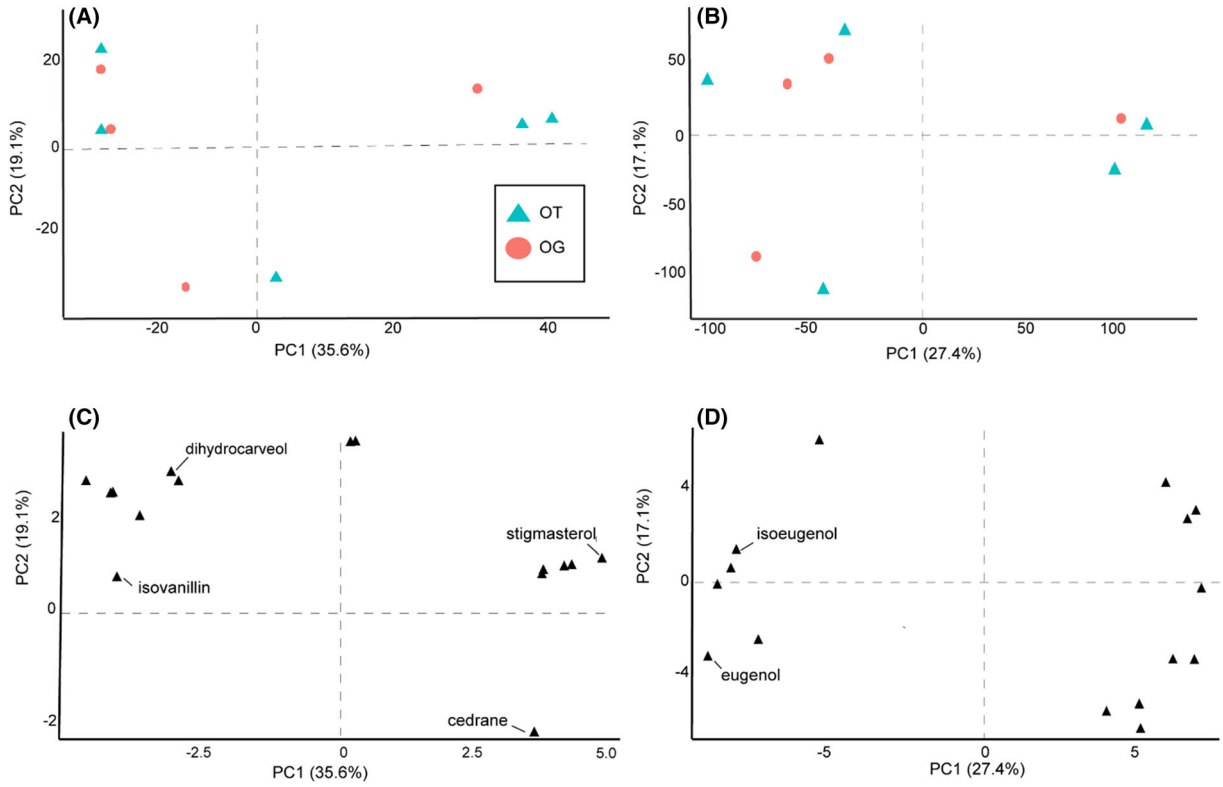


FIGURE 5 Unsupervised principal component analysis (PCA) comparing *Ocimum* tenuiflorum (OT) and *Ocimum gratissimum* (OG) samples using the full metabolite profile from the gas chromatography-single-quadrupole (GC-SQ) (A) or gas chromatography-orbitrap (GC-Orbitrap) (B) and identification of the most variable features from the loadings plot of the GC-SQ (C) and GC-Orbitrap (D) data. Each symbol represents the average peak area for the triplicate extractions of each sample. The samples in each quadrant are the same with both datasets.

be noted only the top 15 compounds with the most variation on PC1 and PC2 were investigated, and of the 30 from the two loadings plots, only six had potential hits through the GNPS spectral matching or via manual search. The GC-SQ data (Figure 5C) shows four potential compounds – dihydrocarveol, isovanillin, cedrane, and stigmasterol; correlated with cluster separation along PC1. The GC-Orbitrap data (Figure 5D) identified isoeugenol and eugenol as distinguishing compounds along PC1, with both more associated with the clusters on the left, and eugenol correlated with the bottom left quadrant. The PCA shows no clear distinction between these species which echoes the results shown in the GC-Orbitrap and GC-SQ data. More information about the essential oil products and better library spectral matching could potentially discern a trend within the PCA.

The lack of an extensive spectral library currently limits GC–HRMS. The NIST database was performed solely on low-resolution GC–MS systems which gives traditional GC–MS systems more identification power. However, as more studies use GC-Orbitrap instruments to generate high-resolution mass spectral data, library resources will improve open-source compound identification. While GC-Orbitrap data may improve semi-targeted compound searches, there are many drawbacks that must be approached before routine use in an industry setting, including increased costs, lack of spectral libraries, and greater expertise required for instrument operation.

ACKNOWLEDGEMENTS

This work was supported in part by the USDA National Institute of Food and Agriculture's Hatch Appropriations (PEN04772) and the Garden Club of America's Anne S. Chatham Fellowship in Medicinal Botany. The authors would like to thank the Huck Institute's Metabolomics Core Facility at the Pennsylvania State University and the Department of Chemistry Mass Spectrometry Core Laboratory at the University of North Carolina Chapel Hill for supporting mass spectrometry instrumentation and data collection.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the MASSive database at https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=870fe9654fbc43cd8223e1b68375a3bc, reference number MSV000091070.

ORCID

Joshua J. Kellogg https://orcid.org/0000-0001-8685-0353

REFERENCES

1. Smith T, May G, Eckl V, Reynolds CM. US sales of herbal supplements increase by 8.6% in 2019. *HerbalGram*. (127):54-69.
2. Thakkar S, Anklam E, Xu A, et al. Regulatory landscape of dietary supplements and herbal medicines from a global perspective. *Regul Toxicol Pharmacol*. 2020;114:104647. doi:10.1016/j.yrtph.2020.104647
3. Ichim MC. The DNA-based authentication of commercial herbal products reveals their globally widespread adulteration. *Front Pharmacol*. 2019;10:1227. doi:10.3389/fphar.2019.01227
4. Kan HX, Jin L, Zhou FL. Classification of medicinal plant leaf image based on multi-feature extraction. *Pattern Recognit Image Anal*. 2017;27(3):581-587. doi:10.1134/S105466181703018X
5. Abraham E, Kellogg JJ. Chemometric-guided approaches for profiling and authenticating botanical materials. *Front Nutr*. 2021;8:24. doi:10.3389/fnut.2021.780228
6. Kellogg JJ, Paine MF, McCune JS, Oberlies NH, Cech NB. Selection and characterization of botanical natural products for research studies: a NaPDI center recommended approach. *Nat Prod Rep*. 2019;36(8):1196-1221. doi:10.1039/C8NP00065D
7. Figueiredo AC, Barroso JG, Pedro LG, Scheffer JJC. Factors affecting secondary metabolite production in plants: volatile components and essential oils. *Flavour Fragr J*. 2008;23(4):213-226. doi:10.1002/ffj.1875
8. Linstrom P. NIST chemistry WebBook, NIST standard reference database 69. Published Online 1997. 10.18434/T4D303
9. Yost RA. The triple quadrupole: innovation, serendipity and persistence. *JMSACL*. 2022;24:90-99. doi:10.1016/j.jmsacl.2022.05.001
10. Peterson AC, McAlister GC, Quarmby ST, Griep-Raming J, Coon JJ. Development and characterization of a GC-enabled QLT-Orbitrap for high-resolution and high-mass accuracy GC/MS. *Anal Chem*. 2010;82(20):8618-8628. doi:10.1021/ac101757m
11. Lacalle-Bergeron L, Izquierdo-Sandoval D, Sancho JV, López FJ, Hernández F, Portolés T. Chromatography hyphenated to high resolution mass spectrometry in untargeted metabolomics for investigation of food (bio)markers. *TrAC*. 2021;135:116161. doi:10.1016/j.trac.2020.116161
12. Rivera-Pérez A, Romero-González R, Garrido Frenich A. Feasibility of applying untargeted metabolomics with GC-Orbitrap-HRMS and Chemometrics for authentication of black pepper (*Piper nigrum* L.) and identification of geographical and processing markers. *J Agric Food Chem*. 2021;69(19):5547-5558. doi:10.1021/acs.jafc.1c01515
13. Rivera-Pérez A, Romero-González R, Garrido Frenich A. Application of an innovative metabolomics approach to discriminate geographical origin and processing of black pepper by untargeted UHPLC-Q-Orbitrap-HRMS analysis and mid-level data fusion. *Food Res Int*. 2021;150(Pt A):110722. doi:10.1016/j.foodres.2021.110722
14. Pandey AK, Singh P, Tripathi NN. Chemistry and bioactivities of essential oils of some Ocimum species: an overview. *Asian Pac J Trop Biomed*. 2014;4(9):682-694. doi:10.12980/APJTB.4.2014C77
15. Shiwakoti S, Saleh O, Poudyal S, Barka A, Qian Y, Zheljazkov VD. Yield, composition and antioxidant capacity of the essential oil of sweet basil and holy basil as influenced by distillation methods. *Chem Biodivers*. 2017;14(4):e1600417. doi:10.1002/cbdv.201600417
16. Pluskal T, Castillo S, Villar-Briones A, Orešič M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform*. 2010;11(1):395. doi:10.1186/1471-2105-11-395
17. Chambers MC, Maclean B, Burke R, et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol*. 2012;30(10):918-920. doi:10.1038/nbt.2377
18. Wang M, Carver JJ, Phelan VV, et al. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat Biotechnol*. 2016;34(8):828-837. doi:10.1038/nbt.3597
19. Keita SM, Vincent C, Schmit JP, Belanger A. Essential oil composition of Ocimum basilicum L., O. gratissimum L. and O. Suave L. in the Republic of Guinea. *Flavour Fragr J*. 2000;15(5):339-341. doi:10.1002/1099–1026(200009/10)15:5%3C339::AID–FFJ922%3E3.0.CO;2–H
20. Matasyoh LG, Matasyoh JC, Wachira FN, Kinyua MG, Muigai T, Mukiama TK. Chemical composition and antimicrobial activity of the essential oil of *Ocimum gratissimum* L. growing in eastern Kenya. *Afr J Biotechnol*. 2007;6:760-765.

21. Vieira RF, Grayer RJ, Paton A, Simon JE. Genetic diversity of Ocimum gratissimum L. based on volatile oil constituents, flavonoids and RAPD markers. *Biochem Syst Ecol*. 2001;29(3):287-304. doi:10.1016/S0305-1978(00)00062-4
22. Joshi R. Chemical composition, in vitro antimicrobial and antioxidant activities of the essential oils of Ocimum gratissimum, O. sanctum and their major constituents. *Indian J Pharm Sci*. 2013;75(4):457-462. doi:10.4103/0250-474X.119834
23. Kicel A, Kurowska A, Kalemba D. Composition of the essential oil of *Ocimum sanctum* L. grown in Poland during vegetation. *J Essent Oil Res*. 2005;17(2):217-219. doi:10.1080/10412905.2005.9698880
24. Pino JA, Rosado A, Rodriguez M, Garcia D. Composition of the essential oil of *Ocimum tenuiflorum* L. grown in Cuba. *J Essent Oil Res*. 1998;10(4):437-438. doi:10.1080/10412905.1998.9700937
25. Saharkhiz MJ, Kamyab AA, Kazerani NK, Zomorodian K, Pakshir K, Rahimi MJ. Chemical compositions and antimicrobial activities of Ocimum sanctum L. essential oils at different harvest stages. *Jundishapur J Microbiol*. 2014;8(1):e13720. doi:10.5812/jjm.13720
26. Zubarev RA, Makarov A. Orbitrap mass spectrometry. *Anal Chem*. 2013;85(11):5288-5296. doi:10.1021/ac4001223
27. Chaleckis R, Meister I, Zhang P, Wheelock CE. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *COBIOT*. 2019;55:44-50. doi:10.1016/j.copbio.2018.07.010
28. Stettin D, Poulin RX, Pohnert G. Metabolomics benefits from Orbitrap GC–MS—comparison of low- and high-resolution GC–MS. *Metabolites*. 2020;10(4):143. doi:10.3390/metabo10040143
29. Mastrangelo A, Ferrarini A, Rey-Stolle F, García A, Barbas C. From sample treatment to biomarker discovery: a tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal Chim Acta*. 2015;900:21-35. doi:10.1016/j.aca.2015.10.001
30. Halket JM, Waterman D, Przyborowska AM, Patel RKP, Fraser PD, Bramley PM. Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. *J Exp Bot*. 2005;56(410):219-243. doi:10.1093/jxb/eri069
31. Schymanski EL, Jeon J, Gulde R, et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol*. 2014;48(4):2097-2098. doi:10.1021/es5002105
32. Wolfender JL, Marti G, Thomas A, Bertrand S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J Chromatogr A*. 2015;1382:136-164. doi:10.1016/j.chroma.2014.10.091
33. Wei X, Koo I, Kim S, Zhang X. Compound identification in GC-MS by simultaneously evaluating the mass spectrum and retention index. *Analyst*. 2014;139(10):2507-2514. doi:10.1039/C3AN02171H
34. Lucero M, Estell R, Tellez M, Fredrickson E. A retention index calculator simplifies identification of plant volatile organic compounds: a retention index calculator. *Phytochem Anal*. 2009;20(5):378-384. doi:10.1002/pca.1137
35. Beale DJ, Pinu FR, Kouremenos KA, et al. Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics*. 2018;14(11):152. doi:10.1007/s11306-018-1449-2
36. You H, Abraham EJ, Mulligan J, et al. Label compliance for ingredient verification: regulations, approaches, and trends for testing botanical products marketed for "immune health" in the United States. *CRFSN*. 2022;1-20. doi:10.1080/10408398.2022.2124230
37. Official Methods of Analysis of AOAC International. *Withanolide Glycosides and Aglycones of Ashwagandha (Withania Somnifera)*. 22nd ed. AOAC International; 2015.
38. Leyva-López N, Gutiérrez-Grijalva E, Vazquez-Olivo G, Heredia J. Essential oils of oregano: biological activity beyond their antimicrobial properties. *Molecules*. 2017;22(6):989. doi:10.3390/molecules22060989
39. Swamy MK, Arumugam G, Kaur R, Ghasemzadeh, Yusoff, Sinniah UR. GC-MS based metabolite profiling, antioxidant and antimicrobial properties of different solvent extracts of Malaysian *Plectranthus amboinicus* leaves. *eCAM*. 2017;2017:1-10. doi:10.1155/2017/1517683
40. Li Y, Kong D, Fu Y, Sussman MR, Wu H. The effect of developmental and environmental factors on secondary metabolites in medicinal plants. *Plant Physiol Biochem*. 2020;148:80-89. doi:10.1016/j.plaphy.2020.01.006
41. Ríos-Rodríguez D, Sahi VP, Nick P. Authentication of holy basil using markers relating to a toxicology-relevant compound. *Eur Food Res Technol*. 2021;247(10):2485-2497. doi:10.1007/s00217-021-03812-z
42. dos Santos E, Leitão MM, Aguero Ito CN, et al. Analgesic and anti-inflammatory articular effects of essential oil and camphor isolated from Ocimum kilimandscharicum Gürke leaves. *J Ethnopharmacol*. 2021;269:113697. doi:10.1016/j.jep.2020.113697
43. Stefan M, Zamfirache M, Padurariu C, Trută E, Gostin I. The composition and antibacterial activity of essential oils in three Ocimum species growing in Romania. *Open Life Sci*. 2013;8(6):600-608. doi:10.2478/s11535-013-0171-8
44. Glauser G, Veyrat N, Rochat B, Wolfender JL, Turlings TCJ. Ultra-high pressure liquid chromatography–mass spectrometry for plant metabolomics: a systematic comparison of high-resolution quadrupole-time-of-flight and single stage Orbitrap mass spectrometers. *J Chromatogr A*. 2013;1292:151-159. doi:10.1016/j.chroma.2012.12.009

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Abraham EJ, Wallace ED, Kellogg JJ. A comparison of high- and low-resolution gas chromatography–mass spectrometry for herbal product classification: A case study with *Ocimum* essential oils. *Phytochemical Analysis*. 2023;34(6):680-691. doi:10.1002/pca.3258

Supplemental Table 1: MZMine Preprocessing parameters

Parameter	GC-SQ	GC-Orbitrap
Mass detection	Noise level: 2.0E2	Noise level: 5.0E5
ADAP chromatogram builder	Min group size in # of scans: 5 Group intensity threshold: 2.0E2 Min highest intensity:2.0E2 *m/z* tolerance: 0.1 *m/z* or 0.0 ppm	Min group size in # of scans: 5 Group intensity threshold: 5.0E5 Min highest intensity:5.0E5 *m/z* tolerance: 0.05 *m/z* or 0.0 ppm
Chromatogram deconvolution	Algorithm: wavelets (ADAP) m/z center calculation: median S/N threshold: 7 S/N estimator: intensity window SN Min feature height: 1,000 Coefficient/area threshold: 30 Peak duration range: 0.00 - 1.00 RT wavelet range: 0.00 - 0.15	Algorithm: wavelets (ADAP) m/z center calculation: median S/N threshold: 10 S/N estimator: intensity window SN Min feature height: 500 Coefficient/area threshold: 100 Peak duration range: 0.00 - 1.00 RT wavelet range: 0.00 - 0.10
Spectral deconvolution – hierarchical clustering	Min cluster distance: 0.05 Min cluster size: 2 Min cluster intensity: 500 Min edge-to-height ratio: 0.2 Min delta-to-height ratio: 0.2 Min sharpness: 30 Shape-similarity tolerance: 18 Choice of model peak based on: sharpness	Min cluster distance: 0.05 Min cluster size: 2 Min cluster intensity: 500 Min edge-to-height ratio: 0.2 Min delta-to-height ratio: 0.2 Min sharpness: 30 Shape-similarity tolerance: 18 Choice of model peak based on: sharpness
ADAP aligner (GC)	Min confidence: 0.1 RT tolerance: 0.15 min m/z tolerance: 0.1 *m/z* or 0.0 ppm Score threshold: 0.75 Score weight: 0.1 RT similarity: RT difference	Min confidence: 0.1 RT tolerance: 0.15 min m/z tolerance: 0.05 *m/z* or 0.0 ppm Score threshold: 0.75 Score weight: 0.1 RT similarity: RT difference
Isotope peak grouper	m/z tolerance: 0.1 *m/z* or 0.0 ppm RT tolerance: 0.25 min	m/z tolerance: 0.05 *m/z* or 0.0 ppm RT tolerance: 0.25 min
Duplicate peak filter	m/z tolerance: 0.1 *m/z* or 0.0 ppm RT tolerance: 0.25 min	m/z tolerance: 0.05 *m/z* or 0.0 ppm RT tolerance: 0.25 min

Supplemental Table 2: GNPS processing parameters

Fragment filter	Window Filter	Precursor ion mass tolerance	Fragment ion tolerance	Molecular network cosine score	Molecular network matched peaks	Max molecular family size
± 17 Da	Top 6 ± 50 Da window	20000 Da	0.5 Da	> 0.7	> 6	100

Supplemental Table 3: Raw peak area data for tentatively identified key *Ocimum gratissimum* compounds in GC-Orbitrap data.

Compound	***m/z***	**RT (min)**	**OG 1.1**	**OG 1.2**	**OG 1.3**	**OG 2.1**	**OG 2.2**	**OG 2.3**	**OG 3.1**	**OG 3.2**	**OG 3.3**	**OG 4.1**	**OG 4.2**	**OG 4.3**
3-hexanol														
a-bergamotene	93.0699	7.40	3.34E+06	2.34E+08	6.12E+07	6.19E+07	4.00E+07	2.44E+07	1.49E+08	1.06E+08	3.15E+05	4.28E+08	2.58E+08	3.75E+07
a-bisabolene	93.0699	7.40	3.34E+06	2.34E+08	6.12E+07	6.19E+07	4.00E+07	2.44E+07	1.49E+08	1.06E+08	3.15E+05	4.28E+08	2.58E+08	3.75E+07
a-bulnesene	107.0731	4.63	1.51E+06	2.75E+04	1.93E+06	1.31E+05	7.23E+04	1.08E+05	1.87E+07	4.65E+06	1.72E+05	4.47E+05	8.32E+05	2.75E+05
a-cadinol														
a-farnese														
alloocimene	121.0682	6.57	1.68E+09	1.72E+09	1.06E+09	2.51E+09	2.62E+08	2.42E+09	3.86E+07	5.50E+02	1.44E+05	1.39E+06	1.51E+09	1.44E+09
a-pinene	93.0700	5.03	1.60E+08	1.16E+05	2.99E+07	5.51E+05	1.51E+05	3.07E+05	3.07E+08	1.55E+05	1.03E+05	9.52E+07	8.24E+05	1.52E+06
aromadendrene	105.0699	6.87	6.49E+07	1.63E+07	6.67E+06	4.25E+06	1.38E+07	2.92E+06	3.46E+07	4.70E+07	3.11E+05	9.74E+07	1.04E+08	1.17E+07
a-terpineol	136.0889	7.92	1.92E+08	7.92E+07	1.81E+08	2.59E+07	5.01E+07	4.04E+07	3.43E+07	4.17E+07	3.31E+05	5.78E+07	7.04E+07	7.60E+07
b-elemene	95.0807	5.02	4.43E+06	8.30E+05	1.32E+06	2.11E+05	1.64E+05	1.07E+05	7.99E+06	1.03E+07	1.98E+04	2.50E+08	4.78E+05	4.53E+05
biscyclogermacrene														
borneol	95.0807	5.02	4.43E+06	8.30E+05	1.32E+06	2.11E+05	1.64E+05	1.07E+05	7.99E+06	1.03E+07	1.98E+04	2.50E+08	4.78E+05	4.53E+05
b-pinene	121.0882	5.09	2.10E+06	2.34E+05	9.48E+06	3.35E+05	3.79E+05	7.30E+04	1.08E+08	5.19E+04	2.30E+04	9.87E+05	2.32E+06	7.93E+03
b-selinene	107.0827	5.14	1.83E+07	1.48E+05	3.19E+06	5.06E+05	6.06E+05	4.65E+05	3.71E+07	1.95E+06	1.84E+04	1.37E+07	2.20E+05	9.39E+05
camphene	121.0882	5.09	2.10E+06	2.34E+05	9.48E+06	3.35E+05	3.79E+05	7.30E+04	1.08E+08	5.19E+04	2.30E+04	9.87E+05	2.32E+06	7.93E+03
camphor	95.0669	5.76	3.46E+08	6.99E+05	1.48E+08	1.65E+07	2.71E+05	1.51E+07	5.91E+05	8.25E+07	5.55E+04	4.70E+08	6.13E+07	7.61E+07
carene														
carvone	108.0717	4.92	3.98E+06	2.10E+04	2.92E+05	2.88E+04	4.84E+04	4.29E+04	4.86E+05	7.45E+05	5.92E+03	1.94E+06	1.98E+05	1.90E+05
caryophyllene	133.0155	3.13	2.72E+05	8.53E+04	1.47E+05	1.03E+05	5.55E+05	1.48E+06	2.67E+05	2.70E+05	6.24E+04	1.92E+05	2.66E+06	2.14E+05
chavicol	167.1010	6.07	6.57E+04	1.82E+04	6.44E+04	3.45E+04	3.30E+04	2.08E+04	8.34E+04	1.22E+05	4.76E+03	3.71E+04	1.31E+05	2.63E+04
copaene														
elemene	81.0699	9.96	3.29E+06	4.47E+07	5.08E+06	2.21E+06	4.21E+06	2.20E+06	2.76E+07	1.73E+08	2.74E+05	2.62E+06	2.91E+06	3.46E+06
elemol	59.0458	7.33	1.32E+06	2.74E+06	1.44E+06	2.90E+04	1.62E+06	2.85E+04	1.88E+06	1.38E+07	2.06E+04	1.14E+06	1.66E+06	1.72E+06
estragole	148.0874	5.74	3.99E+07	1.41E+05	2.62E+04	8.05E+04	2.98E+04	5.21E+04	7.99E+07	2.00E+07	8.21E+03	6.01E+07	1.21E+06	7.16E+05
eucalyptol (1-8,cineole)	81.0699	9.87	1.11E+07	2.85E+06	7.87E+04	2.46E+05	2.58E+06	4.12E+06	2.76E+07	5.96E+07	2.74E+05	5.11E+07	4.83E+07	5.09E+07
eugenol	164.0827	6.39	3.65E+09	2.49E+09	2.58E+03	4.95E+09	2.59E+09	2.26E+09	7.43E+08	9.39E+06	9.43E+05	2.78E+09	2.01E+09	2.52E+09
fenchone	81.0699	9.87	1.11E+07	2.85E+06	7.87E+04	2.46E+05	2.58E+06	4.12E+06	2.76E+07	5.96E+07	2.74E+05	5.11E+07	4.83E+07	5.09E+07
geraniol	139.0898	9.96	1.50E+05	1.58E+05	3.89E+04	3.23E+04	3.27E+05	5.34E+04	2.45E+07	1.64E+05	1.50E+04	3.67E+05	1.68E-06	8.28E+05
germacrene-D	119.0856	7.61	2.41E+08	2.42E+06	1.24E+08	1.03E+08	2.09E+08	1.65E+08	1.22E+06	2.40E+07	3.49E+05	1.95E+08	1.10E+08	1.55E+08

limonene	67.0542	5.98	3.95E+05	4.72E+06	3.74E+05	2.24E+07	2.34E+07	1.46E+07	9.53E+07	2.18E+08	3.38E+05	1.84E+07	6.83E+06	5.22E+06
linalool	93.0699	4.30	7.57E+05	4.47E+04	2.32E+05	1.28E+05	1.53E+05	1.43E+05	5.81E+06	7.86E+06	1.25E+05	1.30E+06	7.55E+05	6.00E+05
menthol	71.0545	9.88	4.79E+06	2.74E+07	8.55E+06	4.42E+05	9.72E+05	5.50E+05	1.29E+06	3.51E+07	1.91E+04	7.99E+06	7.15E+06	1.19E+07
methyl cinnamate	94.0716	5.60	2.91E+06	2.81E+05	1.32E+06	5.07E+06	4.83E+05	1.98E+07	5.77E+07	4.06E+06	5.16E+04	4.58E+06	1.72E+06	4.28E+04
myrcene	93.0699	4.30	7.57E+05	4.47E+04	2.32E+05	1.28E+05	1.53E+05	1.43E+05	5.81E+06	7.86E+06	1.25E+05	1.30E+06	7.55E+05	6.00E+05
ocimene	93.0700	5.03	1.60E+08	1.16E+05	2.99E+07	5.51E+05	1.51E+05	3.07E+05	3.07E+08	1.55E+05	1.03E+05	9.52E+07	8.24E+05	1.52E+06
sabinene	93.0700	4.90	8.81E+07	1.46E+05	1.94E+07	3.59E+05	1.02E+05	2.37E+05	2.18E+08	9.96E+05	1.03E+05	1.26E+08	1.41E+06	1.02E+06
tau-cadinol	161.1162	7.02	5.21E+07	3.74E+06	4.62E+05	2.64E+05	1.14E+07	6.01E+06	1.35E+06	2.28E+05	1.94E+04	3.92E+08	6.03E+07	2.07E+08
terpinene	136.0888	8.01	1.90E+08	4.45E+08	1.80E+08	2.58E+07	5.01E+07	4.04E+07	1.28E+08	4.12E+07	3.26E+05	2.55E+08	3.83E+08	4.08E+08
terpinolene	149.0726	6.58	2.39E+09	1.45E+05	2.54E+09	3.73E+09	3.27E+09	2.10E+09	7.21E+08	2.89E+06	4.73E+05	1.15E+09	3.28E+04	2.15E+09
y-cadinene	188.0845	9.46	2.58E+04	5.06E+04	1.15E+04	1.54E+03	1.63E+03	1.94E+03	8.08E+06	7.89E+06	6.85E+03	1.92E+05	1.76E+05	1.78E+05
y-muurolene														

Supplemental Table 4: Raw peak area data for tentatively identified key *Ocimum tenuiflorum* compounds in GC-Orbitrap data.

Compound	***m/z***	**RT (min)**	**OT 1.1**	**OT 1.2**	**OT 1.3**	**OT 2.1**	**OT 2.2**	**OT 2.3**	**OT 3.1**	**OT 3.2**	**OT 3.3**	**OT 4.1**	**OT 4.2**	**OT 4.3**	**OT 5.1**	**OT 5.2**	**OT 5.3**
3-hexanol																	
a-bergamotene	93.0699	7.40	3.62E+07	3.61E+08	4.02E+08	5.27E+08	5.92E+08	1.24E+07	1.74E+08	1.59E+08	2.59E+06	7.10E+07	1.66E+08	3.14E+08	1.62E+06	3.70E+05	5.98E+07
a-bisabolene	93.0699	7.40	3.62E+07	3.61E+08	4.02E+08	5.27E+08	5.92E+08	1.24E+07	1.74E+08	1.59E+08	2.59E+06	7.10E+07	1.66E+08	3.14E+08	1.62E+06	3.70E+05	5.98E+07
a-bulnesene	107.0731	4.63	1.54E+06	1.27E+07	1.38E+08	2.67E+05	4.57E+05	3.97E+05	1.41E+06	3.75E+06	2.63E+05	2.58E+05	1.46E+05	5.43E+05	4.30E+04	2.91E+04	6.74E+05
a-cadinol																	
a-farnese																	
alloocimene	121.0682	6.57	3.31E+08	8.53E+05	3.97E+08	1.41E+09	1.61E+09	1.93E+09	6.48E-04	1.46E+06	9.46E+07	4.41E+07	2.01E+06	1.16E+09	2.80E+07	2.85E+07	1.25E+09
a-pinene	93.0700	5.03	1.40E+09	2.35E+07	2.16E+09	1.54E+06	2.97E+06	1.13E+07	1.57E+09	2.16E+06	1.17E+06	2.61E+05	9.23E+04	7.97E+06	1.63E+04	2.37E+04	1.41E+06
aromadendrene	105.0699	6.87	3.48E+07	6.20E+07	1.78E+08	3.88E+07	4.82E+06	1.44E+07	2.61E+07	3.30E+07	2.29E+07	1.35E+07	7.69E+06	8.64E+08	1.16E+06	3.30E+05	4.32E+07
a-terpineol	136.0889	7.92	2.61E+07	3.54E+07	7.00E+07	8.26E+06	3.12E+06	2.59E+06	2.70E+07	3.80E+07	1.56E+08	4.40E+07	6.08E+08	3.45E+08	1.19E+06	1.28E+06	8.60E+06
b-elemene	95.0807	5.02	3.05E+07	4.76E+06	4.81E+07	1.14E+07	8.91E+05	8.22E+06	3.22E+07	5.44E+05	5.29E+06	2.24E+05	2.36E+05	6.64E+05	8.36E+04	1.02E+04	5.07E+05
biscyclogermacrene																	
borneol	95.0807	5.02	3.05E+07	4.76E+06	4.81E+07	1.14E+07	8.91E+05	8.22E+06	3.22E+07	5.44E+05	5.29E+06	2.24E+05	2.36E+05	6.64E+05	8.36E+04	1.02E+04	5.07E+05
b-pinene	121.0882	5.09	4.72E+08	9.46E+06	5.00E+06	2.17E+05	4.58E+05	2.08E+05	2.48E+06	8.68E+05	2.61E+05	8.60E+04	9.99E+05	5.41E+06	9.09E+03	4.40E+03	4.96E+04
b-selinene	107.0827	5.14	4.87E+06	6.62E+06	2.38E+08	5.92E+05	7.28E+05	4.55E+05	1.74E+08	1.33E+06	1.28E+06	4.82E+05	4.06E+04	2.35E+06	1.46E+04	1.18E+04	7.45E+05
camphene	121.0882	5.09	4.72E+08	9.46E+06	5.00E+06	2.17E+05	4.58E+05	2.08E+05	2.48E+06	8.68E+05	2.61E+05	8.60E+04	9.99E+05	5.41E+06	9.09E+03	4.40E+03	4.96E+04
camphor	95.0669	5.76	6.35E+05	2.85E+07	7.01E+07	2.61E+04	4.44E+06	6.70E+06	9.38E+07	1.43E+07	7.82E+01	1.90E+06	3.80E+05	1.41E+08	7.25E+04	3.14E+04	6.22E+07
carene																	
carvone	108.0717	4.92	3.34E+07	5.28E+06	6.06E+07	1.73E+05	1.63E+04	3.00E+04	5.61E+07	2.80E+06	3.91E+06	9.15E+04	1.35E+04	1.60E+06	6.70E+03	7.24E+03	1.40E+06
caryophyllene	133.0155	3.13	3.91E+04	1.72E+05	1.26E+05	5.92E+05	4.89E+04	1.01E+05	9.90E+04	1.45E+05	2.31E+05	1.94E+06	1.97E+05	1.18E+05	2.13E+05	7.58E+04	1.28E+05
chavicol	167.1010	6.07	3.56E+04	3.50E+04	5.72E+04	1.01E+05	4.66E+04	5.27E+04	4.79E+04	2.89E+05	3.97E+04	4.66E+04	8.92E+03	1.68E+05	4.10E+03	4.13E+03	1.98E+05

copaene																	
elemene	81.0699	9.96	3.55E+06	5.28E+07	1.79E+07	3.09E+06	1.76E+07	9.70E+06	2.17E+07	1.02E+07	6.32E+06	4.83E+07	5.72E+07	1.82E+07	6.79E+04	7.93E+04	1.38E+07
elemol	59.0458	7.33	7.61E+05	1.13E+06	1.38E+06	1.25E+05	2.03E+06	1.47E+05	3.54E+06	1.52E+05	3.49E+05	1.06E+06	4.11E+06	1.85E+06	2.97E+04	9.89E+03	1.84E+06
estragole	148.0874	5.74	1.81E+07	1.58E+07	4.85E+08	1.13E+06	2.07E+04	1.44E+04	4.57E+07	2.13E+07	2.29E+06	1.41E+05	4.41E+04	4.38E+06	2.36E+03	4.29E+03	5.14E+05
eucalyptol (1-8,cineole)	81.0699	9.87	5.97E+06	5.41E+07	1.96E+07	3.64E+06	1.77E+07	2.57E+06	2.87E+07	1.47E+06	2.54E+06	1.46E+07	1.85E+08	1.82E+07	4.73E+04	7.93E+04	1.53E+06
eugenol	164.0827	6.39	5.09E+08	4.72E+08	5.09E+08	5.97E+03	3.00E+09	3.51E+09	2.30E+08	1.47E+08	1.49E+08	5.37E+07	7.00E+07	2.37E+09	1.19E+02	1.20E+08	3.19E+09
fenchone	81.0699	9.87	5.97E+06	5.41E+07	1.96E+07	3.64E+06	1.77E+07	2.57E+06	2.87E+07	1.47E+06	2.54E+06	1.46E+07	1.85E+08	1.82E+07	4.73E+04	7.93E+04	1.53E+06
geraniol	139.0898	9.96	1.24E+06	1.18E+06	4.71E+05	6.00E+04	1.79E+05	4.85E+04	2.88E+06	2.35E+07	2.38E+07	1.29E+06	9.28E+05	1.38E+05	4.76E+03	2.28E+03	7.71E+04
germacrene-D	119.0856	7.61	2.40E+06	8.71E+06	2.34E+08	1.72E+07	2.54E+07	5.83E+08	1.75E+08	1.96E+08	4.05E+08	4.26E+07	2.06E+08	2.84E+08	1.56E+06	9.22E+05	3.26E+07
limonene	67.0542	5.98	1.38E+08	1.38E+08	1.55E+08	3.92E+05	5.66E+05	2.68E+07	1.78E+08	9.96E+07	5.79E+07	5.96E+06	3.88E+05	6.83E+06	3.81E+04	1.92E+05	4.46E+05
linalool	93.0699	4.30	1.51E+06	8.16E+06	1.06E+07	3.45E+05	1.36E+06	2.66E+05	8.11E+06	4.83E+06	3.38E+06	5.17E+05	2.97E+04	9.30E+04	3.39E+04	4.69E+04	1.39E+05
menthol	71.0545	9.88	4.71E+05	3.21E+07	7.35E+06	2.77E+05	1.95E+06	5.41E+06	5.41E+06	1.18E+07	1.19E+07	2.48E+07	2.91E+04	1.08E+07	1.45E+04	3.90E+03	8.10E+06
methyl cinnamate	94.0716	5.60	1.56E+07	1.97E+06	8.28E+07	9.52E+06	1.61E+07	7.08E+06	1.20E+08	5.41E+06	1.09E+07	5.37E+05	5.27E+04	1.33E+06	8.02E+03	4.32E+03	7.66E+05
myrcene	93.0699	4.30	1.51E+06	8.16E+06	1.06E+07	3.45E+05	1.36E+06	2.66E+05	8.11E+06	4.83E+06	3.38E+06	5.17E+05	2.97E+04	9.30E+04	3.39E+04	4.69E+04	1.39E+05
ocimene	93.0700	5.03	1.40E+09	2.35E+07	2.16E+09	1.54E+06	2.97E+06	1.13E+07	1.57E+09	2.16E+06	1.17E+06	2.61E+05	9.23E+04	7.97E+06	1.63E+04	2.37E+04	1.41E+06
sabinene	93.0700	4.90	1.25E+09	1.57E+07	1.32E+09	1.54E+06	5.25E+05	5.96E+05	1.33E+09	2.41E+06	1.62E+06	1.32E+05	9.23E+04	5.37E+06	1.63E+04	6.00E+04	1.02E+06
tau-cadinol	161.1162	7.02	4.29E+05	4.71E+06	5.70E+07	1.40E+07	2.44E+06	3.21E+07	1.24E+06	8.76E+05	3.86E+05	4.52E+06	3.44E+06	2.58E+08	4.28E+05	1.27E+05	1.50E+07
terpinene	136.0888	8.01	2.55E+07	1.27E+08	6.94E+07	8.06E+06	3.62E+07	1.46E+07	1.12E+08	2.17E+08	1.55E+08	4.20E+08	8.01E+07	3.45E+08	1.19E+06	1.28E+06	2.18E+08
terpinolene	149.0726	6.58	5.42E+06	5.82E+08	8.20E+06	2.02E+09	2.45E+09	2.89E+09	3.33E+08	4.43E+06	3.15E+06	3.50E+06	6.38E+05	1.82E+09	4.68E+07	1.50E+06	3.20E+04
y-cadinene	188.0845	9.46	1.89E+06	1.05E+07	4.05E+06	6.14E+03	2.89E+04	3.37E+03	4.36E+06	1.21E+07	1.05E+07	1.42E+04	4.20E+04	9.06E+03	3.98E+02	4.63E+02	1.88E+04
y-muurolene																	

Supplemental Table 5: Raw peak area data for tentatively identified key *Ocimum gratissimum* compounds in GC-SQ data.

Compound	***m/z***	**RT (min)**	**OG 1.1**	**OG 1.2**	**OG 1.2**	**OG 2.1**	**OG 2.2**	**OG 2.3**	**OG 3.1**	**OG 3.2**	**OG 3.3**	**OG 4.1**	**OG 4.2**	**OG 4.3**
3-hexanol	73.1	4.34	2.99E+03	4.29E+03	4.36E+03	4.24E+03	4.14E+03	4.23E+03	3.85E+03	5.91E+03	4.79E+03	3.68E+03	3.44E+03	6.10E+03
a-bergamotene	204.2	8.08	3.68E+01	3.73E+03	3.23E+03	1.32E+03	3.49E+02	1.57E+02	4.73E+02	0.00E+00	2.22E+03	1.47E+04	5.03E+03	1.13E+04
a-bisabolene	119.1	8.13	5.35E+02	7.23E+03	4.74E+03	2.22E+03	8.07E+02	4.16E+02	6.20E+02	3.53E+02	1.34E+03	3.98E+03	1.43E+03	5.99E+03
a-bulnesene	93.1	7.63	2.25E+02	3.44E+03	9.05E+03	1.85E+02	9.50E+01	7.21E+01	2.13E+03	2.80E+02	8.11E+03	5.32E+04	6.95E+03	4.03E+04
a-cadinol	204.2	8.79	1.33E+02	1.13E+04	1.27E+03	1.55E+02	2.77E+02	3.74E+02	7.02E+02	5.63E+02	6.90E+03	4.25E+03	4.12E+03	1.28E+04
a-farnese														
alloocimene														
a-pinene														
aromadendrene	105.1	8.04	3.66E+02	1.06E+03	2.60E+03	2.78E+03	5.92E+02	4.00E+02	8.11E+02	2.10E+02	1.63E+03	5.70E+03	6.47E+03	1.25E+04
a-terpineol	93.1	6.25	2.06E+03	2.90E+03	4.62E+03	1.85E+02	3.45E+01	9.31E+01	2.01E+03	6.46E+01	1.99E+03	6.95E+02	8.19E+01	2.66E+02

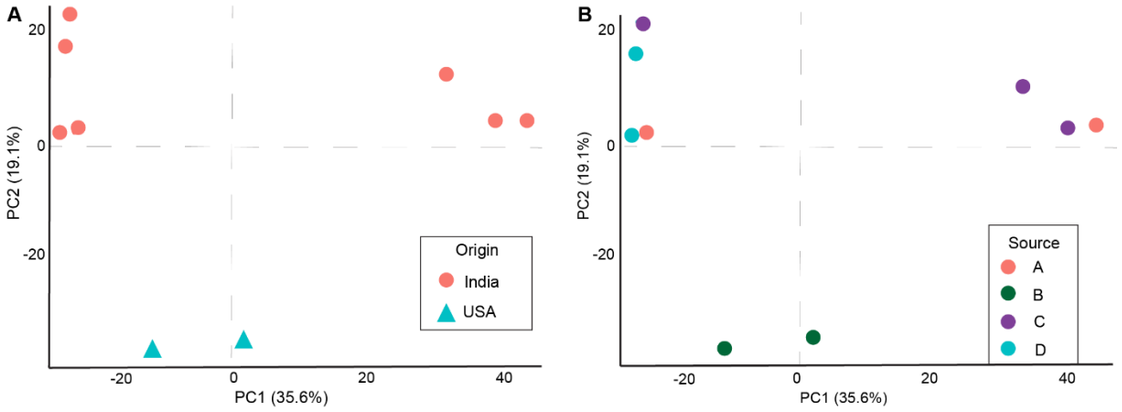
b-elemene	81.1	7.32	9.43E+02	2.48E+03	1.98E+03	3.09E+04	2.59E+04	3.06E+04	2.06E+03	5.40E+02	3.19E+03	3.60E+04	7.32E+03	3.07E+04
biscyclogermacrene														
borneol														
b-pinene	69.1	8.09	3.90E+03	1.88E+04	4.68E+03	5.32E+02	2.52E+02	1.21E+02	1.17E+04	1.25E+04	1.42E+04	7.16E+02	3.18E+02	7.01E+03
b-selinene	93.1	8.83	2.13E+03	1.47E+04	5.51E+03	3.36E+03	9.84E+03	6.53E+03	8.32E+03	1.13E+04	1.52E+04	2.79E+04	2.74E+04	8.41E+04
camphene	121.1	7.98	1.48E+02	2.01E+03	8.14E+02	1.64E+02	3.81E+02	1.39E+02	4.70E+03	6.50E+03	5.05E+03	4.70E+03	1.35E+03	6.48E+03
camphor														
carene	93.1	8.13	1.01E+03	1.77E+04	1.55E+04	1.73E+03	1.29E+03	8.46E+02	4.07E+03	2.80E+02	1.53E+04	9.28E+03	1.50E+03	1.51E+04
carvone														
caryophyllene	91.1	8.47	5.58E+03	3.52E+04	2.21E+04	7.01E+03	2.18E+04	1.40E+04	1.88E+04	4.21E+02	4.67E+04	5.07E+04	4.75E+04	1.56E+05
chavicol	134.1	6.53	2.56E+04	6.38E+04	3.18E+04	0.00E+00	5.25E+01	0.00E+00	2.11E+04	1.69E+02	2.51E+04	1.52E+03	6.53E+01	6.69E+02
copaene														
elemene	81.1	7.32	9.43E+02	2.48E+03	1.98E+03	3.09E+04	2.59E+04	3.06E+04	2.06E+03	5.40E+02	3.19E+03	3.60E+04	7.32E+03	3.07E+04
elemol	189.2	8.78	6.82E+01	2.80E+03	2.09E+02	4.12E+01	9.91E+01	8.25E+01	2.71E+02	2.15E+02	2.37E+03	3.15E+03	3.42E+03	1.23E+04
estragole	148.1	6.29	6.43E+03	3.77E+02	3.87E+04	0.00E+00	0.00E+00	3.60E+01	1.76E+03	0.00E+00	2.45E+03	3.84E+03	1.95E+01	2.03E+01
eucalyptol (1-8,cineole)														
eugenol	164.1	7.20	6.22E+04	9.18E+04	8.20E+04	1.46E+06	1.23E+06	1.42E+06	8.26E+04	7.08E+02	1.29E+05	5.31E+05	4.33E+05	1.08E+06
fenchone														
geraniol	69.1	6.69	2.29E+04	3.80E+04	4.92E+04	7.89E+01	2.33E+02	1.90E+02	1.88E+04	4.46E+03	3.16E+04	2.20E+03	2.40E+02	4.86E+02
germacrene-D	161.1	8.84	1.51E+02	1.71E+04	7.06E+02	1.25E+03	1.78E+03	7.03E+02	7.67E+03	5.55E+03	9.55E+03	2.18E+03	2.82E+03	9.84E+03
limonene	136.1	8.29	7.62E+01	1.99E+03	1.48E+03	3.58E+02	3.68E+02	1.10E+03	2.27E+02	7.77E+01	4.22E+02	1.43E+02	1.61E+02	6.29E+03
Linalool	71.1	5.57	1.12E+05	2.31E+03	3.39E+05	8.71E+01	8.03E+01	3.94E+01	1.55E+04	1.13E+02	1.81E+04	7.34E+03	3.64E+01	1.99E+02
menthol														
methyl cinnamate														
myrcene	93.1	6.40	2.91E+03	4.10E+03	7.45E+03	1.85E+02	3.45E+01	9.31E+01	2.82E+03	6.46E+01	1.99E+03	4.11E+02	5.47E+01	7.43E+01
ocimene	93.1	8.05	1.05E+03	1.78E+04	1.49E+04	1.73E+03	2.06E+02	2.52E+02	4.17E+03	2.80E+02	1.38E+04	1.23E+04	1.84E+03	1.46E+04
sabinene														
tau-cadinol	161.1	8.84	1.51E+02	1.71E+04	7.06E+02	1.25E+03	1.78E+03	7.03E+02	7.67E+03	5.55E+03	9.55E+03	2.18E+03	2.82E+03	9.84E+03
terpinene	93.1	8.13	0.00E+00	2.01E+03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.60E+04	4.70E+03	0.00E+00	0.00E+00	0.00E+00
terpinolene														
y-cadinene														
y-muurolene	105.1	8.79	2.50E+03	1.10E+04	2.64E+03	3.40E+03	1.01E+04	3.61E+03	5.03E+03	8.20E+03	9.35E+03	9.79E+03	8.74E+03	3.44E+04

Supplemental Table 6: Raw peak area data for tentatively identified key *Ocimum tenuiflorum* compounds in GC-SQ data.

Compound	***m/z***	**RT (min)**	**OT 1.1**	**OT 1.2**	**OT 1.3**	**OT 2.1**	**OT 2.2**	**OT 2.3**	**OT 3.1**	**OT 3.2**	**OT 3.3**	**OT 4.1**	**OT 4.2**	**OT 4.3**	**OT 5.1**	**OT 5.2**	**OT 5.3**
3-hexanol	73.1	4.34	3.46E+03	4.14E+03	4.66E+03	1.90E+03	4.16E+03	4.72E+03	3.84E+03	4.58E+03	5.31E+03	5.48E+03	5.80E+03	5.22E+03	4.33E+03	4.25E+03	3.96E+03
a-bergamotene	204.2	8.08	7.20E+02	1.60E+02	1.49E+03	4.53E+03	4.76E+03	1.06E+04	4.69E+02	2.10E+02	2.70E+02	1.36E+02	9.61E+01	1.38E+04	2.67E+03	3.34E+03	7.36E+02
a-bisabolene	119.1	8.13	7.89E+02	3.56E+02	1.83E+03	6.95E+03	1.07E+04	1.49E+04	1.58E+03	2.02E+03	2.75E+03	4.75E+02	1.23E+03	7.79E+03	1.82E+03	4.19E+02	2.91E+02
a-bulnesene	93.1	7.63	3.01E+03	1.54E+02	1.00E+04	2.89E+03	1.34E+03	4.44E+02	5.40E+03	8.93E+02	4.69E+02	6.90E+01	8.32E+01	9.50E+04	7.82E+03	3.00E+03	1.71E+03
a-cadinol	204.2	8.79	3.13E+03	1.28E+04	6.66E+03	1.99E+02	8.76E+03	3.09E+03	5.79E+02	6.46E+03	2.92E+03	1.44E+04	1.81E+04	9.41E+03	6.98E+03	4.44E+03	6.40E+03
a-farnese																	
alloocimene																	
a-pinene																	
aromadendrene	105.1	8.04	1.20E+03	6.28E+02	9.48E+03	5.59E+03	3.97E+03	1.13E+04	7.59E+02	9.50E+02	8.77E+02	2.21E+02	1.47E+03	1.49E+04	7.54E+03	4.79E+03	2.96E+03
a-terpineol	93.1	6.25	1.56E+02	7.58E+01	1.74E+02	2.75E+03	3.60E+01	9.46E+01	2.87E+03	9.39E+01	5.76E+02	2.14E+01	3.90E+01	2.39E+02	1.70E+02	1.89E+02	2.10E+02
b-elemene	81.1	7.32	5.19E+03	1.27E+04	1.06E+04	1.77E+04	1.50E+04	2.41E+04	1.92E+03	9.01E+02	2.24E+03	2.31E+03	6.70E+02	6.05E+04	2.76E+04	4.27E+03	3.04E+04
biscyclogermacrene																	
borneol																	
b-pinene	69.1	8.09	3.21E+02	4.84E+02	4.48E+02	2.17E+04	1.47E+04	4.47E+04	7.64E+03	2.85E+04	2.18E+04	1.25E+03	6.46E+02	2.83E+03	3.53E+02	2.45E+02	1.81E+02
b-selinene	93.1	8.83	1.70E+04	6.18E+04	2.90E+04	2.73E+03	1.90E+04	7.75E+03	6.25E+03	1.60E+04	1.48E+04	5.35E+04	3.48E+04	4.37E+04	3.05E+04	2.51E+04	1.35E+04
camphene	121.1	7.98	5.39E+02	3.39E+02	2.81E+03	3.70E+03	1.19E+03	1.30E+04	3.45E+03	1.18E+04	1.09E+04	7.32E+01	2.75E+02	1.60E+04	4.60E+03	1.23E+03	9.21E+02
camphor																	
carene	93.1	8.13	5.17E+02	5.12E+03	4.82E+03	1.83E+04	3.03E+04	5.35E+04	6.12E+03	2.60E+03	2.43E+03	7.82E+02	3.76E+03	1.35E+04	2.67E+03	2.18E+03	5.95E+02
carvone																	
caryophyllene	91.1	8.47	1.60E+04	4.04E+04	3.77E+04	2.24E+03	2.60E+04	1.35E+04	1.40E+04	2.03E+04	2.32E+04	2.13E+03	1.76E+04	5.35E+04	3.09E+04	2.28E+04	2.16E+04
chavicol	134.1	6.53	3.33E+02	3.33E+02	4.73E+02	6.16E+04	6.13E+04	6.68E+04	1.93E+04	3.36E+04	3.13E+04	2.48E+01	0.00E+00	1.53E+03	1.94E+02	6.12E+01	3.80E+02
copaene																	
elemene	81.1	7.32	5.19E+03	1.27E+04	1.06E+04	1.77E+04	1.50E+04	2.41E+04	1.92E+03	9.01E+02	2.24E+03	2.31E+03	6.70E+02	6.05E+04	2.76E+04	4.27E+03	3.04E+04
elemol	189.2	8.78	3.06E+03	8.88E+03	6.14E+03	2.23E+02	7.74E+02	2.15E+02	2.14E+02	1.90E+03	2.12E+03	1.31E+04	1.35E+04	7.78E+03	5.61E+03	5.07E+03	5.33E+03
estragole	148.1	6.29	5.63E+02	2.52E+01	2.64E+02	2.35E+03	0.00E+00	0.00E+00	2.65E+04	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.17E+02	1.07E+02	0.00E+00	0.00E+00
eucalyptol (1-8,cineole)																	
eugenol	164.1	7.20	1.21E+06	6.16E+05	1.56E+06	8.91E+05	7.17E+05	1.15E+06	3.21E+04	1.86E+04	2.44E+04	7.64E+03	8.60E+03	4.90E+05	1.30E+06	1.14E+06	1.41E+06
fenchone																	
geraniol	69.1	6.69	1.37E+02	9.23E+01	1.85E+02	6.42E+01	7.01E+01	1.78E+02	6.52E+04	3.86E+04	2.83E+04	1.54E+02	7.55E+01	3.26E+02	1.69E+02	1.05E+02	2.39E+02
germacrene-D	161.1	8.84	8.69E+02	3.01E+03	2.76E+03	4.01E+02	1.53E+03	1.33E+03	3.38E+02	1.38E+04	1.50E+04	5.62E+03	6.23E+03	7.06E+03	1.81E+03	3.95E+03	2.80E+03
limonene	136.1	8.29	7.48E+02	1.79E+02	4.02E+02	5.67E+02	1.17E+03	1.81E+03	6.85E+02	6.35E+02	1.39E+03	8.33E+01	1.85E+02	4.43E+02	1.64E+03	5.41E+01	1.23E+02
Linalool	71.1	5.57	4.94E+03	3.79E+01	1.14E+03	3.22E+01	1.07E+02	4.32E+01	1.49E+05	7.51E+01	1.54E+02	7.65E+01	1.05E+02	9.67E+02	4.87E+02	4.69E+01	1.38E+02
menthol																	

methyl cinnamate																	
myrcene	93.1	6.40	1.56E+02	1.22E+02	1.74E+02	2.75E+03	1.99E+02	3.32E+02	2.82E+03	9.39E+01	5.76E+02	2.14E+01	2.78E+01	2.39E+02	1.70E+02	4.24E+01	1.48E+02
ocimene	93.1	8.05	5.17E+02	3.83E+03	9.60E+03	2.22E+04	2.75E+04	5.30E+04	2.27E+03	2.60E+03	2.43E+03	8.02E+02	3.76E+03	2.21E+04	2.67E+03	4.71E+03	3.17E+03
sabinene																	
tau-cadinol	161.1	8.84	8.69E+02	3.01E+03	2.76E+03	4.01E+02	1.53E+03	1.33E+03	3.38E+02	1.38E+04	1.50E+04	5.62E+03	6.23E+03	7.06E+03	1.81E+03	3.95E+03	2.80E+03
terpinene	93.1	8.13	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3.45E+03	1.18E+04	1.09E+04	0.00E+00	6.50E+03	5.05E+03	4.60E+03
terpinolene																	
y-cadinene																	
y-muurolene	105.1	8.79	4.98E+03	1.78E+04	9.85E+03	7.82E+02	7.81E+03	4.48E+03	1.14E+04	1.01E+04	3.59E+04	2.20E+04	2.80E+04	1.42E+04	9.95E+03	7.42E+03	8.80E+03

Supplemental Figure 1: Unsupervised PCA comparing *O. tenuiflorum* and *O. gratissimum* samples' geographic origin (A) and supplier source (B) variation using the full metabolite profile from the GC-SQ data.



Supplemental Figure 2: Unsupervised PCA comparing *O. tenuiflorum* and *O. gratissumum* samples' geographic origin (A) and supplier source (B) variation using the full metabolite profile from the GC-Orbitrap data.