



Comprehensive analysis of commercial biopesticides using UHPLC and GC-HRMS: Targeted, suspect and unknown component determination

Alba Reyes-Ávila, Roberto Romero-González, F. Javier Arrebola-Liébanas, Antonia Garrido French*

Research Group "Analytical Chemistry of Contaminants", Department of Chemistry and Physics, Research Centre for Mediterranean Intensive Agrosystems and Agrifood Biotechnology (CIAMBITAL), Agrifood Campus of International Excellence (ceiA3), University of Almería, 04120 Almería, Spain

ARTICLE INFO

Keywords:

Commercial biopesticides
Targeted analysis
Untargeted analysis
Chromatography
HRMS

ABSTRACT

15 commercial biopesticides (CBs), based on vegetable extracts and oils, were characterized. Ultra-high-performance liquid chromatography (UHPLC) and gas chromatography (GC) coupled to a hybrid high-resolution mass spectrometry (HRMS) analyser, such as quadrupole (Q)-Orbitrap, were used. Both targeted and untargeted (suspect and unknown modes) analyses were carried out. For the targeted analysis, a database was built, encompassing volatile and non-volatile compounds commonly found in vegetable extracts and oils. This database comprised 27 LC-amenable compounds and 31 GC-amenable compounds. 17 targeted compounds were quantified, and 101 compounds were tentatively identified by untargeted analysis. CBs based on essential oils, such as orange oil or cinnamon extract contained the highest number of detected compounds. Monoterpenes (limonene or linalool), and sesquiterpenes (δ -cadinene or γ -elemene) were mainly found. Moreover, some co-formulants such as dibutyl phthalate or butylated hydroxytoluene (BHT) were also detected. The concentration of targeted natural compounds ranged from 2 mg/L (linalool) to 450 g/L (*trans*-cinnamaldehyde).

1. Introduction

Pesticides have been used for decades to eradicate pests and weeds in crops worldwide. However, their abusive and indiscriminate use has led to insect resistance and environmental contamination in soil and groundwater [1]. In recent years, efforts have been made to restrict the uncontrolled use of pesticides and employ more specific compounds to targeted specific pests. In addition, natural pesticides or biopesticides are emerging as an alternative to conventional pesticides [2]. According to the U.S. Environmental Protection Agency (EPA), biopesticides are certain types of pesticides derived from such natural materials as animals, plants, bacteria, and certain minerals [3]. There are different types of biopesticides that are classified according to their extraction sources and the type of compound used in their preparation. The most commonly used are derived from plants, including essential oils or plant extracts. Essential oils are mostly composed of volatile compounds such as monoterpenes, and their composition varies depending on the plant's origin. They usually have various functions, such as attracting or repelling insects due to their high content of natural aromatic substances, including volatile and non-volatile organic compounds [4,5].

However, the availability of commercial biopesticides (CBs) is limited due to the stringent marketing restrictions they face. For instance, in the European Union (EU), the number of restrictions is higher compared to other countries such as United States (USA) [6]. This is because CBs in the EU are evaluated under the same regulations as synthetic pesticides. In addition, the number of active substances in biopesticides recognized by EU is much lower than in other countries, resulting in a lower variety of CBs that can be manufactured. Therefore, prior research is needed to determine the most suitable compounds or mixtures of compounds for effectively eliminating different pests and producing more precise CBs. Some studies have explored the use of compounds present in essential oils or plant extracts to combat various types of pests, such as cabbage looper [7], *Spodoptera littoralis* larvae [8,9], mosquitoes [10,11], and other insects [12–15]. The efficacy of binary mixtures of aromatic compounds against some insects was also evaluated. It has been seen that the synergistic or antagonistic effects are influenced by the molecular structure, type, and position of the functional groups, and by the relative proportions of individual substances in the mixture [9].

On the other hand, the labelling of CBs does not provide much

* Corresponding author.

E-mail address: agarrido@ual.es (A. Garrido French).

<https://doi.org/10.1016/j.microc.2023.109020>

Received 27 April 2023; Received in revised form 16 June 2023; Accepted 21 June 2023

Available online 25 June 2023

0026-265X/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

information about the compounds present in CBs, and only the active principle is commonly specified. For example, in CBs based on essential oils, only the percentage of essential oils is provided. Nevertheless, it would be necessary to enhance the knowledge about the composition of CBs to prevent potential harmful effects on the environment and crops when they are applied. In addition, CBs against more specific pests could be applied with greater precision.

In relation to the characterization of CBs, several articles have focused on determining the amount of only one or two major compounds in commercial products based on plant extracts. To analyze less volatile compounds like alkaloids (matrine) [16] or limonoids (azadirachtin) [17] ultra-high-performance liquid chromatography (UHPLC) has been employed. On the other hand, gas chromatography (GC) was used for the analysis of volatile compounds, such as carvacrol [18] or dimethyl disulfide [19]. The detectors used were photodiode array detector [17] or ultraviolet detector (UV) [16,20] when UHPLC was utilized, and flame ionization detector (FID) [18,19,21] for GC. Considering that these classical detectors are less selective than mass spectrometry (MS) analysers, the use of MS detection could increase the reliability of the analysis. Up to now, classical detectors have been mainly used, so the application of MS could improve the detection of this type of components. In this sense, the use of high-resolution mass analyzers is a novelty as it has not been explored in previous studies. This enables the precise detection of compounds with improved sensitivity compared to the detectors commonly used (FID or UV). Moreover, a wider variety of compounds can be determined, from major compounds to those present at trace levels, increasing the scope of the analysis in comparison with previous studies [16–21], where only one or two of the major compounds were determined. For this purpose, high resolution mass spectrometry (HRMS) was employed in this study, using a Q-Orbitrap analyzer, carrying out both targeted and non-targeted analysis, with suspect and unknown modes, on different types of CBs based on plant extracts. In addition, to achieve a comprehensive view of CB composition, two chromatographic methodologies based on UHPLC and GC were used. Furthermore, different injection and sample preparation modes, such as direct injection (DI), headspace (HS) and solid-phase microextraction (SPME), were evaluated when GC was used. This made possible a more comprehensive analysis comparing the compounds obtained by the different techniques used.

2. Materials and methods

2.1. Materials

Fifteen commercial biopesticides based on essential oils and plant extracts were acquired (Table S1). These CBs were: Prevam® (ORO AGRI; Palmela, Portugal), Cureneem (BIO flower; Tárrega, Spain), Evomax (Biagro; Massalfassar, Spain), Evo Plant, Spruzit RTU (NEUDROFF; Emmerthal, Germany), NeemPro® (TRIFOLIO-M GmbH; Lahnau, Germany), cinnamon extract (jBQ®), field horsetail (ASOCOA®), nettle extract (ASOCOA®) and two others with unknown origin. Cinna, Notrip, Mimset and Scar were provided by Hortalan (El Ejido, Spain).

Methanol (MeOH, ≥99.9%), ethyl acetate (EtOAc, ≥99.7%), acetonitrile (≥99.9%), dimethyl sulfoxide (DMSO, ≥99.7%) and acetone (≥99.8%), all HPLC grade, were provided from Honeywell (Charlotte, USA); water (H₂O, LiChrosolv®), ethanol (EtOH, HPLC grade) and formic acid (FA, LC-MS, 99.0%) supplied from Merck (Darmstadt, Germany), J.T. Baker (Poland) and Fischer Scientific (Hampton, USA), respectively.

Analytical standards of camphor, citronellal and 1,8-cineole were provided by Alfa Aesar (Ward Hill, USA); ricinine and patchouli alcohol by Biosynth Carbosynth (Berkshire, UK); azadirachtin and α -solamargine by Chengdu Biopurify Phytochem (Chengdu, China); biphenyl, veratridine and *trans*-cinnamaldehyde by Dr. Ehrenstorfer (Augsburg, Germany); tomatine, α -solanine and α -(-)-thujone by Extrasynthese (Genay, France); nerol, α -phellandrene and myrcene by Fluka;

cevadine and sabinene by Phytolab (Vestenbergsgreuth, Germany); and rotenone, linalyl acetate, carvacrol, R(-)-carvone, β -citronellol and thymol by Tokyo Chemical Industry (Tokyo, Japan). Moreover, (1S)-(-)- β -pinene, geraniol, acetyleugenol, (+)-pulegone, pyrethrum extract, linalool, *trans*-anethole, eugenol, estragole, isoborneol, isoeugenol, (\pm)-menthol, (-)-nicotine, γ -terpinene, m-cymene, myristicin, oleamide and (R)-(+)-limonene were purchased from Sigma Aldrich (Saint Louis, USA).

Stock solutions of these compounds were prepared at 1000 mg/L in EtOAc, except for α -solamargine, which was dissolved in DMSO; tomatine, α -solanine and azadirachtin in MeOH; and linalool, linalyl acetate, nerol and oleamide in EtOH. Then, solutions at 10 and 1 mg/L in EtOAc were prepared from the stock solution. All solutions were stored at -18 °C.

2.2. UHPLC-Q-Orbitrap-MS method

The chromatographic equipment was a Vanquish™ Flex Quaternary LC coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, USA). The column utilized was a ZORBAX Eclipse Plus C8 (2.1 × 100 mm, 1.8 μ m) provided by Agilent (Santa Clara, USA). Other columns used were a Hypersil GOLD™ aQ (2.1 × 100 mm, 1.9 μ m) and a Hypersil GOLD Phenyl (2.1 × 100 mm, 1.9 μ m) by Thermo Fisher Scientific. The column temperature was set at 30 °C. The chromatographic conditions were carried out at a flow rate of 0.2 mL/min and an injection volume of 10 μ L. The mobile phase was composed of an aqueous solution of formic acid (0.1%) as aqueous phase (phase A) and, MeOH as organic phase (phase B). Elution was carried out using a gradient mode: constant composition of 5% B from 0 to 2 min; increase up to 100% B from 2 min to 16 min; constant composition of 100% B from 16 min to 26 min; decrease to 5 % B from 26 to 27 min. Finally, keep constant this composition for 3 min to equilibrate the column. The total running time was 30 min. The acquisition mode used was full scan and data-dependent acquisition (DDA), in positive and negative ionization modes. The mass/charge (m/z) range used in full scan was 74–1100 with a resolution of 70,000 full width at half maximum (FWHM), and an automatic gain control (AGC) value was set at 10^6 . DDA was used to monitor fragments of the compounds, and the resolution was 35,000 FWHM with an AGC value equal to 10^5 . Minimum AGC target value was $8 \cdot 10^3$. To ensure that MS/MS spectrum of targeted compounds is obtained, an inclusion list was used where the m/z of the precursor ion and the ionization mode were recorded for each compound. For electrospray interface conditions, heater temperature was 305 °C and capillary temperature, 300 °C. The auxiliary and sheath gas used were N₂ (95%), the spray voltage was 4 kV, and the S-lens radio frequency level was 50 (arbitrary units).

Samples were diluted to a ratio of 1:1,000,000 (v/v). For this, 40 μ L of each CB was taken and diluted in 4 mL of water, resulting in a 1:100 (v/v) dilution. Then, a 10 μ L aliquot of that dilution was taken and dissolved in 990 μ L of MeOH, achieving a 1:10,000 (v/v) dilution. This step was repeated, resulting in a final dilution of 1:1,000,000 (v/v). For Cureneem, EtOAc was used instead of water to dissolve it. Before injection of the samples, they were filtered using a nylon filter (13 mm, 0.22 μ m; LLG-Labware, Meckenheim, Germany). To ensure the confidence of the results, three replicates of each CB were prepared.

2.3. GC-Q-Orbitrap-MS method

The chromatographic equipment consisted of a TRACE™ 1310 GC system with a TriPlus™ RSH autosampler (Thermo Scientific™) coupled to a Q-Exactive Orbitrap (Thermo Fisher Scientific) mass spectrometer. A non-polar column (phenyl arylene polymer) J&W DB-5 ms (30 m × 0.25 mm × 0.25 μ m), provided by Agilent Technologies, was chosen for the chromatographic separation. Helium was used as carrier gas with a constant flow rate of 1 mL/min. For MS conditions, full scan in positive mode was used, with a 30–450 m/z range. A 70 eV positive

electron ionization (EI) was employed. The resolution was 70,000 FWHM, and an AGC value equal to 10^6 . Three replicates of each CB were prepared for each injection mode.

2.3.1. Direct injection (DI)

The samples were diluted to a 1:1,000,000 (v/v) dilution. The dilution process followed a similar procedure as described in Section 2.2, but using EtOAc instead of MeOH. However, for Evomax, Evo Plant, Notrip and the two unknown CBs, DMSO was used instead of EtOAc to dissolve them. For MIMSET, acetone was used. Samples were filtered before injection with a nylon filter (13 mm, 0.22 μm ; LLG-Labware).

The injection volume was 1 μL . For chromatographic conditions, initial oven temperature was set at 60 $^{\circ}\text{C}$, and it was maintained for 2 min. It was then increased at a rate of 6 $^{\circ}\text{C}/\text{min}$ to 220 $^{\circ}\text{C}$, and it was kept for 2 min. Finally, it was raised to 280 $^{\circ}\text{C}$ at a rate of 20 $^{\circ}\text{C}/\text{min}$, and it was held for 4 min. The total running time was 37 min.

2.3.2. Headspace (HS)

The samples were diluted to a 1:1,000 (v/v) dilution. To achieve this, 10 μL of each CB was taken and diluted in 10 mL of water, achieving a 1:1,000 (v/v) dilution. However, for CURENEEM, EtOAc was used instead of water to dissolve it.

The samples were then incubated for 20 min at 60 $^{\circ}\text{C}$, and they were agitated every 10 s. The injection volume was set to 1 mL. For chromatographic conditions, the initial oven temperature was 60 $^{\circ}\text{C}$, and it was maintained for 2 min. Then it was increased at a rate of 6 $^{\circ}\text{C}/\text{min}$ to 220 $^{\circ}\text{C}$, and it was kept for 20 min. Finally, it was raised to 280 $^{\circ}\text{C}$ at a rate of 20 $^{\circ}\text{C}/\text{min}$, where it was held for 4 min. The total running time was 60 min.

2.3.3. Solid phase microextraction (SPME)

In 20 mL headspace vials, 100 μL of each sample was placed and no further dilutions were necessary. The samples were pre-incubated for 5 min with the agitator temperature set at 70 $^{\circ}\text{C}$. The fiber extraction time was 10 min, and the desorption time was 20 min. For chromatographic conditions, initial oven temperature was 60 $^{\circ}\text{C}$, and it was maintained for 2 min. Then it was increased at a rate of 6 $^{\circ}\text{C}/\text{min}$ to 200 $^{\circ}\text{C}$. Finally, it was raised to 280 $^{\circ}\text{C}$ at a rate of 50 $^{\circ}\text{C}/\text{min}$, where it was held for 4 min. The total running time was 31 min. The fiber used was a 50/30 μm DVB/CAR/PDMS by Agilent.

2.4. Data analysis

All data were acquired using the Xcalibur Sequence Setup software. The data were analysed using Xcalibur 3.0, including Qual and Quan Browser functionalities. For the creation of a home-made database, the mzVault™ 2.3 SP1 and TraceFinder 4.0 programs were used. For the analysis of unknown compounds, the Compound Discoverer™ 3.3 program was utilized. These software were provided by Thermo Fisher Scientific. The error mass was set at 5 ppm in all cases. In addition, the NIST (National Institute of Standards and Technology) MS Search 2.2 library was used.

The selected settings for Compound Discoverer were as follows: mass tolerance 5 ppm, min peak intensity 50000, intensity tolerance 30%, S/N threshold 3, intensity threshold 0.1% and retention time (RT) tolerance 0.1 min. For the UHPLC-Q-Orbitrap-MS method, $[\text{M} + \text{H}]^+$, $[\text{M} + \text{H}_2\text{O}]^+$, $[\text{M} + \text{Na}]^+$, $[\text{M} - \text{H}]^-$ and $[\text{M} - \text{H} + \text{FA}]^-$ were selected as preferred adducts. ChemSpider, Mass List (LCMS Co-formulant PPP, Natural Products Atlas 2020_06, Lipid Maps Structure Database and EFS HRAM Compound Database), mzCloud and mzVault were selected as library search. The libraries in GC-MS methods were NIST library (mainlib, NISTDEMO and replib), GC-Orbitrap Contaminants Library, GC-Orbitrap Other Environments, GC-Orbitrap PCBs and GC-Orbitrap Pesticides.

3. Results and discussion

3.1. Method optimization

For UHPLC-Q-Orbitrap-MS, the optimization of different chromatographic parameters was needed. First, different types of stationary phases were tested, such as a C8, a C18 and a phenyl column. The C8 column gave the best results because a better separation of the chromatographic peaks and a higher sensitivity were obtained. The flow rate was tested at 0.3 mL/min, but many of the compounds eluted at the same time and had similar fragment ions, making it challenging to separate them effectively. In consequence, a flow rate at 0.2 mL/min was selected. In addition, MeOH was used as the mobile phase because when acetonitrile was tested, the intensity of the peaks was considerably lower.

The chromatographic methods used in GC-Q-Orbitrap-MS were optimized in previous studies of the research group. Only the volume of CBs used in SPME-GC was evaluated to obtain a good peak sensitivity, without losing information about the compounds at lower concentrations. Thus, volumes of 1 mL, 100 μL , and 10 μL were tested. When 1 mL of CB was evaluated, the major compounds (limonene in CB1 and CB11, and *trans*-cinnamaldehyde in CB2 and CB10) saturated the fiber, showing very large peaks in the chromatogram that interfered with the detection of compounds that eluted at similar RTs. On the contrary, at a sample volume of 10 μL , some minor compounds disappeared, or their sensitivity decreased significantly. A volume of 100 μL was considered optimal, as it resulted in more well-defined and sensitive peaks, while also preventing saturation of the fiber by the major compounds.

3.2. Sample treatment

In UHPLC-Q-Orbitrap-MS and DI-GC, it was necessary to apply a high dilution of the CBs to avoid saturation of the detector signal when injecting a highly concentrated sample. In UHPLC-Q-Orbitrap-MS, Cureneem was diluted in EtOAc due to its low solubility in water. On the other hand, in DI-GC mode, some biopesticides (CB11-CB15) had to be dissolved in DMSO due to their low solubility in the different tested organic solvents (EtOAc, acetone and hexane). Moreover, CB9 had to be dissolved in acetone since it was not soluble in EtOAc. In HS-GC, a lower dilution (1:1,000 v/v) was made in water to prevent EtOAc from interfering with the signal of the more volatile compounds.

3.3. Home-made database

To identify the targeted compounds belonging to the different CBs, a database was built, collecting information of the different compounds that could be present in them [5,22]. For this purpose, a previous characterization of the targeted analytes was carried out by UHPLC-Q-Orbitrap-MS and GC-Q-Orbitrap-MS (DI, HS and SPME injection modes). The characterization was carried out by injecting 100 $\mu\text{g}/\text{L}$ of each standard. For UHPLC-Q-Orbitrap-MS, RT, characteristic fragments of each analyte and adducts were included in the home-made database. On the other hand, for GC-Q-Orbitrap-MS, the same information was used, but instead of adducts, the Kovats retention index was collected. Moreover, a spectral database was built with the MS/MS spectra generated for each compound, applying mzVault program. A library was created for the compounds characterized by UHPLC-Q-Orbitrap-MS. Then, this library was incorporated into Compound Discoverer software and used to confirm the targeted compounds identified in the different CBs.

A total of 27 compounds were characterized by UHPLC-Q-Orbitrap-MS (Table S2). The fragment ions of the MS/MS spectra, obtained for each compound by DDA, were compared with those found in databases, such as mzCloud, and bibliography [23–25], confirming that the chromatographic peaks obtained corresponded to these compounds. Some of the characterized compounds were monoterpenes with the same

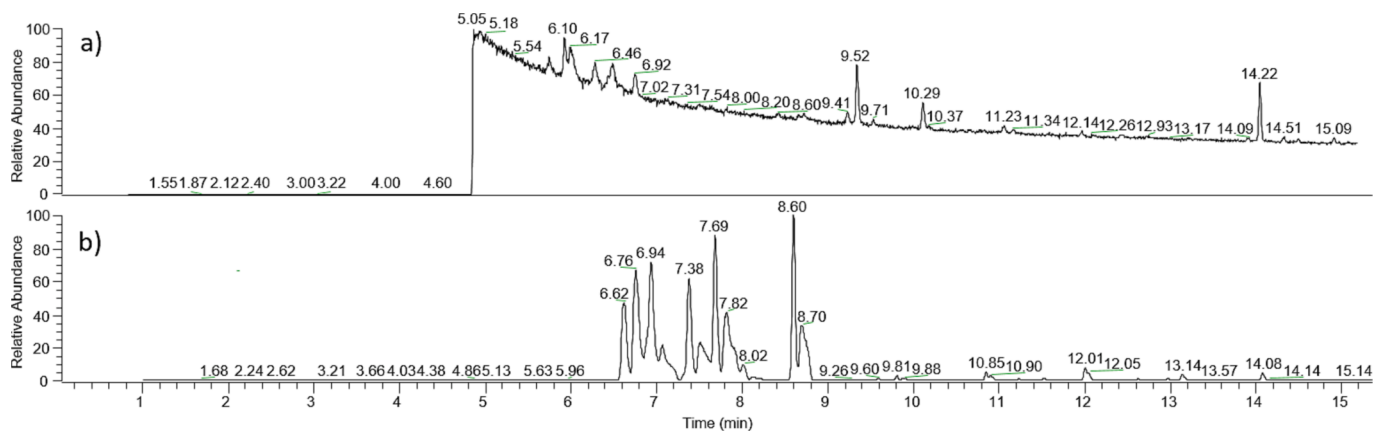


Fig. 1. HS-GC chromatogram of volatile compounds (200 µg/L) in: a) ethyl acetate solvent and b) water solvent.

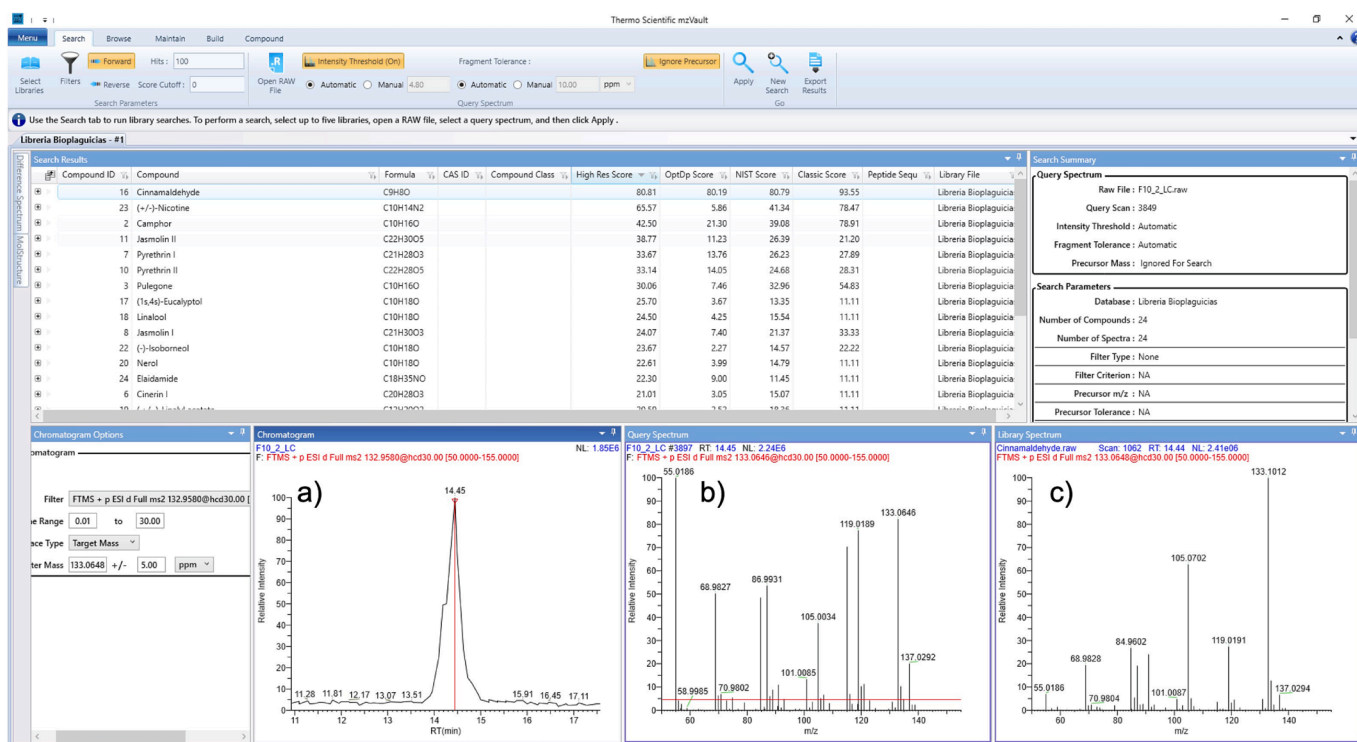


Fig. 2. a) Extracted chromatogram for ion m/z 133.0648 for *trans*-cinnamaldehyde, and MS/MS spectrum obtained in: b) the commercial biopesticide CB10, and c) *trans*-cinnamaldehyde standard stored in the mZVault library.

molecular formula, $C_{10}H_{18}O$. It was observed that the adduct formed from these monoterpenes was $[M-H_2O + H]^+$ having a precursor ion with a m/z of 137.1325 and the m/z of their fragments were 95.0855 and 81.0699. In addition, linalyl acetate ($C_{12}H_{20}O_2$) formed an adduct $[M-CH_3OOH + H]^+$, which shared the same m/z values and fragments as the previously mentioned compounds. These adducts have also been described by Politi et al [26]. However, they could be detected separately due to their different RTs, being 17.56 (1,8-cineole), 17.68 (nerol), 17.70 (geraniol), 17.74 (linalool), 17.83 (isoborneol), 18.10 (citronellal), and 18.83 min (linalyl acetate). The other targeted compounds formed an adduct $[M + H]^+$, except azadirachtin A that presented other different adducts at 15.40 min, namely $[M-H]^-$, $[M-H + FA]^-$ for negative ionization mode, and $[M + Na]^+$ for positive ionization mode. The formation of the adduct $[M-H + FA]^+$ was possible due to the use of formic acid in the aqueous mobile phase.

On the other hand, a total of 31 compounds were characterized by

three injection modes in GC-Q-Orbitrap-MS (Table S3). For the characterization of the compounds by HS-GC, water was used as solvent because EtOAc interfered with the signal of the compounds, observing that their signal decreased or simply did not appear in the chromatogram (Fig. 1). As happened in UHPLC-Q-Orbitrap-MS, $C_{10}H_{18}O$ monoterpenes presented the same fragment ions at m/z 95.0855 or 121.1012 in some cases. Although they had the same fragmentation pattern, it has been possible to detect them separately because their different RTs, which were 9.68 (linalool), 10.89 (citronellal), 11.36 (isoborneol), and 12.66 min (nerol). The same behaviour was observed for monoterpenes with the molecular formula $C_{10}H_{16}$, as all of them presented a fragment at m/z 91.0542 and 93.0699, whose RTs were 6.67 (sabinene), 6.76 (β -pinene), 6.94 (myrcene), 7.38 (α -phellandrene), 7.94 (limonene), and 8.66 min (γ -terpinene). On the other hand, as can be seen in Table S2 and Table S3, some of the characterized compounds can be detected using both methods, allowing for a potential double confirmation of the

Table 1
Targeted and suspect compounds detected in the commercial biopesticides analyzed by UHPLC-Q-Orbitrap-MS^{a,b}.

Compound	RT (min)	Concentration (g/L)				
		CB2	CB3	CB5	CB7	CB10
Target compound						
Azadirachtin A	15.40		6.85			
<i>trans</i> -Cinnamaldehyde	14.41	454.3				396.8
Pyrethrin I	19.84				0.14	
Pyrethrin II	18.83				0.05	
Suspect compound						
6-desacetyl-nimbin ^c	17.30			9.17		
Azadirachtin B ^c	15.81		5.54			
Azadirachtin N ^c	15.46		0.51			
Azadiradione ^c	17.52			74.69		
Gedunin ^c	17.90			15.31		
Nimbin ^c	17.79			11.97		
Nimbinin ^c	18.61			24.22		
Salannin ^c	18.24			18.93		

^a Abbreviations: RT: retention time.

^b Codes of commercial biopesticides in Table S1.

^c Semi-quantification has been carried out using azadirachtin A as standard.

analytes. For example, *trans*-cinnamaldehyde, camphor, pulegone, nicotine or acetyleugenol were characterized by both UHPLC and GC approaches.

3.4. Targeted analysis

For the targeted analysis, the “.raw” files of each CB, including their replicates, were analysed applying the LC and GC home-made databases developed previously, using TraceFinder software. In relation to UHPLC-Q-Orbitrap-MS method, the detected compounds were confirmed based on their RT and MS/MS spectrum with a mass error < 5 ppm. The spectral library mzVault, generated from the standards, provided the necessary spectra to confirm the targeted compounds by matching the MS/MS spectra of the CBs with those stored in the library, showing in Fig. 2 the example of *trans*-cinnamaldehyde. The UHPLC-Q-Orbitrap-MS method was useful to detect less volatile targeted compounds. *trans*-Cinnamaldehyde, pyrethrins and azadirachtin A were therefore identified with a confidence level 1, according to Schymanski et al. [27]. Due to the fact that there was no matrix effect of the different biopesticides (high dilution of CBs was performed), quantification was carried out using solvent calibration curves. For the quantification, a calibration curve for each compound in the four working modes used (UHPLC, DI-GC, HS-GC and SPME-GC) was prepared. The linearity of the calibration

curves and the limit of quantification were obtained by preparing standard solutions at concentrations of 1, 2, 5, 10, 25, 50, 100 and 200 µg/L in EtOAc for UHPLC-Q-Orbitrap-MS and DI-GC, and in H₂O for HS-GC. In the case of SPME-GC, higher concentrations were necessary to detect the analytes: 1, 2, 3, 5, 6, 8, 9 and 10 mg/L in EtOAc.

Then, targeted compounds were identified and quantified in the studied CBs, showing the results in Table 1 and Table 2. In CBs derived from cinnamon extract (CB2 and CB10), *trans*-cinnamaldehyde has been found by UHPLC-Q-Orbitrap-MS method at high concentrations: 454.30 g/L for CB2, and 396.83 g/L for CB10 (Table 1). This compound is a phenylpropanoid commonly present in cinnamon essential oils [28,29], and it has antibacterial, antifungal, and insecticidal properties [30,31]. In CB7, only pyrethrin I and pyrethrin II were detected at very low concentration as specified on the product label (0.18 g/L). These two pyrethrins are the most abundant in pyrethrum extract, which explains the absence of other pyrethrins (cinerin I and II, and jasmolin I and II). In addition, azadirachtin A was found in CBs based on neem extract (CB3). Azadirachtin is a triterpenoid present in *Azadirachta Indica* (neem), being one of its main bioactive compounds. It is well known that this compound has a high insecticidal action, so it has been used for pest control since ancient times. In CB3, azadirachtin A was found at 6.85 g/L.

On the other hand, three different injection modes (DI, HS and SPME) were utilized in GC-Q-Orbitrap-MS to compare the results obtained in the tested CBs and determine which mode may be more suitable. This was useful to detect a greater number of volatile compounds present in the CBs. In this case, instead of mzCloud, the NIST library has been used as the spectral database to compare the MS spectrum obtained in the samples. A total of 17 target compounds were identified using this method as can be seen in Table 2. The results obtained by HS-GC and SPME-GC were practically the same, while by DI-GC only *trans*-cinnamaldehyde, limonene, menthol and linalool were detected. In fact, these four targeted compounds were the only analytes found by the three GC injection modes in CBs.

In the biopesticide based on orange oil (CB1), the main compound found was limonene, which is commonly detected in this type of oil. It was also the predominant compound in the biopesticide CB11. Limonene is a monocyclic monoterpene with insecticidal properties [7,9,10]. Its concentrations were 39.14 and 52.38 g/L, respectively (Table 2).

CB1 contained a total of eight targeted compounds, while CB11 had up to 13 different compounds as can be seen in Table 2. The compounds of both biopesticides have been detected by HS-GC and SPME-GC at similar concentrations. This indicates that both methods are effective in detecting these compounds, providing additional confirmation of their

Table 2
Targeted compounds detected in commercial biopesticides obtained by SPME-GC-Q-Orbitrap-MS^{a,b}.

Compound	RI	RT (min)	Concentration (g/L)					
			CB1	CB2	CB3	CB5	CB10	CB11
Sabinene	974	6.39	2.73					6.00
Myrcene	991	6.94	8.59	0.004				10.50
m-Cymene	1023	7.69		0.024				
Limonene	1030	7.94	39.14	0.019	0.050			52.38
γ-Terpinene	1060	8.66	0.029					0.026
Linalool	1099	9.68	0.325	0.002				0.633
Citronellal	1153	10.89	0.029					
Isoborneol	1157	11.36				0.025		
Menthol	1175	11.63			0.051			7.54
Citronellol	1228	12.68						0.048
Pulegone	1237	13.02						0.010
Carvone	1242	13.11	0.544					0.538
Linalyl acetate	1257	13.17						0.022
<i>trans</i> -Cinnamaldehyde	1270	13.84		386.6			371.8	
<i>trans</i> -Anethole	1286	14.11						0.012
Thymol	1291	14.18	0.036		7.47			0.033
Eugenol	1357	15.57		0.065				0.043

^a Abbreviations: RI: Kovats retention index.; RT: retention time.

^b Codes of commercial biopesticides in Table S1.

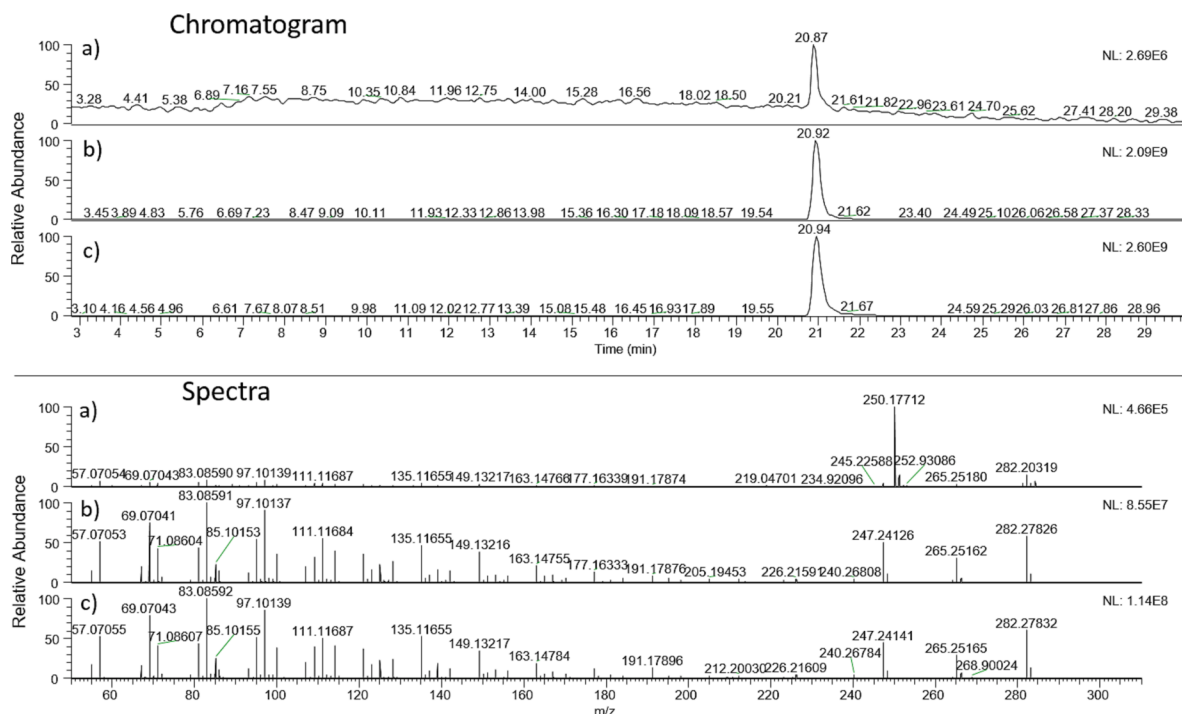


Fig. 3. UHPLC-Q-Orbitrap chromatogram and MS/MS spectra: a) unfiltered methanol, b) filtered methanol, and c) 100 µg/L oleamide standard (unfiltered vial).

presence in the biopesticides. *trans*-Cinnamaldehyde was also detected in CBs based on cinnamon extract (CB2 and CB10). This result agrees with that obtained by UHPLC-Q-Orbitrap-MS. In addition, the concentration obtained by DI-GC was similar to that obtained by UHPLC-Q-Orbitrap-MS, indicating that DI-GC could be useful for the identification of major compounds in CBs.

Other targeted compounds that were found included monoterpenes, especially myrcene and linalool, which were detected in some CBs. CB11 was the biopesticide with the highest number of targeted compounds, with a total of 13, including menthol (7.54 g/L) and sabinene (6.00 g/L), in addition to limonene. There are studies that have investigated the insecticidal action of menthol on insects such as mosquitoes [13,32]. Therefore, this biopesticide presents a wide variety of compounds that could kill or repel different insects.

3.5. Untargeted analysis

3.5.1. Suspect mode

In relation to the detection of suspect compounds, an analysis was performed using several LC and GC home-made databases, built previously by the research group [33], using the TraceFinder program. These databases contained compounds such as pesticides and co-formulants that may be found in this type of CBs. Compounds were tentatively confirmed with a mass error < 5 ppm and with at least two fragment ions, with a mass error < 10 ppm. Among the tentatively identified compounds, dibutyl phthalate was found in almost all CBs by UHPLC-Q-Orbitrap-MS method as can be seen in Table S4. When data obtained by GC-Q-Orbitrap-MS method was evaluated, butylated hydroxytoluene (BHT) was found in CB1, CB7 and CB8; 1-methylnaphthalene in CB3 and CB5; and biphenyl in CB2, CB5 and CB10 (Table S4). Biphenyl was confirmed through the injection of the standard, so it was identified at confidence level 1. These detected compounds are also present in conventional plant protection products as additives. Moreover, BHT has antioxidant activity and is used in food, cosmetic products, and plastics [34,35].

On the other hand, different types of azadirachtins were usually detected in neem extracts. For this reason, a bibliographic search was

carried out to collect information on these compounds, including their adducts and fragment ions [36–38]. Once this information was obtained, these compounds were searched for in CBs based on neem extract (CB3 and CB5) using Qual Browser program. Different types of azadirachtin were detected, such as azadirachtin B, azadirachtin N, azadiradione, nimbin or salannin. Their presence was confirmed by monitoring several adducts ($[M + Na]^+$, $[M - H]^-$, $[M - H + FA]^-$, $[M + H]^+$) at the same RT for each compound. For example, for azadirachtin B, the adducts $[M + Na]^+$, $[M - H]^-$, $[M - H + FA]^-$ were found at a RT of 15.81 min, while for nimbin, the detected ions were $[M + Na]^+$ and $[M + H]^+$ at 17.79 min. Then, a semi-quantification approach of the detected azadirachtins was carried out using azadirachtin A as standard. In CB3, azadirachtin B (5.54 g/L) had a similar concentration than azadirachtin A (6.85 g/L) as can be seen in Table 1. In CB5 the concentrations of azadirachtins ranged from 9.17 g/L (6-desacetyl-nimbin) to 74.69 g/L (azadiradione).

3.5.2. Unknown mode

Regarding the unknown analysis, Compound Discoverer program was used. The “raw” files obtained from each CB were processed according to the specifications described in Section 2.4. Afterwards, different filters were applied to reduce the number of obtained compounds and improve data processing. The filters applied were used to remove the background (peaks observed in blank samples); to remove compounds that had not an assigned name, those whose name was only a dot or those starting with the word [similar]; and to remove compounds with a mass error > 5 ppm or < -5 ppm. In addition, for data obtained by UHPLC-Q-Orbitrap-MS, there must be an MS/MS spectrum in DDA for the precursor ion of the selected compound, the peak area was > 100,000,000 in any file, and PQQ:FWMH2Base was < 0.2 to eliminate those peaks whose resolution and peak shape are not adequate to be considered valid. Chromatographic peaks with a value above 0.2 were not considered to have a valid shape. The compounds that were tentatively confirmed exhibited good peak shape, and the *m/z* of their precursor ions matched those provided by several databases such as ChemSpider, mzCloud or Lipid Maps Structure Database. For GC-Q-Orbitrap-MS, search index and reverse search index > 500, and the peak area was > 100,000 were selected. Moreover, it has been

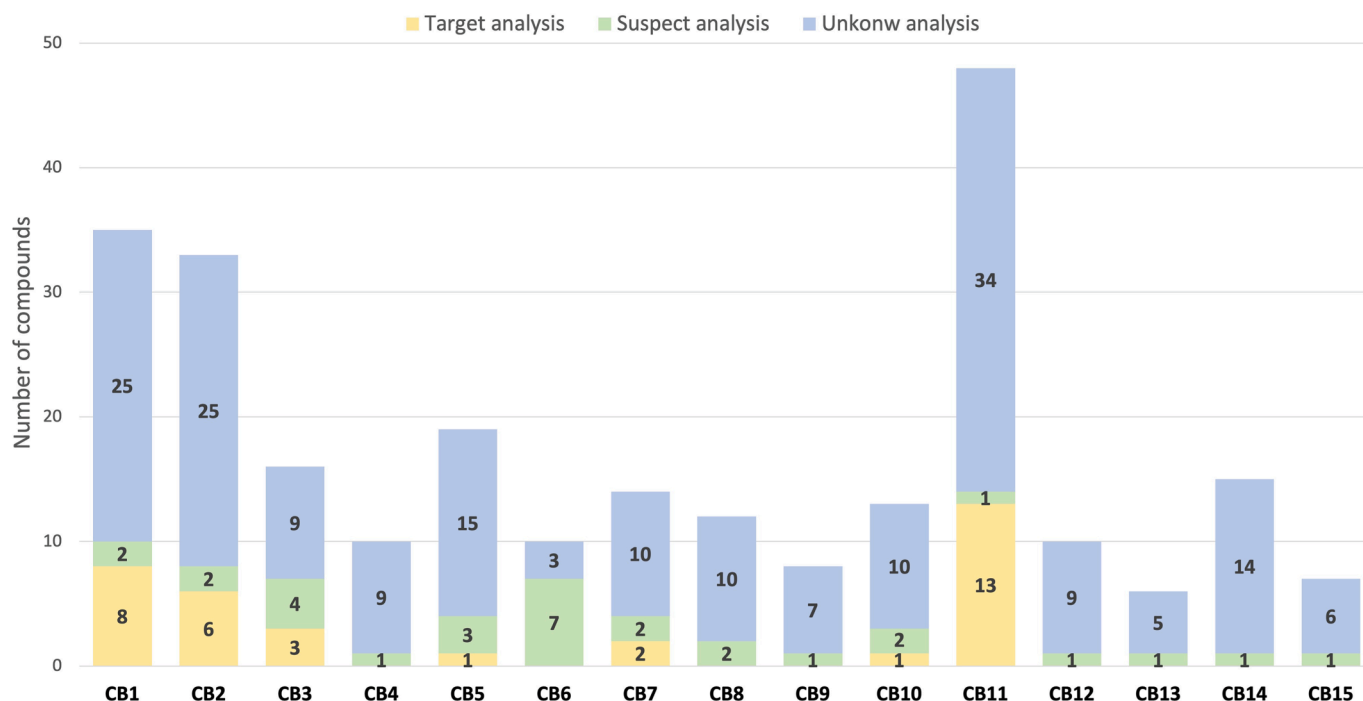


Fig. 4. Total number of compounds found by all the techniques used for each commercial biopesticide.

considered that the RT of compounds coincided with the Kovats retention index that they should present at that RT.

In UHPLC-Q-Orbitrap-MS method, it was observed that oleamide was found at a high concentration in all the CBs. Oleamide had a confidence level 1 because its MS/MS spectrum and RT were confirmed using a reference standard. Therefore, samples were further diluted to enable quantification of oleamide due to the high signal obtained. However, as the dilution increased, the concentration did not decrease; instead it remained the same or even increased. After a bibliographic search, it was found that this compound, apart from being a natural compound that can be found in some plants, is also used as a lubricant in the manufacture of plastics [39,40]. To rule out that the presence of this compound did not come from the plastics used, an experiment was designed analysing MeOH, and another sample of MeOH filtered with the same syringe and nylon filter used in the filtering stage of the CBs (Fig. 3). As can be seen, the chromatographic peak found in filtered methanol at 20.9 min corresponded to that obtained from the oleamide standard. In addition, MS/MS spectrum exactly matches the MS/MS spectrum of the standard, unlike MS/MS spectrum obtained from the unfiltered MeOH at that RT. Based on these findings, it was concluded that the oleamide leached from the plastic of the syringe and was being washed away by the organic solvent used during the preparation of the samples. The presence of this compound in different laboratory materials has also been described in the bibliography [41]. In addition, other fatty acid amides such as lauroyl diethanolamide, lauramide, linoleoyl ethanolamide, myristamide, palmitoleamide, linoleamide, palmitamide, stearamide and erucamide were also found in the filtered methanol. All these compounds are also used as lubricants. Therefore, as these compounds were found in the blanks, they were not considered as compounds present in CBs. On the other hand, oleamide and myristamide have also been detected in most of the biopesticides by DI-GC mode. This could confirm that they appear due to the filtering stage of the samples, which was only carried out for the DI-GC technique, and not for HS-GC and SPME-GC.

A total of 7 compounds have tentatively identified by the UHPLC-Q-Orbitrap-MS method (Table S5). The MS/MS spectrum of some compounds coincided with the registered in the mzCloud spectral database. These compounds (palmitamide and phytosphingosine) had a

confidence level of 2 because their MS/MS spectra fit with the mass spectral library under the same acquisition parameters (DDA, HCD (higher energy collisional dissociation), and 30 eV). Phytosphingosine was found in most CBs. This sphingolipid is naturally present in many plants and its antifungal activity against the soil microbiome has been described by Li et al [42]. Additionally, 2,5-di-*tert*-butyl-1,4-hydroquinone present in many of the CBs is an antioxidant used to prevent oxidation in oils.

A total of 87 compounds were tentatively identified by GC-Q-Orbitrap-MS method as can see in Table S5. Most of them are natural compounds, with a notable abundance of terpenes, such as monoterpenes (α -thujene, camphene...) and sesquiterpenes (δ -cadinene, α -copaene, valencene...). These are present in most essential oils derived from plants and flowers, and different additives such as diisobutyl phthalate or 2,4-di-*tert*-butyl-phenol (antioxidant no.33) were also found. Some of the untargeted compounds found in CB1 (orange oil) and CB11 were limonene oxidation products, such as carveol, *cis*-limonene epoxide, α -limonene diepoxide or α -terpineol [43]. Other compounds found in CB1 were naturally occurring compounds that are present in orange oils, such as α -copaene, caryophyllene, δ -cadinene, octanal, terpinolene, terpinene-4-ol, decenal, undecanal, *cis*- β -farnesene, germacrene D and valencene [44,45]. Many of them are also present in CB11. Compounds derived from *trans*-cinnamaldehyde, such as α -methyl cinnamaldehyde or methyl cinnamate, have been found in CB2. In this CB, a total of 27 untargeted compounds was found, and most of them are present in cinnamon extracts such as benzaldehyde, camphene, benzyl alcohol, aromadendrene, or caryophyllene oxide [28,29,46,47]. β -Caryophyllene and its oxide show a good insecticidal activity against different types of insects such as *S. zeamais* [48] or *Spodoptera frugiperda* [49]. On the other hand, some of the antioxidants found were 2,4-di-*tert*-butyl-phenol, which has a structure like BHT, antioxidant no.33 and diphenylamine.

A wide variety of natural compounds have been found in CB11, that come from most of the essential oils that comprise this biopesticide. However, its labeling does not specify the specific type of vegetable extract or essential oil present in the CB. Because the compounds detected in CB11 were very similar to those in CB1, it is possible that one of the plant extracts used in this biopesticide is derived from oranges or

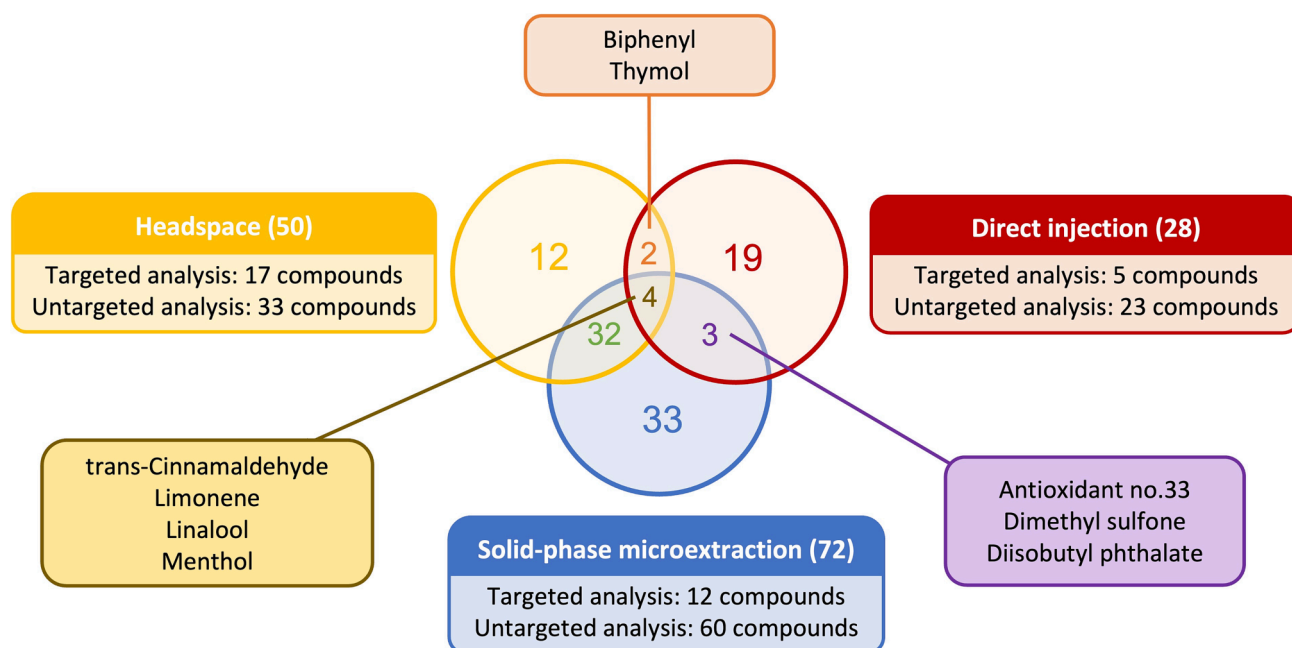


Fig. 5. Number of compounds obtained in each GC injection mode by targeted and untargeted analysis. Total number of compounds for each mode in parenthesis.

another plant belonging to the same family, which may contain similar compounds.

A summary of the compounds detected using the different approaches is shown in Fig. 4. As it can be seen, the CBs with the highest number of compounds were CB1, CB2 and CB11. This outcome is expected since they are based on oils of orange, cinnamon and vegetable extracts and essential oils, respectively, which have a greater number of volatile compounds in their extracts. However, a smaller number of compounds were found in CB6, CB13 and CB15. Limited information is available regarding the composition of these three CBs (CB13–CB15) and their analysis did not yield significant findings. Therefore, it is possible that they are mainly composed of minerals (Mn, N, or Zn) or oxides such as potassium oxide, in higher concentrations.

SPME-GC has proven to be the most efficient mode by detecting a large number of compounds (72 compounds). Although the number found by DI-GC and HS-GC modes was lower (28 and 50 compounds respectively), they have confirmed the presence of many of the compounds that were also detected by SPME-GC and by one of the other two modes (Fig. 5). Only four compounds, including limonene or *trans*-cinnamaldehyde, were successfully detected by all three injection modes in the same CBs. SPME-GC and HS-GC modes share a larger number of compounds, 32 analytes, such as terpinolene, carveol, α -copaene or BHT. On the other hand, only 3 compounds (dimethyl sulfone, antioxidant no.33 and diisobutyl phthalate) were detected simultaneously by SPME-GC and DI-GC, while only 2 compounds (thymol and biphenyl) were detected by DI-GC and HS-GC.

4. Conclusion

Due to the use of the different methods and analyses in the CBs, a great variety of different compounds were found, from natural substances such as limonene or *trans*-cinnamaldehyde to additives such as BHT. The use of both types of analyses (targeted and untargeted) provides extensive information on the composition of CBs, encompassing compounds commonly found in biopesticides as well as other substances. On the other hand, comparing all the chromatographic methods used, HS-GC and SPME-GC yielded the identification of a higher number of compounds, including those detected by suspected and unknown strategies. However, between these two modes, SPME-GC yielded better

results. In addition, there is no interference from the different fatty acid amides preventing the filtering of the samples as occurs in DI-GC and UHPLC. On the other hand, UHPLC is a complementary technique for those CBs that contain fewer volatile compounds that are not detected by GC, such as pyrethrins or azadirachtins. A comparison of the MS/MS spectra and fragments of standards with those found in the CBs enhances the reliability of the identification. This enables the confirmation of these compounds with a confidence level 1. Moreover, the use of spectral libraries databases such as mzCloud and NIST library are a good tool to find and confirm suspect and unknown compounds through MS and MS/MS spectra.

CRediT authorship contribution statement

Alba Reyes-Ávila: Formal analysis, Investigation, Validation, Software, Writing – original draft, Visualization. **Roberto Romero-González:** Data curation, Investigation, Software, Methodology, Supervision, Writing – review & editing. **F. Javier Arrebola-Liébanas:** Methodology, Software, Supervision, Writing – review & editing. **Antonia Garrido Frenich:** Conceptualization, Data curation, Resources, Writing – review & editing, Funding acquisition, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Authors gratefully acknowledge to the University of Almería, the Andalusian Ministry of Economic Transformation, Industry, Knowledge, and Universities and FEDER for financial support (project reference: UAL2020-FQM-B1943).

- [38] L. Song, J. Wang, Q. Gao, X. Ma, Y. Wang, Y. Zhang, H. Xun, X. Yao, F. Tang, Simultaneous determination of five azadirachtins in the seed and leaf extracts of *Azadirachta indica* by automated online solid-phase extraction coupled with LC-Q-TOF-MS, *Chem. Cent. J.* 12 (2018) 85, <https://doi.org/10.1186/s13065-018-0453-y>.
- [39] C. Tang, Q. Huang, Y. Zhu, Y. Liang, Y. Zeng, X. Luo, B. Mai, X. Peng, Nontarget analysis and characterization of alkylamides in electrical product plastics by gas chromatography-positive chemical ionization quadrupole-orbitrap high-resolution mass spectrometry and quasi-molecular ion screening and anchoring algorithm, *J. Chromatogr. A*. 1682 (2022), 463466, <https://doi.org/10.1016/j.chroma.2022.463466>.
- [40] M.A. Farajzadeh, M. Ebrahimi, A. Ranji, E. Feyz, V. Bejani, A.A. Matin, HPLC and GC methods for determination of lubricants and their evaluation in analysis of real samples of polyethylene, *Microchim. Acta*. 153 (2006) 73–78, <https://doi.org/10.1007/s00604-005-0429-1>.
- [41] U. Jug, K. Naumoska, V. Metličar, A. Schink, D. Makuc, I. Vovk, J. Plavec, K. Lucas, Interference of oleamide with analytical and bioassay results, *Sci. Rep.* 10 (2020) 2163, <https://doi.org/10.1038/s41598-020-59093-1>.
- [42] C. Li, Q. Tian, M.K.u. Rahman, F. Wu, Effect of anti-fungal compound phytosphingosine in wheat root exudates on the rhizosphere soil microbial community of watermelon, *Plant Soil*. 456 (1-2) (2020) 223–240.
- [43] L.L. Zhang, Y. Zhang, J.N. Ren, Y.L. Liu, J.J. Li, Y.N. Tai, S.Z. Yang, S.Y. Pan, G. Fan, Proteins differentially expressed during limonene biotransformation by *Penicillium digitatum* DSM 62840 were examined using iTRAQ labeling coupled with 2D-LC-MS/MS, *J. Ind. Microbiol. Biotechnol.* 43 (2016) 1481–1495, <https://doi.org/10.1007/S10295-016-1826-7>.
- [44] C. Torres-Alvarez, A. Núñez González, J. Rodríguez, S. Castillo, C. Leos-Rivas, J. G. Báez-González, Chemical composition, antimicrobial, and antioxidant activities of orange essential oil and its concentrated oils, *CyTA – J. Food*. 15 (2017) 129–135, <https://doi.org/10.1080/19476337.2016.1220021>.
- [45] A. Verzera, A. Trozzi, G. Dugo, G. Di Bella, A. Cotroneo, Biological lemon and sweet orange essential oil composition, *Flavour, Fragr. J.* 19 (2004) 544–548, <https://doi.org/10.1002/ffj.1348>.
- [46] V.S. Rana, C.B. Devi, M. Verdeguer, M.A. Blázquez, Variation of terpenoids constituents in natural population of *Cinnamomum tamala* (L.) leaves, *J. Essent. Oil Res.* 21 (2009) 531–534, <https://doi.org/10.1080/10412905.2009.9700237>.
- [47] R. Wang, R. Wang, B. Yang, Extraction of essential oils from five cinnamon leaves and identification of their volatile compound compositions, *Innov. Food Sci. Emerg. Technol.* 10 (2009) 289–292, <https://doi.org/10.1016/j.ife.2008.12.002>.
- [48] S. Ma, R. Jia, M. Guo, K. Qin, L. Zhang, Insecticidal activity of essential oil from *Cephalotaxus sinensis* and its main components against various agricultural pests, *Ind. Crop. Prod.* 150 (2020), 112403, <https://doi.org/10.1016/j.indcrop.2020.112403>.
- [49] N.C. Cárdenas-Ortega, M. Martín González-Chávez, R. Figueroa-Brito, A. Flores-Macías, D. Romo-Asunción, D.E. Martínez-González, V. Pérez-Moreno, M. A. Ramos-López, Composition of the essential oil of *Salvia ballotiflora* (Lamiaceae) and its insecticidal activity, *Molecules*. 20 (2015) 8048–8059, <https://doi.org/10.3390/molecules20058048>.