Collaborative peer validation of a harmonized SPME-GC-MS method for analysis of selected volatile compounds in virgin olive oils

Postprint of Food Control Volume 135, May 2022, 108756

DOI https://doi.org/10.1016/j.foodcont.2021.108756

Collaborative peer validation of a harmonized SPME-GC-MS method for analysis of selected volatile compounds in virgin olive oils

Ramón Aparicio-Ruiz[1], Clemente Ortiz Romero[1], Enrico Casadei[7], Diego L. García-González[1], Maurizio Servili[2], Roberto Selvaggini[2], Florence Lacoste[3], Julien Escobessa[3], Stefania Vichi[4], Beatriz Quintanilla-Casas[4], Pierre-Alain Golay[5], Paolo Lucci[6], Erica Moret[6], Enrico Valli[7]*, Alessandra Bendini[7], Tullia Gallina Toschi[7]

1- Instituto de la Grasa (CSIC), Sevilla, Spain.

2 - Department of Agricultural, Food and Environmental Sciences, Università degli Studi di Perugia, Perugia, Italy.

3 - ITERG (Institut des Corps Gras), Canejan, France.

4 - Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, Campus de l'Alimentació de Torribera, Universitat de Barcelona, Santa Coloma de Gramenet, Spain.

5 - Nestlé Research Center, Lausanne, Switzerland.

6 - Department of Agri-Food, Animal and Environmental Science Università degli Studi di Udine, Udine, Italy.

7- Department of Agricultural and Food Sciences, Alma Mater Studiorum - Università di Bologna, Cesena, Italy.

*Corresponding author: Dr. Enrico Valli, PhD. E-Mail: enrico.valli4@unibo.it; Tel: +39 0547 338116. Department of Agricultural and Food Sciences, Alma Mater Studiorum - Università di Bologna, piazza Goidanich, 60, 47521 Cesena (FC), Italy.

Abstract

The requirement for developing an instrumental method for analysis of volatile compounds responsible for the aroma that supports the work of the sensory panel test of virgin olive oils is a matter of great importance. In this paper, five laboratories participated in a collaborative study within the EU H2020 OLEUM project to develop a peer interlaboratory study of a harmonized SPME-GC-MS method for determination of volatile compounds in virgin olive oil responsible for positive attributes (e.g. fruity) and the main sensory defects. Linearity ($R^2 > 0.94$) and repeatability (mean relative standard deviation, RSD% = 7.60%) were satisfactory. Reproducibility results were uneven depending on the compound. The lowest RSD% values were found for (*Z*)-3-hexenyl acetate (19.19%), 1-hexanol (13.26%), and acetic acid (17.47%). The limits of quantification were < 0.07 mg/kg for all compounds except for (*E*)-2-decenal and pentanoic acid. The study of different quantification methods revealed that the correction of the calibration curves using the internal standard led to a slightly worse repeatability, but better accuracy and reproducibility. The results obtained by five laboratories are preparatory towards a trial proper validation study, already planned in OLEUM project, involving external labs participating on a voluntary basis.

Keywords: virgin olive oil; volatile compounds; sensory analysis; SPME-GC-MS; collaborative trial validation.

Declaration of competing interest: None.

1. Introduction

The analysis of Volatile Organic Compounds (VOCs) in combination with suitable statistical tools has been identified as the best approach for supporting the current official method of sensory evaluation of positive and negative attributes (García-González et al., 2011; Morales et al., 2013; Cecchi et al., 2019; Valli et al., 2020; Aparicio-Ruiz et al., 2019; Valli et al., 2020). Today, the evaluation of these attributes (presence/absence and their intensity) is carried out according to a method known as panel test (IOC, 1987 and subsequent amendments) which is the official method to classify virgin olive oils (VOOs) according to their sensory characteristics (EEC, 1991 and subsequent amendments). In fact, VOO is the only edible oil product with international regulations requiring official sensory analysis carried out by panelists to verify commercial categories (Garcia-Vico et al., 2017; García-González et al., 2018). However, the panel test is subject to some weaknesses and limitations (García-González & Aparicio, 2004; García-González et al., 2007; Aparicio-Ruiz et al., 2019). Thus, debated classifications are sometimes observed as well as misalignments in the classification carried out by different panels (Barbieri et al., 2020a). These problems have promoted the investigation of instrumental tools to support the daily work of panelists and to overcome other known drawbacks, such as the length and cost of the sensory analysis procedure and the limited number of panels (Aparicio-Ruiz et al., 2019; Romero et al., 2015; Casadei et al., 2021) in addition to the recommendations for managing a panel in emergency circumstances, such as a pandemic (IOC, 2020). To mitigate these drawbacks, an instrumental method based on the analysis of VOCs is required with the objective of providing additional analytical information to reinforce VOO classification into quality categories. These methods can be based on untargeted approaches with the aid of chemometric classification (García-González & Aparicio, 2004; Quintanilla-Casas et al., 2020; Garrido-Delgado et al. 2011, Valli et al., 2020; Barbieri et al., 2020b) or targeted determination of individual volatile markers as they are key odorants of VOO aroma (Aparicio et al., 2012; Morales et al., 2013; Servili et al., 2015; Cecchi et al., 2019; Casadei et al.,

2021). In the targeted determination, prior to proposing a classification scheme based on concentration ranges or decision rules, it is necessary to evaluate the performance of the method in quantitative terms with an interlaboratory perspective. Thus, in addition to intra-lab validation studies (Romero et al., 2015; Aparicio-Ruiz et al., 2018; Cecchi et al., 2019), the aim is to propose a daily routine method that is focused on detection of a minimum number of selected diagnostic markers. Moreover, before proposing this method as routine quality control, an inter-lab study was regarded as necessary to check the results when slightly different conditions are applied (e.g. different column brands, different GC instrument and MS equipment). This study would allow the evaluation of the expected errors when results from different laboratories are compared.

Although several analytical solutions have been proposed for VOO quality control, to date the regulatory bodies are unwilling to adopt them, partially due to the lack of a harmonized protocol that is accepted and internationally applied and the lack of inter-lab performance evaluation. One of the main sources of variability in the methods is the extraction technique to concentrate volatile compounds (Morales et al., 2013). In the last years, methods based on SPME are gaining importance in relation to other approaches because of their simplicity and efficiency in extraction, not only in VOO analysis (Vichi et al., 2003; Morales 2013), but also in the quality control of other foods (Giuffrida et al., 2005; Jimenez-Alvarez et al., 2008a, 2008b). Kanavouras et al. (Kanavouras, Kiritsakis & Hernandez, 2005; Kanavouras & Hernandez, 2006) compared the isolation capability between Tenax trapping and HS-SPME. They observed that a larger amount of volatile compounds was isolated when applying the first technique, while the second was quicker and led to a more rapid descriptive analysis of oxidized VOOs. On the other hand, Servili et al. (2004) compared the Head-Space Analysis (HSA) of volatile compounds in olive oils using SPME-GC/MS, electronic nose and Proton Transfer Reaction (PTR)-MS in terms of their capacity to classify VOOs according to the variety, geographical origin and ripening stage of the fruit.

Another source of variability in the analytical methods is the detector. In this regard, recently, another comparative study was carried out on two SPME-GC methods: SPME-GC-mass spectrometry (MS) and SPME-GC-Flame Ionization Detector (FID) (Aparicio-Ruiz et al., 2018). The results and the experience working with both detectors highlighted that the two options provide advantages, and thus it is necessary to evaluate the performance of methods based on the two detectors. FID is a robust and low-cost option, and commonly used in all the labs working on quality control of VOO. On the other hand, MS facilitates the identification of volatile compounds, which is particularly advantageous in VOO aroma given the presence of a large number of volatile compounds (Morales et al., 2013; Cecchi et al., 2021). On the other hand, control labs and producers demand simplicity in the analysis and they require methods that are affordable in accordance to their facilities, and GC-MS instruments are not always available in all the labs also due to the high cost. With the aim of developing analytical instrumental methods to support the panel test, the European Union has encouraged the development, harmonization and validation of such methods through the Horizon 2020 funded project OLEUM (Casadei et al., 2021). Within the frame of this project, a harmonized method with two possible detectors has been developed (SPME-GC-FID and SPME-GC-MS) to analyze volatile compounds in VOOs. The harmonization includes the definition and set up of all the possible variables that were identified as sources of errors, such as GC column, SPME fiber composition and length, vial volume, and internal standard, as well as the calibration and quantification procedures (Casadei et al., 2021). The performance of the method based on SPME-GC-FID has been evaluated in a peer interlaboratory study by three different laboratories involved in the OLEUM project (Casadei et al., 2021). With the same objective, in the present work, five laboratories, all being active partners in the OLEUM project, carried out an inter-lab evaluation of the SPME-GC-MS joint protocol. The validation was carried out by each laboratory following the same analytical conditions and on the same samples, in order to make the results obtained by each laboratory comparable in a harmonized procedure and methodology, as previously done with FID (Casadei et al., 2021). Aside from the detector, the analytical variables are the same as those used in

SPME-GC-FID, as well as the analyzed samples and the time frame given to the labs to provide their data. For these reasons, the outcomes of this work are also comparable with the results obtained by Casadei et al. (Casadei et al., 2021). Although the primary objective of this investigation is not to compare the results from SPME-GC-FID and SPME-GC-MS, the discussion on the detector is relevant and the use of a different detector means that it can be considered as another method, requiring also studying the validation parameters. Furthermore, some conclusions comparing the analytical parameters of both methods will be herein provided.

2. Materials and Methods

2.1. Chemicals

Table 1 shows the VOCs studied in this work. The pure standards of these compounds were purchased from Merck KGaA (Darmstadt, Germany). The CAS number and purity of each of the standards are also shown in Table 1. Additionally, a mixture of *n*-alkanes from 8 to 20 carbon atoms (~ 40 mg/L each, in *n*-hexane) and 4-methyl-2-pentanol (purity ≥98%) were also purchased from the same supplier for calculation of the linear retention indexes (LRI) and its use as internal standard (IS), respectively. The LRI values determined in this work matched with many reported LRI for VOCs in VOOs (Guclu et al., 2016; Da Ros et al., 2019; Zhou et al., 2019).

2.2. Samples

For this study, a set of 15 samples were selected for the peer inter-laboratory validation study of the SPME-GC-MS method. The selection was carried out to possibly cover the natural ranges of concentration normally present in VOOs and were the same samples used in a previous study on SPME-GC-FID performance (Casadei et al., 2021). These samples were sensory evaluated in the course of the OLEUM project by six panels (Barbieri et al., 2020a) to have accurate information on their commercial categories. Thus, these samples were categorized as 3 extra virgin (EV), 6 virgin (V), and 6 lampante (L) olive oils. In Vs and Ls, 6 oils were graded as rancid, 3 as fusty/muddy

144 sediment, 2 as musty-humid-earthy and 1 as winey-vinegary according to the main perceived defect
145 reported by the panelists.

146 2.3. Internal standard (IS) solution and sample preparation

147 The IS solution was prepared as described by Casadei et al. (2021). For this purpose, 4-methyl-
148 2-pentanol, the IS used in this work, was diluted in refined olive oil to have an approximate
149 concentration of 50 mg/kg. The weights during this preparation were used to calculate the exact
150 concentration. The sample was also prepared following the procedure by Casadei et al. (2021) in
151 which 0.1 g of the IS solution was added to 1.9 g of the VOO sample to have an approximate
152 concentration of 2.5 mg/kg. The exact concentration was also calculated by considering the weights
153 in the preparation.

154 *2.4. Gas chromatographic coupled to mass spectrometer analysis*

155 The sample, placed in a 20 mL vial closed with a septum (polytetrafluoroethylene), was left
156 for 10 min at 40 ºC under agitation to allow for equilibration of the volatiles in the headspace. After
157 that, the SPME fiber was exposed to the headspace for 40 min at 40 °C. The fiber was then inserted
158 into the injector port of the GC. Table 2 describes the specific characteristics of the analysis carried
159 out by the five labs that applied the joint protocol: University of Udine, University of Perugia, ITERG,
160 University of Barcelona, and Nestlé Research Center, coded as Laboratory 1-5 respectively. The
161 volatiles adsorbed by the fiber were thermally desorbed in the hot injection port of a GC for 5 min at
162 250 °C with the purge valve off (splitless mode) and injected into a capillary column of a gas
163 chromatograph with a mass spectrometry detector. The capillary column was of a polar phase based
164 on polyethylene glycol (PEG) (e.g. ZB-WAX or TR-WAX), length 60 m, internal diameter 0.25 mm
165 and coating 0.25 – 0.50 µm. The specific column brand and characteristics of each lab are shown in
166 Table 2. The transfer line temperature was set at 260 ºC. The temperature of the ion source was set
167 according to the technical specifications of each instrument. The carrier gas used by the five labs was
168 helium, although this parameter was not specified in the harmonized protocol to open the possibility

that other labs can use hydrogen if their facility is configured for that. All the labs used an autosampler although this accessory was not considered mandatory in the protocol provided since the analysis (extraction and injection) can be carried out manually. The oven temperature was held at 40 °C for 10 min and then programmed to increase by 3 °C/min to a final temperature of 200 °C. A cleaning step was added at the end of the oven programmed temperature by all participants (20 °C/min to 250 °C for 5 min) to ensure that the column was ready for the next analysis.

2.6. Identification and quantification of VOCs

Linear Retention Index (LRI) and standards were used for identification (Casadei et al., 2021) in addition to mass spectrometry (MS databases of each lab shown in Table 2). Table 1 shows the characteristic m/z of each compound to be used in the integration with the extracted ion chromatogram mode. The positive ionization mode was used in the 5 labs. Figure 1 shows the chromatogram of L and V samples.

The quantification of selected VOCs was carried out by the three quantification methods described by Casadei et al. (2021), named QM1, QM2, and QM3. These three methods were applied by the five labs using the same Excel files for the calculations. QM1 and QM2 used the calibration curves with the equations $A_{Analyte}/A_{IS} = m_{QM1} \cdot C_{Analyte}$ and $A_{Analyte} = m_{QM2} \cdot C_{Analyte}$, respectively; where $A_{Analyte}$ is the area corresponding to the analyte, A_{IS} is the area corresponding to the IS used in building the calibration curves and m_{QM1} is the slope of the calibration curve. QM3 was based in the equation $(A_{Analyte}/A_{IS}) = (m_{Analyte}/m_{IS}) \cdot (C_{Analyte}/C_{IS})$; where $A_{Analyte}$ is the area corresponding to the analyte, A_{IS} is the area corresponding to the IS, m_{IS} is the slope of the calibration curve built for IS, $m_{Analyte}$ is the slope of the calibration curve built for the analyte, $C_{Analyte}$ is the concentration corresponding to the analyte, and C_{IS} is the concentration of the IS in the sample (Kalua, Bedgood, & Prenzler, 2006).

2.7 Calibration curves

The quantification for each VOC in the headspace of VOOs was carried out by using calibration curves that were built as linear regression (intercept equal to 0), for the 18 VOCs described

194 in Table 1. These calibration curves were prepared using standard mixtures (SMs), as reported in
195 Casadei et al., 2021 (Casadei et al., 2021), instead of preparing dilutions for each single compound.
196 The two mixtures, coded as SM-A and SM-B (Table 1), were prepared to have a concentration of
197 10,000 mg/kg for each VOCs, and were used to have subsequent dilutions, coded as SM1 (200
198 mg/kg), SM2 (20 mg/kg) and SM3 (2 mg/kg). SM1 was prepared by adding 5 g of refined olive oil
199 in a 20 mL vial. Next, 0.2 g of SM-A or SM-B was added and more refined olive oil was added to
200 reach a total of 10 g. In order to prepare SM2, 1 g of SM1 was added to 5 g of refined olive oil. SM3
201 was likewise prepared by adding 1 g of SM2 to 5 g refined olive oil. The necessary weights of refined
202 oil and these three standard mixtures to obtain these concentrations are described by Casadei et al.
203 (2021).

204 The concentrations used for calibration curves were 0.05, 0.10, 0.15, 0.20, 0.25, 0.5, 1.00,
205 1.50, 2.00, 2.50, 5.00, and 10.00 mg/kg for the compounds included in SM-A. The calibration curves
206 for the compounds in SM-B were the same but adding three new points (15.00, 20.00, and 25.00
207 mg/kg), since most of these compounds were present in VOO at higher concentration.

208 The refined olive oil used in the calibration curves and in the IS solution was analyzed by
209 SPME-GC-MS for checking absence of volatile compounds that can interfere with the analyses.

210 *2.8 Peer inter-laboratory validation of the method*

211 The parameters considered were those in accordance with ISO 78-2 and ISO 5725 (ISO, 2016,
212 2019): repeatability, reproducibility, linearity, recovery, precision, limits of detection (LOD) and
213 quantification (LOQ), which were compared in order to have a peer inter-laboratory validation of the
214 method. This study was carried out for each of the 18 quantified VOCs.

215 *2.8.1 Linearity*

Linearity was evaluated using the calibration curve for each VOC (section 2.7). The regression coefficient (R^2) was considered for each calibration curve, built as linear regression passing through the origin of the axes.

2.7.2 Repeatability

The repeatability of the method was studied in terms of intra-day precision with a single operator and instrument in each of the laboratories. With this purpose, one L sample (with rancid as main perceived defect) selected from the 15 samples was analyzed seven times in a single batch; the relative standard deviation (RSD%) was calculated for each of the 18 analytes.

2.7.3 Reproducibility

For reproducibility, the study was based on the analysis of the 15 samples. These samples were analyzed in duplicate by the five laboratories. The relative standard deviation of the concentrations provided by the involved labs was calculated.

2.7.4 Recovery

Recovery was calculated by analyzing the two standard mixtures, SM-A and SM-B, diluted in refined olive oil to reach 5 mg/kg. For each of the 18 analytes, the following formula was applied:

$$R_{ap} = \frac{C}{C_{ref}} \times 100$$

Where R_{ap} was the apparent recovery, C is the concentration determined with QM1, QM2 or QM3 (see section 2.6), and C_{ref} is the actual concentration calculated from the exact weights in the dilution of SM-A and SM-B to reach the target concentration of 5 mg/kg.

2.7.5 Precision associated with the internal standard

To calculate the precision associated with the IS, the relative standard deviation (RSD) of the chromatographic area of the IS (4-methyl-2-pentanol) determined in the repeatability study (see section 2.7.2) was used. In fact, the precision should not only consider variability in the instrumental measurement, but also in the addition of the IS. The precision ($RSD\%_{Area\ IS}$) was calculated using the formula:

$$RSD\%_{Area\ IS} = \frac{\delta_{Area\ IS}}{\overline{X}_{Area\ IS}} \times 100$$

Where $\delta_{Area\ IS}$ is the standard deviation of the chromatographic areas assigned to the IS and $\overline{X}_{Area\ IS}$ is the average of these areas.

2.7.6 Limits of detection (LODs)

LOD was defined as the minimum amount or concentration of each compound that can be reliably detected. Since several procedures to calculate LOD and LOQ are available in the literature, in this investigation different calculation methods were applied, all being based on the slope of the calibration curves (m) and the standard errors of the regression ($SE_{regression}$) and intercept ($SE_{intercept}$) (Desimoni & Brunetti, 2015; Shrivastava & Gupta, 2011) through the following equations (henceforth, calculation methods 1-4):

1) LOD = 3.3 x ($SE_{regression}/m_{QM1}$), using the ratio $Area_{Analyte}/Area_{IS}$ as the variable Y of the regression and where SE is the standard error of the regression.

2) LOD = 3.3 x ($SE_{intercept}/m$), using the ratio $Area_{Analyte}/Area_{IS}$ as the variable Y of the regression with intercept different from zero.

3) LOD = 3.3 x ($SE_{intercept}/m$), using the $Area_{Analyte}$ as the variable Y of the regression with intercept different from zero.

These three methods were applied in the five laboratories to extract the LODs. Additionally, a fourth method (henceforth calculation method 4) based on the following equation was applied:

4) LOD = 3.3 x (δ_{Areas}/m_{QM1}), where δ_{Areas} (standard deviation) is referred to three replicated areas at low concentration (0.05 mg/kg).

2.7.7 Limits of determination or quantification (LOQs)

LOQ was calculated through the same calculation methods applied for LOD, but applying a factor of 10 instead of 3.3, both based on the calibration curves (see methods 1-4 listed in the section 2.7.6) and the additional calculation of S/N. In the latter, a S/N of 10 is generally accepted to be sufficient to allow for quantification of the analyte.

2.8 Data processing and statistical analysis

Data processing and calculations were carried out with Microsoft® spreadsheet program 2016 (Microsoft Corp., Redmond, WA). Outlier detection was performed with Grubbs' test (Grubbs, 1950). Analysis of variance ($p<0.05$) was carried out with Statistica (StatSoft Inc., Tulsa, OK).

3. Results and Discussion

The performance of the method was assessed through evaluation of several parameters (Aparicio-Ruiz et al., 2021), as explained in the following paragraphs. These parameters were studied for each of the 18 selected VOCs (Table 1). The rationale of the selection of these VOCs was described by Casadei et al. (2021). Thus, these compounds were considered the most suited markers to define the sensory characteristics, both fruity and defects (fermentative and non-fermentative) of VOOs. This number of compounds was considered large enough to represent the primary sensory attributes and low enough to be affordable, considering that several concentration levels need to be assessed for each of the analytes. Moreover, the presentation of the parameters for each of the VOCs is followed by a discussion on comparison of results with those related to the parallel SPME-GC-FID approach (Casadei et al., 2021) with the view to evaluate the advantages, disadvantages and/or opportunities offered by the two detectors.

283 In assessment of these parameters, data obtained by the laboratories were reported in an Excel
284 file to avoid errors and ensure that they were computed using the same procedure.

285 3.1 *Linearity*

286 Linearity was studied for the two types of calibration curves described in section 2.7 (QM1
287 and QM2). The study of regression performance (mainly R^2 coefficient and the standard deviation of
288 the regression) for these two quantification strategies allowed assessment of the effect of the IS in
289 linearity, since both quantification methods differs in the use of the IS to correct the calibration curves.
290 Table 3 shows the mean values of the R^2 for the 18 volatile compounds reported by the five labs. R^2
291 coefficients were higher than 0.94 for the 18 selected volatile compounds. The coefficients provided
292 by the labs were homogeneous and no large differences between them were detected. Thus, the
293 standard deviations of R^2 for the five labs had a maximum of 0.058 and 0.072 for QM1 and QM2
294 respectively. The R^2 data were significantly higher ($p>0.05$) for QM1 for ethyl acetate, ethanol, ethyl
295 propanoate, 3-methyl-1-butanol, while R^2 were higher for QM2 in the case of (*E*)-2-heptenal, 6-
296 methyl-5-hepten-2-one, nonanal, (*E*,*E*)-2,4-hexadienal, and pentanoic acid. However, the effect of
297 the IS was more evident in the improvement of linearity in QM1 for the aforementioned compounds.
298 The diverse effect of the use of the IS in different compounds can be explained by the degree of the
299 competition phenomena in the IS absorption to the fiber in relation to the analytes. This effect can be
300 peculiar in some cases producing some deviation of the linearity if the competition phenomena and
301 the affinity to the fiber are different for the IS and the analyte. Figure S1 shows the calibration curves
302 of ethyl propanoate and (*E*,*E*)-2,4-hexadienal as examples of two compounds in which the IS had an
303 evident effect on linearity. Although these are two extreme cases that were not seen in all the labs and
304 the effect of IS on linearity was not always so obvious, the mean R^2 (Table 3) showed a clear effect
305 of linearity for these two compounds. Thus, in the case of ethyl propanoate, the correction by the IS
306 (QM1) produced a better linearity (R^2 for QM1 and QM2 were 0.994 and 0.939, respectively), while

307 in the case of (*E*,*E*)-2,4-hexadienal, better linearity was obtained when the calibration was made
308 without the correction applied by the IS (R^2 for QM1 and QM2 were 0.975 and 0.997, respectively).

309 The compound that provided the worst linearity in terms of R^2 was (*E*)-2-decenal (R^2 for QM1
310 and QM2 were 0.942 and 0.966). On the other hand, the best linearity (R^2>0.990) was observed for
311 (*E*)-2-hexenal, acetic acid, 1-hexanol, ethyl propanoate, hexanal, octane, and 3-methyl-1-butanol.

312 The typical errors and slopes of each compound were also studied in the case of QM1, where
313 the concentration is calibrated against the ratio of $Area_{Analyte}/Area_{IS}$, and the latter ratio allows
314 comparison between labs and instruments. The slopes for each compound are shown in Table S1 and
315 Figure S2. The slope was particularly high for ethyl acetate and ethyl propanoate, with a mean slope
316 of 0.666 and 0.508, respectively (Table S1). However, the standard deviation of these mean slopes
317 (0.655 and 0.552 for ethyl acetate and ethyl propanoate, respectively) demonstrates the wide variety
318 between labs. Thus, for example, Figure 2 shows the calibration curves of the five labs for ethyl
319 propanoate. The different slopes can indicate the different sensitivities of the MS detector for this
320 compound. Excluding octane, ethanol and acetic acid, for the remainder of the compounds, the slope
321 values were lower than 0.1 (Table S2). In terms of typical error, the highest mean errors were found
322 for ethyl acetate and ethanol (0.231 and 0.184, respectively), with also a large difference between
323 labs.

324 3.2 *Repeatability*

325 The repeatability of the method was studied for each of the compounds quantified by each one
326 of the three quantification methods (QM1, QM2 and QM3). Table 4 shows the repeatability expressed
327 as mean RSD%. Considering the results for QM1, the volatile compounds with RSD% higher than
328 10% were ethyl propanoate, nonanal, and (*E*)-2-decenal. The RSD% value for the latter compound
329 was particularly high (17.23%), probably due to the low concentration in the sample studied (0.002
330 mg/kg). The average RSD% for the 18 compounds was 7.60%, although it was 6.16% when the three
331 aforementioned compounds were omitted. Regarding the other two quantification methods, QM2 and

QM3, the RSD% values were generally lower compared with QM1. However, significant differences were found only for the acids (acetic, propanoic and pentanoic acids) between the RSD% values from QM1 and QM2, in (*Z*)-3-hexenyl acetate and (*E*)-2-decenal between the RSD% values from QM1 and QM3, and in the (Z)-3-hexenyl acetate and 1-hexanol between the RSD% values from QM2 and QM3 (Table 4).

The RSD% values of the duplicates of the 15 VOOs were also examined to check if the repeatability RSD% shown in Table 4 agreed with the variability observed in the duplicates, considering that the 15 samples included a wide range of qualities and concentration values. These RSD% values are shown in Figure S3. The highest RSD% values corresponded to ethyl propanoate (8.38 ± 7.58%), nonanal (14.18 ± 13.82%), 1-octen-3-ol (11.20 ± 10.36%), (*E*,*E*)-2,4-hexadienal (9.71 ± 8.55%), (*E*)-2-decenal (10.83 ± 8.31%), and pentanoic acid (12.32 ± 11.85%). These results confirmed the lower repeatability for ethyl propanoate, nonanal and (*E*)-2-decenal.

The repeatability values presented here may have been benefited using an autosampler, which could reduce the error in the manipulation of the sample in the extraction time and injection of the fiber. However, the use of an autosampler was not considered strictly necessary since the analysis can be carried out manually and not all the labs are equipped with an autosampler. In consequence, it was not described in the protocol. Besides, a validation with a higher number of labs, some of them including manual sample preparation, would allow to evaluate the effect of using autosampler. Lastly, internal testing in the labs has proved that the benefits are not enough to include a specification of the use of this accessory, although the workload reduction is clearly an advantage.

3.3 *Reproducibility*

The reproducibility was studied by analyzing 15 samples in duplicate by each lab, including the three quality categories. Table 5 shows the mean RSD% for each VOC for the first quantification method (QM1). The concentration ranges determined by the labs for each sample are also shown in Table 5. Outliers were removed by Grubbs' test (alpha = 0.05). The higher RSD% values (> 40%)

357 corresponded to 6-methyl-5-hepten-2-one (43.20%), nonanal (46.05%), and (*E*,*E*)-2,4-hexadienal
358 (63.46%). Octane (38.50%) and ethyl propanoate (38.96%) also showed RSD% close to 40%. In the
359 case of ethyl propanoate, these values can be explained by the low concentration values (<0.05 in
360 most cases). The lowest RSD% values (< 20%) were found for (*Z*)-3-hexenyl acetate (19.19%), 1-
361 hexanol (13.26%), and acetic acid (17.47%). Table 5 shows the RSD% values when the quantification
362 methods QM2 and QM3 were applied. The RSD% values for QM1 were generally lower compared
363 with those found for QM2 and QM3. Thus, RSD% average values for the 18 compounds were
364 30.89%, 48.02% and 55.41%. The comparison of RSD% values for QM1 and QM2 revealed a
365 correction effect of the IS when results from different labs are compared, while the intra-lab
366 repeatability RSD% was similar or lower for QM2 in which no IS correction was applied (Table 5).
367 The reproducibility RSD% values of QM1 were significantly lower ($p<0.05$) than the values obtained
368 with QM2 for 10 of the 18 compounds: octane, ethyl acetate, 3-methyl-1-butanol, (*E*)-2-hexenal, (*Z*)-
369 3-hexenyl acetate, (*E*)-2-heptenal, 6-methyl-5-hepten-2-one, 1-hexanol, propanoic acid, and
370 pentanoic acid (Table 5). Regarding QM3, the RSD% values were also significantly higher than those
371 obtained with QM1 for 8 compounds. These results highlight that QM1 was the best method in terms
372 of reproducibility.

373 Although 3-methyl-1-butanol and (*E*)-2-hexenal eluted very close to each other, no apparent
374 effect was observed in the RSD% for reproducibility (25.95% and 19.55%, respectively, for QM1)
375 and repeatability (5.09% and 4.15%, respectively, for QM1). Only when these two compounds are
376 simultaneously present at high concentration, resolution problems can be given. However, (*E*)-2-
377 hexenal is typically present at high concentration in fresh EV oils while 3-methyl-1-butanol is present
378 at high concentration in V and L oils with fermentative defects (e.g. winey-vinegary defect). Thus, in
379 most of the cases, only one of the two compounds is predominant, although the identification requires
380 special attention to identify possible resolution problems.

381 3.4 *Recovery*

Recovery, among other parameters, is another important criterion to consider on which is the best quantification method. Table 6 shows the mean recovery values (%) for each of the selected volatile compounds obtained with the three quantification methods (QM1, QM2, and QM3). The recovery values derived from the ratio of the actual concentrations, obtained considering the exact weights in the dilution of SM-A and SM-B to reach the target concentration (5 mg/kg), with the calculated ones determined with the three quantification methods. The mean recovery values were 94%, 105% and 179% for QM1, QM2, and QM3, respectively. These results are comparable with the same values obtained in a parallel peer inter-laboratory validation work carried out with FID detector and three labs: 89%, 115%, and 181% for QM1, QM2, and QM3, respectively (Casadei et al., 2021). From the three quantification methods, QM1 provided the best recovery (close to 100%) among the three calculation methods, followed by QM2. Thus, the mean recovery values ranged from 72% to 106% for QM1 while they ranged from 71% to 150% for QM2. In another work, a method based on dynamic headspace thermal desorption (DHS-TD) combined to GC-MS was developed to identify and simultaneously quantify 51 VOCs in EVs and the recoveries obtained ranged from 50.9% to 113.9% (Reboredo-Rodríguez et al., 2012). However, this study was carried out with a different sampling and therefore the recovery values are not fully comparable (Oliver-Pozo et al, 2019). Following the analysis of the results in the present study, QM2 showed better results for nonanal and acetic acid compared to QM1. These results point out that the IS exerted a negative effect by introducing more error in the quantification for these two compounds, while the use of IS reduced quantification errors in terms of accuracy in the remainder of the compounds. Nevertheless, a dependent analysis of variance ($p < 0.05$) showed that there were no significant differences between the recovery values obtained with QM1 and QM2. In the case of QM3, a significant difference with respect to QM1 was observed for (*E*)-2-decenal. Furthermore, the high standard deviation for the recovery values obtained for QM3 for all the compounds points out the higher variation of the values between labs when this quantification methodology is applied. Thus, the standard deviation varied between 5-67% for QM1, 11-80% for QM2, and 29-221% for QM3.

Analyzing the differences between compounds, and focusing on recovery values for QM1, the highest errors (difference of recovery values with respect to 100%) in quantification were observed for (*E*)-2-hexenal, (*Z*)-3-hexenyl acetate, 1-octen-3-ol, acetic acid, and nonanal and (*E*)-2-decenal, which were particularly noticeable in the latter compound. Thus, the deviation of 100% recovery in this compound was around 28% (Table 6), while in the other 5 compounds this error was always below 20%. With respect to the other compounds, the deviation from R_{ap} = 100% was always lower than 10%. Only ethanol, ethyl propanoate, hexanal, (*E*)-2-heptenal, and 6-methyl-5-hepten-2-one were affected by a slight overestimation (R_{ap} >100%), while the remainder were affected by underestimation (R_{ap} <100%).

In general, the different recoveries obtained for the selected compounds can be partially explained by a low or higher adsorption on the fiber and by competition phenomena with other compounds that have a higher affinity for the fiber polymers (Oliver-Pozo, Aparicio-Ruiz, Romero, & García-González, 2015). These phenomena may influence the linearity of the calibration curves, especially when the compounds are present at high concentrations. With the aim of evaluating the impact on quantification of the possible lack of linearity at the points of high concentrations (>10 mg/kg), the analytes were quantified again using a calibration curve at low concentrations (0.05-2.5 mg/kg) and the recovery values were compared when the entire concentration range was used in the calibration (0.05-10.00/25.00 mg/kg) (Table 6). In the case of the recovery values calculated from QM3, no significant differences were observed when comparing the recoveries obtained from the two concentration ranges. The lack of a significant difference may be partially explained by the high variation of recovery values for QM3 between the 5 labs. This variation was shown by the standard deviation found for QM3 recoveries, which was higher compared with those for QM1 and QM2 (Table 6). On the contrary, in the case of QM1, significantly different recovery values were obtained for ethyl acetate and (*E*)-2-decenal, whereas significant differences were found for octane, ethyl acetate, ethyl propanoate, propanoic acid and (*E*)-2-decenal for QM2. Regarding the mean of the

433 mean recovery values, they were 94 % and 130% for QM1 when the entire concentration range and
434 the low concentration range were used respectively. These two values were 105% and 101% for QM2
435 and 179% and 176% for QM3. These results show that the calibration with lower concentrations did
436 not produce better results in general terms since significant differences were found for only some
437 compounds.

438 *3.5 Precision associated with the IS*

439 Since the IS influences quantification, the RSD% of the chromatographic areas corresponding the IS
440 was studied for each of the participant labs by analyzing the 15 samples for the reproducibility study
441 (N = 15 for each lab). The RSD% ranged from 4.02% to 15.44% for the five labs, the mean RSD%
442 being 9.66%. This error could be attributed to instrumental error or to competition phenomena in the
443 absorption to the SPME fiber rather than to the human error by adding 0.1 g of the IS solution to the
444 sample. A study made by adding 0.1g of this solution by one operator for 60 times (N = 60) revealed
445 a RSD% value in the measured weights of only 0.66%. The lowest values of the IS chromatographic
446 areas corresponded to L and V olive oils category in which high intensity of defects were identified
447 and consequently the higher concentration of compounds can produce competition phenomena
448 (Oliver-Pozo et al., 2015). Thus, two samples coded as S5 and S15 (Table 5) were characterized with
449 significantly lower values of IS chromatographic areas, and these two samples were two L oils with
450 a high median of defect (5.2 and 5.4, respectively, for fusty/muddy sediment defect). Without these
451 two samples, the average RSD% was 7.15% (ranging from 4.06% to 11.46%).

452 *3.6 Limits of detection (LOD)*

453 Three methodologies were studied to obtain the limits of detection in the calibration curves
454 built by each of the VOCs. The first method (calculation method 1, section 2.7.6) used standard error
455 of the regression and the calibration equations having an intercept forced to zero. The other two
456 methods, referred to as calculation methods 2 and 3, used calibration equations having an intercept,
457 and the standard deviation of this intercept was used in the calculation of the LOD. Method 2 used

458 the chromatographic area of the analyte divided by the area of the IS as instrument output, while
459 method 3 used the chromatographic area of the analyte. The objective of applying different methods
460 was to check the consistency of the LOD obtained through different procedures and to check which
461 results best matched with the actual observations of the signals at low concentrations (Aparicio-Ruiz
462 et al., 2018). The LOD values calculated with these methods are shown in Table 7 as means and
463 ranges of the values obtained from the laboratories involved. The values were > 0.10 mg/kg for all
464 compounds. Method 1 produced higher values than methods 2 and 3. Thus, the LOD obtained from
465 calculation method 2 ranged from 0.10 to 0.59 mg/kg, while the LODs from method 1 were higher
466 than 1.00 mg/kg for 9 compounds.

467 The highest values of LODs in the three methods were found for hexanal, 1-hexanol, 1-octen-
468 3-ol, (*E*,*E*)-2,4-hexadienal, acetic acid, and (*E*)-2-decenal (e.g. > 1.5 mg/kg for calculation method
469 1). The lowest values were found for octane, ethyl acetate, ethyl propanoate, 3-methyl-1-butanol, and
470 propanoic acid (e.g. < 0.65 mg/kg for calculation method 1). However, it was observed that
471 concentrations which were lower than the calculated LODs produced clearly detectable signals as
472 observable peaks in the chromatogram with measurable chromatographic areas. Thus, the LOD
473 values obtained with these methods did not match the perceived signals when analyzing compounds
474 in the low concentration range of the calibration curve (0.05-0.25 mg/kg). In the low concentrations,
475 the signals were always detected and linearity was observed. Table S2 shows the regression
476 coefficients (R^2) when low concentrations were considered (0.05, 0.10, 0.15, 0.20, 0.25 mg/kg). All
477 compounds showed R^2 values >0.90 in this range of the calibration, except for nonanal and (*E*)-2-
478 decenal (0.613 and 0.629, respectively), since they were barely detected at low concentration (0.05
479 mg/kg) by three of the five laboratories. On the contrary, two labs obtained R^2 values > 0.95 for these
480 two compounds. In addition, the calculated standard deviation of the R^2 presented low values, being
481 < 0.11 for all the compounds except nonanal and (*E*)-2-decenal (0.436 and 0.431, respectively). These
482 results show that the response of the detector for nonanal and (*E*)-2-decenal may differ depending on

483 the characteristics of the mass detector. The low LODs in these two compounds is also affected by
484 the low adsorption to the SPME fiber compared with other compounds. Thus, Figure S4 shows the
485 chromatograms of SMA and SMB (Table 2) diluted at a concentration of 20 mg/kg. Nonanal and (*E*)-
486 2-decenal showed a chromatographic area that were 10 times lower than the other compounds. Table
487 S3 also shows the values of the slope and intercept when a regression equation is built with the low
488 concentration range. The mean values of the slope ranged from 0.001 to 0.959, which shows a
489 different sensitivity of the detector depending on the compounds. On the other hand, the intercept
490 values were close to zero in all cases, ranging from -0.033 to 0.014, pointing out a lack of impurities
491 or noise.

492 The results described above illustrate the need to calculate LOD values that are in accordance
493 with observations when the analytes are analyzed at low concentrations. Thus, an additional method
494 (calculation method 4) based on the standard deviation of the areas for three replicates of the analyses
495 of the analytes at low concentration (0.05 mg/kg) was applied. This methodology provided more
496 representative values when it was applied in the peer validation study for SPME-GC-FID method
497 (Casadei et al., 2021). The LOD values were in the range 0.01-0.18 mg/kg. The lowest LODs (0.01
498 mg/kg) corresponded to octane, 3-methyl-1-butanol, (*E*)-2-hexenal, (*Z*)-3-hexenyl acetate, 1-hexanol,
499 1-octen-3-ol, (*E,E*)-2,4-hexadienal, acetic acid, and propanoic acid, while the highest LOD (0.18
500 mg/kg) corresponded to (*E*)-2-decenal. The comparison of these LOD values and the concentrations
501 calculated in the 15 samples (Table 5) revealed that many samples showed concentration values lower
502 than the LODs in the case of ethyl propanoate, (*E*)-2-decenal and pentatonic acid. However, these
503 problems did not fully explain the reproducibility RSD% for these compounds, since their values
504 (38.96, 36.65, 27.11% respectively when QM1 is applied) were not the highest (Table 5).

505 3.7 *Limits of determination or quantification (LOQ)*

506 The LOQ values calculated with the three methods are shown in Table 8. The values were
507 high (> 1.0 mg/kg in most of the cases) and did not correspond with the clearly distinguishable signals

508 and high linearity observed in the chromatographic areas when the analyte was present at low
509 concentrations (< 0.25 mg/kg) (Table S2). In the case of method 1, the LOQs were around 5 mg/kg
510 for hexanal, 1-hexanol, 1-octen-3-ol, acetic acid and (*E*)-2-decenal. However, with calculation
511 method 4, LOQs were in the range of 0.01-0.53 mg/kg. Considering this method, the lowest LOQs
512 (<0.03 mg/kg) corresponded to 1-hexanol, (*Z*)-3-hexenyl acetate, propanoic acid, octane, (*E*)-2-
513 hexenal, and 1-octen-3-ol. The highest LOQs (> 0.07 mg/kg) corresponded to ethyl propanoate,
514 hexanal, ethyl acetate, ethanol, nonanal, pentanoic acid and (*E*)-2-decenal. Of the latter, (*E*)-2-decenal
515 was the compound that showed a LOQ clearly above the minimum concentration used in the
516 calibration curves. The quantification procedure was strictly applied in this case as well since the aim
517 of the work was a strict application of the method and the evaluation of its performance. Regarding
518 the LOQ values for the compounds and the concentrations calculated in the 15 samples, the latter
519 were lower or close to the LOQ in most samples for ethyl propanoate, 1-octen-3-ol, (*E*)-2-decenal
520 and pentanoic acid. However, as stated above, this did not seem to affect the RSD% values for
521 reproducibility (Table 5). On the contrary, the highest RSD% value (63.46% when QM1 was applied)
522 was found for (*E*,*E*)-2,4-hexadienal (Table 5), which could be explained by the fact that its
523 concentrations was close to the LOQ limit, even if all the concentrations were higher than the LOD.
524 This could lead to some difficulties in integration and result in higher errors.

525 3.8 *Comparative study of validations using SPME-GC-FID and SPME-GC-MS methods*

526 A comparative study of the SPME-GC-FID method carried out by three labs (Casadei et al.,
527 2021) and the present SPME-GC-MS (applied by five labs) was made considering the values of the
528 parameters studied in each validation for the set of 18 VOCs. Both studies were carried out on the
529 same samples and with exactly the same procedure.

530 In terms of linearity, the mean values of R^2 were slightly higher for MS (0.983) than for FID
531 (0.977). In addition, higher R^2 values were observed using QM1 with respect to QM2, both in FID
532 and MS, which indicates a general improvement of the calibration results when the IS is added and

used in quantification of the concentration of compounds. The VOCs that showed the highest linearity in their calibrations were ethyl propanoate and 3-methyl-1-butanol for FID (R^2 = 0.998) and octane, hexanal and 3-methyl-1-butanol for MS (R^2 = 0.996). The lowest linearity was observed for (*E*)-2-heptenal in FID (R^2 = 0.936) and for (*E*)-2-decenal in MS (R^2 = 0.942). In general terms, compounds presenting high R^2 values for the labs that used FID matched with those that presented high linearity for the labs using MS. The same was observed for compounds with less linearity.

Regarding repeatability, MS presented lower mean RSD% values in each of the three QMs applied (7.60% for QM1, 6.00% for QM2 and 5.70% for QM3 in MS; compared to 11.52%, 8.18% and 9.65% in FID, respectively). Therefore, QM1 gave the highest mean RSD% value, both in FID and MS, and the best repeatability was obtained by applying QM2 in FID and QM3 in MS. The RSD% values considering the three QMs ranged between 3.60% and 15.62% for FID and between 2.21% and 17.23% for MS. Thus, the performance of the methods in terms of repeatability was similar when using the two detectors. The VOCs that showed the best repeatability (lower mean RSD% value considering the three QMs) were acetic acid and propanoic acid with FID (5.18% and 5.74%, respectively) and (*Z*)-3-hexenyl acetate and (*E*)-2-hexenal with MS (3.76% and 3.83%, respectively). Ethyl propanoate and 1-octen-3-ol had the highest mean values of RSD% in FID (13.80% and 13.29%, respectively), whereas ethyl propanoate, again, and hexanal (11.37 % and 10.14%, respectively) had the worst repeatability in MS validation.

Considering the reproducibility of the method, both for FID and MS showed similar or better RSD% values with QM1 compared with QM2 and QM3. However, the advantage of using QM1 is more evident in the method using MS. Thus, the mean RSD% values of the 18 VOCs for QM1, QM2 and QM3 were 38.79%, 39.18% and 37.66% for FID and 31.77%, 48.02% and 55.41% for MS, respectively. On the other hand, of the 18 selected compounds, the use of IS in the quantification showed to have a positive effect in reproducibility (lower RSD% for QM1 compared to QM2) in 7 compounds in FID and 16 compounds in MS. Considering only QM1, the mean RSD% for the 18

VOCs quantified was lower in MS than in FID, ranging between 12.05% (octane) and 121.99% (ethyl propanoate) for FID; and between 13.26% (1-hexanol) and 63.46% ((*E*,*E*)-2,4-hexadienal) for MS. However, excluding this anomalous value of RSD% in ethyl propanoate in the validation with FID, the mean RSD% for the rest of VOCs would be 32.59% and the maximum value of RSD% would be 48.06% for 1-hexanol. For 6 compounds (octane, ethyl acetate, 3-methyl-1-butanol, nonanal, (*E*,*E*)-2,4-hexadienal, and propanoic acid), the RSD% value was lower in the method using FID compared to MS, although 3 compounds (octane, ethyl acetate, (*E*,*E*)-2,4-hexadienal) had a clear difference, with the RSD% for FID being approximately one half. For the rest of compounds (12), the RSD% were lower for MS, and in 3 (ethyl propanoate, 1-hexanol, acetic acid) the RSD% was the half as low or even less compared to the method using FID.

When comparing the recovery between the two methods, mean values closer to 100% were observed in the laboratories that used MS for QM1 and QM2 (94% and 105% with MS vs. 89% and 115% with FID, respectively). QM3 had very high recovery values in both validations (mean values of 181% and 179% for FID and MS, respectively). Even though, as stated, the quantification with QM1 provided very similar average recovery results compared to QM2 in both validations, the mean deviation from 100% was substantially lower for QM1 in the laboratories using MS (7.70% applying QM1 vs. 16.40% with QM2). The compound with the best recovery using QM1 was 6-methyl-5-hepten-2-one in FID (99%), and 3-methyl-1-butanol and 1-hexanol (100%) in MS. The compound with deviation greater from 100% was (*E*)-2-decenal, in both FID (160%) and MS (72%).

Precision, expressed as the RSD% of the chromatographic areas corresponding to the IS (4-methyl-2-pentanol) ranged from 4.52% to 9.65% (mean 7.56%) in the validation with FID. Using MS, the RSD% ranged from 4.02% to 15.44% for the five labs, with a mean RSD% of 9.66%. As observed, the obtained values were low, which suggested good precision for both FID and MS validations. Although one of the sources of errors is the competition phenomena of the IS in the adsorption to the fiber, particularly in L oils with high median of defect, the difference in the mean

583 RSD% obtained in the studies centered in FID and MS detectors is not due these phenomena since
584 both studies were carried out with the same samples and the same procedure, so the competition
585 phenomena occurred at the same degree. Since the difference is not too high, it can be attributed to
586 the inherent error of the different instruments.

587 The LOD values of the 18 VOCs was calculated using 4 different methods. In both the
588 validations with FID and MS, calculation method 4 had lower and more representative values for this
589 parameter with respect to the other methods, and thus was the method of choice. In both cases, the
590 values coincided with the visual analysis of peaks for most of the VOCs in the calibration
591 chromatograms. On the other hand, the laboratories that used MS obtained mean values of LOD that
592 were lower than the laboratories using FID (0.03 mg/kg and 0.08 mg/kg with calculation method 4,
593 respectively). The compound with the lowest LOD in both validations was 1-hexanol (<0.005 mg/kg
594 in FID and 0.01 mg/kg in MS), while the one with the highest value for this parameter was (*E*)-2-
595 decenal (0.64 mg/kg in FID and 0.18 mg/kg in MS), for both types of detectors.

596 For the LOQ, the same conclusions as for the LOD were reached since the difference between
597 the two limits is only a factor of 3. In fact, the LOQ values were about 3 times greater than those
598 obtained in the calculation of the LOD, ranging between 0.01 mg/kg (1-hexanol) and 1.93 mg/kg
599 ((*E*)-2-decenal) in the validation with FID and between 0.01 mg/kg and 0.53 mg/kg (for the same two
600 VOCs) in validation with MS.

601 This study carried out with the same samples allowed the comparison of the interlaboratory
602 performance of the SPME-GC method with two possible detectors, FID and MS. Although the values
603 of the analytical quality parameter were different for these two detectors, showing an influence of the
604 detection principle on the quantification of the analyte, we did not observe a unanimous advantage
605 for one of them that would make the other to be discarded. Then, other considerations as the
606 availability of the detector or the lab configuration can be also taken into account when deciding on
607 one of the two approaches. MS clearly offers the advantage of an easy identification of volatile

compounds, which is particularly interesting in virgin olive oil given the high complexity of volatile composition, and for that reason MS can be regarded as a first choice provided that the available funds is not a constraint. However, the identification becomes a routine work when FID is applied after a previous training using the two standard mixtures developed (SM-A and SM-B). An international validation with more labs with no previous experience in volatile analysis would also provide useful information on the robustness of both options when they are implemented in control labs with no special requirements and the written protocol is directly applied without previous training.

4. Conclusions

The purpose of this investigation was the peer validation study of a SPME-GC-MS method for analysis of selected VOCs to support sensory analysis in quality control of VOOs. This represents a further step forwards in the quali-quantitative evaluation of diagnostic volatile markers under the same analytical conditions of a method using FID as a detector. After that, the proficiency of the method was also evaluated through a proper inter-laboratory trial with the active involvement of several external laboratories with a consolidated expertise in the olive oil analytical sector.

From this peer inter-laboratory study, method performance parameters obtained in each laboratory were investigated, compared and discussed with the aim to highlight similarities and eventual differences, as well as to calculate mean values and dispersion of the results. The quantification of the selected VOCs was carried out on the same samples by applying three different quantification methods (QMs): from analysis of all the dataset it turned out that the most promising method was QM1 using a calibration based on the IS and the external calibration curve ($A_{Analyte}/A_{IS}$ vs. $C_{Analyte}$). Although QM1 showed slightly worse repeatability than the other methods, it had better accuracy and reproducibility. This finding was also observed for the FID method, even if with MS it was more evident. In general, satisfactory results were obtained for linearity, recovery, precision and repeatability parameters, although reproducibility has a rather high RSD% (>40%) for some

633 compounds (ethyl propanoate, 6-methyl-hepten-2-one, and (*E*,*E*)-2,4-hexadienal). Further
634 investigation in a validation study with more labs including more diversity of instruments and GC
635 columns brands and the use of manual injection would serve to assess the effect of these variables on
636 the method performance.

637 The results of this work also serve to optimize future application of the method and to have
638 an accurate knowledge of the errors. The first interlaboratory experiences carried within OLEUM
639 project revealed that the RDS% values for reproducibility were higher than 100% in many cases when
640 the analytical variables were not harmonized. In the results showed in this study, some compounds
641 provided RSD% higher than 35%. When proposing concentration limits and ranges for each category,
642 these errors need to be considered as well together with other aspects, like the odor thresholds and
643 the masking effect between aromas. On the other hand, the management of the concentration limits
644 and the associated errors is influenced by the specific classification criteria. Then, in the particular
645 cases of the differentiation between L and non-L, L oils show clearly high concentrations of volatile
646 markers, while the differentiation between EV and non-EV is based on the absence or the presence
647 of some volatile markers even at very low concentrations.

648 This study compared the performance characteristics of the method when applied with FID or
649 MS. Given that these two options provide advantages and disadvantages, and that they are
650 alternatively available in the labs working in olive oil analysis, knowledge on their performance is
651 needed. Only at the end of a full validation process with the involvement of a large number of
652 laboratories participating on a voluntary basis, it will be possible to conduct a study aimed at
653 individuating the concentration ranges of variability, as well as a proposal of limits, for the selected
654 volatile compounds (especially those related to sensory defects) in relation to the different quality
655 grades of VOOs. Moreover, also considering the pros that - for the samples analyzed herein - the
656 sensory evaluation was performed by 6 different panels, the concentrations obtained could be related

with the presence of sensory defects or positive attributes (fruity), thus being useful to define the ranges/limits for the selected markers in order to support the panel test.

Funding: This work was supported by the Horizon 2020 European Research project OLEUM "Advanced solutions for assuring the authenticity and quality of olive oil at a global scale", which has received funding from the European Commission within the Horizon 2020 Programme (2014–2020), grant agreement no. 635690. The information expressed in this article reflects the authors' views; the European Commission is not liable for the information contained herein.

CRediT authorship contribution statement

Ramón Aparicio-Ruiz: Conceptualization, Methodology, Formal analysis, Software, Data curation, Writing - original draft, Writing - review & editing. **Clemente Ortiz-Romero**: Formal analysis, Methodology, Validation, Writing - review & editing, Software, Data curation. **Enrico Casadei**: Conceptualization, Methodology, Validation, Formal analysis, Writing - original draft, Writing - review & editing. **Diego L. García González**: Conceptualization, Methodology, Validation, Data curation, Writing - original draft, Writing - review & editing, Supervision. **Maurizio Servili**: Conceptualization, Methodology, Validation, Data curation, Writing - review & editing, Supervision. **Roberto Selvaggini**: Conceptualization, Methodology, Formal analysis, Data curation, Writing - review & editing. **Florence Lacoste**: Conceptualization, Methodology, Validation, Data curation, Writing - review & editing, Supervision. **Julien Escobessa**: Methodology, Formal analysis, Data curation, Writing - review & editing. **Stefania Vichi**: Formal analysis, Methodology, Validation, Data curation, Writing – review & editing. **Beatriz Quintanilla-Casas:** Formal analysis, Methodology, Data curation, Writing - review & editing. **Pierre Alain Golay**: Conceptualization, Methodology, Formal analysis, Data curation, Writing - review & editing. **Paolo Lucci**:

Conceptualization, Methodology, Validation, Data curation, Writing - review & editing, Supervision. **Erica Moret**: Conceptualization, Methodology, Formal analysis, Data curation, Writing - review & editing. **Enrico Valli**: Conceptualization, Methodology, Validation, Data curation, Writing - original draft, Writing - review & editing, Supervision. **Alessandra Bendini**: Conceptualization, Methodology, Validation, Data curation, Writing - review & editing, Supervision. **Tullia Gallina Toschi**: Conceptualization, Methodology, Validation, Project administration, Writing - review & editing, Supervision.

Acknowledgements: The authors are grateful to the sensory panel leaders and tasters involved in the OLEUM project (Eurofins, Germany; ITERG, France; IPTPO, Croatia; UNIBO, Italia; UZZK, Turkey; ZRS/UP, Slovenia) and to the numerous companies that provided the commercial virgin olive oils for this investigation. The authors would like to express their gratitude to Prof. Lanfranco Conte for his contribution in terms of discussion and ideas on the herein presented method.

694 **References**

695 Aparicio, R., Morales, M. T., & García-González, D. L. (2012). Towards new analyses of aroma and
696 volatiles to understand sensory perception of olive oil. *European Journal of Lipid Science and*
697 *Technology*, *114*, 1114–1125. https://doi.org/10.1002/ejlt.201200193.

698 Aparicio-Ruiz, R., García-González, D. L., Morales, M. T., Lobo-Prieto, A., & Romero, I. (2018).
699 Comparison of two analytical methods validated for the determination of volatile compounds in virgin
700 olive oil: GC-FID vs GC-MS. *Talanta*, *187*, 133–141. https://doi.org/10.1016/j.talanta.2018.05.008

701 Aparicio-Ruiz, R., Morales, M. T., & Aparicio, R. (2019). Does authenticity of virgin olive oil
702 sensory quality require input from chemistry? *European Journal of Lipid Science and Technology*,
703 *121*, 1900202. https://doi.org/10.1002/ejlt.201900202.

704 Aparicio-Ruiz, R., Ortiz Romero, C., Casadei, E., García González D. L., Servili M., Selvaggini R.,
705 et al. (2021). OLEUM Project. Data of a harmonized SPME-GC-MS method for the analysis of
706 selected volatile compounds in virgin olive oils [Dataset]. CSIC,
707 http://dx.doi.org/10.20350/digitalCSIC/13965.

708 Barbieri, S., Bubola, K. B., Bendini, A., Bučar-Miklavčič, M., Lacoste, F., Tibet, U., et al. (2020a).
709 Alignment and proficiency of virgin olive oil sensory panels: The OLEUM approach. *Foods*, *9*, 355.
710 https://doi.org/10.3390/foods9030355.

711 Barbieri, S., Cevoli, C., Bendini, A., Quintanilla-Casas, B., García-González, D. L., & Gallina Toschi,
712 T. (2020b). Flash gas chromatography in tandem with chemometrics: a rapid screening tool for
713 quality grades of virgin olive oils. *Foods*, *9*, 862. https://doi.org/10.3390/foods9070862.

714 Desimoni, E., & Brunetti, B. (2015). About estimating the limit of detection by the signal to noise
715 approach. *Pharmaceutica Analytica Acta*, *6*, 1000355. https://doi.org/10.4172/2153-2435.1000355.

716 Casadei, E., Valli, E., Aparicio-Ruiz, R., Ortiz Romero, C., García González, D. L., Vichi, S., et al.
717 (2021). Peer inter-laboratory validation study of a harmonized SPME-GC-FID method for the
718 analysis of selected volatile compounds in virgin olive oils. *Food Control,* 123, 107823.
719 https://doi.org/10.1016/j.foodcont.2020.107823.

720 Cecchi, L., Migliorini, M., & Mulinacci, N. (2021). Virgin olive oil volatile compounds:
721 Composition, sensory characteristics, analytical approaches, quality control, and authentication.
722 *Journal of Agricultural and Food Chemistry, 69,* 2013–2040.
723 https://doi.org/10.1021/acs.jafc.0c07744.

724 Cecchi, L., Migliorini, M., Giambanelli, E., Rossetti, A., Cane, A., Melani, F., et al. (2019).
725 Headspace solid-phase microextraction−gas chromatography−mass spectrometry quantification of
726 the volatile profile of more than 1200 virgin olive oils for supporting the panel test in their
727 classification: comparison of different chemometric approaches. *Journal of Agricultural and Food*
728 *Chemistry, 67*, 9112−9120. https://doi.org/10.1021/acs.jafc.9b03346.

729 Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet A., et al. (2020). Olive oil quality and
730 authenticity: A review of current EU legislation, standards, relevant methods of analyses, their
731 drawbacks and recommendations for the future. *Trends in Food Science & Technology, 105,*
732 485−493. https://doi.org/10.1016/j.tifs.2019.02.025.

733 Da Ros, A., Masuero, D., Riccadonna, S., Brkić Bubola, K., Mulinacci, N., Mattivi, F., Lukić, I., &
734 Vrhovsek, U. (2019). Complementary Untargeted and Targeted Metabolomics for Differentiation of
735 Extra Virgin Olive Oils of Different Origin of Purchase Based on Volatile and Phenolic Composition
736 and Sensory Quality. Molecules (Basel, Switzerland), 24(16), 2896.
737 https://doi.org/10.3390/molecules24162896

738 European Commission Regulation (1991). On the characteristics of olive oil and olive residue oil and
739 on the relevant methods of analysis, and subsequent amendments. *Official Journal of European*
740 *Community 11* (L248), 1–102, 2568/91.

741 Gallina Toschi, T., Valli, E., Conte, L., García-González, D. L., Maquet, A., Brereton, P., et al.
742 (2017). EU project OLEUM: Better solutions to protect olive oil quality and authenticity. *Agro Food*
743 *Ind. Hi-Tech, 28,* 2-3. https://www.teknoscienze.com/tks_article/eu-project-oleum-better-solutions-
744 to-protect-olive-oil-quality-and-authenticity/

745 García González D.L., Aparicio, R., & Aparicio-Ruiz, R (2018). Olive oil. In: J.F. Morin & Michèle
746 Lees (Eds.), FoodIntegrity Handbook: A guide to food authenticity issues and analytical solutions
747 (pp. 335-358). Eurofins Analytics France. https://doi.org/10.32741/fihb.18.oliveoil.

748 García-González, D. L., Aparicio, R. (2004). Classification of different quality virgin olive oils by
749 metal-oxide sensors. *European Food Research Technology., 218,* 484–487.
750 https://doi.org/10.1007/s00217-003-0855-4.

751 García-González, D. L., Aparicio, R. (2010). Research in olive oil: Challenges for the near future.
752 *Journal of Agricultural and Food Chemistry, 58,* 12569-12577. https://doi.org/10.1021/jf102735n

753 García-González, D. L., Tena, N., & Aparicio, R. (2007). Characterization of olive paste volatiles to
754 predict the sensory quality of virgin olive oil. *European Journal of Lipid Science and Technology,*
755 *109,* 663-672. https://doi.org/10.1002/ejlt.200700056.

756 García-González, D. L., Vivancos, J., & Aparicio, R. (2011). Mapping brain activity induced by
757 olfaction of virgin olive oil aroma. *Journal of Agricultural and Food Chemistry, 59,* 10200–10210.
758 https://doi.org/10.1021/jf202106b.

759 Giuffrida, F., Golay, P-A., Destaillats, F., Hug, B., & Dionisi, F. (2005). Accurate determination of
760 hexanal in beef bouillons by headspace solid-phase microextraction gas-chromatography mass-

761 spectrometry. *European Journal of Lipid Science and Technology, 107*, 792-798.
762 https://doi.org/10.1002/ejlt.200500240.

763 Guclu G, Sevindik O, Kelebek H, Selli S. Determination of Volatiles by Odor Activity Value and
764 Phenolics of cv. Ayvalik Early-Harvest Olive Oil. Foods. 2016; 5(3):46.
765 https://doi.org/10.3390/foods5030046

766 International Olive Council (IOC). (1987). *Sensory analysis of olive oil method for the organoleptic*
767 *assessment of virgin olive oil*. IOOC/T.20/Doc. no. 3.

768 International Olive Council (IOC). (2020). *Guidelines for the management of virgin olive oil tasting*
769 *panels in the event of a pandemic*. COI/MPP/Doc. No 1/Rev 1 November 2020.

770 International Organization for Standardization (ISO). (2016). Chemistry - layouts for standards - *Part*
771 *2. Methods of chemical analysis, 78–2*, 1999.

772 International Organization for Standardization (ISO). (2019). Accuracy (trueness and precision) of
773 measurement methods and results - *Part 2. Basic method for the determination of repeatability and*
774 *reproducibility of a standard measurement method, 5725–2*, 2019.

775 Jimenez-Alvarez D., Giuffrida F., Golay P-A, Cotting C., Destaillats F., Dionisi F., et al. (2008b).
776 Profiles of volatile compounds in milk containing fish oil analyzed by HS-SPME-GC/MS, *European*
777 *Journal of Lipid Science and Technology, 110,* 277-283. https://doi.org/10.1002/ejlt.200700148.

778 Jimenez-Alvarez, D., Giuffrida, F., Golay, P-A., Cotting, C., Lardeau, A., & Keely, B.J. (2008a).
779 Antioxidant activity of oregano, parsley, and olive mill wastewaters in bulk oils and oil-in-water
780 emulsions enriched in fish oil. *Journal of Agricultural and Food Chemistry, 56,* 7151-7159.
781 https://doi.org/10.1021/jf801154r.

782 Kalua, C. M., Allen, M. S., Bedgood, D. R., Bishop, A. G., Prenzler, P. D., & Robards, K. (2007).
783 Olive oil volatile compounds, flavour development and quality: A critical review. *Food Chemistry, 100*, 273–286. https://doi.org/10.1016/j.foodchem.2005.09.059.

785 Kalua, C. M., Bedgood, D. R., & Prenzler, P. D. (2006). Development of a headspace solid phase
786 microextraction-gas chromatography method for monitoring volatile compounds in extended time -
787 course experiments of olive oil. *Analytica Chimica Acta, 556*, 407–414.
788 https://doi.org/10.1016/j.aca.2005.09.050.

789 Kanavouras, A., & Hernandez, R. J. (2006). The analysis of volatiles from thermally oxidized virgin
790 olive oil using dynamic sorption-thermal desorption and solid phase micro-extraction techniques.
791 *International Journal of Food Science and Technology, 41*, 743–750. https://doi.org/10.1111/j.1365-
792 2621.2005.01019.x

793 Kanavouras, A., Kiritsakis, A., & Hernandez, R. J. (2005). Comparative study on volatile analysis of
794 extra virgin olive oil by dynamic headspace and solid phase micro-extraction. *Food Chemistry, 90*,
795 69–79. https://doi.org/10.1016/j.foodchem.2004.03.025.

796 Morales, M. T., Aparicio-Ruiz, R., & Aparicio, R. (2013). Chromatographic methodologies:
797 Compounds for olive oil odor issues. In R. Aparicio-Ruiz, J. Harwood (Eds.), *Handbook of Olive Oil:*
798 *Analysis and Properties* (pp. 261-309). Springer.

799 Morales, M. T., Luna, G., & Aparicio, R. (2005). Comparative study of virgin olive oil sensory
800 defects. *Food Chemistry, 91*, 293−301. https://doi.org/10.1016/j.foodchem.2004.06.011.

801 Oliver-Pozo, C., Aparicio-Ruiz, R., Romero, I., & García-González, D. L. (2015). Analysis of volatile
802 markers for virgin olive oil aroma defects by SPME-GC/FID: Possible sources of incorrect data.
803 *Journal of Agricultural and Food Chemistry, 63*, 10477−10483.
804 https://doi.org/10.1021/acs.jafc.5b03986.

Oliver-Pozo, C., Trypidis, D., Aparicio, R., García-González, D. L., Aparicio-Ruiz, R. (2019). Implementing dynamic headspace with SPME sampling of virgin olive oil volatiles: Optimization, quality analytical study, and performance testing. *Journal of Agricultural and Food Chemistry*, *67*, 2086–2097. https://doi.org/10.1021/acs.jafc.9b00477.

Reboredo-Rodríguez, P., González-Barreiro, C., Cancho-Grande, B., & Simal-Gándara, J. (2012). Dynamic headspace/GC-MS to control the aroma fingerprint of extra-virgin olive oil from the same and different olive varieties. *Food Control*, *25*, 684–695. https://doi.org/10.1016/j.foodcont.2011.12.005.

Romero, I., García-González, D. L., Aparicio-Ruiz, R., & Morales, M. T. (2015). Validation of SPME-GCMS method for the analysis of virgin olive oil volatiles responsible for sensory defects. *Talanta*, *134*, 394–401. https://doi.org/10.1016/j.talanta.2014.11.032.

Romero, I., García-González, D. L., Aparicio-Ruiz, R., & Morales, M. T. (2017). Study of volatile compounds of virgin olive oils with 'frostbitten olives' sensory defect. *Journal of Agricultural and Food Chemistry, 65*, 4314−4320. https://doi.org/10.1021/acs.jafc.7b00712.

Salas, J. J., Sánchez, C., García-González, D. L., & Aparicio, R. (2005). Impact of the suppression of lipoxygenase and hydroperoxide lyase on the quality of the green odor in green leaves. *Journal of Agricultural and Food Chemistry, 53*, 1648−1655. https://doi.org/10.1021/jf040331l.

Serrano, A., de la Rosa, R., Sánchez-Ortiz, A., & León, L. (2020) Genetic and environmental effect on volatile composition of extra virgin olive oil. *European Journal of Lipid Science and Technology*, *122*, 1–10. https://doi.org/10.1002/ejlt.202000162.

Servili, M., Esposto, S., Taticchi, A., Urbani, S., Di Maio, I., Veneziani, G., & Selvaggini, R. (2015). New approaches to virgin olive oil quality, technology, and by-products valorization. *European Journal of Lipid Science and Technology*, *117*, 1882–1892. https://doi.org/10.1002/ejlt.201500138.

828 Shrivastav, A., & Gupta, V. P. (2011). Methods for the determination of limit of detection and limit
829 of quantification of the analytical methods. *Chronicles of Young Scientists*, *2*, 21-25.
830 https://doi.org/10.4103/2229-5186.79345.

831 Van den Dool, H., & Kratz, P. D. (1963). A generalization of the retention index system including
832 linear temperature programmed gas-liquid partition chromatography. *Journal of Chromatography A*,
833 *11*, 463–471. https://doi.org/10.1016/S0021-9673(01)80947-X.

834 Vichi, S., Castellote, A. I., Pizzale, L., Conte, L. S., Buxaderas, S., & López-Tamames, E. (2003).
835 Analysis of virgin olive oil volatile compounds by headspace solid-phase microextraction coupled to
836 gas chromatography with mass spectrometric and flame ionization detection. *Journal of*
837 *Chromatography A*, *983* (1–2), 19–33. https://doi.org/10.1016/S0021-9673(02)01691-6.

838 Zhou, Q., Liu, S., Liu, Y., & Song, H. (2019). Comparative Analysis of Volatiles of 15 Brands of
839 Extra-Virgin Olive Oils Using Solid-Phase Micro-Extraction and Solvent-Assisted Flavor
840 Evaporation. Molecules (Basel, Switzerland), 24(8), 1512.
841 https://doi.org/10.3390/molecules24081512

842

Figure captions

Figure 1. Chromatogram of volatile compounds of a lampante olive oil and a virgin olive analysed by SPME-GC-MS. The correspondence of the codes with the volatile compounds is shown in Table 1.

Figure 2. Calibration curves of ethyl propanoate built for the quantification method 2 (QM2).

Table 1. Selected volatile compounds, CAS numbers, purities of chemical standards, linear retention times (LRT), characteristic m/z in the mass spectra and sensory defects or positive attribute to which they are related.

Code	Volatile compound	CAS number	Purity of the chemical standard (%)[a]	LRI[b]	SM_x[c]	Characteristic m/z	Related defect/attribute[d]
1	Octane	111-65-9	99.7	802 ± 1.85	A	85	Fusty/muddy sediment
2	Ethyl acetate	141-78-6	99.8	899 ± 0.84	A	43	Winey-vinegary
3	Ethanol	64-17-5	99.9	933 ± 1.06	B	31, 45	Winey-vinegary, fusty/muddy sediment
4	Ethyl propanoate	105-37-3	99.7	954 ± 7.23	A	57	Frostbitten olives
5	Hexanal	66-25-1	98	1080 ± 8.02	B	44	Rancid
6	3-Methyl-1-butanol	123-51-3	98.5	1210 ± 4.53	A	55, 70	Fusty/muddy sediment
7	(*E*)-2-Hexenal	6728-26-3	97.0	1215 ± 9.18	B	69, 83	Fruity
8	(*Z*)-3-Hexenyl acetate	3681-71-8	98.0	1312 ± 4.96	B	67, 82	Fruity
9	(*E*)-2-Heptenal	18829-55-5	95	1321 ± 10.08	A	83	Musty-humid-earthy, rancid
10	6-Methyl-5-hepten-2-one	110-93-0	97.0	1337 ± 10.00	A	108	Fusty/muddy sediment
11	1-Hexanol	111-27-3	99.9	1356 ± 4.79	B	56	Fruity
12	Nonanal	124-19-6	95	1392 ± 9.21	B	98	Rancid
13	1-Octen-3-ol	142-83-6	98.0	1453 ± 6.70[e]	B	81	Musty-humid-earthy
14	(*E*,*E*)-2,4-Hexadienal	3391-86-4	95.0	1401 ± 10.71[e]	A	57	Rancid
15	Acetic acid	64-19-7	99.8	1475 ± 35.27	B	60	Winey-vinegary
16	Propanoic acid	79-09-4	99.8	1547 ± 46.54	A	74	Fusty/muddy sediment, musty-humid-earthy
17	(*E*)-2-Decenal	3913-81-3	95.0	1644 ± 10.39	A	70	Rancid
18	Pentanoic acid	109-52-4	99.8	1759 ± 32.92	A	60, 73	Rancid

[a] Minimum purity as expressed by the supplier.
[b] LRI: Linear Retention Index, Relative Retention Time indicative parameter. Mean ± error from two labs that reported the results (UNIUD and UNIPG).
[c] SM: Standard mixture containing each volatile compound (SM-A: low concentration range 0.05-10.00 mg/kg; SM-B: high concentration range 0.20-25.00 mg/kg).
[d] Main perceived defect/attribute when the volatile compound is at high concentrations (above its odor threshold). Some compounds may be related to more than one defect/attribute. More information can be found in Casadei et al. (2021), Morales et al. (2005, 2013).
[e] The order of these two compounds may be altered depending on the column brand and/or column film thickness.

Table 2. Characteristics of the GC-MS instruments used in each lab during the peer inter-laboratory validation study.

	Laboratory 1	**Laboratory 2**	**Laboratory 3**	**Laboratory 4**	**Laboratory 5**
GC Instrument (Equipment)	7890B, Agilent Technologies[1]	78900A, Agilent Technologies[1]	7890B, Agilent Technologies[1], equipped with a "Multimode Injector" (MMI) 7693A	6890N, Agilent Technologies[1]	HP6890, Agilent Technologies[1]
Autosampler	PAL RSI 85, CTC Analytics AG[2]	Combipal, CTC Analytics[2]	PAL3 RSI 120, CTC Analytics AG[2]	Combi-PAL, CTC Analytics AG[2]	MPS (MultiPurpose Sampler), GERSTEL GmbH & Co.KG[3]
Sample agitation	250 rpm Agitator on time (s): 5; Agitator off time (s): 2	400 rpm (continuous)	No agitation applied	250 rpm (continuous)	250 rpm Agitator on time (s): 3; Agitator off time (s): 90
GC column	DB-WAX, Agilent J&W[1], CA. 60 m; I.D. 0.25 mm; film thickness 0.25 µm	ZB-WAX Zebron, Phenomenex[4], 60 m x 250 µm x 0.25 µm,	HP-INNOWax, Agilent Technologies[1], 60 m; i.d. 0.25 mm; film thickness 0.25 µm	Supelcowax-10, Supelco[5], 60 m; I.D. 0.25 mm; film thickness 0.25 µm.	DB-WAX Ultra Inert, Agilent J&W[1], length 60 m, i.d. 0.25 mm; film thickness 0.5 µm
MS instrument (equipment)	5977A, Agilent Technologies[1], single quadrupole mass spectrometer,	5975C, Agilent Technologies[1], single quadrupole mass spectrometer,	5977B, Agilent Technologies[1], single quadrupole mass spectrometer with EI Extractor (XTR) source	5975C, Agilent Technologies[1], inert XL quadrupolar analyser	MSD5975, Agilent Technologies[1], single quadrupole mass spectrometer
MS database	NIST v14[6]	NIST MS Search 2.0[6]	NIST v14[6]	Wiley6[7]	NIST v14[6]
GC-MS Interface Temp.	280°C	275 °C	260 °C	280ºC	220°C
Ion source temperature	175°C	230°C	200°C	230ºC	200°C
Mass range m/z	31-350 m/z	30-300 m/z	25–350 m/z	35-300 m/z	29-350 m/z
Quadrupole temperature	150°C	150°C	190°C	150ºC	150°C
Scan rate	1.6 scans/s	5.1 scan/s	4.3 scan/s	5.1 scans/s	2.0 scans/s

Note: [1], Agilent Technologies, Santa Clara, CA, USA ; [2], CTC Analytics AG, Zwingen, Switzerland; [3], GERSTEL GmbH & Co.KG, Mülheim an der Ruhr, Germany; [4], Torrance, CA, USA; [5], Bellefonte, PA, USA; [6], Gaithersburg, MD; [7], Hoboken, NJ, USA.

Table 3. Linearity expressed as R^2 (mean and standard deviation of the five labs) computed from the calibration curves used in the quantification methods 1 and 2 (QM1, QM2) for the 18 volatile compounds.

Volatile compounds	QM1	QM2
Octane	0.996±0.003	0.966±0.038[a]
Ethyl acetate	0.982±0.023[a]	0.906±0.078[a]
Ethanol	0.984±0.011[a]	0.953±0.047[a]
Ethyl propanoate	0.994±0.008	0.939±0.053[a]
Hexanal	0.996±0.003	0.980±0.021
3-methyl-1-butanol	0.996±0.002	0.941±0.068
(*E*)-2-Hexenal	0.990±0.009[b]	0.994±0.007[b]
(*Z*)-3-Hexenyl acetate	0.987±0.012[b]	0.992±0.006[b]
(*E*)-2-Heptenal	0.976±0.027[b]	0.997±0.001
6-Methyl-5-hepten-2-one	0.975±0.025[b]	0.997±0.001
1-Hexanol	0.993±0.006	0.992±0.005
Nonanal	0.976±0.024	0.990±0.007
1-Octen-3-ol	0.983±0.019	0.993±0.005
(*E*,*E*)-2,4-Hexadienal	0.975±0.027	0.997±0.002
Acetic acid	0.993±0.005	0.989±0.011
Propanoic acid	0.983±0.028[b]	0.995±0.005
(*E*)-2-Decenal	0.942±0.057[b]	0.966±0.025[b]
Pentanoic acid	0.969±0.032[b]	0.993±0.008[b]

[a] Certain saturation at high concentrations in data provided by some of the involved labs.
[b] Certain lower sensitivity (lower slope) at low concentrations in data provided by some of the involved labs.

Table 4. Repeatability expressed as mean RSD%.

Volatile compounds	RSD% (Mean±SD)		
	QM1	QM2	QM3
Octane	6.77±4.33[a]	7.95±4.11	6.47±4.91
Ethyl acetate	6.99±3.49	4.77±0.21	5.75±4.02
Ethanol	9.51±2.72	6.21±2.14	6.52±1.94
Ethyl propanoate	15.27±15.87[a]	15.55±15.63	15.13±17.34
Hexanal	5.49±3.67	4.84±2.00	4.53±1.94
3-Methyl-1-butanol	5.09±1.80	5.63±2.58	2.88±2.44
(*E*)-2-Hexenal	4.15±1.74	2.99±0.40	2.21±1.30
(*Z*)-3-Hexenyl acetate	5.23±0.55[c]	4.86±0.84[d]	3.11±0.61[cd]
(*E*)-2-Heptenal	5.38±0.76	4.75±4.23	3.31±3.61
6-Methyl-5-hepten-2-one	5.05±1.17	5.82±0.89	4.40±0.07
1-Hexanol	3.89±1.46	4.12±0.72[d]	2.39±0.34[d]
Nonanal	11.84±7.33[a]	9.89±3.96	7.36±9.39
1-Octen-3-ol	6.98±1.59	5.40±0.98	5.84±3.03
(*E,E*)-2,4-Hexadienal	8.51±2.99	4.20±0.72	6.79±5.13
Acetic acid	7.87±0.47[b]	3.48±2.59[b]	5.48±3.09
Propanoic acid	5.70±0.19[b]	2.35±1.56[b]	3.32±2.08
(*E*)-2-Decenal	17.23±5.08[c]	12.00±2.77	13.86±5.10[c]
Pentanoic acid	5.83±0.27[b]	3.17±0.58[b]	2.83±1.86

[a] One outlier has been removed (Grubbs test $p<0.05$).
[b] Significant difference ($p<0.05$) between QM1 and QM2.
[c] Significant difference ($p<0.05$) between QM1 and QM3.
[d] Significant difference ($p<0.05$) between QM2 and QM3.

Table 5. Reproducibility values for the SPME-GC-MS method expressed as the mean of the RSD%, calculated for each of the 15 analyzed samples (S1-S15). The concentration ranges (minimum and maximum values) and the mean RSD% values are also shown.

Compounds	Concentration range (mg/kg) in samples (S) - *Minimum (first row)/Maximum (second row)*															RSD% QM1[a]	RSD% QM2[a]	RSD% QM3[a]
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S12**	**S13**	**S14**	**S15**			
Octane	0.02^{f}	0.06	0.06	0.08	1.18	0.06	0.02^{f}	0.53	0.03	0.12	0.07	<0.01^{ef}	0.96	0.02^{f}	0.20	38.50^{bc}	68.01^{bd}	53.92^{cd}
	0.07	0.23	0.17	0.18	3.08	0.15	0.06	1.17	0.06	0.24	0.12	0.02^{f}	1.37	0.06	0.54			
Ethyl acetate	0.02^{ef}	0.11	<0.01^{ef}	0.65	0.62	0.82	0.51	0.16	0.09	0.70	0.29	0.03^{f}	0.14	0.11	0.16	28.17^{bc}	71.28^{bd}	51.93^{cd}
	0.04^{f}	0.22	0.01^{f}	0.92	0.72	1.65	0.94	0.28	0.17	0.92	0.53	0.06^{f}	0.37	0.19	0.34			
Ethanol	0.14	0.37	0.07^{f}	4.64	18.16	5.60	9.52	3.09	1.72	4.41	16.67	1.21	12.01	4.03	1.67	32.33^{c}	40.07^{d}	52.52^{cd}
	0.40	1.17	0.31	12.92	24.60	11.46	14.13	5.27	3.64	11.43	25.26	2.55	18.52	6.43	4.94			
Ethyl propanoate	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	0.01^{ef}	0.02^{f}	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	0.01^{ef}	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	38.96^{c}	48.81	69.72^{c}
	<0.01^{ef}	<0.01^{ef}	<0.01^{ef}	0.02^{f}	0.03^{f}	0.01^{ef}	<0.01^{ef}	0.01^{ef}	<0.01^{ef}	0.03^{f}	0.01^{ef}	<0.01^{ef}	0.01^{ef}	<0.01^{ef}	0.01^{ef}			
Hexanal	0.70	4.33	2.74	1.26	2.23	0.92	0.43	2.26	0.60	0.45	0.62	0.51	0.79	0.80	1.53	23.04^{c}	25.83^{d}	53.85^{cd}
	1.35	7.47	4.04	2.36	3.42	1.60	1.01	4.13	1.28	0.80	1.05	1.54	1.03	1.14	3.29			
3-Methyl-1-butanol	0.01^{f}	0.02^{f}	0.04	0.20	2.56	0.14	0.12	0.13	0.05	0.12	0.56	0.02^{f}	0.21	0.05	0.38	25.95^{bc}	64.65^{bd}	41.51^{cd}
	0.02^{f}	0.05	0.07	0.40	2.84	0.37	0.24	0.22	0.12	0.26	0.76	0.04	0.37	0.06	0.83			
(*E*)-2-Hexenal	9.02	11.01	0.84	6.48	2.20	5.21	3.72	3.32	3.05	1.90	1.42	9.38	2.09	22.73	18.16	19.55^{bc}	23.07^{bd}	46.91^{cd}
	16.98	16.83	1.53	9.34	4.65	7.71	6.01	4.81	4.74	2.82	2.57	15.93	3.31	43.32	23.85			
(*Z*)-3-Hexenyl acetate	<0.01^{ef}	0.23	1.56	0.63	0.09	0.20	2.59	1.16	2.78	1.15	0.17	0.62	0.20	1.78	0.09	19.18^{bc}	30.57^{bd}	62.04^{cd}
	0.01^{f}	0.39	2.70	0.77	1.08	3.07	4.56	1.80	5.19	1.55	0.27	0.90	0.29	3.03	0.21			
(*E*)-2-Heptenal	0.05	0.21	0.04^{f}	0.07	0.27	0.01^{ef}	0.02^{f}	0.16	0.02^{f}	0.02^{f}	0.02^{f}	0.01^{ef}	0.07	0.03^{f}	0.13	24.89^{b}	63.16^{bd}	36.16^{d}
	0.10	0.40	0.20	0.17	0.73	0.07	0.26	0.48	0.14	0.05	0.07	0.05	0.53	0.17	0.34			
6-Methyl-5-hepten-2-one	0.01^{ef}	0.28	0.16	0.02^{f}	0.24	0.01^{ef}	<0.01^{ef}	0.24	0.02^{f}	0.02^{f}	0.09	<0.01^{ef}	0.26	0.01^{ef}	0.03^{f}	43.20^{b}	65.10^{bd}	61.64^{d}
	0.04^{f}	0.39	0.36	0.04^{f}	0.78	0.05^{f}	0.03^{f}	0.50	0.08	0.07	0.54	0.03^{f}	0.79	0.06	0.16			
1-Hexanol	0.14	0.27	1.33	0.61	1.65	1.72	1.10	0.68	0.36	1.01	0.21	0.42	1.84	0.80	1.03	13.26^{bc}	27.71^{bd}	59.96^{cd}
	0.30	0.89	2.72	0.82	2.01	2.46	1.54	0.69	0.53	1.24	0.32	0.94	4.15	1.54	1.21			
Nonanal	0.59	0.76	0.48	0.15	5.29	0.12	0.03^{f}	2.83	0.26	0.11	0.36	0.07^{f}	0.48	0.03^{f}	0.46	46.05	42.51	53.70
	1.54	4.80	1.75	1.53	8.65	1.17	0.94	5.41	0.83	1.57	0.94	0.35	1.36	0.58	2.52			
1-Octen-3-ol	0.01^{f}	0.03	0.02^{f}	0.01^{f}	0.06	0.01^{f}	<0.01^{ef}	0.03	<0.01^{ef}	<0.01^{ef}	0.02^{f}	<0.01^{ef}	0.02^{f}	<0.01^{ef}	0.02^{f}	31.48^{c}	38.87^{d}	64.07^{cd}
	0.01^{f}	0.05	0.03	0.02^{f}	0.18	0.01^{f}	0.01^{f}	0.05	0.01^{f}	0.01^{f}	0.03	<0.01^{ef}	0.04	0.01^{f}	0.07			
(*E*,*E*)-2,4-Hexadienal	0.06	0.05	0.03^{f}	0.02^{f}	0.01^{f}	0.03^{f}	0.03^{f}	0.02^{f}	0.06	0.12	0.01^{f}	0.14	0.04	0.27	0.08	63.46^{c}	69.01^{d}	105.47^{cd}
	0.58	0.62	0.14	0.31	0.53	0.51	0.25	0.20	0.83	0.46	0.06	1.16	0.12	1.20	1.03			
Acetic acid	0.19	1.20	0.30	2.46	3.94	9.63	0.79	0.89	0.37	3.99	0.62	0.27	0.38	0.42	0.26	17.47^{c}	22.81^{d}	71.83^{cd}
	0.45	3.67	0.62	6.52	8.95	25.06	1.98	2.12	0.62	12.75	1.68	0.58	0.84	0.75	0.72			
Propanoic acid	0.39	1.80	0.37	0.46	0.05	0.04	<0.01^{ef}	0.22	<0.01^{ef}	0.22	<0.01^{ef}	0.01^{f}	0.03	0.01^{f}	0.12	26.69^{b}	51.03^{bd}	25.19^{d}
	0.70	2.93	0.82	0.92	0.17	0.11	0.03	0.44	0.07	0.44	0.07	0.06	0.15	0.10	0.33			

Table cont.

(*E*)-2-Decenal	0.25[f]	0.02[ef]	0.04[ef]	0.08[ef]	0.49[f]	0.10[ef]	0.03[ef]	0.30[f]	0.04[ef]	0.08[ef]	0.04[ef]	0.03[ef]	0.01[ef]	0.09[ef]	0.20	36.65[c]	54.33	61.52[c]
	0.98	1.09	0.28[f]	0.14[ef]	3.57	0.13[ef]	0.06[ef]	2.26	0.06[ef]	0.25[f]	0.09[ef]	0.03[ef]	2.14	0.09[ef]	1.18			
Pentanoic acid	0.85	0.22	0.02[ef]	0.08[f]	0.05[f]	0.03[ef]	0.01[ef]	0.02[ef]	<0.01[ef]	0.11[f]	0.01[ef]	<0.01[ef]	<0.01[ef]	0.01[ef]	0.01[ef]	27.11[b]	57.61[bd]	25.51[d]
	2.08	0.48	0.18	0.22	0.13[f]	0.09[f]	0.01[ef]	0.09[f]	0.04[ef]	0.18	0.02[ef]	0.05[f]	0.06f	0.02[ef]	0.04[ef]			

[a] Relative Standard Deviation (%) calculated as mean of RSD% for each compound among the involved labs by removing outliers.
[b] RSD% values obtained for QM1 and QM2 showed significant differences ($p<0.05$).
[c] RSD% values obtained for QM1 and QM3 showed significant differences ($p<0.05$).
[d] RSD% values obtained for QM2 and QM3 showed significant differences ($p<0.05$).
[e] Concentration is below the LOD (Table 7).
[f] Concentration is below the LOQ (Table 8).

Table 6. Mean and standard deviation values of recovery (R_{ap}) calculated from the results of the labs involved using the three types of quantification methods (QMs). The recovery values are shown when the entire concentration range and low concentration range were applied in the calibration curves.

Volatile compounds	**Whole concentration range (0.05-10.00/25.00 mg/kg)[a]**			**Low concentration range (0.05-2.5 mg/kg)**		
	QM1	**QM2**	**QM3**	**QM1**	**QM2**	**QM3**
Octane	92±21	90±42	135±123	93±28	68±38[c]	117±82
Ethyl acetate	99±22	94±46	118±79	74±10[c]	54±31[c]	94±28
Ethanol	104±67	131±80	138±104	71±39	71±45	108±85
Ethyl propanoate	101±18	96±44	128±87	86±12	64±37[c]	103±39
Hexanal	106±11	150±67	266±221	119±42	114±53	188±142
3-Methyl-1-butanol	100±9	93±35	139±106	94±13	68±39	108±33
E-2-Hexenal	88±9	118±37	224±152	144±63	129±55	223±167
(*Z*)-3-Hexenyl acetate	88±5	121±54	248±180	159±82	139±60	267±227
(*E*)-2-Heptenal	102±25	92±21	157±96	152±56	92±23	180±139
6-Methyl-5-hepten-2-one	105±28	94±21	163±97	154±59	93±22	181±131
1-Hexanol	100±7	140±69	269±206	143±58	135±69	238±202
Nonanal	82±16	107±26	224±140	155±74	136±54	247±195
1-Octen-3-ol	86±8	121±53	252±175	166±80	147±63	283±246
(*E,E*)-2,4-Hexadienal	95±13	89±25	147±102	148±54	90±22	180±146
Acetic acid	84±26	105±11	208±146	125±72	115±72	157±104
Propanoic acid	94±25	88±37	119±44	111±26	76±36[c]	115±26
(*E*)-2-Decenal	72±21	71±32	109±29[b]	158±34[c]	110±53[c]	160±126
Pentanoic acid	99±16	92±22	184±172	184±87	114±47	223±250

[a] The highest concentration depended on the compound (see Table 2).
[b] Recovery values found for QM1 and QM3 showed significant differences ($p<0.05$). Non-significant differences were found between the recovery values of QM1 and QM2, and between QM2 and QM3 for all the compounds.
[c] Recovery values found for low concentration range and the whole concentration range showed significant differences ($p<0.05$).

Table 7. Mean values of LOD (mg/kg) for each VOC by applying four calculation methods; the ranges are also shown in parentheses for the first three methods.

Volatile Compounds	Calculation Method 1	Calculation Method 2	Calculation Method 3	Calculation Method 4
Octane	0.64 (0.18-0.89)	0.23 (0.06-0.31)	0.72 (0.06-1.27)	0.01
Ethyl acetate	0.44 (0.42-0.48)	0.19 (0.17-0.24)	0.43 (0.17-0.68)	0.03
Ethanol	1.29 (1.07-1.56)	0.45 (0.38-0.55)	0.54 (0.51-0.58)	0.03
Ethyl propanoate	0.25 (0.17-0.30)	0.10 (0.07-0.12)	0.22 (0.07-0.49)	0.02
Hexanal	1.69 (1.42-2.13)	0.53 (0.45-0.67)	1.43 (0.22-2.50)	0.02
3-Methyl-1-butanol	0.62 (0.38-0.84)	0.22 (0.13-0.29)	0.62 (0.29-0.90)	0.01
(*E*)-2-Hexenal	0.96 (0.07-1.64)	0.38 (0.03-0.64)	0.13 (0.05-0.27)	0.01
(*Z*)-3-Hexenyl acetate	1.00 (0.17-1.73)	0.39 (0.06-0.68)	0.15 (0.07-0.30)	0.01
(*E*)-2-Heptenal	0.92 (0.34-1.48)	0.32 (0.12-0.52)	0.16 (0.16-0.16)	0.02
6-Methyl-5-hepten-2-one	1.12 (0.72-1.55)	0.39 (0.25-0.54)	0.18 (0.10-0.24)	0.02
1-Hexanol	1.69 (0.73-2.22)	0.53 (0.23-0.70)	0.93 (0.47-1.18)	0.01
Nonanal	1.33 (0.21-2.09)	0.52 (0.08-0.83)	0.50 (0.10-0.76)	0.03
1-Octen-3-ol	1.58 (0.57-2.47)	0.53 (0.19-0.83)	0.52 (0.25-0.69)	0.01
(*E,E*)-2,4-Hexadienal	0.87 (0.34-1.73)	0.31 (0.12-0.61)	0.12 (0.08-0.17)	0.01
Acetic acid	1.83 (0.85-2.63)	0.59 (0.28-0.85)	0.92 (0.59-1.18)	0.01
Propanoic acid	0.58 (0.27-1.18)	0.20 (0.10-0.41)	0.36 (0.11-0.51)	0.01
(*E*)-2-Decenal	1.60 (1.19-2.40)	0.56 (0.42-0.84)	0.57 (0.41-0.68)	0.18
Pentanoic acid	0.98 (0.31-1.42)	0.34 (0.11-0.50)	0.19 (0.14-0.25)	0.05

Table 8. Mean values of the LOQ (mg/kg) for each volatile compound by applying four calculation methods; the ranges are shown in parentheses for the first three methods.

Volatile Compounds	Calculation Method 1	Calculation Method 2	Calculation Method 3	Calculation Method 4
Octane	1.95 (0.56-2.69)	0.68 (0.20-0.95)	2.18 (0.19-3.85)	0.03
Ethyl acetate	1.35 (1.26-1.45)	0.58 (0.50-0.73)	1.31 (0.52-2.07)	0.08
Ethanol	3.91 (3.24-4.72)	1.38 (1.14-1.65)	1.64 (1.54-1.74)	0.09
Ethyl propanoate	0.74 (0.52-0.92)	0.30 (0.21-0.37)	0.67 (0.20-1.47)	0.07
Hexanal	5.11 (4.30-6.46)	1.62 (1.37-2.04)	4.34 (0.68-7.58)	0.07
3-Methyl-1-butanol	1.89 (1.14-2.55)	0.66 (0.40-0.89)	1.89 (0.87-2.72)	0.04
(*E*)-2-Hexenal	2.90 (0.22-4.97)	1.14 (0.09-1.95)	0.38 (0.15-0.82)	0.03
(*Z*)-3-Hexenyl acetate	3.03 (0.50-5.24)	1.20 (0.19-2.06)	0.46 (0.21-0.91)	0.02
(*E*)-2-Heptenal	2.79 (1.04-4.48)	0.97 (0.36-1.57)	0.48 (0.47-0.49)	0.05
6-Methyl-5-hepten-2-one	3.41 (2.19-4.70)	1.19 (0.77-1.64)	0.55 (0.30-0.74)	0.06
1-Hexanol	5.11 (2.23-6.73)	1.62 (0.70-2.13)	2.82 (1.42-3.59)	0.01
Nonanal	4.02 (0.65-6.33)	1.58 (0.25-2.50)	1.52 (0.30-2.31)	0.09
1-Octen-3-ol	4.80 (1.73-7.47)	1.61 (0.58-2.52)	1.57 (0.76-2.09)	0.03
(*E,E*)-2,4-Hexadienal	2.65 (1.03-5.25)	0.93 (0.36-1.84)	0.37 (0.25-0.51)	0.04
Acetic acid	5.53 (2.58-7.98)	1.79 (0.84-2.58)	2.79 (1.78-3.57)	0.04
Propanoic acid	1.75 (0.82-3.57)	0.61 (0.29-1.25)	1.11 (0.34-1.54)	0.02
(*E*)-2-Decenal	4.85 (3.62-7.28)	1.69 (1.27-2.54)	1.72 (1.24-2.07)	0.53
Pentanoic acid	2.96 (0.94-4.29)	1.03 (0.33-1.50)	0.59 (0.43-0.76)	0.15

FIGURE 1

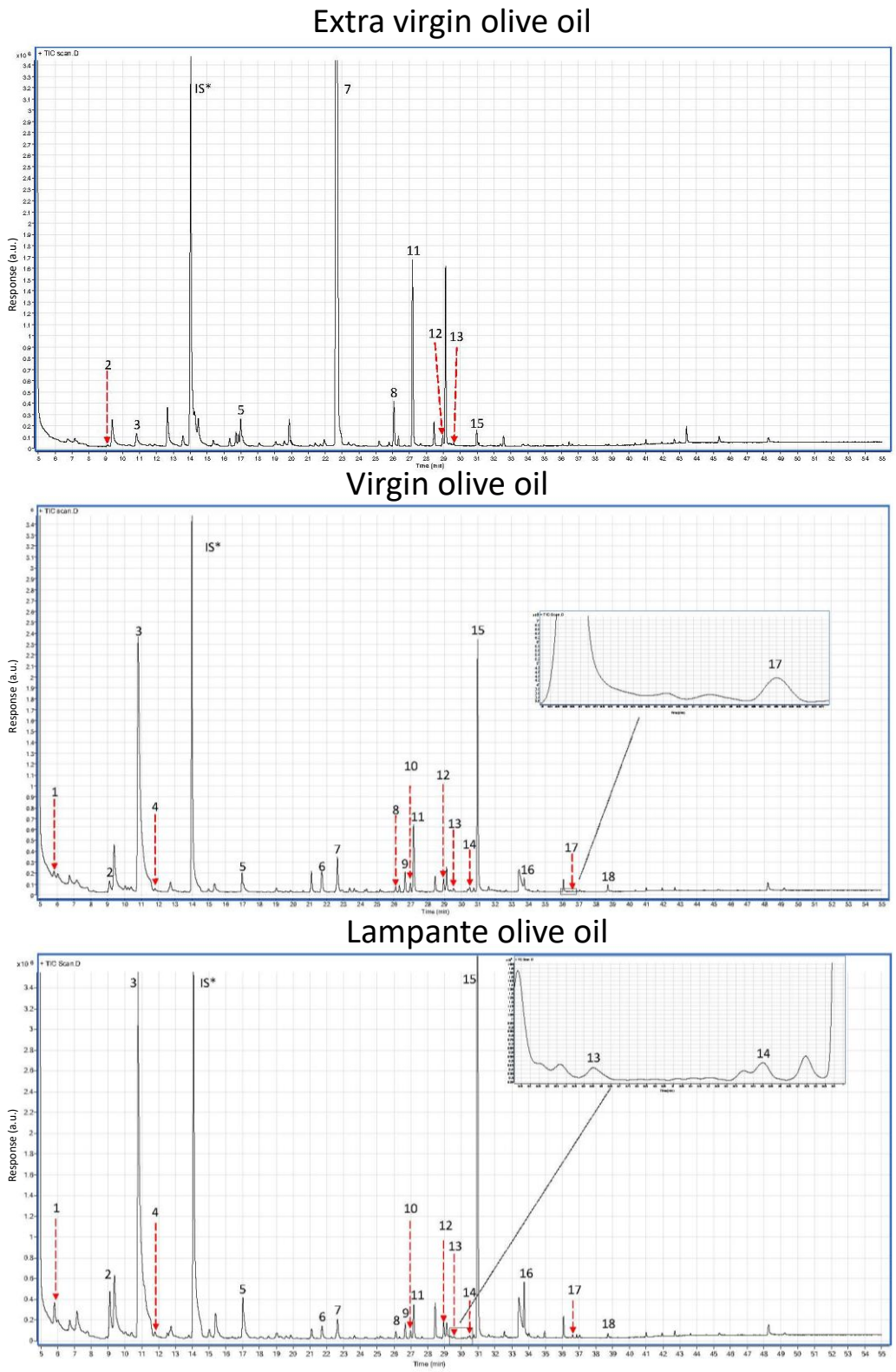

Figure 1. Chromatogram of volatile compounds of an extra virgin olive oil, virgin olive oil, and a lampante olive oil analyzed by SPME-GC-MS. The correspondence of the codes with the volatile compounds is shown in Table 1.

FIGURE 2

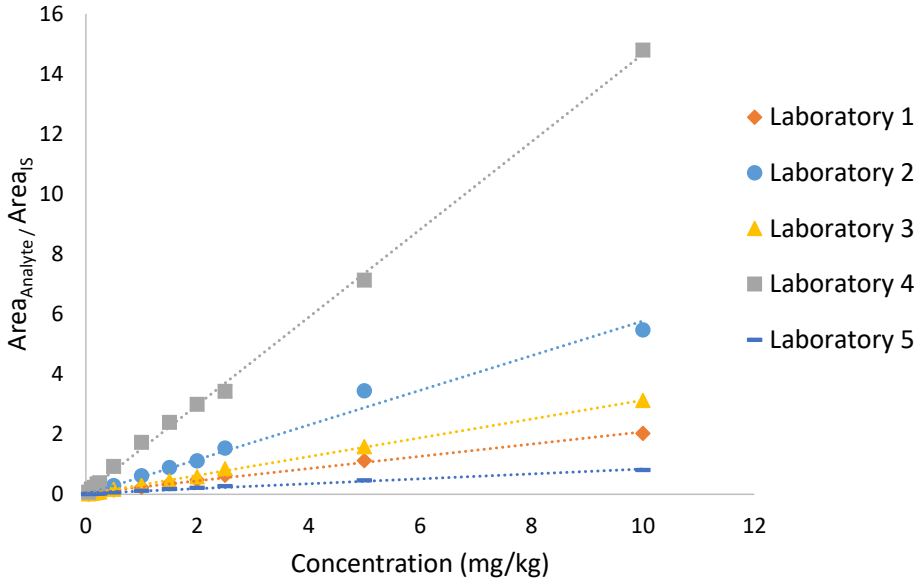

Figure 2. Calibration curves of ethyl propanoate for quantification method 2 (QM2).

SUPPLEMENTARY INFORMATION

Collaborative validation trial of a harmonized SPME-GC-MS method for analysis of selected volatile compounds in virgin olive oils

Ramón Aparicio-Ruiz[1], Clemente Ortiz Romero[1], Diego L. García González[1], Maurizio Servili[2], Roberto Selvaggini[2], Florence Lacoste[3], Julien Escobessa[3], Stefania Vichi[4], Beatriz Quintanilla-Casas[4], Pierre Alain Golay[5], Paolo Lucci[6], Erica Moret[6], Enrico Valli[7]*, Alessandra Bendini[7], Tullia Gallina Toschi[7]

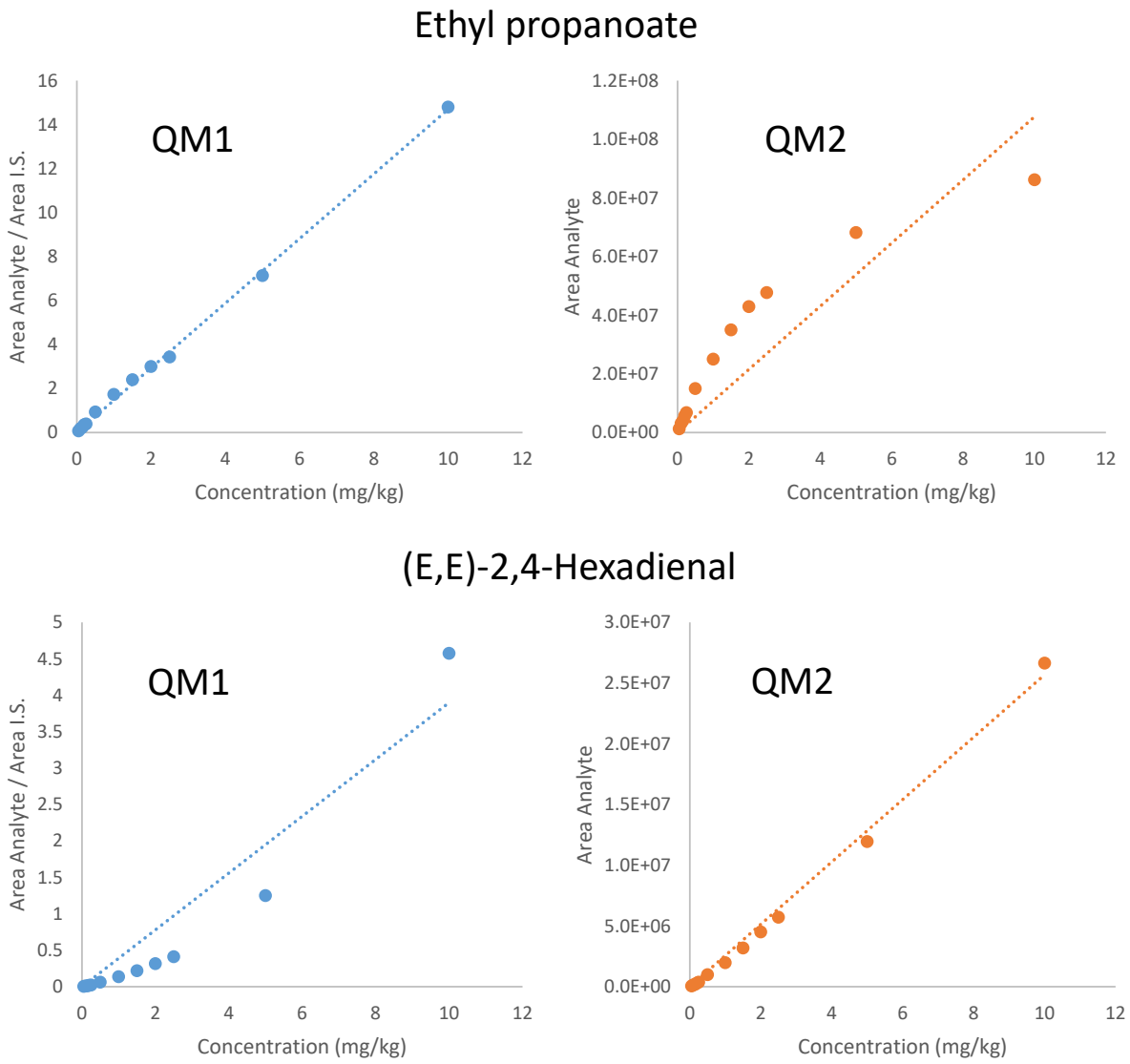

Figure S1. Calibration curves of ethyl propanoate and (*E*,*E*)-2,4,hexadienal built in quantification methods 1 and 2 (QM1, QM2).

Table S1. Slope and typical error (mean±standard deviation) of the regression equation built for the calibration curves (QM1).

Volatile compounds	Slope (Mean±SD)	Typical error (Mean±SD)
Octane	0.201±0.186	0.052±0.066
Ethyl acetate	0.666±0.655	0.231±0.238
Ethanol	0.159±0.136	0.184±0.163
Ethyl propanoate	0.508±0.552	0.067±0.035
Hexanal	0.099±0.089	0.064±0.045
3-Methyl-1-butanol	0.091±0.057	0.022±0.016
(*E*)-2-Hexenal	0.064±0.047	0.053±0.043
(*Z*)-3-Hexenyl acetate	0.072±0.064	0.068±0.070
(*E*)-2-Heptenal	0.037±0.030	0.024±0.030
6-Methyl-5-hepten-2-one	0.036±0.026	0.023±0.024
1-Hexanol	0.092±0.046	0.085±0.069
Nonanal	0.004±0.003	0.006±0.005
1-Octen-3-ol	0.073±0.044	0.108±0.104
(*E*,*E*)-2,4-Hexadienal	0.073±0.061	0.044±0.038
Acetic acid	0.135±0.080	0.112±0.079
Propanoic acid	0.052±0.028	0.022±0.024
(*E*)-2-Decenal	0.002±0.002	0.002±0.003
Pentanoic acid	0.058±0.034	0.041±0.041

Table S2. Linearity in the low concentration range of the calibration curve (0.05-0.25 mg/kg) (QM1).

Volatile compound	R^2	Slope	Intercept
Octane	0.972±0.030	0.152±0.101	0.002±0.003
Ethyl acetate	0.978±0.026	0.959±0.979	0.004±0.007
Ethanol	0.963±0.051	0.246±0.240	-0.003±0.014
Ethyl propanoate	0.975±0.034	0.532±0.543	0.002±0.004
Hexanal	0.964±0.034	0.093±0.075	0.001±0.002
3-Methyl-1-butanol	0.969±0.030	0.112±0.076	-0.001±0.001
(*E*)-2-Hexenal	0.941±0.107	0.044±0.040	-0.001±0.001
(*Z*)-3-Hexenyl acetate	0.987±0.009	0.055±0.063	-0.001±0.001
(*E*)-2-Heptenal	0.984±0.021	0.017±0.009	0.000±0.000
6-Methyl-5-hepten-2-one	0.980±0.018	0.019±0.011	0.000±0.000
1-Hexanol	0.979±0.028	0.065±0.042	0.000±0.000
Nonanal	0.613±0.436	0.001±0.000	0.001±0.001
1-Octen-3-ol	0.976±0.020	0.039±0.028	-0.033±0.058
(*E*,*E*)-2,4-Hexadienal	0.986±0.019	0.051±0.034	-0.001±0.001
Acetic acid	0.977±0.019	0.132±0.089	0.014±0.018
Propanoic acid	0.975±0.021	0.044±0.031	0.000±0.001
(*E*)-2-Decenal	0.629±0.431	0.000±0.000	0.000±0.000
Pentanoic acid	0.908±0.109	0.020±0.014	0.001±0.001

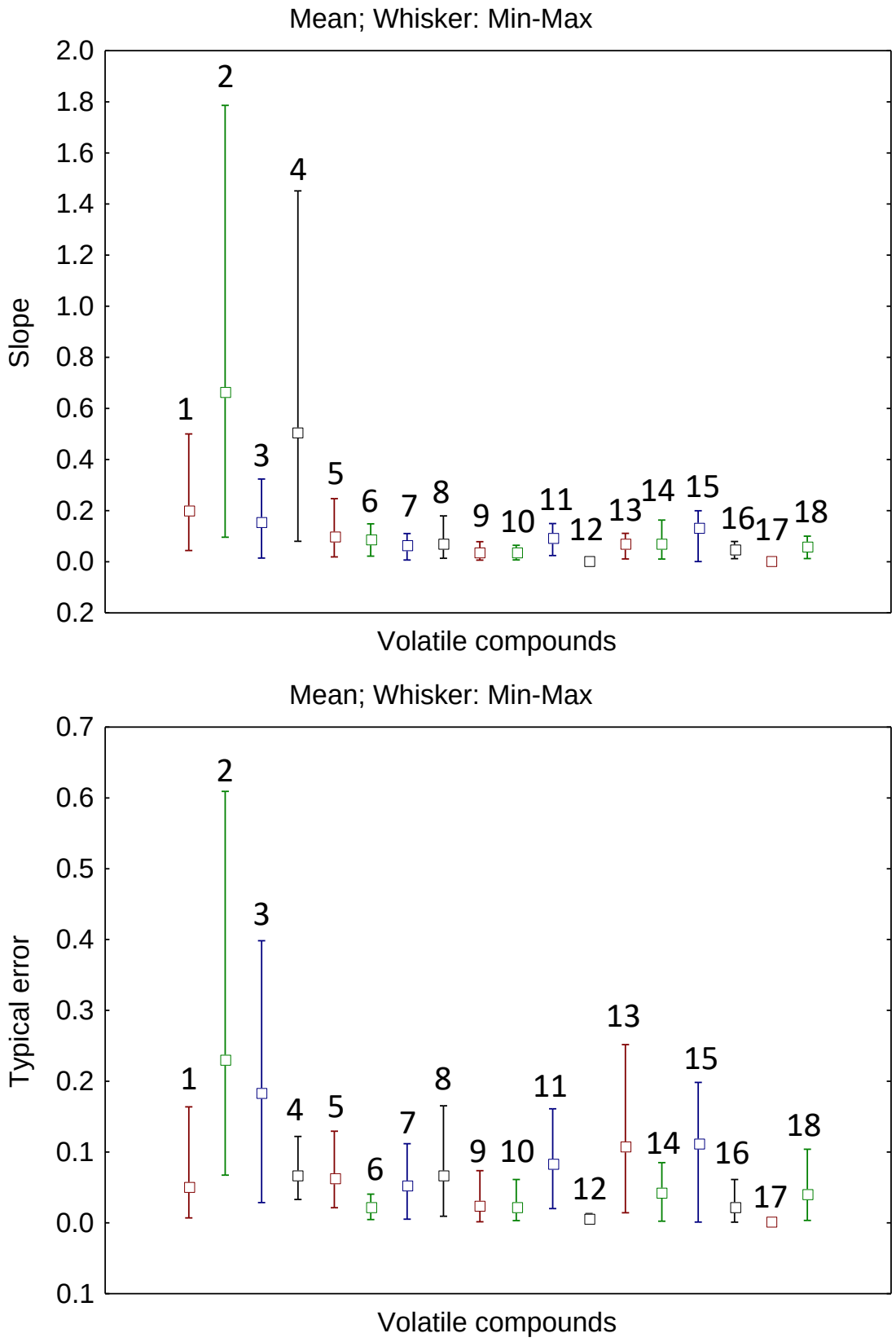

Figure S2. Box and whisker plots of the slope and typical error (mean ± standard deviation) of the regression equation built for the calibration curves (QM1). The volatile compound codes correspond to Table 1.

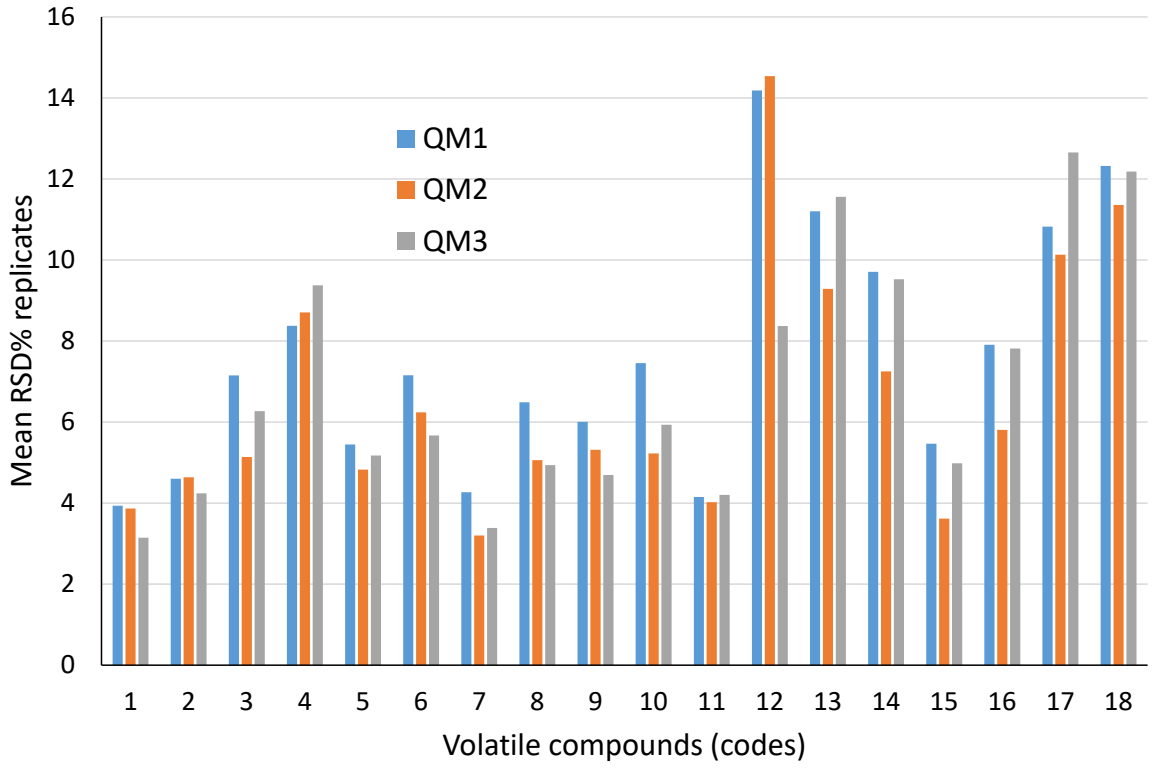

Figure S3. Mean relative standard deviation (RSD%) computed from the duplicates of the 15 samples analyzed by the 5 laboratories. The volatile compound codes correspond to Table 1.

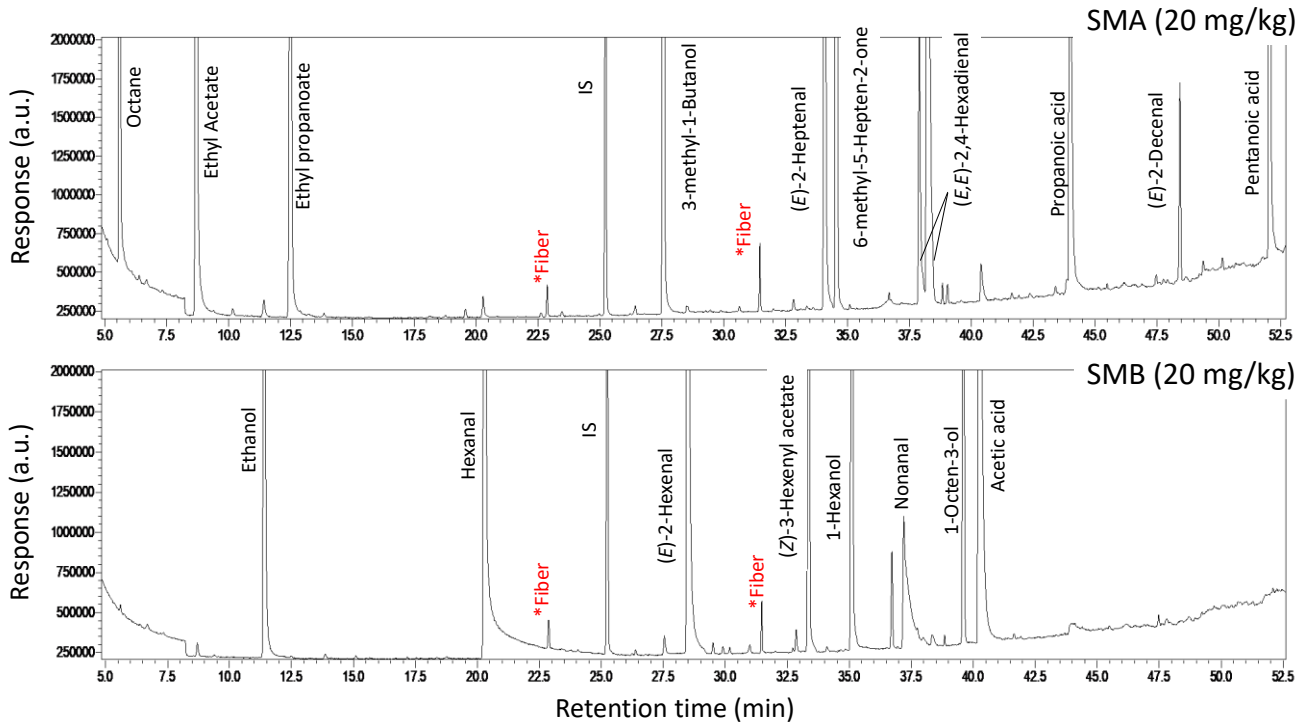

Figure S4. Chromatograms of the standard mixtures SMA and SMB built for calibration (calibration point 20 mg/kg). Note: *Compounds deriving from the SPME divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fiber.