Comparative Evaluation of Terpenes and LOX-derived Volatile Compounds for Clustering and authentication of Monovarietal Extra Virgin Olive Oils Using HS-SPME-GC-MS and Chemometric Analysis

Tommaso Ugolini[1]
Fabrizio Melani[2]
Bruno Zanoni[1]
Marzia Migliorini[3]
Serena Trapani[3]
Nadia Mulinacci[2]
Lorenzo Cecchi[1] ✉

[1] Università di Firenze
Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali (DAGRI)
Firenze, Italia

[2] Università di Firenze
Dipartimento di NEUROFARBA, Sezione Nutraceutica e Farmaceutica
Firenze, Italia

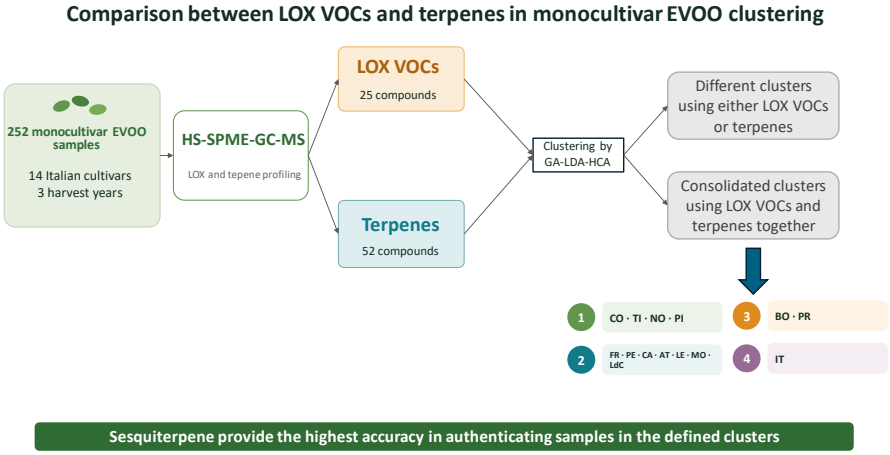
[3] Carapelli Firenze S.p.A.,
Via Leonardo da Vinci 31,
Tavarnelle Val di Pesa,
Firenze, Italia

✉ CORRESPONDING AUTHOR
Via Donizetti 6 – 50144, Firenze
+39 0552755517
Lo.cecchi@unifi.it

Recieved: February 26, 2026
Accepted: September 18, 2026

The topic of clustering and authenticating monovarietal extra virgin olive oils based on their chemical features has gained increasing attention in recent years, mainly due to the increasing interest of a segment of consumers in this product and to the increasing number of concourses for high-quality monovarietal extra virgin olive oil. The aim was to compare the authenticating and clustering capability of the two classes of volatile compounds mainly associated with the positive attributes of monocultivar EVOOs: volatile compounds from the lipoxygenase pathway and terpenes. A total of 252 samples from 14 of the main Italian cultivars were collected over three consecutive production years and two validated HS-SPME-GC-MS methods were adopted to analyse their volatile and terpene profiles. Suitable approaches were adopted for data processing, including Genetic Algorithm (GA), Linear Discriminant Analysis (LDA), and Hierarchical Cluster Analysis (HCA). After application of the same data processing approach to the same sample-set, different chemical clusters of cultivars were obtained using either LOX VOCs or terpenes, pointing out a different behaviour of these groups of volatile molecules in the different cultivars, while terpenes were proven to be more reliable in authenticating the samples in the defined clusters. This was confirmed by the clusters obtained when LOX VOCs and terpenes were used together, obtaining the same chemical clusters of cultivars as those obtained when only terpenes were used: i) Frantoio, Pendolino, Casaliva, Ascolana Tenera, Leccino, Moraiolo, Leccio del Corno; ii) Coratina, Tonda Iblea, Nocellara, Picholine; iii) Bosana, Peranzana; iv) Itrana. Finally, sesquiterpenes showed the best performance in authenticating samples in the defined clusters (accuracy 90%, even improved during validation with external samples), followed by LOX VOCs, while monoterpenes showed the worst results, suggesting poor association with the samples' botanical origin. The findings of this study indicate that for cultivar clustering and authentication purposes, terpenes, and particularly sesquiterpenes, are the most reliable and efficient molecules to be used, allowing for a reduce analytical effort towards this goal.

Keywords: quality control; sesquiterpenes; monoterpenes; extra virgin olive oil; machine learning; pentene dimers



1. INTRODUCTION

Olive oil is unique among food products in requiring not only chemical analyses but also sensory evaluation to ensure accurate labeling according to its commercial category. According to European regulations, virgin olive oils are classified as extra virgin (EVOO), virgin olive oils (VOO), or lampante virgin olive oils (LVOO), based on the results of the "Panel Test", a standardized sensory assessment that evaluates the intensity of positive attributes (fruitiness, bitterness, and pungency) and detects the presence of sensory defects [1,2]. EVOO represents the highest quality category, characterized by the absence of sensory defects and the presence of desirable attributes, including oils that may originate from various cultivars and geographical areas, leading to diverse sensory profiles beyond the standard positive traits (i.e., fruitiness, bitterness, pungency).

Consumers are generally willing to pay a premium for high-quality EVOOs, especially those with distinct characteristics linked to the declaration of cultivar or geographical origin, such as organic, monovarietal, or PDO/PGI-certified EVOOs [3-6]. As a result, the interest of both consumers and companies in monovarietal extra virgin olive oil (MEVOO) has been growing in recent years [7-9]. However, several challenges still hinder their broader diffusion: (i) the prevalence of mixed-cultivar olive groves, (ii) the issue of the so-called "cross-contamination" during processing, (iii) increased vulnerability to fraud due to their high market value, and (iv) limited consumer awareness regarding the sensory characteristics of different monovarietal oils [7,10,11]. In this context, supporting producers in maintaining the quality and distinctiveness of MEVOOs is essential, alongside providing tools to combat fraud and to help consumers recognize cultivar-specific characteristics.

MEVOOs are characterized by sensory profiles that reflect the cultivar of origin, although some cultivars may share similar sensory or chemical features [12-14]. The different sensory profiles are primarily attributed to variations in specific chemical classes, notably Volatile Organic Compounds originating from the lipoxygenase pathway (LOX-derived VOCs), terpenes and phenolic compounds [15-20]. While no official chemical/statistical method currently exists for MEVOO authentication [8,11], several studies have proposed promising approaches based on statistics applied to chemical composition data [21-24]. Among these, terpenes and LOX-derived VOCs in particular have shown potential for cultivar differentiation [12,14,16,17,25-28]. In some of these articles, the main obstacle hindering the application of the related approach is the limited number of samples or cultivars of origin.

Recent literature has highlighted the relevance of specific molecular classes in explaining the chemical differences among MEVOOs [12,14]. Although complete classification of samples in individual cultivars has not always been achieved, MEVOOs have been grouped into distinct consistent clusters, independent of olive-growing conditions or processing and storage variables. VOCs have proven effective in distinguishing among MEVOO clusters [12]. A total of 320 MEVOO samples, collected from nine cultivars across four production years and major Mediterranean-producing countries, were analyzed for 72 VOCs using HS-SPME-GC-MS. Chemometric tools, including Agglomerative Hierarchical Clustering and Partial Least Squares Discriminant Analysis (PLS-DA), allowed for 94% correct classification of samples into three cultivar clusters: (i) Arbequina and Nocellara; (ii) Coratina, Frantoio, Leccino, and Moraiolo; (iii) Hojiblanca, Picual, and Koroneiki. Among the VOCs, ten compounds, eight of which were LOX-derived, were the most significant. Similarly, volatile hydrocarbon compounds (VHCs), such as terpenes and pentene dimers, also demonstrated strong discriminatory ability [14]. In this case, 225 MEVOO samples from 14 Italian cultivars were analyzed across three production years. Sixty VHCs were quantified using HS-SPME-GC-MS. Through Agglomerative Hierarchical Clustering and Linear Discriminant Analysis (LDA), 94% of samples were correctly classified into four cultivar clusters: (i) Ascolana Tenera, Casaliva, Leccio del Corno, Moraiolo, Pendolino, Frantoio, and Leccino; (ii) Coratina, Picholine, Tonda Iblea, and Nocellara; (iii) Bosana and Peranzana; (iv) Itrana. Interestingly, the positions of Coratina and Nocellara varied depending on whether VOC or VHC profiles were used, suggesting that different compound classes yield slightly different clustering outcomes. However, these two studies were based on different sample sets, cultivars, and statistical approaches. Therefore, confirmation is needed by working with the same sample set and applying the same data analysis tools to both LOX-related VOCs and terpenes.

The objective of this study is to compare the performance of LOX-related VOCs and terpenes in authenticating and clustering MEVOOs from the major Italian cultivars. To this end, 252 MEVOO samples from 14 cultivars, collected over three consecutive production years, were analyzed using HS-SPME-GC-MS. The resulting data were processed using Genetic Algorithm (GA), Linear Discriminant Analysis (LDA), and Hierarchical Cluster Analysis (HCA).

2. EXPERIMENTAL PART

2.1. CHEMICALS AND STANDARDS PREPARATION

Ultrapure water was obtained using a Milli-Q system (Millipore SA, Milsheim, France). All commercial standards used for VOC analysis by HS-SPME-GC-MS were acquired from Merck (Darmstadt, Germany), comprising 8 internal and 72 external standards. For analysis of terpenes and pentene dimers (VHCs) by HS-SPME-GC-MS, a total of 14 commercial standards were used, sourced as follows: (i) the pentene dimer 3-ethyl-1,5-octadiene (mixture of (*Z*) and (*E*) form, 95%), 3,7-decadiene (mixture of (*E*,*E*), (*E*,*Z*), and (*Z*,*Z*) forms, 95%), and the C10 diene (*E*,*E*)-2,8-decadiene were purchased from ChemSampCo (Trenton, NJ, USA); (ii) the monoterpenes *d*-limonene (97%), α-pinene (98%), δ-3-carene (99%), β-myrcene (90%), α-thujone (> 96%), and the sesquiterpenes β-elemene (98%), (+)-valencene (80%), β-caryophyllene (80%), (+)-cyclosativene (99%) and (+)-ledene (95%) were obtained from Merck Life Science S.r.l. (Milano, Italy); (iii) the sesquiterpene α-copaene (95%) was sourced from MuseChem (Fairfield, NJ, USA). Internal standards included (*E*,*E*)-2,8-decadiene, α-thujone, and (+)-ledene. An additional set of compounds was included for qualitative analysis of VHCs: (i) α-caryophyllene, (*E*)-β-ocimene, α-cedrene, α-phellandrene, α-terpinene, α-terpineol, eucalyptol, terpinolene, β-pinene, linalool, longifolene, and farnesene (isomeric mixture), kindly provided by IBBR-CNR (Florence Division, Florence, Italy); (ii) allo-ocimene (isomeric mixture), acquired from Merck Life Science S.r.l. (Milan, Italy).

Internal and external standard solutions for both VOC and VHC analyses were prepared using the Multiple Internal Standard Normalization (MISN) method. The standards were dissolved in refined olive oil, previously confirmed to be free of interfering substances. Calibration curve solutions for VOCs and VHCs were prepared by diluting the standard mixtures, maintaining a constant concentration of internal standards while varying the concentrations of external standards, following validated procedures [29,30]. All solutions were stored at -20 °C until analysis by HS-SPME-GC-MS.

2.2. SAMPLES

The MEVOO samples (n = 159) of Italian cultivars were collected over two consecutive production years: 2020–2021 and 2021–2022, and were promptly analyzed after collection. Table I presents the sample set categorized by botanical origin and production year. This sample set is representative of the Italian MEVOO production landscape, as it includes the main cultivars from various regions across the country. A further group of MEVOO samples (n = 93) of the same cultivars was collected in the following olive oil production year (2022-2023) to be used as the test set

Table I - The collected monovarietal extra virgin olive oils in 2020 and 2021 production years

Cultivar (Code)	Total	Production year	
		2020-21	2021-22
Ascolana Tenera (**AT**)	12	5	7
Bosana (**BO**)	7	1	6
Coratina (**CO**)	24	15	9
Casaliva (**CA**)	6	6	0
Frantoio (**FR**)	15	11	4
Itrana (**IT**)	6	3	3
Leccio del Corno (**LdC**)	6	2	4
Leccino (**LE**)	19	8	11
Moraiolo (**MO**)	14	10	4
Nocellara (**NO**)	13	5	8
Pendolino (**PE**)	7	5	2
Peranzana (**PR**)	15	7	8
Picholine (**PI**)	6	4	2
Tonda Iblea (**TI**)	9	7	2
Total	**159**	**89**	**70**

2.3. ANALYSIS FOR SAMPLE COMMERCIAL CLASSIFICATION

The chemical legal parameters of the MEVOO samples for virgin olive oil classification were analyzed in accordance with Regulations (EU) 2022/2104 and (EU) 2022/2105 [1,2].

2.4. HS-SPME-GC-MS ANALYSIS OF VOLATILE ORGANIC COMPOUNDS

The LOX-related VOCs other than pentene dimers were analyzed through Headspace-Solid Phase Microextraction combined with Gas Chromatography–Mass Spectrometry (HS-SPME-GC-MS), following the procedure described by Fortini et al. (2017) [30]. The quantification of individual VOCs was performed using a previously validated protocol based on the MISN (Multiple Internal Standard Normalization) technique. In brief, 4.3 grams of olive oil and 0.1 grams of an internal standard solution, prepared using refined olive oil devoid of interfering substances, were introduced into a 20-mL screw-cap vial. VOCs were extracted by exposing a 1-cm DVB/CAR/PDMS fiber (50/30 µm, Supelco, Belfonte, PA, USA) to the headspace for 20 minutes at 45°C. The adsorbed VOCs were thermally desorbed in the injection port of an Agilent 6890 N gas chromatograph (Santa Clara, CA, USA) at 260°C for 1.7 minutes under splitless conditions.

Chromatographic separation was conducted on an HP-Innowax capillary column (50 m × 0.2 mm i.d., 0.4 µm film thickness), while detection was carried out using a 5975-series quadrupole mass spectrometer detector (Agilent Technologies, Palo Alto, CA, USA). The GC oven temperature was programmed to start at 40°C, increase to 156°C at a rate of 4°C/

min, and then reach 260°C at 10 °C/min, with a final hold at 260°C for 6 minutes. Helium served as the carrier gas at a flow rate of 1.2 mL/min. The transfer line and ion source temperatures were maintained at 250°C and 230°C, respectively. Mass spectrometry was performed in scan mode (m/z range: 30-350 Th, scan rate: 1500 Th/s), using an electron ionization energy of 70 eV.
Identification of VOCs was achieved by comparing their retention times and mass spectra with those of authentic reference standards. For each analyte, the most suitable internal standard was chosen, and quantification was based on a six-point linear calibration curve (plotting area ratio against concentration ratio) constructed from the respective pure compound. Additional methodological specifics are available in [30].

2.5. HS-SPME-GC-MS ANALYSIS OF VOLATILE HYDROCARBON COMPOUNDS

The profile of VHCs, including monoterpenes, sesquiterpenes, and pentene dimers, in all samples was determined using a recently validated HS-SPME-GC-MS protocol designed for the concurrent analysis of these classes of volatile hydrocarbons [29]. In short, approximately 3.2700 g of sample and 0.1000 g of an internal standard (IntStd) solution were placed into a 20 mL screw-cap vial. Following a 5-minute equilibration period, VHCs were extracted by exposing a 1-cm SPME fiber (50/30 µm DVB/CAR/PDMS, Agilent, Palo Alto, CA, USA) to the vial headspace at 90°C for 65 minutes under continuous agitation (400 rpm). Desorption was performed in the injection port of a 6890 N GC system, interfaced with a 5975 MS detector (Agilent Technologies), for 1.7 minutes. The fiber was subsequently conditioned at 260°C for 20 minutes.
Separation of analytes was achieved using an HP-Innowax capillary column (50 m × 0.2 mm i.d., 0.4 µm film thickness) under the same oven temperature program applied to the analysis of VOCs. The other GC and MS conditions were the same as those applied to VOC analysis.
In total, 60 volatile hydrocarbons were identified: 7 pentene dimers (PD), 24 monoterpenes (MT, including 9 cyclic, 8 acyclic, 2 C11 homoterpenes, 3 oxygenated, and 2 unidentified compounds) and 29 sesquiterpenes (ST, comprising 22 cyclic, 6 acyclic, and 1 oxygenated ST). When commercially available standards were accessible (see “Chemicals and Standard Preparation”), identification was based on matching mass spectra, retention times, and retention indices. In cases where standards were not available, identification was tentative and relied on (i) comparison of the mass spectra with those in the NIST08/Wiley98 library (minimum match threshold: 80%) and (ii) the calculation of retention indices using a standard C9–C30 n-alkane mixture analyzed under the same conditions, which were then compared with literature values [31]. Quantitative analysis was conducted by constructing an 8-point calibration curve for each compound included in the ExtStd solution, following the multiple internal standard normalization method [30]. Each analyte was paired with the most appropriate internal standard, consistent with approaches previously applied to virgin olive oil VOC analysis. Additional methodological details can be found in the referenced publication.

2.6. DATA ANALYSIS

Various data analysis techniques were implemented using R software (version 4.5.0, The R Foundation for Statistical Computing), including the Shapiro–Wilk test, Kruskal–Wallis test, Dunn’s test, Linear Discriminant Analysis (LDA), and Hierarchical Cluster Analysis (HCA). Additionally, a Genetic Algorithm (GA) was employed to identify optimal combinations of variables.
Sample classification was performed using a supervised approach such as GA-LDA, as follows: Linear Discriminant Analysis (LDA) constructs a discriminant function that maximally separates predefined classes based on a selected subset of variables. Since previous studies [32] emphasize the need for variable selection when using LDA, it was combined with a Genetic Algorithm (GA). GA is a natural selection-inspired metaheuristic approach, particularly suitable for problems involving numerous heterogeneous variables across many samples. The GA iteratively evolves a population of solutions (variable combinations), selecting the most promising ones based on a fitness function. The process begins with a randomly generated population of individuals (variable sets), which are evaluated and ranked. The top-performing individuals are selected and modified through genetic operators, crossover (recombination) and mutation, to generate a new population (generation). This cycle repeats for several generations until the difference in score (obtained from a fitness function) between the new and old generations is below a threshold value.
The parameters used in the GA were:

- Number of variables per individual: 6
- Population size: 100
- Number of elite individuals: 50
- Crossover probability: 1 (applied to all elite individuals)
- Mutation probability: random
- Score convergence threshold value: 0.005

LDA served as the fitness function, assessing each variable combination based on the Class Error Average (i.e., the mean percentage of misclassified samples across all classes). This metric was derived from a confusion matrix obtained through 5-fold cross-validation. Overall classification accuracy was also calculated.
To investigate the clustering structure among MEVOO

samples, the unsupervised technique Hierarchical Cluster Analysis (HCA) was applied using normalized confusion matrices from GA-LDA as similarity measures. HCA was carried out using two distance metrics (Euclidean and Manhattan) and Average, Complete, and Ward2 linkage criteria.
After assessing that data were non-normally distributed through the Shapiro-Wilk test, the Kruskal-Wallis and the Dunn's post hoc tests were applied to identify the chemical variables that significantly differentiated the cultivars, and to evaluate significant differences among cultivars at $p < 0.05$.

3. RESULTS AND DISCUSSION

3.1. CHARACTERIZATION OF THE LOX VOCS PROFILE OF MONOCULTIVAR EVOOS

The characterization of the profile of pentene dimers and terpenes of the 14 cultivars under study (i.e., Ascolana Tenera, Bosana, Casaliva, Coratina, Frantoio, Itrana, Leccino, Leccio del Corno, Moraiolo, Nocellara, Pendolino, Peranzana, Picholine, Tonda Iblea) has already been described and can be found in [14]. The LOX VOCs profile of these cultivars is given in Table II, in which the median values of each molecule in each cultivar are reported. With values ranging from 8644 mg/kg for Nocellara to 38124 mg/kg for Casaliva, (*E*)-2-hexenal was the most abundant LOX VOC in all the 14 cultivars, in agreement with previous literature [12,20,33,34]. Low median values (9272-11863 mg/kg) were observed for Tonda Iblea, Bosana and Ascolana Tenera, with a slightly higher amount in Moraiolo (13345 mg/kg). Higher and similar values were observed for Peranzana (17458 mg/kg) and Picholine (17993 mg/kg), and for Frantoio, Itrana, Leccino and Leccio del Corno (20103-22506 mg/kg); finally, the highest values were observed for Casaliva (38124 mg/kg), followed by Pendolino (32850 mg/kg) and Coratina (28706 mg/kg). Following (*E*)-2-hexenal, different LOX VOCs were the most abundant compounds in different groups of cultivars: for some cultivars, especially those with the lowest amounts of (*E*)-2-hexenal, quite high values of (*Z*)-3-hexenal, (*Z*)-3-hexenyl acetate and (*Z*)-3-hexenol were observed, suggesting that in these cultivars a lower expression of the enzyme isomerase (catalyzing isomerization of (*Z*)-3-hexenal to (*E*)-2-hexenal) occurs [12]. This was the case for Itrana, which for (*Z*)-3-hexenal, (*Z*)-3-hexenyl acetate and (*Z*)-3-hexenol showed median values of 6.292 mg/kg, 3.014 mg/kg, and 1.471 mg/kg respectively, and Tonda Iblea, with values of 5.167 mg/kg, 1.840 mg/kg and 1.627 mg/kg for those molecules. Nocellara also showed significant amounts of (*Z*)-3-hexenal (5.543 mg/kg), and it was characterized by the highest amount of (*Z*)-3-hexenol (2.418 mg/kg) but a quite low amount of (*Z*)-3-hexenyl acetate, behaviour not so different from that of Ascolana Tenera. Picholine showed significant amounts of (*Z*)-3-hexenal, but not so high amounts of (*Z*)-3-hexenyl acetate and (*Z*)-3-hexenol. Finally, Bosana, Peranzana, and Picholine showed a similar pattern, with the greatest content of (*Z*)-3-hexenyl acetate, followed by slightly lower amounts of (*Z*)-3-hexenal and low amounts of (*Z*)-3-hexenol. A similar behaviour in the Bosana and Peranzana cultivars was already observed when studying the terpene profile of these cultivars [14]. The remaining group of cultivars (i.e., Casaliva, Coratina, Frantoio, Leccino, Leccio del Corno, Moraiolo, Pendolino) was strongly characterized by (*E*)-2-hexenal; they not only showed high levels of this molecule, but also showed low levels of all other LOX VOCs (i.e., below 1 mg/kg); for some of them (i.e., Casaliva and Moraiolo), the highest values among the other LOX VOCs were for (*Z*)-3-hexenyl acetate (approx. 0.9 mg/kg both), while for the majority (i.e., Coratina, Frantoio, Leccino, Leccio del Corno) the highest values were for 1-penten-3-one.

3.2. COMPARATIVE ASSESSMENT OF LOX VOCS AND TERPENES IN CULTIVAR CLUSTERING

In two previous research studies based on different sample sets, cultivars, and statistical approaches (i.e., ANOVA-HCA-PLS/DA and Kruskal–Wallis-PCA-LDA-HCA, respectively), different clusters of cultivars were obtained using LOX VOCs or terpenes, respectively [12,14]. To verify the reliability of this finding that suggested that different clusters are obtained using either LOX VOCs or terpenes, in this research, the same sample set and the same statistical approach (i.e., GA-LDA-HCA) were applied when using either LOX VOCs or terpenes. Therefore, two data matrices were created: the first with the 159 samples (see Table I) in the rows and the 25 LOX VOCs in the columns; the second with the 159 samples in the rows and the 52 terpenes in the columns.
The LOX VOCs were: 3-Pentanone, 1-Penten-3-one, Hexanal, (*E*)-2-Pentenal, 1-Penten-3-ol, (*Z*)-3-Hexenal, (*E*)-2-Hexenal, Hexyl acetate, (*E*)-2-Pentenol, (*Z*)-2-Penten-1-ol, (*Z*)-3-Hexenyl acetate, (*E*)-2-Hexenyl acetate, 1-Hexanol, (*E*)-3-Hexenol, (*Z*)-3-Hexenol, (*E*)-2-Hexenol, (*Z*)-2-Hexenol, (*E*,*E*)-2,4-Hexadienal, 3,4-diethyl-1,5-hexadiene (RS+SR), 3,4-diethyl-1,5-hexadiene (meso), (*Z*)-3-ethyl-1,5-octadiene, (*E*)-3-ethyl-1,5-octadiene, (*Z*,*Z*)-3,7-decadiene, (*Z*,*E*)-3,7-decadiene, (*E*,*E*)-3,7-decadiene. The terpenes were 24 monoterpenes (i.e., α-Pinene, β-Pinene, (*Z*,*Z*)-2,6-Dimethyl-1,3,5,7-octatetraene ((*Z*,*Z*)-cosmene), monoterpene 3, δ-3-Carene, β-Myrcene, α-Phellandrene, α-Terpinene, *d*-Limonene, Eucalyptol, (*Z*,*Z*)-Alloocimene, β-Phellandrene, (*Z*,*E*)-2,6-Dimethyl-1,3,5,7-octatetraene ((*Z*,*E*)-cosmene), γ-terpinene, β-*trans*-ocimene, monoterpene 4, (*Z*)-4,8-Dimethylnona-1,3,7-triene, *p*-Mentha-1,4(8)-diene (terpinolene), (*E*)-4,8-Dimethylnona-1,3,7-triene,

Table II - Median Values (mg/kg) and *p*-values from Kruskal-Wallis test of LOX-related VOCs in Monovarietal Extra Virgin Olive Oils from different Cultivars. Different letters in each row indicate significant differences among cultivars for that molecule at $p < 0.05$, according to Dunn's test

LOX-related VOC	*p*-value	Ascolana Tenera	Bosana	Casaliva	Coratina	Frantoio	Itrana	Leccino	Leccio del Corno	Moraiolo	Nocellara	Pendolino	Peranzana	Picholine	Tonda Iblea
1-Penten-3-ol	2.1×10^{-06}	0.470 a	0.502 abc	0.538 abc	0.660 b	0.560 abc	0.633 abc	0.462 ac	0.444 ac	0.451 a	0.518 ac	0.614 bc	0.644 b	0.595 abc	0.574 abc
(*E*)-2-Penten-1-ol	3.3×10^{-05}	0.060 abcd	0.052 abcd	0.057 abc	0.075 acd	0.060 abcd	0.097 d	0.056 b	0.055 abc	0.056 ab	0.079 d	0.059 abcd	0.081 cd	0.066 abcd	0.077 cd
(*Z*)-2-Penten-1-ol	1.4×10^{-12}	0.405 ab	0.377 abc	0.345 a	0.582 d	0.362 a	0.578 cd	0.340 a	0.326 a	0.343 a	0.484 cd	0.409 abc	0.529 cd	0.507 cd	0.482 bcd
1-Hexanol	1.4×10^{-06}	0.433 abc	0.330 abd	0.562 c	0.603 c	0.347 abd	0.266 abd	0.417 ac	0.340 abcd	0.265 d	0.431 ac	0.384 abcd	0.538 ac	0.251 bd	0.464 ac
(*E*)-3-Hexenol	1.7×10^{-05}	0.017 abc	0.008 ad	0.014 abd	0.012 ad	0.009 d	0.011 abcd	0.010 d	0.009 d	0.011 ad	0.025 c	0.011 abd	0.014 abd	0.010 d	0.019 bc
(*Z*)-3-Hexenol	3.3×10^{-15}	1.476 ab	0.533 cde	0.649 cde	0.650 cd	0.438 e	1.471 abf	0.437 e	0.523 ce	0.790 cdf	2.418 a	0.728 cdef	0.921 abdf	0.802 bcdf	1.627 ab
(*E*)-2-Hexenol	2.5×10^{-07}	0.427 abcd	0.255 abce	0.711 acd	0.633 d	0.377 abcd	0.230 abce	0.472 ad	0.431 abcd	0.210 be	0.306 bce	0.509 abcd	0.348 abc	0.274 abcde	0.114 e
(*Z*)-2-Hexenol	3.2×10^{-06}	0.005 abc	0.004 ab	0.007 ac	0.008 c	0.006 ac	0.003 abc	0.007 ac	0.006 ac	0.004 b	0.005 abc	0.006 ac	0.005 ab	0.004 ab	0.004 ab
Hexanal	5.3×10^{-06}	0.900 ab	0.785 abc	0.786 abc	0.845 ab	0.575 ac	0.647 abc	0.480 c	0.541 ac	0.584 ac	1.635 b	0.577 abc	0.944 ab	0.713 abc	1.452 b
(*E*)-2-Pentenal	3.2×10^{-12}	0.081 abc	0.065 adef	0.070 abdef	0.075 ad	0.048 ef	0.131 bc	0.057 def	0.050 def	0.044 e	0.105 c	0.051 def	0.093 abc	0.072 abcdf	0.125 c
(*Z*)-3-Hexenal	1.1×10^{-15}	2.333 abc	2.129 abc	0.612 ade	0.774 a	0.473 de	6.292 bc	0.399 d	0.293 de	0.786 ae	5.543 b	0.749 ade	1.975 ac	3.913 bc	5.167 b
(*E*)-2-Hexenal	6.2×10^{-13}	11.863 abc	11.080 ab	38.124 d	28.706 d	22.506 cd	20.103 acd	21.998 cd	22.369 acd	13.345 ab	8.644 b	32.850 d	17.458 abc	17.993 abcd	9.272 b
(*E*,*E*)-2,4-Hexadienal	1.6×10^{-10}	0.374 abcd	0.187 e	0.433 abcf	0.363 abc	0.275 ade	0.734 bf	0.184 de	0.173 de	0.235 ade	0.574 bf	0.470 abcdf	0.354 acde	0.515 bcf	0.645 f
Hexyl acetate	1.8×10^{-15}	0.154 abc	0.684 a	0.389 ab	0.023 d	0.095 bce	0.194 abc	0.030 d	0.063 cde	0.076 cde	0.086 cde	0.055 cde	0.534 a	0.019 de	0.491 a
(*Z*)-3-Hexenyl acetate	1.4×10^{-13}	0.921 ab	2.532 a	0.897 abc	0.367 cd	0.493 bcd	3.014 a	0.189 d	0.263 cd	0.907 ab	0.484 bc	0.388 bcd	3.031 a	0.322 bcd	1.840 a
(*E*)-2-Hexenyl acetate	3.6×10^{-09}	0.002 a	0.022 bcd	0.034 b	0.001 ac	0.008 bcd	0.005 acd	0.001 a	0.004 acd	0.000 a	0.000 a	0.001 a	0.012 bd	0.000 a	0.002 a
1-Penten-3-one	5.3×10^{-06}	0.574 abc	0.640 abcd	0.662 abcd	0.874 ade	0.632 abc	1.331 de	0.562 b	0.612 abcd	0.638 bc	1.025 acde	0.603 abcd	0.853 ade	0.805 abcde	1.112 e

(*E*,*Z*)-Alloocimene, (*E*,*E*)-Alloocimene, (*E*,*E*)-2,6-Dimethyl-1,3,5,7-octatetraene ((*E*,*E*)-cosmene), Linalool, α-Terpineol) and 28 sesquiterpenes (i.e., α-Cubebene, Ylangene, (+)-Cyclosativene, α-Copaene, *cis*-Bergamotene, (+)-Sativene, *trans*-Bergamotene, β-Elemene, β-Copaene, β-Caryophyllene, Sesquisabinene, (*E*)-β-Farnesene, *cis*-α-Bisabolene, γ-Curcumene, γ-muurolene, (*Z*,*E*)-α-Farnesene, Eremophyllene, α-Muurolene, (*E*,*E*)-α-Farnesene, δ-Cadinene, β-Sesquiphellandrene, α-Curcumene, Sesquiterpene 2, *trans*-Geranylacetone, *cis*-calamenene, (*Z*,*E*,*E*)-allofarnesene, (*E*,*Z*,*E*)-allofarnesene, *trans*-β-Nerolidol).

The GA-LDA-HCA method was applied to each of the two matrices to group the cultivars in clusters with either similar terpene profiles or LOX VOCs profiles. The approach used to run the HCA consisted of performing several steps of clustering by reducing the total number of clusters by two at a time, until obtaining 4 clusters (i.e., 5 steps were performed to reduce the total number from 14 cultivars to 4 clusters). The clusters obtained with the two groups of molecules are summarized in Table III. As we can see, very different clusters were obtained with the two groups of molecules; in both cases, the same clusters were obtained using both the Euclidean and Manhattan distances. To preliminarily validate the obtained clusters, the membership of the 159 samples in the obtained clusters was authenticated using the respective class of molecules and the LDA method in 5-fold cross-validation mode. The accuracy of the results was 83.1% with the LOX VOCs and 92.9% with the terpenes. In summary, the initial hypothesis was confirmed by obtaining different clusters using either LOX VOCs or terpenes. Terpenes were shown to be more reliable than LOX VOCs in cultivar clustering and authentication.

3.3. OVERALL CLUSTERING OF CULTIVARS USING LOX VOCS AND TERPENES TOGETHER

Since different clusters were obtained using the different groups of volatile molecules, it can be assumed that the most reliable clustering is obtained using all the molecules together, because in this way all information is used.

Therefore, a new data matrix was created with the 159 samples in the rows and all 77 molecules (i.e., 25 LOX VOCs and 52 terpenes) in the columns; it was used as the database to apply the GA-LDA-HCA method again to group the cultivars in clusters this time with similar profiles in terms of the combination of terpenes and LOX VOCs. The HCA approach was the same as that described above, reducing the total number of clusters by two at a time, until obtaining 4 clusters. The obtained clusters are summarized in Table IV. Also in this case, the same clusters were obtained using either the Euclidean or Manhattan distance. Interestingly, the clusters were the same as those obtained using only terpenes (see section 3.2), suggesting that terpenes are more strongly associated with the botanical origin of the different cultivar clusters compared to LOX VOCs. The application of the LDA method in 5-fold cross-validation mode for allocating the 159 samples in the obtained clusters resulted in an accuracy of 94.4%, indicating that the

Table III - Clusters of cultivars obtained applying the GA-LDA-HCA approach with either Euclidean or Manhattan distance using LOX VOCs or Terpenes (comprising Mono- and Sesquiterpenes)

	Clusters from LOX VOCs		**Clusters from Terpenes**	
	Euclidean	Manhattan	Euclidean	Manhattan
Cluster 1	LdC, LE, FR, PE, MO	LdC, LE, FR, PE, MO	CO, TI, NO, PI	CO, TI, NO, PI
Cluster 2	BO, PR, PI, CO	BO, PR, PI, CO	FR, PE, CA, AT, LE, MO, LdC	FR, PE, CA, AT, LE, MO, LdC
Cluster 3	AT, TI, NO, IT	AT, TI, NO, IT	BO, PR	BO, PR
Cluster 4	CA	CA	IT	IT

AT, Ascolana Tenera; BO, Bosana; CA, Casaliva; CO, Coratina; FR, Frantoio; IT, Itrana; LE, Leccino; LdC, Leccio del Corno; MO, Moraiolo; NO, Nocellara; PE, Pendolino; PR, Peranzana; PI, Picholine; TI Tonda Iblea

Table IV - Clusters of cultivars obtained applying the GA-LDA-HCA approach with either Euclidean or Manhattan distance using the combination of LOX VOCs and Terpenes.

	Clusters from LOX VOCs + Terpenes	
	Euclidean	Manhattan
Cluster 1	FR, PE, CA, AT, LE, MO, LdC	FR, PE, CA, AT, LE, MO, LdC
Cluster 2	CO, TI, NO, PI	CO, TI, NO, PI
Cluster 3	BO, PR	BO, PR
Cluster 4	IT	IT

AT, Ascolana Tenera; BO, Bosana; CA, Casaliva; CO, Coratina; FR, Frantoio; IT, Itrana; LE, Leccino; LdC, Leccio del Corno; MO, Moraiolo; NO, Nocellara; PE, Pendolino; PR, Peranzana; PI, Picholine; TI Tonda Iblea

Table V - Confusion matrices obtained applying *GA-LDA* method to test the ability of (**A**) LOX-derived VOCs, (**B**) Monoterpenes (MT), (**C**) Sesquiterpenes (ST) and (**D**) terpenes (MT + ST) in allocating samples in the 4 clusters of cultivars defined as described in section 3.3

A. GA-LDA - LOX VOCs

Cluster	Samples	1	2	3	4
1	79	**69 (87.3 %)**	7 (8.9 %)	3 (3.8 %)	0
2	52	9 (17.3 %)	**39 (75.0 %)**	2 (3.8 %)	2 (3.8 %)
3	22	3 (13.6 %)	3 (13.6 %)	**16 (72.7 %)**	0
4	6	0	3 (50 %)	0	**3 (50 %)**
Total	**159**	Accuracy = 79.9 %			
		Class Error Average = 28.7 %			

B. GA-LDA - Monoterpenes (MT)

Cluster	Samples	1	2	3	4
1	79	**61 (77.2 %)**	16 (20.3 %)	2 (2.7 %)	0
2	52	8 (15.4 %)	**42 (80.8 %)**	2 (3.8 %)	0
3	22	8 (36.4 %)	6 (27.3 %)	**8 (36.4 %)**	0
4	6	4 (66.7 %)	2 (33.3 %)	0	**0**
Total	**159**	Accuracy = 69.8 %			
		Class Error Average = 51.4 %			

C. GA-LDA - Sesquiterpenes (ST)

Cluster	Samples	1	2	3	4
1	79	**72 (91.1 %)**	4 (5.1 %)	1 (1.3 %)	2 (2.5 %)
2	52	6 (11.5 %)	**46 (88.5 %)**	0	0
3	22	3 (13.6 %)	0	**19 (86.4 %)**	0
4	6	0	0	0	**6 (100 %)**
Total	**159**	Accuracy = 89.9 %			
		Class Error Average = 8.5 %			

D. GA-LDA - Terpenes (MT + ST)

Cluster	Samples	1	2	3	4
1	79	**73 (92.4 %)**	3 (3.8 %)	0	3 (3.8 %)
2	52	6 (11.5 %)	**46 (88.5 %)**	0	0
3	22	1 (4.5 %)	0	**21 (95.5 %)**	0
4	6	0	0	0	**6 (100 %)**
Total	**159**	Accuracy = 91.8 %			
		Class Error Average = 5.9 %			

combination of terpenes and LOX VOCs is very effective for authenticating the samples in the defined clusters.

According to the adopted classification method (i.e., GA-LDA), (*Z*)-2-pentenol, α-cubebene, *cis*-bergamotene, an unknown sesquiterpene, (*E*)-4,8-dimethylnona-1,3,7-triene and (+)-sativene were the molecules that contributed most. They comprise 1 LOX VOC, 1 monoterpene, and 5 sesquiterpenes, suggesting that sesquiterpenes are the most discriminating group of molecules. This finding preliminarily agrees with previous literature where sesquiterpenes prevailed over monoterpenes in discriminating cultivar clusters [14]; the capability of the different groups of molecules to authenticate the botanical origin of samples according to the defined clusters was further investigated as described in the next section.

3.4. COMPARATIVE ASSESSMENT OF LOX VOCS, MONO- AND SESQUITERPENES IN CULTIVAR AUTHENTICATION

Different data matrices were created. In all of them, the 159 samples were placed in the rows and were labelled with the cluster each sample belonged to. The columns were filled with: (i) the 25 LOX VOCs; (ii) the 24 monoterpenes; (iii) the 28 sesquiterpenes; (iv) the 52 terpenes (i.e., the sum of mono- and sesquiterpenes). The GA-LDA method was applied once for each of these matrices in 5-fold cross-validation mode for allocating the 159 samples in the obtained clusters, thus allowing evaluation of the capability of each group of molecules to classify the monocultivar EVOO samples in the correct cluster of cultivars. For each group of molecules, the GA-LDA method was run several times, and the average results from the 10 best combinations of variables as defined by the GA are reported in the confusion matrices in Table V.

Overall, it is immediately evident that terpenes are more effective in authenticating the botanical origin of the samples according to the defined clusters. In fact, only relatively good results were obtained using LOX VOCs, with an accuracy of 79.9%, and a class error average (CEA) of 28.7% mainly due to a high rate of misclassification for cluster 4 (50%) followed by clusters 3 (27.3%) and 2 (25.0%). However, much worse results were obtained working with monoterpenes, with an accuracy of 69.8%, and a CEA of 51.4%, with a high rate of misclassification for cluster 4 (100%) and clusters 3 (63.6%). These data indicate that monoterpenes are poorly associated with the samples' botanical origin, contributing poorly to their discrimination into the defined clusters, and that LOX VOCs showed partial association with the defined clusters. On the contrary, sesquiterpenes exhibited very good results with an accuracy of approx. 90%, and a CEA of 8.5%, with values of misclassification never greater than 13.6% (cluster 3), and even reaching 0% for cluster 4. α-Cubebene, (+)-cyclosativene, *cis*-bergamotene, and (*Z*,*E*,*E*)-allofarnesene were the molecules that made the greatest contribution to explaining the sesquiterpenes discrimination capability. The superior performance of sesquiterpenes may reflect their closer association with cultivar origin and their relative stability against extraction conditions [14,17,35], whereas LOX VOCs are enzymatically formed during crushing and malaxation and are more strongly modulated by technological and fruit-related factors, including ripening stage and phytosanitary status [36,37]. A slight increase in the performance was obtained when all terpenes (mono- and sesquiterpenes) were used together: accuracy increased to 91.8%, and CEA was only 5.9%, with values of misclassification never greater than 11.5% (cluster 2), and confirming 100% accuracy for cluster 4. Summarizing, terpenes, especially sesquiterpenes, are the class of volatile molecules more associated with the varietal origin of the MEVOOs. Their capability in discriminating among cultivars has already been highlighted in the literature [14,17,35], but this is the first article in which their discrimination capability is confirmed working with clusters of the main Italian cultivars and in which their discrimination capability was demonstrated to be greater in comparison with the LOX VOCs typical of high-quality EVOOs

3.5. VALIDATION OF THE CLASSIFICATION MODEL WITH EXTERNAL VALIDATION-SET

To validate the developed classification models, a test set of 93 samples of the same cultivars from the 2022-2023 production year was used. Applying an LDA using the different groups of variables (i.e., LOX VOCs, monoterpenes, sesquiterpenes, and terpenes all together) obtained from the previous GA-LDA method allowed the 93 samples to be allocated to the clusters defined above, thus verifying the classification capability of the proposed models. For each model, the LDA was applied using the 10 best combinations of 6 variables. A classification threshold of

Table VI - Results of validation of the GA-LDA classification models with different classes of molecules, in terms of accuracy (% on classified samples) and Class Error Average (%)

	LOX VOCs + terpenes	LOX VOCs	MT	ST	Terpenes (MT + ST)
Not classified samples (%)	5.4	21.5	34.4	6.5	4.3
Accuracy (% on classified)	88.6	75.3	77.0	90.8	93.3
CEA (%)	9.8	36.0	41.6	9.8	7.3

60% was adopted, only above which a sample was classified as belonging to that cluster; therefore, a few samples remained unclassified. The average results from the 10 best combinations of variables as defined by the GA are reported in Table VI.
Again, the best results were obtained using terpenes, especially sesquiterpenes, with 90.8% accuracy and 9.8% CEA. Interestingly, the combination of LOX VOCs with terpenes gave worse results than terpenes alone, confirming again that these two classes of molecules give diverse information on MEVOOs from different cultivars. Also in this validation step, monoterpenes and LOX VOCs gave the worst results. This validation step confirms very good results for sesquiterpenes, also confirming that this class of molecules is more associated with the botanical origin of MEVOOs from cultivars belonging to different clusters

4. CONCLUSIONS

In this study, the potential of two main groups of EVOO volatile compounds (namely terpenes (including monoterpenes and sesquiterpenes) and lipoxygenase-related compounds) was compared for the authentication of monovarietal extra virgin olive oils belonging to 14 of the main Italian cultivars. Given the recent increase in the production of and interest in these high-quality oils and the growing number of national and international competitions dedicated to monovarietal EVOOs, it has become increasingly important to have reliable tools that, on the one hand, allow the clustering of cultivars with similar characteristics into specific clusters, and on the other, authenticate the botanical origin of these oils. The work carried out in this research, also thanks to the sampling of about 250 oils over three production years, the use of validated chemical methods (using HS-SPME-GC-MS), and appropriate data-analysis tools (GA, LDA, HCA), made it possible to: i) complete the characterization of the volatile profile of these cultivars, highlighting differences in lipoxygenase-related compounds among them; ii) define cultivar clusters based solely on volatile-profile traits; iii) establish that the most reliable and effective class of volatile compounds for clustering and authenticating Italian monovarietal EVOOs is sesquiterpenes; iv) determine that monoterpenes are less closely associated with the botanical origin of the oils than sesquiterpenes and lipoxygenase-related compounds. The results of this research therefore show that, for the purpose of authenticating Italian monovarietal EVOOs, the analysis of sesquiterpenes alone is sufficient, thereby reducing the analytical effort to the assessment of these compounds only.

Acknowledgment
This research was presented at the “Congresso Nazionale Società Chimica Italiana, Milano 26-30 August 2024” as a poster presentation entitled: “Combination of HS-SPME-GC-MS analysis of terpenes and pentene dimers with chemometrics for characterization and clustering of Italian monovarietal extra virgin olive oils”. The poster was awarded the prize conferred by the Rivista Italiana delle Sostanze Grasse

REFERENCES

[1] European Commission. Commission Delegated Regulation (EU) 2022/2104 of 29 July 2022 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards marketing standards for olive oil, and repealing Commission Regulation (EEC) No 2568/91 and Commission Implementing Regulation (EU) No 29/2012. Off. J. Eur. Union., L284/1-22, (2022).
[2] European Commission. Commission Implementing Regulation (EU) 2022/2105 of 29 July 2022 laying down rules on conformity checks of marketing standards for olive oil and methods of analysis of the characteristics of olive oil. Off. J. Eur. Union., L284/23-48, (2022).
[3] A. Bajoub, A. Bendini, A. Fernández-Gutiérrez, A. Carrasco-Pancorbo. Olive oil authentication: A comparative analysis of regulatory frameworks with especial emphasis on quality and authenticity indices, and recent analytical techniques developed for their assessment. A review. Critical Reviews in Food Science and Nutrition, 58(5), 832-857, (2018)
[4] L. Conte, A. Bendini, E. Valli, P. Lucci, S. Moret, A. Maquet, ... T.G. Toschi. Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. Trends in Food Science & Technology, 105, 483-493, (2020).
[5] A. Stasi, F. Diotallevi, A. Marchini, G. Nardone, G. Italian extra-virgin olive oil: Impact on demand on being market leaders, private labels or small producers. Rev. Financ. Econ, 13, 39-54, (2018).
[6] C. Delgado, J.X. Guinard. How do consumer hedonic ratings for extra virgin olive oil relate to quality ratings by experts and descriptive analysis ratings? Food Quality and Preference, 22(2), 213-225, (2011)
[7] L. Guerrini, F. Corti, P. Masella, L. Calamai, G. Angeloni, A. Spadi, ... A. Parenti. Cross-Batch Contamination in a Continuous Horizontal Decanter Centrifuge during Virgin Olive Oil Pro-

duction. European Journal of Lipid Science and Technology, 125(2), 2200174, (2023).
[8] C. Montealegre, M.L. Marina Alegre, C. García-Ruiz, C. Traceability markers to the botanical origin in olive oils. Journal of Agricultural and Food Chemistry, 58(1), 28-38, (2010).
[9] Olive Oil Times, 2023. How Monovarietal Olive Oils Promote Ecological Farming, Safeguard Landscapes (https://www.oliveoiltimes.com/world/how-monovarietal-olive-oils-promote-ecological-farming-safeguard-landscapes/124357). Latest access, 29th October, 2024
[10] E. Casadei, E. Valli, F. Panni, J. Donarski, J.F. Gubern, P. Lucci, ... T.G. Toschi. Emerging trends in olive oil fraud and possible countermeasures. Food Control, 124, 107902, (2021).
[11] I.G. Dias, A. Fernandes, A.C.A. Veloso, A.A.S.C. Machado, J.A. Pereira, A.M. Peres. Single-cultivar extra virgin olive oil classification using a potentiometric electronic tongue. Food Chemistry, 160, 321–329. (2014).
[12] L. Cecchi, M. Migliorini, E. Giambanelli, A. Cane, B. Zanoni, V. Canuti, ... F. Melani. Is the volatile compounds profile a suitable tool for authentication of virgin olive oils (Olea europaea L.) according to cultivars? A study by using HS-SPME-GC-MS and chemometrics. Food Control, 139, 109092 (2022)
[13] L. Cecchi, A. Parenti, M. Bellumori, M. Migliorini, N. Mulinacci, L. Guerrini. Clustering monovarietal extra virgin olive oil according to sensory profile, volatile compounds, and k-mean algorithm. European Journal of Lipid Science and Technology, 124, 2200038. (2022).
[14] T. Ugolini, F. Mattagli, F. Melani, B. Zanoni, M. Migliorini, S. Trapani, ... L. Cecchi. HS-SPME-GC-MS and Chemometrics for the Quality Control and Clustering of Monovarietal Extra Virgin Olive Oil: A 3-Year Study on Terpenes and Pentene Dimers of Italian Cultivars. Journal of Agricultural and Food Chemistry, 72(19), 11124-11139, (2024)
[15] C. Campestre, G., Angelini, C., Gasbarri, F. Angerosa. The compounds responsible for the sensory profile in monovarietal virgin olive oils. Molecules, 22(11), 1833. (2017).
[16] S. Kesen., H. Kelebek., S. Selli. Characterization of the Key Aroma Compounds in Turkish Olive Oil from different Geographic Origins by Application of Aroma Extract Dilution Analysis (AEDA). J. Agric. Food Chem, 62(2), 391-401, (2014).
[17] B. Torres-Cobos, B. Quintanilla-Casas, A. Romero, A. Ninot, R.M. Alonso-Salces, T.G. Toschi, ... S. Vichi. Varietal authentication of virgin olive oil: Proving the efficiency of sesquiterpene fingerprinting for Mediterranean Arbequina oils. Food Control, 128, 108200, (2021)
[18] T. Cecchi, B. Alfei. Volatile profile of monovarietal extra virgin olive oils via HS-SPME-GC-MS: newly identified compounds, flavors molecular markers, and terpenic profile. Food Chemistry, 141, 2025–2035, (2013)
[19] D. Ocakoglu, F. Tokatli, B. Ozen, F. Korel. Distribution of simple phenols, phenolic acids and flavonoids in Turkish monovarietal extra virgin olive oils for two harvest years. Food Chemistry, 113(2), 401-410. (2009)
[20] M. Žanetić, M. Jukić Špika, M.M. Ožić, K. Brkić Bubola. Comparative study of volatile compounds and sensory characteristics of dalmatian monovarietal virgin olive oils. Plants, 10(10), 1995. (2021)
[21] V. Skiada, P. Katsaris, M.E. Kambouris, V. Gkisakis, Y. Manoussopoulos. Classification of olive cultivars by machine learning based on olive oil chemical composition. Food chemistry, 429, 136793. (2023)
[22] D.F. Al Riza, N. Kondo, V.K. Rotich, C. Perone, C., F. Giametta, F. Cultivar and geographical origin authentication of Italian extra virgin olive oil using front-face fluorescence spectroscopy and chemometrics. Food Control, 121, 107604, (2021)
[23] A. Maléchaux, R. Garcia, Y. Le Dréau, A. Pires, N. Dupuy, M.J. Cabrita. Chemometric discrimination of the varietal origin of extra virgin olive oils: usefulness of 13C distortionless enhancement by polarization transfer pulse sequence and 1H nuclear magnetic resonance data and effectiveness of fusion with mid-infrared spectroscopy data. Journal of Agricultural and Food Chemistry, 69(14), 4177-4190, (2021)
[24] B. Torres-Cobos, L. Bontempo, A. Roncone, B. Quintanilla-Casas, M. Servili, F. Guardiola, ... A. Tres. Ground-breaking comparison of target stable isotope ratios vs. emerging sesquiterpene fingerprinting for authenticating virgin olive oil origin. Food Chemistry, 478, 143655. (2025).
[25] E. Pouliarekou, A. Badeka, M. Tasioula-Margari, S. Kontakos, F. Longobardi, M.G. Kontominas. Characterization and classification of Western Greek olive oils according to cultivar and geographical origin based on volatile compounds. Journal of Chromatography A, 1218(42), 7534-7542, (2011)
[26] I. Kosma, A. Badeka, K. Vatavali, S. Kontakos, M. Kontominas. Differentiation of Greek extra virgin olive oils according to cultivar based on volatile compound analysis and fatty acid composition. European Journal of Lipid Science and Technology, 118(6), 849-861, (2016)
[27] I. Lukić, M. Lukić, M. Žanetić, M. Krapac, S.

Godena, K. Brkić Bubola. Inter-varietal diversity of typical volatile and phenolic profiles of Croatian extra virgin olive oils as revealed by GC-IT-MS and UPLC-DAD analysis. Foods, 8(11), 565, (2019)

[28] N. Martins, N.T., Jiménez-Morillo, F. Freitas, R. Garcia, M.G. da Silva, M.J. Cabrita. Revisiting 3D van Krevelen diagrams as a tool for the visualization of volatile profile of varietal olive oils from Alentejo region, Portugal. Talanta, 207, 120276, (2020)

[29] L. Cecchi, S. Orlandini, D. Balli, B. Zanoni, M. Migliorini, E: Giambanelli..., N. Mulinacci. Analysis of Volatile Hydrocarbons (Pentene Dimers and Terpenes) in Extra Virgin Olive Oil: Optimization by Response Surface Methodology and Validation of HS-SPME-GC-MS Method. Journal of Agricultural and Food Chemistry, 72(5), 2813-2825. (2024).

[30] M. Fortini, M. Migliorini, C. Cherubini, L. Cecchi, L. Calamai. Multiple internal standard normalization for improving HS-SPME-GC-MS quantitation in virgin olive oil volatile organic compounds (VOO-VOCs) profile. Talanta, 165, 641-652. (2017).

[31] H.A.N.D. Van Den Dool, P.D. Kratz. A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. Journal of chromatography. (1963).

[32] R. Aparicio-Ruiz, M.T. Morales, R. Aparicio. Does authenticity of virgin olive oil sensory quality require input from chemistry? European Journal of Lipid Science and Technology, 121(12), 1900202. (2019)

[33] A.G.A.D., Carvalho, L., Olmo-García, B.R.A., Gaspar, A., Carrasco-Pancorbo, V.N., Castelo-Branco, A.G., Torres. Evaluating quality parameters, the metabolic profile, and other typical features of selected commercial extra virgin olive oils from Brazil. Molecules, 25(18), 4193. (2020).

[34] Q., Zhou, S., Liu, Y., Liu, H., Song. Comparative analysis of volatiles of 15 brands of extra-virgin olive oils using solid-phase micro-extraction and solvent-assisted flavor evaporation. *Molecules*, *24*(8), 1512. (2019).

[35] R. Bortolomeazzi, P. Berno, L Pizzale, L.S. Conte. Sesquiterpene, Alkene, and Alkane Hydrocarbons in Virgin Olive Oils of Different Varieties and Geographical Origins. Journal of Agricultural and Food Chemistry, 49 (7), 3278–3283. (2001)

[36] C.M. Kalua, M.S. Allen, D.R. Bedgood Jr., A.G. Bishop, P.D. Prenzler, K. Robards. Olive oil volatile compounds, flavour development and quality: A critical review. *Food Chemistry*, 100(1), 273–286 (2007).

[37] M. Servili, R. Selvaggini, A. Taticchi, S. Esposto, G.F. Montedoro. Volatile compounds and phenolic composition of virgin olive oil: Optimization of temperature and time of exposure of olive pastes to air contact during the mechanical extraction process. *Journal of Agricultural and Food Chemistry*, 51, 7980–7988 (2003).