

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Profiling food volatiles by Comprehensive Two-dimensional Gaschromatography coupled with Mass Spectrometry: advanced fingerprinting approaches for comparative analysis of the volatile fraction of roasted hazelnuts (*Corylus avellana* L.) from different origin

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/78813> since 2016-12-01T14:23:10Z

Published version:

DOI:10.1016/j.chroma.2010.07.006

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This Accepted Author Manuscript (AAM) is copyrighted and published by Elsevier. It is posted here by agreement between Elsevier and the University of Turin. Changes resulting from the publishing process - such as editing, corrections, structural formatting, and other quality control mechanisms - may not be reflected in this version of the text. The definitive version of the text was subsequently published in [*JOURNAL OF CHROMATOGRAPHY A* Volume: 1217 Issue: 37 Pages: 5848-5858 Published: SEP 10 2010, <http://dx.doi.org/10.1016/j.chroma.2010.07.006>].

You may download, copy and otherwise use the AAM for non-commercial purposes provided that your license is limited by the following restrictions:

- (1) You may use this AAM for non-commercial purposes only under the terms of the CC-BY-NC-ND license.
- (2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.
- (3) You must attribute this AAM in the following format: Creative Commons BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>), <http://dx.doi.org/10.1016/j.chroma.2010.07.006>

1 **Profiling food volatiles by Comprehensive Two-dimensional**
2 **Gaschromatography coupled with Mass Spectrometry: advanced fingerprinting**
3 **approaches for comparative analysis of the volatile fraction of roasted hazelnuts**
4 **(*Corylus avellana* L.) from different origins.**

5
6 Chiara Cordero^{*a}, Erica Liberto^a, Carlo Bicchi^a, Patrizia Rubiolo^a, Peter Schieberle^b, Stephen E.
7 Reichenbach^c, Qingping Tao^d.

8
9 ^a Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via P. Giuria
10 9, I-10125 Torino, Italy

11 ^b Deutsche Forschungsanstalt für Lebensmittelchemie, Lise-Meitner-Straße 34, 85354 Freising,
12 Germany

13 ^c Computer Science and Engineering Department, University of Nebraska , 1400 R Street, Lincoln,
14 NE 68588-0115, USA

15 ^d GC Image, LLC, PO Box 57403, Lincoln NE 68505-7403, USA

16
17
18 *** Address for correspondence:**

19 Dr. Chiara Cordero - Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di
20 Torino, Via P. Giuria 9, I-10125 Torino, Italy – e-mail: chiara.cordero@unito.it ; phone: +39 011
21 670 7662; fax: +39 011 2367662

22

23 **Abstract**

24 This study examined how advanced fingerprinting methods (i.e., non-targeted methods) provide
25 reliable and specific information about groups of samples based on their component distribution on
26 the GCxGC chromatographic plane. The volatile fractions of roasted hazelnuts (*Corylus avellana*
27 L.) from nine different geographical origins, comparably roasted for desirable flavor and texture,
28 were sampled by Headspace-Solid Phase Micro Extraction (HS-SPME) and then analyzed by
29 GCxGC-qMS. The resulting patterns were processed by: (a) “chromatographic fingerprinting”, i.e.,
30 a pattern recognition procedure based on retention-time criteria, where peaks correspondences were
31 established through a comprehensive peak pattern covering the chromatographic plane; and (b)
32 “comprehensive template matching” with reliable peak matching, where peak correspondences
33 were constrained by retention time and MS fragmentation pattern similarity criteria. Fingerprinting
34 results showed how the discrimination potential of GCxGC can be increased by including in sample
35 comparisons and correlations all the detected components and, in addition, provide reliable results
36 in a comparative analysis by locating compounds with a significant role. Results were completed by
37 a chemical speciation of volatiles and sample profiling was extended to known markers whose
38 distribution can be correlated to sensory properties, geographical origin, or the effect of thermal
39 treatment on different classes of compounds. The comprehensive approach for data interpretation
40 here proposed may be useful to assess product specificity and quality, through measurable
41 parameters strictly and consistently correlated to sensory properties and origin.

42
43 **Key-words:** GCxGC, fingerprint analysis, comprehensive template matching fingerprinting,
44 roasted hazelnut, *Corylus avellana* L. volatile fraction, key-aroma markers.

45

46

47 **1. Introduction**

48 The term *fingerprint*, in its general meaning, refers to the “*impression of a fingertip on any*
49 *surface...an ink impression of the lines upon the fingertip taken for the purpose of identification*”
50 and/or “*something that identifies: as (a) a trait, trace, or characteristic revealing origin or*
51 *responsibility; (b) analytical evidence (as a spectrogram) that characterizes an object or substance;*
52 *in particular the chromatogram or electrophoretogram obtained by cleaving a protein by enzymatic*
53 *action and subjecting the resulting collection of peptides to two-dimensional chromatography or*
54 *electrophoresis.*” [1]. For chromatographers, this definition evokes the intrinsic potential of the bi-
55 dimensional separation patterns, obtained by comprehensive methods, for sample characterization,
56 differentiation, discrimination and, as a consequence, classification on the basis of the peculiar
57 component distribution over the 2D plane. In particular, comprehensive two-dimensional gas
58 chromatography (GCxGC) has proven to be a powerful tool for sample profiling, i.e., the
59 exhaustive analysis of a complex mixture to characterize its chemical composition. GCxGC yields
60 highly informative separation patterns because of its great practical peak capacity, sensitivity, and
61 structure-retention patterns for chemically related groups of substances, produced by applying two
62 different separation principles one for each chromatographic dimension. However, the improvement
63 in information causes a large and complex dataset for each sample, consisting of bi-dimensional
64 retention data, detector responses and MS spectra requiring suitable data mining (*a*) to interpret the
65 higher level of information and (*b*) to extract useful and consistent data on sample compositional
66 characteristics.

67 Different approaches have been investigated to link raw data (i.e. separation data) with the
68 chemical composition of samples, and their effectiveness has been demonstrated for different fields
69 of application [2-6]. GCxGC approaches are commonly classified into two main groups: targeted
70 and non-targeted methods [2]. Non-targeted methods often are based on chemometric techniques or
71 on image processing procedures [2-6], but the multi-dimensionality of the GCxGC separation may
72 only partially be exploited. The MS fragmentation pattern is a critical point for several approaches
73 because it includes a number of variables (i.e. m/z fragments and intensities) whose control is
74 difficult. On the other hand, interpretation of fragmentation patterns may be crucial for analyte
75 identification and quantification. This is an area of active research. Cross-sample analyses with
76 GCxGC include oil spill identification [7], metabolomic analysis of mouse tissue [8], chemical
77 profiles of illicit drug samples [9], investigation of changes in cocoa bean volatiles caused by
78 moisture damage [6], and profiles of impurities in a chemical weapon precursor [4]. Extracted
79 features have been compared and analyzed using methods such as Fisher Ratio, PCA, and machine
80 learning algorithms. An important problem in cross-sample analysis is feature matching, i.e.,

81 matching the same features across samples. For example, datapoint-to-datapoint analyses have been
82 reported but that approach is subject to problems related to retention-time variability.
83 Comprehensive matching of all peaks across complex chromatograms can account for retention-
84 time variability but is intractable, even with mass spectrometry, so peaks are sometimes matched
85 selectively rather than comprehensively. The challenge of automated comprehensive comparisons is
86 addressed in this paper.

87 This study investigated a) how advanced fingerprinting approaches can fully exploit the
88 informative content of GCxGC-qMS patterns (¹D and ²D retention times, detector responses, and
89 MS spectra) and can profitably be applied to complex food samples investigations, and b) which
90 advantages they provide, by including in the discrimination process all the separation dimensions
91 and maintaining intact the informative content. The food matrix here investigated is hazelnuts
92 (*Corylus avellana* L.), which, besides their economic value [11] and potential health benefits
93 [12,13], have a unique and distinctive flavor [14-20] and a crispy and crunchy texture [21] induced
94 by a technological thermal treatment. Roasting is the key step in industrial hazelnut processing,
95 inducing several chemical reactions on specific precursors, present at different concentrations in the
96 raw material. It produces a mixture consisting of several groups of compounds (i.e., furans,
97 pyrazines, ketones, alcohols, aldehydes, esters, pyrroles, thiophenes, sulfur compounds, aromatic
98 compounds, phenols, pyridines, thiazoles, oxazoles, lactones, alkanes, alkenes, and acids among the
99 others) whose complexity is challenging to explore, even with GCxGC-qMS. Roasting has to be
100 monitored because sensory properties are influenced, on one hand, by the quali-quantitative
101 distribution of aroma markers resulting from the thermal treatment due to lipid-oxidation, Maillard
102 reactions, and Strecker degradation, and, on the other hand, by the geographical origin through
103 primary and secondary metabolites, in particular terpenoids.

104 The number of volatiles effectively contributing to the aroma of a food is rather limited
105 and complex analytical procedures are required to detect, identify, and possibly quantify odour
106 active components occurring at trace level, sometimes below ppts (ng/Kg), for a reliable
107 characterization of the overall aroma. This is particularly true for analytes with very low odour-
108 thresholds, called “key-aroma” markers, whose concentration-in-the-food-matrix/odor threshold
109 ratio (also defined as Odor Activity Value, OAV) is ≥ 1 [22]. GCxGC sensitivity was demonstrated
110 to be crucial in characterizing the aroma profile of Arabica coffee samples, enabling study of the
111 quali-quantitative distribution of key-aroma markers [23].

112 The potential of novel advanced fingerprinting methods are shown here to: (a) reveal
113 samples compositional peculiarities, (b) delineate fingerprints with different discrimination
114 potential, and (c) locate compounds (known and unknown) comparatively important for

115 geographical origin and characteristics of technological treatment assessments. Fingerprinting
116 results are additionally validated and confirmed through known markers, in particular aroma
117 compounds, identified by GC-O and Aroma Extract Dilution Analysis (AEDA) [24], and other
118 markers whose distribution greatly influence sample sensory properties or indicate the extent of
119 thermal treatments, storage time, and conditions.

120

121 **2. Materials and Methods**

122 **2.1. Reference Compounds and Solvents**

123 Standard samples of *n*-alkanes (from *n*-C9 to *n*-C25) and pure reference compounds were
124 supplied by Sigma-Aldrich (Milan, Italy). Standard stock solution of *n*-dodecane, the internal
125 standard (ISTD) was prepared in acetone at 1000 µg/mL, stored at -18°C, and used to prepare
126 standard working solutions in concentrations ranging from 70 to 7 µg/mL, likewise stored at -18°C.
127 Solvents (acetone, cyclohexane, *n*-hexane, dichloromethane) were all HPLC-grade from Riedel-de
128 Haen (Seelze, Germany).

129

130 **2.2 Hazelnut samples**

131 Commercially representative samples of *Corylus avellana* L. (harvest years 2007 and 2008)
132 from different cultivars/varieties and geographical origins were analyzed. Monovarieties from Italy
133 were “Tonda Gentile Romana” (named *Romana*), “Nocciola di Giffoni” (*Giffoni*), “Nocciola del
134 Piemonte” (*Piemonte*) and “Mortarella”, while Turkish hazelnuts from “Akçakoca”, “Giresun”,
135 “Ordu”, and “Trabzon” regions were blends of different cultivars. Akçakoca hazelnuts are
136 composed mainly by *Tombul*, *Mincane*, *Foşa* and *Cakildak* cultivars; Giresun by *Tombul* and
137 *Kalinkara*; Ordu by *Tombul*, *Palaz* and *Kalinkara*; and Trabzon by *Mincane*, *Tombul*, and *Foşa*.
138 The “Cile” sample is representative of an experimental plantation of Mediterranean varieties of
139 *Corylus avellana* L in Cile. Raw hazelnuts were selected on the basis of their dimensions (caliber
140 within 12-13 cm) and submitted to roasting in an industrial plant at different time/temperature ratios
141 consistent with their desirable final sensory characteristics. Roasted samples were then hermetically
142 sealed under vacuum in non-permeable polypropylene/aluminum/polyethylene packages and stored
143 at -20°C until their chemical analysis. Hazelnuts were supplied by Nocciole Marchisio Cortemilia
144 (CN), Italy.

145

146 **2.3 Isolation of the volatiles by Solvent Assisted Flavor Evaporation – SAFE extraction**

147 Roasted hazelnuts (100 g) were frozen in liquid nitrogen and then grinded by a commercial
148 blender (Moulinette, Quelle, Nürnberg, Germany). The hazelnut powder (50 g) was extracted for 3

149 h at 40°C with diethyl ether (600 mL) under constant stirring, dried over anhydrous sodium sulfate,
150 and concentrated to 200 mL using a Vigreux column (50 cm x 1 cm internal diameter). The
151 concentrate then was submitted to Solvent Assisted Flavor Evaporation (SAFE) [25-27] to remove
152 the nonvolatile fraction, the resulting distillate was reduced to 200 µL by means of a Vigreux
153 column, and the odor-active compounds were evaluated by Aroma Extract Dilution Analysis,
154 AEDA [28].

155

156 **2.4 GC-O/FID and Aroma Extract Dilution Analysis (AEDA)**

157 GC analyses were performed on a Trace GC-Ultra gas chromatograph (Thermo Fischer
158 Instruments, Mainz, Germany) with a SE-54 (5% phenyl - 95% polydimethylsiloxane), and a FFAP
159 (100% polyethylene glycol) column both 30 m x 0.32 mm ID, 0.25 µm df (J&W Scientific, Folsom,
160 CA (USA)). Samples were introduced by cold on-column injection at 40°C. After 2 min, the
161 temperature of the oven was raised at 6°C/min to 240°C and held for 5 minutes. Analyses were
162 performed at constant pressure (90 KPa) with helium as carrier gas. The linear retention indices
163 (I_s^T) were calculated using *n*-alkanes as reference.

164 The Flavor Dilution (FD) factors [25] of the odorants were determined by AEDA. An
165 aliquot of each distillate (0.5 µL of 200 µL) was submitted to GC analysis on the FFAP column, the
166 effluent was split to both the FID and the sniffing port (1:1 by vol.), and the odor-active regions and
167 the odor qualities were assigned by three assessors (GC-O). The extract was stepwise diluted with
168 diethyl ether (1:1 by vol) and aliquots of the diluted solutions (0.5 µL) were again evaluated by
169 three assessors.

170

171 **2.5. Headspace Solid Phase Microextraction (HS-SPME) devices and sampling conditions**

172 The SPME device and fibers were from Supelco (Bellefonte, PA, USA). A
173 Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) df 50/30 µm, 2 cm length
174 fiber was chosen and conditioned before use as recommended by the manufacturer. Roasted
175 hazelnuts (1.0 g) were ground, immediately sealed in a 20 mL vial and equilibrated for 20 min at
176 50°C before sampling. The Internal Standard loading procedure onto the SPME fibre [29,30] was as
177 follows: the SPME device was manually inserted into a 20 mL sealed vial containing 4 mL of ultra-
178 pure water to which 2 µL of *n*-undecane (ISTD) standard working solution at 7.0 µg/mL was added.
179 The fiber was then exposed to the headspace at 50°C for 20 min. After ISTD loading, the fiber was
180 exposed to the matrix headspace at 50°C for another 20 min. The vial was vibrated for 10 s every 5
181 min with an electric engraver (Vibro-Graver V74, Burgess Vibrocrafters Inc, Brayslake, IL) to
182 speed up the analyte equilibration process between headspace and fiber coating. Only that part of

183 the vial in which the solid sample was present was heated, in order to keep the SPME fiber as cold
184 as possible, to improve the vapor phase/fiber coating distribution coefficient. After sampling, the
185 SPME device was immediately introduced into the GC injector for thermal desorption for 10 min at
186 250°C. Each experiment was carried out in triplicate: the resulting Relative Standard Deviation
187 (RSD%) referred to the identified analytes on the normalized 2D-peak volumes was always below
188 15%.

189

190 **2.6. GCxGC-qMS analyses**

191 GCxGC analyses were performed on an Agilent 6890 GC unit coupled with an Agilent 5975 MS
192 detector operating in EI mode at 70 eV (Agilent, Little Falls, DE, USA). The transfer line was set at
193 280°C. A Standard Tune option was used and the scan range was set at m/z 35–250 with the *fast*
194 *scanning* option applied (10000 amu/s) to obtain a number of data points for each chromatographic
195 peak suitable to make its identification and quantitation reliable. The system was provided with a
196 two-stage thermal modulator (KT 2004 loop modulator from Zoex Corporation, Houston, TX,
197 USA) cooled with liquid nitrogen and, with the hot jet pulse time set at 400 ms, a modulation time
198 of 4 s was applied to all experiments. A 1.0 m x 100 μ m ID fused silica capillary loop was used.
199 The column set consisted of a ¹D CW20M column (100% polyethylene glycol) (30 m x 0.25 mm
200 ID, 0.25 μ m df) coupled with a ²D OV1701 column (86% polydimethylsiloxane, 7% phenyl, 7%
201 cyanopropyl) (1 m x 0.1 mm ID, 0.10 μ m df) from MEGA (Legnano (Milan)-Italy).

202 One micro liter of the *n*-alkanes sample solution was automatically injected into the GC
203 instrument with an Agilent ALS 7683B injection system under the following conditions: injector:
204 split/splitless; mode: split; split ratio: 1/100; and injector temperature: 280°C. The HS-SPME
205 sampled analytes were recovered through thermal desorption of the fiber for 10 min directly into
206 the GC injector under the following conditions: injector: split/splitless; mode: split; split ratio: 1/50;
207 injector temperature: 250°C; carrier gas: helium at a constant flow of 1.0 mL/min (initial head
208 pressure 280 KPa); temperature program: from 50°C (1 min) to 260°C (5 min) at 2.5°C/min;
209 modulation period: 4 s.

210 Data were acquired by Agilent MSD ChemStation ver D.02.00.275 (Agilent Technologies,
211 Little Falls, DE, USA) and processed using GC Image GCxGC Software, version 2.0 (GC Image,
212 LLC, Lincoln NE, USA).

213

214 **3. Results and Discussion**

215 This study develops an integrated approach based on advanced fingerprinting methods and
216 extended target analysis to provide information on the quali-quantitative distribution of volatiles in

217 hazelnut samples (*Corylus avellana* L.) of different varieties and geographical origin, submitted to
218 thermal treatment.

219 In the first part, samples were submitted to non-targeted data-processing methods, i.e.,
220 fingerprint analysis, that demonstrated high specificity and sensitivity in revealing compositional
221 differences and similarities between samples by extending the discrimination potential to the entire
222 chromatographic profile [31,32]. In the second part, fingerprinting results were analyzed in depth
223 by identifying analytes and correlating their distribution with sample sensory properties, thermal
224 stress, and geographical origin in view of sample quality assessment.

225

226 **3.1 Hazelnut volatiles advanced fingerprinting**

227 **3.1.1 General concepts**

228 A new, effective, specific, and reliable non-targeted analysis approach for complex samples
229 was adopted [32] for a comparative analysis of two-dimensional chromatographic data. This
230 approach does not rely on sample chemical speciation, but instead relies on the information
231 provided by the GCxGC separation (i.e. analyte relative retention, detector response and MS
232 fragmentation patterns) in toto. This approach, known as “template-based fingerprinting”, is
233 inspired by biometric fingerprinting [32]. Most existing automatic biometric fingerprint verification
234 systems are based on the fact that human fingertips have unique characteristics, e.g., ridge
235 bifurcations and endings that can be localized and extracted from inked impressions or detailed
236 images of the fingertip. These characteristics are called “minutiae features” and are cross-matched
237 with a set of stored templates [33,34].

238 A GCxGC separation pattern is composed of a number of 2D peaks spread over a two-
239 dimensional plane. Each peak reasonably corresponds to a single compound, is potentially
240 informative, and can be treated as a separate *minutiae* for a comparative pattern analysis, as is done
241 for fingertip features. The goal of chromatographic fingerprinting is to catalog features of a
242 chromatogram comprehensively, quantitatively, and in a way comparable across the samples.

243 This task can be performed in two ways: (a) by locating *minutiae features* extracting
244 information from analytes distribution over the GCxGC chromatographic plane (i.e.,
245 *chromatographic fingerprinting*), or (b) by considering each individual 2D peak, together with its
246 time coordinates, detector response, and MS fragmentation, as a potential *fingerprint minutiae* and
247 including it in the sample template that can be used for a direct plot comparison (i.e.,
248 *comprehensive template matching fingerprinting*).

249

250 **3.1.2 Chromatographic Fingerprinting**

251 The first approach aimed to locate and detect fingerprint *minutiae* in each GCxGC pattern of
252 hazelnuts volatiles using features of a cumulative chromatogram from all analyzed samples to
253 compare patterns reliably and/or to reveal differences between samples. First, a cumulative
254 chromatogram was formed by summing all of the chromatograms of the set, with retention times
255 alignment applied only where necessary [32]. Then, 2D chromatogram areas containing *features*
256 were detected and treated as fingerprint *minutiae* to form the so called “consensus template”, i.e.,
257 the collection of *minutiae* from the sample set. **Figure 1** shows the cumulative chromatogram for
258 the nine samples of roasted hazelnuts and the regions of detected peaks used for fingerprinting. In
259 this analysis, the number of chromatographic features was 411.

260 The *features* from the *consensus template* then were copied into each individual 2D
261 chromatogram with the least-squares-optimal retention-times transformation (geometric scaling and
262 translation) determined from peak matching. This elaboration keeps coherent the pattern of the
263 *minutiae* in the retention-times plane and compensate for retention times shifts. The response in
264 each *feature* (i.e. Total Ion Current absolute abundance) was computed by summing the response at
265 all datapoints in it. The result was a fingerprint obtained by grouping all the cumulative *minutiae*
266 that reliably matched across the sample set and a semi-quantitative distribution based on an average
267 percent response corresponding to each *feature* (i.e. the response within the *feature* divided by the
268 response within the entire chromatogram). The fingerprinting results (**Table 1**), obtained by
269 applying the cumulative fingerprint on each sample chromatogram, are useful for a preliminary
270 analysis to focus the attention on those regions of the chromatogram, in which the detector response
271 varied significantly, thereby indicating analytes with a highly informative role in this comparative
272 process.

273 Fingerprint *minutiae* were sifted in various ways to generate tables of potentially significant
274 *features*. In this application the first 20 *minutiae* with the largest average response, i.e., the response
275 within the *feature* divided by the response within the entire chromatogram, were ranked. **Table 1**
276 lists the first 20 *minutiae* corresponding to the regions of the chromatogram with the largest average
277 percent responses, presumably produced by compounds that are the major constituents of the
278 sample. The cumulative results of the chromatographic fingerprinting are summarized at the bottom
279 of the table as number of matched *features* with the *consensus template*, together with the percent
280 of matching.

281 Cumulative results, in particular the percent of matching *features*, can be interpreted as an
282 indication of similarity between samples, since they are obtained by matching the *consensus*
283 *template*, formed by all the fingerprint *minutiae* collected from the cumulative chromatogram (i.e.
284 the cumulative GCxGC plot obtained by summing chromatograms from the nine hazelnut

285 varieties), with each single pattern of the sample set. As a general observation, Piemonte hazelnuts
286 show the lowest matching percentage, 68.4%, with only 281 *features* over 411 corresponding to the
287 template *minutiae*, while Cile (73.7%), Ordu (73.7%), Akçakoca (73.0%) and Romana (71.3%)
288 samples showed similar matching rates.

289 Results based on comprehensive chromatographic *features* have some limits as for example,
290 they may define features incompletely (e.g., placing two important chromatographic peaks in the
291 same fingerprint *feature*) or incorrectly (e.g., splitting a chromatographic peak into two fingerprint
292 *features*) or worse establish inconsistent correspondences between peaks with different identities.
293 On the other hand, this approach diminishes errors for mis-matched features related to unavoidable
294 errors in detecting peaks, unmixing coeluting peaks, and distinguishing coincident peaks with the
295 same retention indices across multiple images. And, the lower specificity of this approach enables
296 an effective and less time consuming classification of samples especially when one has to process
297 unknown patterns and consequently the need is to “scan” comprehensively all the chromatographic
298 plane to find informative relevant variable regions.

299

300

301 **3.1.3 Comprehensive template matching fingerprinting**

302 The specificity of the fingerprinting process is clearly improved when positive matches are
303 limited to those peaks resulting from the same analyte within a set of samples. Complex
304 chromatograms, such as those from roasted hazelnuts volatiles (**Figure 1**), may include hundreds of
305 peaks and the identification of which peaks in a pair (or in a set) of chromatograms correspond for
306 both relative retention (i.e., time position) and identity (MS fragmentation) is fundamental.

307 “Template-matching fingerprinting” was used successfully in previous investigations to
308 identify target analytes in two-dimensional chromatograms [35]. This approach, implemented with
309 the possibility to extend correspondences to the MS fragmentation pattern similarity, was, thus,
310 adopted for a non-targeted analysis to try to reliably match as many peaks as possible in a set of
311 chromatograms. The procedure first detects peaks in a source chromatogram to create a template
312 that records the retention times, detector responses, and MS fragmentation patterns. Next, on one of
313 the chromatograms to be compared, the matching algorithm determines the geometric
314 transformation in the retention-times plane that best fits the expected peak pattern in the template
315 and, in addition, evaluates the mass spectral match factor for the corresponding peaks. The
316 correspondence is established, if a peak is detected within the retention-times window around the
317 corresponding transformed template peak, also showing an MS fragmentation pattern with a proper

318 match factor [36,37]. The effectiveness of the algorithm adopted for the template transformation has
319 been extensively discussed in previous work [36,38].

320 This operation, applied to the entire set of sample chromatograms, generates a *consensus template*
321 of non-targeted peaks that can be matched across all pairs of chromatograms within the set.

322 The following procedure was applied to establish reliable peak correspondences across the set of
323 chromatograms:

324 1. Each chromatogram was baseline corrected in agreement with a specific algorithm whose
325 peculiarities are discussed in detail in a previous paper [39].

326 2. 2D-peaks were detected. For explanatory purposes, the set of chromatograms denoted A, B,...I,
327 are considered in which the detected peaks in chromatogram A are denoted $A(i)$ where i is a
328 unique peak ID.

329 3. A template was created for the first chromatogram. For each peak in the chromatogram, a peak
330 was added to the chromatogram template together with its expected retention times. For
331 example, the template for chromatogram A will have an expected peak denoted $a(i)$ at the
332 retention times of the detected peak $A(i)$.

333 4. For each peak in the template a rule was added to constrain MS matching using a CLICTM
334 expression [31] such as:

335 $\text{Match}(\langle \text{ms} \rangle) > \text{match_factor}$

336 where " $\langle \text{ms} \rangle$ " is the average mass spectrum of the template peak. The match function computes
337 the match factor between the template spectrum and the detected peak spectrum, and the
338 corresponding match-factor value should be the highest match-factor determined by considering
339 all other peaks in the source chromatogram for the template by using the NIST MS Search
340 algorithm [40]. In other words, the match-factor with the peak that has the most similar mass
341 spectrum is determined and "accept" only those with a value higher than that in the rule [32].

342 5. Next, the template was matched to the detected peaks in the next chromatogram of the set. For
343 example, when the template from chromatogram A is matched to the detected peaks in
344 chromatogram B, template peak $a(i)$ either matches some peaks $B(j)$ or not. Then, for each
345 unmatched peak in the chromatogram B, a template peak was added to the template, e.g.,
346 template peak $b(j)$ for peak $B(j)$.

347 6. Step 5 is repeated for every chromatogram, producing a comprehensive template with a peak for
348 every detected peak in the set of chromatograms.

349 The comprehensive template was matched to each chromatogram and the set of peaks that matched
350 at least for two chromatograms in the set, were included in a *consensus template*.

351 The automatic processing of samples, possible with the implemented tools present in the last
352 software release, takes on average 2 minutes for each chromatogram (9-12 MB each data file) and
353 outputs are given in different file formats.

354 Each peak in the *consensus template* was listed together with its expected retention times (i.e.,
355 averages of the retention times of the corresponding peaks in the set of individual templates), the
356 mass spectrum (i.e. the average of the mass spectra) and the match factor value for the rule (i.e. the
357 average of the match factor values). In the example, if $A(i)$, $B(j)$, and $C(k)$ are matched peaks, then
358 the consensus template peak denoted is $t(i,j,k) = \text{Average}(A(i), B(j), C(k))$.

359 **Figure 2** illustrates a GCxGC plot of Italian hazelnuts from Piedmont (i.e., Piemonte), with
360 the locations of all 422 peaks in the *consensus template*. The subset of 196 template peaks with
361 matches in all nine chromatograms are shown with white filled circles. **Table 2** lists the first twenty
362 2D-peaks that reliably matched across the set and were present in all nine varieties. Template peaks
363 are listed in decreasing order of average normalized volume together with their retention times (^1D
364 $\text{min} - ^2\text{D}$ s) and relative standard deviation (RSD%). The first column indicates the peak numbering
365 (M_i) and, where possible, the identity of the specific analyte. The largest value on each row is in
366 bold while the smallest is in italics. Cumulative results are summarized at the bottom of the table as
367 number of matched peaks with the *consensus template* together with percent matching. Again, the
368 number of matched peaks over the reference template, composed by 422 peaks that reliably
369 matched across the set, indicate the degree of similarity of each sample pattern with the *consensus*
370 *template*. In this case, matching results indicate unequivocally those peaks (i.e., analytes) that are
371 present in, at least, two samples within the set and whose variation can be considered as a
372 diagnostic tool for a better pattern discrimination or to correlate sample composition with known
373 chemical descriptors. It is interesting to note that Piemonte hazelnuts still showed the smallest
374 matching percentage, 46.4%, indicating here again a lower degree of similarity with the *consensus*
375 *template*. On the other hand, results visualized in **Figure 3**, are in agreement with those reported by
376 the chromatographic fingerprinting, except for the Akçakoca and Ordu varieties. Differences
377 between samples are larger than those reported from simple pattern recognition (i.e.,
378 chromatographic fingerprinting) and demonstrate that constraining positive correspondences to MS
379 fragmentation similarity greatly improved the sensitivity and specificity of the method.

380 Because one of the goals of this study was also to evaluate abilities, and limits, of
381 fingerprinting techniques in sample profiling with a focus on technological and aroma markers, the
382 last step in data elaboration was the identification of discriminating analytes. *Minutiae features*
383 significantly varying across samples were first examined then, on the basis of template-based

384 fingerprinting results, reliably matched peaks were located on each sample profile and analytes
385 identified. Results are summarized in **Table 3**.

386 The list reports 79 analytes with a certain discrimination potential, confirmed by
387 fingerprinting elaboration, and with a known role in defining sensory properties, as indicators of the
388 intensity of thermal treatments or as components of vegetable origin (terpenoids) characteristic of
389 the un-roasted hazelnut volatile fraction. Data interpretation can now be based on a limited number
390 of known targets, thus affording a more effective and realistic discrimination process. It is
391 interesting to observe that, with the exception of *features* 10, 17 and 18 (see **Table 1** for *feature*
392 numbering), the two fingerprinting approaches gave univocal results in indicating regions whose
393 response variation over the sample set was high in both, chromatographic fingerprinting, and/or
394 template matching of 2D peaks with MS. On the other hand, reliable peak matching provided more
395 definitive results, because it also revealed peaks that were present in few samples (data not shown)
396 representing a valuable qualitative diagnostic tool, in this case identifying marker analytes whose
397 presence could be ascribed to specific geographical origins.

398 Terpenoids such as α -pinene, sabinene and limonene were detected in all hazelnuts
399 patterns, but β -pinene, δ -3-carene, α - and γ -terpinene, and *trans*-sabinene hydrate were present in
400 few samples and, in particular, δ -3-carene and *trans*-sabinene hydrate showed a high variability.
401 Moreover, it has to be stressed that the reliability of a comparative analysis on samples, whose
402 volatiles distribution is conditioned by several variables: botanical origin, pedo-climatic harvest
403 conditions, post-harvest storage and roasting time/temperature ratios, has to be proved and up-dated
404 constantly. In this perspective, the fingerprinting procedure appears to be a valuable methodology
405 because of its potential to directly compare samples patterns and easily extract information on
406 analytes distribution, including minor components. Results on technological markers and aroma
407 compounds will be discussed in the next section.

408

409 **3.2 Sample profiling: aroma and technological markers.**

410 Comprehensive template fingerprinting results were also used to define a more specific profile for
411 each sample based on aroma and/or technological marker distribution, to be used as an additional
412 informative tool for sample discrimination. The aim of this extended target analysis of the sample
413 pattern was to see whether the comparatively significant analytes detected by the fingerprinting
414 methods can be correlated to known markers and, in consequence, to sample properties, thus
415 concurring to define their overall quality. Markers were identified on the basis of their linear
416 retention indexes (I^T_S) and MS-EI fragmentation pattern similarity (fixed acceptable value above
417 850 referred to Identity Spectrum Match factor resulting from the NIST Identity Spectrum Search

418 algorithm - NIST MS Search 2.0) with compounds collected in commercial and in-house databases
419 or, where possible, with authentic standard confirmation.

420 The extended list of markers in **Table 3** consists of: (a) analytes with the highest ranking
421 in the template-based fingerprinting procedure (classification based on decreasing order of SD on
422 average normalized volumes) and (b) analytes whose sensory, technological, and botanical
423 significance is already known [13,15,26].

424 The results derived from the distribution of aroma markers are interesting. Several potent
425 odorants were detected in the GCxGC patterns of the roasted hazelnuts under study. These
426 compounds, isolated by Solvent Assisted Flavor Evaporation (SAFE) extraction from raw and
427 roasted hazelnuts and identified by GC-O, and in particular with the AEDA screening technique
428 [18], showed high Flavor Dilution (FD) factors indicating their prominent role in defining the
429 characteristic aroma of the final product. This group of odorants, 56 in the raw and 57 in roasted
430 hazelnuts, showed FD factors above 19 and can be defined as “key-aroma” compounds [25,26].
431 **Table 3** reports the list of identified analytes together with *feature* numbering (F_i), derived by
432 chromatographic fingerprinting, identification number (#ID), compound name, Odor Quality [41]
433 for the sub-set of 16 key-aroma markers of roasted hazelnuts (indicated with an asterisk), 1D and 2D
434 retention times and average normalized volumes for the nine geographical origins. Markers were
435 identified on the basis of their linear retention indices (I_s^T) and EI-MS spectra compared to those of
436 authentic standards.

437 The distribution of potent odorants in the four Italian (i.e., *Romana*, *Giffoni*, *Mortarella*,
438 *Piemonte*), standard roasted hazelnut samples is visualized in the histogram of **Figure 4**. This
439 profiling confirms the perceivable differences of the overall sensory impact provided by roasted
440 samples of different origin [18,20,24,42]. In particular: 2- and 3-methylbutanal (4 and 7) and 2,3-
441 pentanedione (12) concur to define the characteristic malty and buttery notes; 5-methyl-4-
442 heptanone, 5-methyl-(Z)-2-hepten-4-one (27) and 5-methyl-(E)-2-hepten-4-one (filbertone) (35) are
443 responsible for the fruity and nutty sensation ; hexanal (13) and octanal (34) are perceived as green
444 and fatty respectively, while secondary lipid-peroxidation products such as (E)-2-heptenal (38), (E)-
445 2-octenal (47), (E)-2-nonenal (59), (E)-2-decenal (70), (E,E)-2,4-decadienal (74) provide fatty
446 sensations. The sweet and caramel like note can be ascribed to the presence of 4-hydroxy-2,5-
447 dimethyl-3(2H)-furanone (79), while phenylacetaldehyde (68) and 2-phenylethanol (76) elicit
448 flowery and honey-like sensations. The highly variable abundance of some markers (e.g. 2- and 3-
449 methylbutanal, hexanal, octanal, nonanal (45) and acetic acid (52)) is extremely informative of this
450 aroma profiling assessment and provides a further valuable interpretation key for sample
451 discrimination.

452 Aroma compounds are characterized by a very high concentration variability in roasted
453 samples, ranging from traces (ng/g) to several percent (g/100g), therefore sample pre-concentration
454 is mandatory for a complete aroma profiling extended to the entire pattern of key-odorants. The
455 literature refers to an average amount in roasted Romana hazelnuts ranging from 7 mg/kg of 3-
456 methylbutanal, the most abundant, to about 2 µg/kg of (*E,E*)-2,4-decadienal [42]. However, thanks
457 to its high sensitivity, GCxGC enabled us to identify and monitor the variation of 16 key-aroma
458 compounds and semi-quantify them by their relative abundance in the sample set. Even though it's
459 well-known that HS-SPME is not representative of the "absolute" composition of the volatile
460 fraction of a sample, after a careful standardization of the sampling procedure, it delivers reliable
461 data, also avoiding long and artefact producing chemical treatments [43].

462 Further interesting groups of markers, useful to evaluate the thermal treatment and/or the
463 post-harvest storage conditions, are compounds formed by the Maillard reaction, the Strecker
464 degradation, and lipid-peroxidation, whose presence can be correlated to known precursors in the
465 raw material. In addition, their abundance reflects the extent of thermal stress or exposure to
466 oxidative conditions. Pyrazines for example, present a homogeneous distribution. The highest
467 variability was registered for 2,5-dimethylpyrazine (41) and 2-ethyl-3,5-dimethyl pyrazine (51),
468 while 2,5-diethyl pyrazine (49) was detected in only one sample, the Piemonte origin. Despite their
469 high odor thresholds and, as a consequence, low impact on sensory properties, alkyl pyrazines
470 formation can successfully be correlated with the extent of thermal treatments representing a very
471 sensitive tool for technological profiling.

472 Secondary products of lipid-peroxidation, such as saturated and unsaturated aldehydes can
473 simultaneously provide information on aroma and technological profile. Lipid oxidation strongly
474 affects shelf life and sensory characteristics of hazelnuts and depends on several factors such as the
475 concentrations of unsaturated fatty acids, enzymatic activity, mineral composition, and amount of
476 antioxidants [44,45]. Prolonged storage of hazelnuts induces the formation of volatile off-flavors,
477 short chain fatty acids, and saturated and un-saturated aldehydes, such as hexanal and octanal, the
478 most abundant lipid oxidation products that can increase up to tenfold their original concentrations
479 [46]. The roasting procedure is also a factor promoting lipid oxidation. The homologous series of
480 saturated aldehydes: hexanal, heptanal, octanal, nonanal, and decanal (the latter detected only in
481 few samples) can, therefore, be diagnostic in this perspective, especially, because of their very high
482 variability within the samples investigated. On the other hand, unsaturated aldehydes such as (*E*)-2-
483 heptenal, (*E*)-2-octenal, (*E*)-2-nonenal, (*E*)-2-decenal and (*E,E*)-2,4-decadienal, present in very low
484 concentrations, were only detected thanks to GCxGC sensitivity, emphasizing its ability to detect
485 trace and minor components and include them in sample profiling. However, it has to be stressed

486 that GC-O screening indicated the homologous series of (Z)-alkenals (i.e., (Z)-2-octenal, (Z)-2-
487 nonenal, (Z)-2-decenal) as the highest impacting odorants responsible for the fatty and deep-fried
488 notes in pan-roasted hazelnuts. This unusual behavior was ascribed to the procedure exposing
489 grinded hazelnut to air before roasting, therefore increasing the possibility for unsaturated fatty
490 acids to react with oxygen [24]. Industrial roasting, performed on fruits protected by kernel, reduces
491 the exposure of fatty fraction to oxidative, degradation and, consequently, reduces the formation of
492 (Z)-alkenals.

493 Aroma and technological marker profiles, extended to a wide range of analytes, are
494 undoubtedly two very powerful diagnostic tools enabling correlation between quality descriptors
495 (aroma and sensory properties) and process variables (post-harvest storage conditions, roasting
496 treatment). Roasted hazelnut volatiles are a challenging fraction to evaluate how fingerprinting
497 methods can guide towards a more profitable speciation of samples, improving the effectiveness of
498 GCxGC targeted analysis.

499

500 **4. Conclusions**

501 Fingerprint analysis, whose results are based on the degree of similarity with a reference
502 template, showed to be effective for sample comparison and classification of roasted hazelnuts.
503 *Chromatographic fingerprinting*, in particular, was (a) effective as a "screening" method to locate
504 informative relevant regions on the separation space, (b) versatile for processing of single channel
505 detectors patterns (GCxGC-FID, GCxGC-ECD etc...) and (c) less time consuming since the
506 automatic processing of raw data took less than 1 min for each chromatogram. It may incompletely
507 delineate features, but may have fewer mismatched features. Feature matching was constrained by
508 retention times and MS fragmentation patterns to obtain consistent correspondences only for those
509 analytes whose spectra referred a fixed degree of similarity with the corresponding template
510 spectrum. The reliable peak matching procedure, implemented in the *comprehensive template*
511 *matching fingerprinting* approach, enabled a successful screening of 2D peak distribution over the
512 sample set, and the extraction of consistent information on analytes that were present in all or a few
513 samples, suggesting the possible discrimination roles they can play in the comparative process. The
514 cumulative matching results (percent matching) obtained with this approach showed, in fact, better
515 specificity and sensitivity in discriminating samples differing for geographical origin than those
516 obtained with chromatographic fingerprinting. The main limit concerns mismatching for those
517 template peaks whose reference MS spectrum is qualitatively unacceptable (intensity below a given
518 S/N) and, as a consequence matching values below the expected threshold.

519 Fingerprint analysis is an important tool to extend the informative potential of GCxGC; in particular
520 in the flavor field, the fingerprint-assisted investigation of the distributions of known and unknown
521 markers of a vegetable matrix can be very useful for the definition of the so-called *product*
522 *signature* in terms of sensory properties, botanical/geographical origin and/or to study the
523 modifications induced by thermal treatments on primary and secondary metabolites.

524

525 **Acknowledgments**

526 This research was carried out within the project entitled: “Sviluppo di metodologie innovative per
527 l’analisi di prodotti agroalimentari” (FIRB Cod.: RBIP06SXMR_002) of the Ministero
528 dell’Istruzione, dell’Università e della Ricerca (MIUR) (Italy).

529 **REFERENCES**

530

- 531 [1] Merriam Webster Dictionary and Thesaurus ISBN: 978-0-87779-640-4
532 [2] K. M. Pierce, J. C. Hoggard, R. E. Mohler, R. E. Synovec, J. Chromatogr. A, 1184 (2008)
533 341.
534 [3] H.G. Schmarr, J. Bernhardt, J. Chromatogr. A, 1217 (2010) 565
535 [4] J. Vial, H. Nocairi, P. Sassi, S. Mallipatu, G. Cognon, D. Thiébaud, B. Teillet, D.N.
536 Rutledge, J. Chromatogr. A 1216 (2009) 2866
537 [5] J. C. Hoggard, J. H. Wahl, R. E. Synovec, G. M. Mong, C. G. Fraga, Anal. Chem. 82 (2010)
538 689
539 [6] B. Wang, A. Fang, J. Heim, B. Bogdanov, S. Pugh, M. Libardoni, X. Zhang, Anal. Chem.
540 82 (2010) 5069
541 [7] E. M. Humston, J. D. Knowles, A. McShea, R. E. Synovec, Chromatogr. A, 1217 (2010)
542 1963
543 [8] R.B. Gaines, G.S. Frysinger, M.S. Hendrick-Smith, J.D. Stuart, J.D. Environ. Science &
544 Technol. 33 (1999) 2106.
545 [9] R.A. Shellie, W. Welthagen, J. Zrostliková, J. Spranger, M. Ristow, O. Fiehn, R.
546 Zimmermann, J. Chromatogr. A. 1086 (2005) 83.
547 [10] Th. Gröger, M. Schäffer, M. Pütz, B. Ahrens, K. Drew, M. Eschner, R. Zimmermann, J.
548 Chromatogr. A. 1200 (2008) 8.
549 [11] F. Acükurt, M. Ozdemir, G. Biringen, M. Loker, Food Chem. 65 (1999) 309.
550 [12] Durak, I. Koksal, M. Kacümaz, S. Buyukkociak; B. M. Y. Cüimen, H. S. Ozturk, Clin.
551 Chim. Acta 284 (1999) 113.
552 [13] D. Stone, Cereal Foods World 45 (2000) 424.
553 [14] T.E. Kinlin, R. Muralidhara, A.O. Pittet, A. Sanderson, J.P. Walradt, J. Agric. Food Chem.
554 20 (1972) 1021.
555 [15] R. Emberger “On the role of nature-identical aroma compounds in food flavoring” (in
556 German); 43. Diskussionstagung des Forschungskreises der Ernährungsindustrie e.V., 1985,
557 41-60.
558 [16] W. Silberzahn, “A GC-MS study on aroma compounds in hazelnuts“ (in German). PhD
559 dissertation, Technical University of Berlin, 1988.
560 [17] S. Langourieux, R. Perren, F. Escher in P. Schieberle, K.H. Engel, (Eds.), Frontiers of
561 Flavour Science, Deutsche Forschungsanstalt für Lebensmittelchemie: Garching, Germany,
562 2000; pp 527-535.
563 [18] S. Saklar, S. Katnas, S. Ungan, Int. J. Food Sci. Technol. 36 (2001) 271.
564 [19] C. Alasalvar, F. Shahidi, K.R. Cadwallader, J. Agric. Food Chem. 51 (2003) 5067.

- 565 [20] C. Alasalvar, A.Z. Odabasi, N. Demir, M.O. Bolaban, F. Shahidi, K.R. Cadwallader, J. Food
566 Sci. 69 (2004) 99.
- 567 [21] S. Saklar, S. Katnas, S. Ungan, Int. J. Food Sci. Technol. 36 (2001) 271.
- 568 [22] W. Grosch, Chem Senses 26 (2001) 533.
- 569 [23] C. Cordero, C. Bicchi, P. Rubiolo. J. Agr. Food Chem. 56 (2008) 7655.
- 570 [24] Burdack-Freitag, P. Schieberle, Changes in the Key odorants of Italian hazelnuts (Tonda
571 romana) induced by roasting. J. Agr. Food Chem. submitted
- 572 [25] P. Schieberle in Goankar A (Ed) Characterization of Food: Emerging Methods. Elsevier,
573 Amsterdam, The Netherlands 1995, pp. 403.
- 574 [26] W. Grosch, Trends Food Sci. Technol. 4 (1993) 68.
- 575 [27] W. Engel, W. Bahr, P. Schieberle, Eur. Food Res. Technol. 209 (1999) 237.
- 576 [28] F. Ullrich, W. Grosch, Z. Lebensm. Unters. Forsch., 184 (1987) 277.
- 577 [29] Y. Wang, J. O'Reilly, Y. Chen, J. Pawliszyn, J Chromatogr. A, 1072 (2005) 13.
- 578 [30] L. Setkova, S. Risticovic, C. M. Linton, G. Ouyang, L.M. Bragg, J. Pawliszyn, Anal Chim
579 Acta 581 (2007) 221.
- 580 [31] C. Cordero, C. Bicchi, M. Galli, S. Galli, P. Rubiolo. J. Sep. Sci. 31 (2008) 3437.
- 581 [32] C. Cordero, E. Liberto, C. Bicchi, P. Rubiolo, S.E. Reichenbach, X. Tian, Q. Tao. J. Chrom.
582 Sci 48 (2010) 251.
- 583 [33] K. Jain, L. Hong, S. Pankanti, R. Bolle, Proceedings of the IEEE, 85 (1997) 1365.
- 584 [34] D. Maio, D. Maltoni, IEEE T. Pattern Anal. 19 (1997) 27.
- 585 [35] S. Reichenbach, M. Ni, V. Kottapalli, A. Visvanathan. Chem. Intell. Lab. Syst. 71 (2004)
586 107.
- 587 [36] S. Reichenbach, P. Carr, D. Stoll, and Q. Tao, J. Chromatogr. A, 1216 (2009) 3458.
- 588 [37] S. Reichenbach, V. Kottapalli, M. Ni, and A. Visvanathan, J. Chromatogr. A 1071 (2005)
589 263.
- 590 [38] M. Ni, S.E. Reichenbach, A. Visvanathan, J. TerMaatc, E. B. Ledford, J. Chromatogr. A
591 1086 (2005) 165.
- 592 [39] S. Reichenbach, M. Ni, D. Zhang, E. Ledford, J. Chromatogr. A 985 (2003) 47.
- 593 [40] S. Stein. NIST Mass Spectral Search Program (Version 2.0f). NIST Mass Spectrometry
594 Data Center, 2008
- 595 [41] M. Rychlik, P. Schieberle, W. Grosch. Compilation of Odor Thresholds, Odor qualities and
596 Retention Indices of Key Food Odorants. Deutsche Forschungsanstalt für
597 Lebensmittelchemie and Institut für Lebensmittelchemie der Technischen Universitat
598 Munchen. (1998) ISBN 3-9803426-5-4.
- 599 [42] A. Burdack-Freitag. PhD thesis, TU Muenchen, Germany. On the formation of key odorants
600 during roasting (in German), 2007.
- 601 [43] C. Bicchi; C. Cordero; E. Liberto; B. Sgorbini, P. Rubiolo. J Chromatogr. A 1184 (2008)
602 220.
- 603 [44] M. Ozdemir, O. Devres, Food Rev. Int. 15 (1999) 309.
- 604 [45] B. Fallico, E. Arena, M. Zappalà, Food Chem. 81 (2003) 569.
- 605 [46] J. L. K. Kinderlerer, S. Johnson, J. Sci. Food Agric. 58 (1992) 89.
- 606
- 607

608 **Captions to Tables :**

609

610 **Table 1:** First 20 *minutiae* with the largest average percent response (i.e., the response within the
611 mesh panel divided by the response within the entire chromatogram) together with *feature*
612 numbering (F_i), average retention times (1D min – 2D s) of the *feature* apex; *feature's* average
613 percent response relative standard deviation (RSD%) and average percent response from hazelnuts
614 of nine origins. The largest value on each row is in bold while the smallest is in italics. Cumulative
615 results, number of matched *features* with the *consensus template*, are expressed as percent of
616 matching.

617

618

619 **Table 2:** First 20 *peaks* that reliably match across the sample set (retention times and MS
620 fragmentation pattern) and present in all samples with the largest variability referred to average
621 normalized volume. Peaks, in decreasing order of average normalized volumes, are listed together
622 with *peak* numbering (P_i), Compound name, retention times (1D min – 2D s); *peak* normalized
623 volumes relative standard deviation (RSD%) and normalized volumes from hazelnuts of nine
624 origins. The largest value on each row is in bold while the smallest is in italics. Cumulative results,
625 number of matched *peaks* with the *consensus template* are expressed as % of matching. Asterisk (*)
626 indicates key-aroma markers (see text for details).

627

628 **Table 3:** List of analytes adopted to characterize the samples: Chromatographic fingerprinting
629 *features* numbering (F_i), Identification number (#ID), Compound name, Odor Quality for key-
630 aroma (*) markers of roasted hazelnuts, 1D and 2D retention times, average normalized volumes for
631 the nine geographical origins (average value of three replicates). Markers were identified on the
632 basis of their linear retention indices (I^T_s) and MS-EI spectra compared with those of authentic
633 standards.

634

635

636

637 **Figure legends:**

638

639 **Figure 1:** Cumulative chromatogram for the nine samples of roasted hazelnuts and the regions of
640 detected peaks used for chromatographic fingerprinting shown as white polygons. The number of
641 chromatographic *features* is 411.

642

643 **Figure 2:** GCxGC-qMS plot of Italian hazelnuts from Piedmont (i.e., Piemonte). Circles indicate
644 the 422 peaks in the *consensus template*. The subset of 196 template peaks with matches in all nine
645 chromatograms are shown with white filled circles.

646

647 **Figure 3:** Fingerprinting results expressed as % of matching with the *consensus template* (i.e.,
648 number of matched peaks divided by the total number of template peaks). Results are referred to
649 chromatographic (-----) and comprehensive template matching fingerprinting with MS approach
650 (—).

651

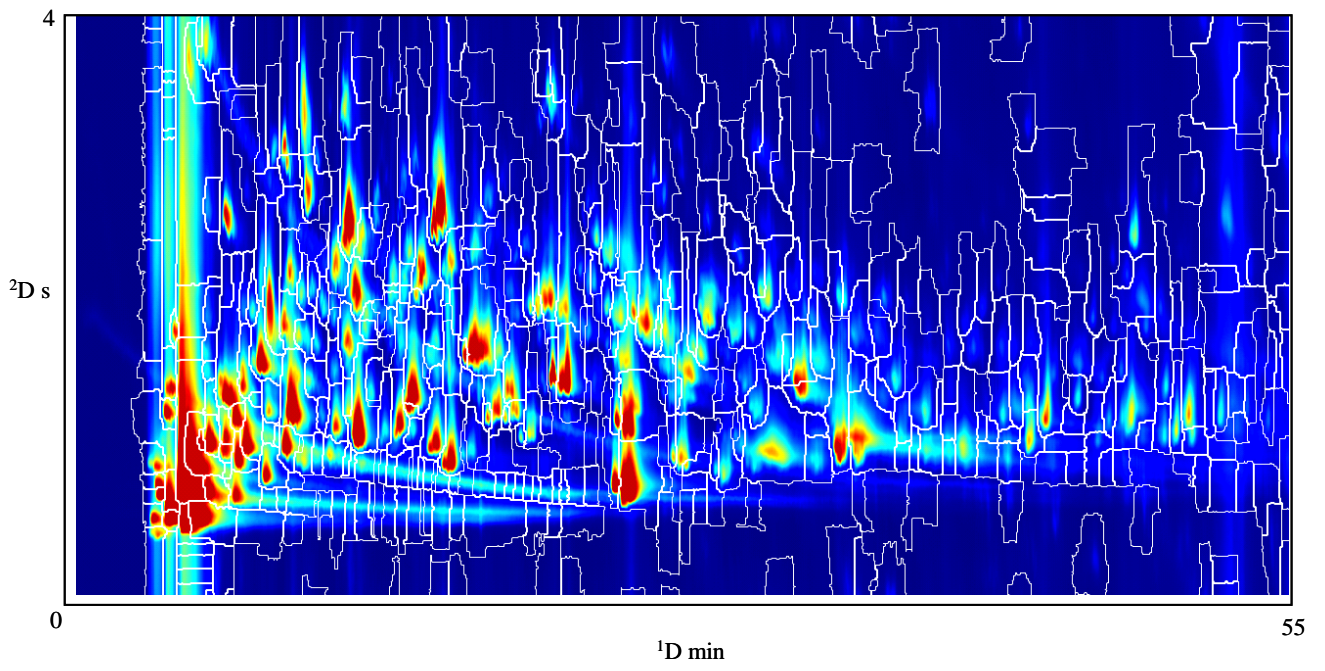
652 **Figure 4:** Key-aroma pattern of the four Italian varieties (i.e., *Romana*, *Giffoni*, *Mortarella*,
653 *Piemonte*) submitted to a standard roasting procedure. Results are reported as normalized 2D-Peak
654 Volume over the ISTD. For analyte ID (*x*-axis) and full data of all investigated samples see Table 3.

655

656

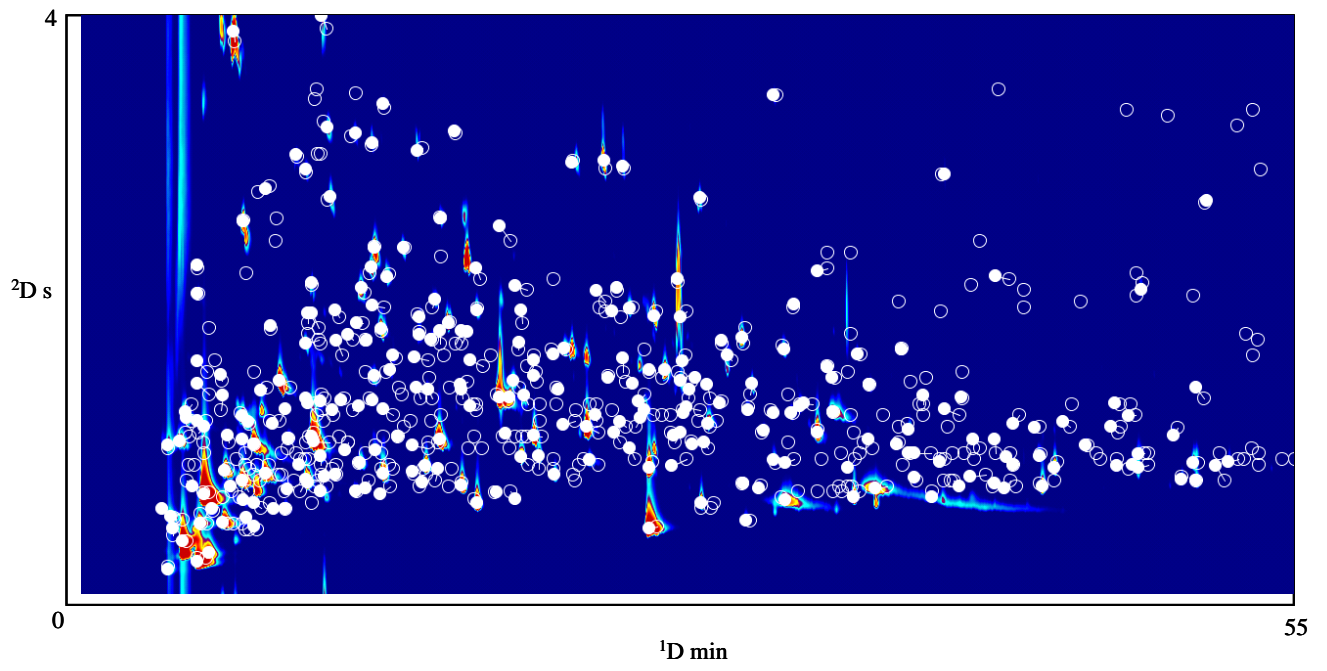
657

658 **Figure 1**
659



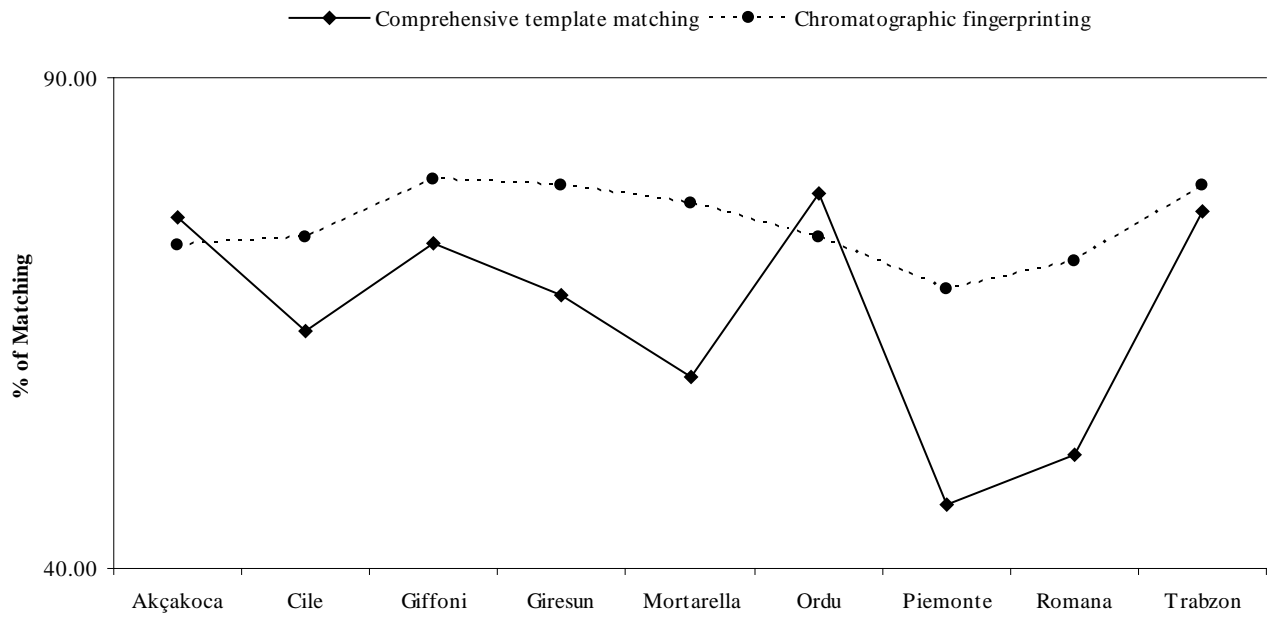
660
661
662
663
664
665

666 **Figure 2**
667



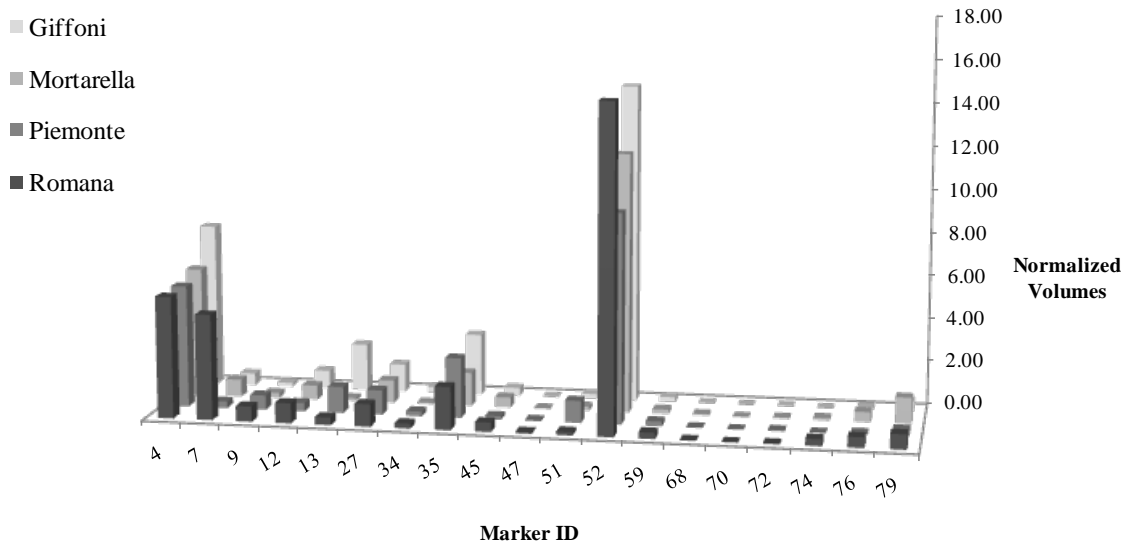
668
669
670
671

672 **Figure 3**
673



674
675
676

677 **Figure 3**



678

Table 1: First 20 *minutiae* with the largest average percent response (i.e., the response within the mesh panel divided by the response within the entire chromatogram) together with *feature* numbering (F_i), average retention times (1D min – 2D s) of the *feature* apex; *feature*'s average percent response relative standard deviation (RSD%) and average percent response from hazelnuts of nine origins. The largest value on each row is in bold while the smallest is in italics. Cumulative results, number of matched *features* with the *consensus template*, are expressed as percent of matching.

Feature n°	1D (min)	2D (s)	RSD%	Average percent response								
				Akçakoca	Cile	Giffoni	Giresun	Mortarella	Ordu	Piemonte	Romana	Trabzon
F1	4.42	0.53	51.65	8.23	22.73	16.86	20.87	29.97	<i>1.17</i>	<i>27.57</i>	35.73	26.28
F2	3.35	0.41	150.51	0.40	0.72	1.39	<i>0.01</i>	0.69	0.19	9.59	5.70	0.89
F3	7.55	1.03	174.64	0.04	0.04	0.04	5.80	<i>0.03</i>	0.04	1.20	3.45	0.03
F4	40.89	0.90	113.61	1.12	3.82	0.36	4.94	0.23	2.24	<i>0.19</i>	0.33	0.67
F5	4.75	0.66	37.55	4.70	2.57	6.87	<i>1.61</i>	4.73	3.38	4.25	3.67	3.79
F6	15.22	1.89	128.65	0.91	1.44	0.22	4.14	0.05	2.00	0.16	<i>0.15</i>	0.33
F7	3.95	0.41	39.04	4.36	1.99	3.60	<i>1.15</i>	3.94	1.96	4.19	2.51	2.59
F8	18.29	0.99	86.97	2.27	2.38	0.58	2.66	<i>0.21</i>	2.29	0.01	0.19	0.82
F9	20.02	2.18	174.40	0.58	0.86	0.08	3.01	<i>0.01</i>	0.25	0.02	0.05	0.12
F10	3.49	0.58	74.49	0.61	1.12	0.78	0.07	1.21	1.95	0.60	<i>0.06</i>	1.85
F11	3.82	0.86	91.16	0.62	1.54	0.50	0.28	0.30	0.32	1.82	<i>0.12</i>	0.43
F12	27.95	1.27	155.76	0.35	0.47	0.12	1.72	0.06	<i>0.01</i>	0.09	0.10	0.17
F13	5.62	0.74	55.33	1.22	1.25	0.47	1.32	0.61	<i>0.16</i>	0.29	0.97	0.72
F14	11.35	1.60	87.37	0.25	0.12	0.32	0.17	0.22	0.29	1.11	0.70	<i>0.17</i>
F15	36.29	0.82	107.10	0.16	0.58	0.10	0.79	<i>0.05</i>	0.34	0.05	0.07	0.11
F16	45.22	0.99	150.42	0.10	0.13	0.07	0.85	<i>0.04</i>	0.23	0.04	0.04	0.06
F17	3.15	0.90	87.27	0.25	0.12	0.55	<i>0.01</i>	0.19	0.40	0.04	0.04	0.41
F18	33.35	0.82	47.79	0.58	0.37	0.51	0.23	0.50	0.56	<i>0.07</i>	0.25	0.25
F19	6.89	0.70	52.61	0.26	0.24	0.35	<i>0.03</i>	0.19	0.18	0.38	0.56	0.33
F20	7.02	0.99	83.74	0.33	0.07	0.17	<i>0.04</i>	0.14	0.15	0.14	0.05	0.49
Chromatographic Fingerprinting results				Akçakoca	Cile	Giffoni	Giresun	Mortarella	Ordu	Piemonte	Romana	Trabzon
Number of matched features (over 411)				300	303	327	325	317	303	281	293	325
Match %				72.99	73.72	79.56	79.08	77.13	73.72	68.37	71.29	79.08

Table 2: First 20 *peaks* that reliably match across the sample set (retention times and MS fragmentation pattern) and present in all samples with the largest variability referred to average normalized volume. Peaks, in decreasing order of average normalized volumes, are listed together with *peak* numbering (*Pi*), Compound name, retention times (¹D min – ²D s); *peak* normalized volumes relative standard deviation (RSD%) and normalized volumes from hazelnuts of nine origins. The largest value on each row is in bold while the smallest is in italics. Cumulative results, number of matched *peaks* with the *consensus template* are expressed as % of matching. Asterisk (*) indicates key-aroma markers (see text for details).

Peak n°	Compound name	¹ D (min)	² D (s)	RSD%	Normalized Volumes								
					Akçakoca	Cile	Giffoni	Giresun	Mortarella	Ordu	Piemonte	Romana	Trabzon
P1	Acetic Acid*	23.16	0.66	38.40	8.29	8.99	15.36	5.77	14.72	5.33	11.94	16.48	9.65
P2	3-Methyl butanal*	4.76	0.66	40.52	4.07	2.38	7.61	2.25	5.98	3.45	5.63	5.62	3.40
P3	2-Propanone	3.96	0.41	43.27	3.77	1.84	3.99	<i>1.61</i>	4.99	2.00	5.55	3.83	2.32
P4	2-Furancarboxaldehyde	23.22	0.90	32.51	2.02	2.44	3.83	1.66	1.99	<i>1.26</i>	2.76	3.14	2.55
P5	Pentanol	13.62	0.86	91.00	1.51	3.92	0.63	4.70	0.43	2.58	<i>0.46</i>	1.00	0.64
P6	Hexanol	18.22	1.03	84.28	1.96	2.21	0.65	3.73	<i>0.27</i>	2.34	0.37	0.51	0.74
P7	5-Methyl-(E)-2-hepten-4-one (<i>Filbertone</i>)*	15.29	1.73	54.05	1.82	0.87	2.84	0.70	1.55	1.38	2.76	1.98	<i>0.40</i>
P8	Octanal*	15.22	1.89	150.25	0.79	1.33	0.25	5.79	0.06	2.04	<i>0.21</i>	0.23	0.30
P9	2-Methylpyrazine	14.22	0.95	33.20	<i>0.87</i>	1.25	2.37	0.87	1.62	1.27	1.85	1.40	1.34
P10	Heptanol	23.02	1.15	134.46	0.71	1.24	0.34	4.70	<i>0.11</i>	1.66	0.31	0.29	0.36
P11		15.82	0.74	38.12	0.65	1.36	1.54	0.78	1.25	<i>0.38</i>	1.07	1.56	1.01
P12	2-Furanmethanol	32.56	0.86	45.64	0.62	0.84	1.60	0.41	1.03	<i>0.31</i>	0.94	1.17	0.87
P13		3.69	0.78	175.24	<i>0.14</i>	0.24	0.22	3.51	0.24	0.17	0.32	0.53	0.22
P14	2,4-Dimethyl-1-heptene	4.42	1.15	70.97	0.39	0.73	0.53	0.78	0.72	<i>0.08</i>	1.64	0.19	0.68
P15	Octanol	27.89	1.32	146.99	0.30	0.44	0.13	2.40	<i>0.07</i>	0.78	0.13	0.16	0.15
P16		3.82	0.86	70.97	0.54	1.43	0.55	0.39	0.38	0.33	0.41	<i>0.18</i>	0.38
P17	Dihydro-2(3H)-Furanone	30.82	1.11	25.02	<i>0.32</i>	0.40	0.70	0.39	0.66	0.53	0.50	0.60	0.49
P18	2-Methyl-1-butanol	11.89	0.82	47.23	0.59	0.29	0.59	<i>0.05</i>	0.70	0.42	0.63	0.78	0.34
P19		5.56	0.53	33.40	0.37	0.28	0.62	0.51	0.61	0.33	0.43	0.67	<i>0.28</i>
P20	3-Methyl-2-pentanone	6.29	1.69	62.72	0.36	0.29	0.18	<i>0.20</i>	<i>0.20</i>	0.79	0.87	0.71	0.32
Comprehensive template matching results					Akçakoca	Cile	Giffoni	Giresun	Mortarella	Ordu	Piemonte	Romana	Trabzon
Number of matched peaks (over 422)					320	271	309	286	251	330	196	218	322
Match %					75.83	64.22	73.22	67.77	59.48	78.20	46.45	51.66	76.30

Table 3: List of analytes adopted to characterize the samples: Chromatographic fingerprinting *features* numbering (*F_i*), Identification number (#ID), Compound name, Odor Quality for key-aroma (*) markers of roasted hazelnuts, ¹D and ²D retention times, average normalized volumes for the nine geographical origins (average value of three replicates). Markers were identified on the basis of their linear retention indices and MS-EI spectra compared with those of authentic standards.

Feature ID	#ID	Compound name	Odor Quality	¹ D (min)	² D (s)	Normalized Volumes									
						Akçakoca	Cile	Giffoni	Giresun	Mortarella	Ordu	Piemonte	Romana	Trabzon	
F2,F7	1	2-Propanone		3.95	0.41	3.77	1.84	3.99	1.61	4.99	2.00	5.55	3.83	2.32	
	2	4-Methyl octane		4.22	1.23	0.23	1.16	1.32	0.78	0.00	2.03	1.25	0.29	0.34	
F1	3	2,4-Dimethyl-1-heptene		4.42	1.15	0.73	0.53	0.78	0.72	0.08	1.64	0.19	0.68	0.73	
F5	4	3-Methylbutanal*	malty	4.75	0.66	4.07	2.38	7.61	2.25	5.98	3.45	5.63	5.62	3.40	
	5	Ethanol		4.95	0.45	0.00	0.05	0.00	0.00	0.04	0.37	0.56	0.08	0.02	
F13	6	2,2-Dimethyl decane		5.28	2.42	0.41	0.00	0.54	0.34	0.00	0.11	0.45	0.00	0.11	
	7	2-Methylbutanal*	malty	5.62	0.78	1.23	2.12	0.55	0.00	0.79	0.48	0.30	4.86	3.32	
	8	3-Methyl-2-pentanone		6.29	0.94	0.36	0.29	0.18	0.20	0.20	0.79	0.87	0.71	0.32	
F19	9	α-Pinene*	terpene-like	6.35	1.70	0.37	0.29	0.17	0.20	0.23	0.76	0.66	0.70	0.32	
	10	(E)-2-Butenal		6.82	0.66	0.23	0.22	0.39	0.08	0.24	0.19	0.55	0.86	0.29	
F20	11	2,3,5-Trimethylfuran		7.02	0.99	0.28	0.07	0.19	0.05	0.18	0.16	0.19	0.11	0.44	
	12	2,3-Pentanedione*	buttery	7.15	0.70	0.40	0.19	0.83	0.00	0.65	0.00	0.38	0.93	0.44	
F3	13	Hexanal*	green	7.75	1.11	8.63	1.30	2.16	18.66	0.14	1.40	1.21	0.35	0.52	
	14	2-Methyl-1-propanol		7.95	0.62	0.09	0.00	0.00	0.00	0.08	0.06	0.00	0.07	0.00	
F14	15	<i>n</i> -Undecane	ISTD	8.15	3.74	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
	16	β-Pinene		8.29	1.97	0.03	0.04	0.00	0.04	0.00	0.04	0.04	0.11	0.00	
	17	Sabinene		8.75	1.93	0.11	0.14	0.08	0.11	0.25	0.11	0.28	0.30	0.22	
	18	2-Pentanol		8.82	0.70	0.62	0.26	0.59	0.00	0.70	0.43	0.23	0.99	0.57	
	19	3,3-Dimethyl-1-butene		9.02	0.86	2.73	2.49	3.30	1.57	3.60	3.27	1.33	3.88	3.83	
	20	4-Heptanone		9.02	1.40	0.24	0.21	0.36	0.06	0.26	0.14	0.06	0.20	0.33	
	21	δ-3-Carene		9.55	2.10	0.11	0.17	0.00	0.10	0.75	0.27	0.02	0.69	0.28	
	22	3-Methyl-4-heptanone		9.75	1.81	0.32	0.19	0.41	0.06	0.33	0.20	0.25	0.25	0.21	
	23	α-Terpinene		10.62	2.18	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.04	
	24	Pyridine		10.69	0.82	0.00	0.00	0.37	0.00	0.00	0.00	0.12	0.00	0.00	
	25	Heptanal		10.95	1.56	0.82	0.66	0.41	4.22	0.10	0.28	0.22	0.12	0.14	
	F14	26	Limonene		11.29	2.10	0.10	3.39	0.09	0.13	0.45	2.17	0.08	0.63	0.66
	F14	27	5-Methyl-(Z)-2-hepten-4-one*	fruity, hazelnut-like	11.49	1.68	1.31	0.38	1.27	0.24	1.04	0.42	1.12	1.04	0.69
		28	2-Methyl-1-butanol		11.89	0.82	0.59	0.29	0.59	0.05	0.70	0.42	0.63	0.78	0.34
	F6	29	2-Pentylfuran		12.62	1.73	0.27	0.68	0.10	0.64	0.09	0.30	0.09	0.10	0.14
		30	γ-Terpinene		13.15	2.18	0.00	0.03	0.00	0.03	0.17	0.02	0.00	0.08	0.30
31		Pentanol		13.62	0.86	1.51	3.92	0.63	4.70	0.43	2.58	0.46	1.00	0.64	
32		2-Methylpyrazine		14.22	0.94	0.87	1.25	2.37	0.87	1.62	1.27	1.85	1.40	1.34	
33		3-Hydroxy-2-butanone		15.15	0.78	0.29	0.32	0.60	0.23	0.71	0.19	0.36	0.56	0.31	
34		Octanal*	fatty, green	15.22	1.89	0.79	1.33	0.25	5.79	0.06	2.04	0.21	0.23	0.30	
35		5-Methyl-(E)-2-hepten-4-one (<i>Filbertone</i>)*	fruity, hazelnut-like	15.29	1.73	1.82	0.87	2.84	0.70	1.55	1.38	2.76	1.98	0.40	
36		1-Hydroxy-2-propanone		15.55	0.70	0.01	0.02	0.02	0.02	0.00	0.02	0.10	0.00	0.02	
37		2,5-Dimethylpyrazine		16.69	1.19	0.87	0.08	1.91	0.62	0.18	0.13	0.45	0.14	0.14	
38		(E)-2-Heptenal		16.75	1.56	0.82	0.66	0.41	4.22	0.1	0.28	0.22	0.12	0.14	

	39	2,6-Dimethyl pyrazine		16.95	1.19	0.25	0.20	0.69	0.30	0.52	0.31	0.68	0.56	0.37
	40	2-Ethylpyrazine		17.15	1.19	0.37	0.29	0.74	0.85	0.53	0.38	0.52	0.57	0.43
	41	2,3-Dimethyl pyrazine		17.75	1.19	0.10	0.13	0.29	0.07	0.17	0.11	0.15	0.26	0.14
F8	42	Hexanol		18.22	1.03	1.96	2.21	0.65	3.73	0.27	2.34	0.37	0.51	0.74
	43	2-Ethyl-6-methyl pyrazine		19.55	1.40	0.14	0.10	0.37	0.11	0.29	0.15	0.31	0.16	0.18
	44	2-Ethyl-5-methyl pyrazine		19.82	1.44	0.42	0.28	0.73	0.29	0.72	0.46	0.24	0.13	0.45
F9	45	Nonanal*	fatty, green	19.95	2.22	0.56	0.85	0.36	4.11	0.46	0.74	0.15	0.43	0.39
	46	2-Ethyl-3-methyl pyrazine		20.52	1.40	0.18	0.15	0.53	0.13	0.47	0.24	0.47	0.28	0.23
	47	(<i>E</i>)-2-Octenal*	fatty, green	21.62	1.85	0.16	0.51	0.03	0.70	0.02	0.24	0.04	0.07	0.10
	48	3-Ethyl-2,5-dimethyl pyrazine		22.35	1.64	0.14	0.10	0.35	0.11	0.31	0.19	0.40	0.17	0.17
	49	2,5-Diethyl pyrazine		22.95	1.68	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.00	0.00
	50	Heptanol		23.02	1.15	0.71	1.24	0.34	4.70	0.11	1.66	0.31	0.29	0.36
	51	2-Ethyl-3,5-dimethyl pyrazine*	earthy	23.06	1.64	0.13	0.23	0.22	0.15	0.18	0.11	1.01	0.16	0.09
	52	Acetic acid*	sour	23.10	1.66	15.36	5.77	14.72	5.33	11.94	16.48	9.65	14.96	9.61
	53	2-Furancarboxaldehyde		23.22	0.90	2.02	2.44	3.83	1.66	1.99	1.26	2.76	3.14	2.55
	54	1-(Acetyloxy)-2-propanone		23.49	1.03	0.14	0.00	0.36	0.00	0.21	0.16	0.12	0.00	0.00
	55	<i>trans</i> -Sabinene hydrate		23.49	1.60	0.13	0.05	0.07	0.13	0.19	0.22	0.00	0.18	0.09
	56	Decanal		24.89	2.47	0.16	0.00	0.00	0.17	0.00	0.05	0.00	0.00	0.04
	57	Pyrrole		25.55	0.78	0.14	0.14	0.52	0.07	0.25	0.10	0.30	0.24	0.21
	58	Benzaldehyde		25.82	1.15	0.46	0.05	0.65	0.26	0.04	0.03	0.37	0.15	0.02
	59	(<i>E</i>)-2-Nonenal*	fatty, green	26.42	2.01	0.27	0.08	0.22	0.21	0.19	0.19	0.21	0.28	0.18
	60	2,4-Dimethyl-3-pentanol		26.69	1.40	0.24	0.07	0.21	0.19	0.14	0.18	0.28	0.20	0.18
	61	Propanoic acid		27.29	0.74	0.14	0.07	0.18	0.23	0.00	0.09	0.09	0.00	0.00
F12	62	Octanol		27.89	1.31	0.30	0.44	0.13	2.40	0.07	0.78	0.13	0.16	0.15
	63	5-Methyl- 2-furancarboxaldehyde		28.22	1.15	0.05	0.09	0.04	0.06	0.05	0.05	0.00	0.00	0.08
	64	3-Methyl propanoic acid		28.69	0.99	0.17	0.08	0.24	0.16	0.23	0.13	0.18	0.23	0.12
F2	65	3-Methyl-2-cyclohexen-1-one		29.22	1.48	0.14	0.08	0.25	0.11	0.21	0.13	0.24	0.20	0.11
F2	66	2,3-Butanediol		29.55	0.82	0.31	0.49	0.42	0.19	0.63	0.23	0.10	0.65	0.13
	67	Dihydro-2(3H)-furanone		30.82	1.11	0.32	0.40	0.70	0.39	0.66	0.53	0.50	0.60	0.49
	68	2-Phenylacetaldehyde*	honey-like	31.35	1.23	0.11	0.03	0.02	0.07	0.05	0.05	0.01	0.02	0.02
	69	Butanoic acid		31.42	0.82	0.12	0.02	0.11	0.26	0.11	0.18	0.02	0.11	0.08
	70	(<i>E</i>)-2-Decenal*	fatty	31.49	2.18	0.03	0.01	0.00	0.00	0.00	0.14	0.00	0.00	0.00
	71	2-Furanmethanol		32.55	0.86	0.62	0.84	1.60	0.41	1.03	0.31	0.94	1.17	0.87
	72	2- and 3- Methyl butanoic acid*	sweaty	33.09	1.07	0.03	0.02	0.04	0.01	0.04	0.07	0.04	0.03	0.02
F15	73	Pentanoic acid		36.29	0.86	0.26	1.07	0.11	1.12	1.20	0.73	0.05	0.18	1.07
	74	(<i>E,E</i>)-2,4-decadienal*	deep-fried	36.95	1.89	0.00	0.51	0.00	0.10	0.10	0.45	0.00	0.31	0.10
F4	75	Hexanoic acid		40.89	0.94	1.08	15.98	0.43	6.55	2.74	11.77	0.19	4.89	0.35
	76	2-Phenylethanol*	honey-like	43.02	1.19	0.18	0.40	0.12	0.20	0.46	0.48	0.07	0.48	0.23
F16	77	3-Acetylpyrrole		45.49	1.03	0.24	0.13	0.35	0.00	0.25	0.18	0.22	0.41	0.17
	78	1H-pyrrole-2-carboxaldehyde		47.49	1.03	0.20	0.11	0.39	0.19	0.25	0.12	0.18	0.33	1.06
	79	4-hydroxy-2,5-dimethyl-3(2H)-furanone *	sweet	48.15	1.02	0.63	0.27	0.51	0.76	1.20	0.37	0.31	0.69	0.25